

Vanderci de Andrade Aguilera
Fabiane Cristina Altino
Conceição de Maria de Araújo Ramos

■ ORGANIZADORAS

ESTUDOS DIALETAIS BRASILEIROS E EUROPEUS:

uma homenagem a João Saramago

Volume 1

 **editora
UFMS**

João Saramago é um açoriano militante e um corvino dedicado, com uma permanente participação nos mais diversos aspetos culturais do seu arquipélago e da sua ilha. Publicou cerca de duas dezenas de artigos ou capítulos de livro sobre os Açores, entre os quais uma mão-cheia sobre o Corvo. Realizou o primeiro estudo dialetométrico aplicado a materiais do arquipélago. O ponto principal na produção científica de João Saramago são, pois, os estudos açorianos, sob várias perspetivas, mas existem outros dois tão ou mais importantes: a atualização metodológica da dialetologia do português, por um lado, e a *expertise* em geografia linguística (particularmente atlas linguísticos), por outro.

Fernando Brissos



ESTUDOS DIALETAIS BRASILEIROS E EUROPEUS:

*uma homenagem a
João Saramago*

Volume 1



**UNIVERSIDADE FEDERAL
DE MATO GROSSO DO SUL**

Reitor

Marcelo Augusto Santos Turine

Vice-Reitora

Camila Celeste Brandão Ferreira Ítavo

Obra aprovada pelo

CONSELHO EDITORIAL DA UFMS

Resolução nº 152-COED/AGECOM/UFMS, de 24 de outubro de 2022.

CONSELHO EDITORIAL

Rose Mara Pinheiro (presidente)

Adriane Angélica Farias Santos Lopes de Queiroz

Andrés Batista Cheung

Alessandra Regina Borgo

Delasnieve Miranda Daspert de Souza

Elizabeth Aparecida Marques

Maria Lígia Rodrigues Macedo

William Teixeira

Dados Internacionais de Catalogação na Publicação (CIP)
(Coordenadoria de Bibliotecas – UFMS, Campo Grande, MS, Brasil)

Estudos dialetais brasileiros e europeus : uma homenagem a João Saramago : volume 1 / organizadoras, Vanderci de Andrade Aguilera, Fabiane Cristina Altino, Conceição de Maria Araújo Ramos. -- Campo Grande, MS : Ed. UFMS, 2022.
168 p. : il. color. ; 21 cm.

Volume 1: Estudos fonéticos e dialetométricos
ISBN 978-65-89995-13-5

1. Linguística – Estudo e ensino. 2. Diatologia. 3. Linguística - Atlas. I. Aguilera, Vanderci de Andrade. II. Altino, Fabiane Cristina. III. Ramos, Conceição de Maria Araújo.

CDD (23) 410.7

Bibliotecária responsável: Tânia Regina de Brito – CRB 1/2.395

Vanderci de Andrade Aguilera
Fabiane Cristina Altino
Conceição de Maria de Araújo Ramos
ORGANIZADORAS

ESTUDOS DIALETAIS BRASILEIROS E EUROPEUS:

*uma homenagem a
João Saramago*

Volume 1

Campo Grande
2022



Projeto Gráfico, Editoração Eletrônica
Secretaria da Editora UFMS

Impressão e Acabamento
Polimpessos Serviços Gráficos LTDA

A revisão linguística e ortográfica
é de responsabilidade dos autores e das organizadoras

ESTUDOS DIALETAIS BRASILEIROS E EUROPEUS: UMA HOMENAGEM A JOÃO SARAMAGO
VOLUME 1 – ESTUDOS FONÉTICOS E DIALETOMÉTRICOS

Organizadoras

Vanderci de Andrade Aguilera • Fabiane Cristina Altino • Conceição de Maria de Araújo Ramos

Apresentação

Felício Wessling Margotti

Autores

*Fernando Brissos • Hans Goebel • Jacyra Andrade Mota • Maria do Carmo Sá Teles de Araújo Rolo •
Maria Luiza de Carvalho Cruz-Cardoso • Silvia Figueiredo Brandão*

Direitos exclusivos
para esta edição



Secretaria da Editora UFMS
Av. Costa e Silva, s/nº | Bairro Universitário
Campo Grande - MS, 79070-900
Fone: (67) 3345-7203
e-mail: sedit.agecom@ufms.br

Editora associada à



Associação Brasileira
das Editoras Universitárias

ISBN: 978-65-89995-13-5
Depósito Legal na Biblioteca Nacional
Impresso no Brasil

PREFÁCIO

João António das Pedras Saramago nasceu em 1952 no município do Corvo, uma minúscula ilha dos Açores, situada a cerca de dois mil km da capital portuguesa. Do alto de suas montanhas rochosas, o horizonte desse lugar tem somente as águas do Atlântico a oferecer. Quis o destino que os sonhos desse ilhéu corvino não ficassem confinados a um território restrito a 17,2 km². Concluídos os primeiros anos de estudo, abalou-se para terras continentais onde, em 1976, licenciou-se em Filologia Românica pela Faculdade de Letras da Universidade de Lisboa, instituição na qual passou a trabalhar como pesquisador e professor. Dois anos depois, para fins de progressão na carreira, concluiu seu doutoramento com a tese “A ilha do Corvo – alguns aspectos linguísticos”, aprovada com distinção e louvor. A trilha estava então traçada, e a Dialectologia estabeleceu-se definitivamente na vida de João Saramago, seja na pesquisa, seja na docência. Vinculado ao Centro de Linguística da Universidade de Lisboa (CLUL), é atualmente a principal referência nos estudos dialetológicos e geolinguísticos de Portugal.

Brilhante nos trabalhos e nos estudos científicos que faz, é, com certeza, o corvino mais instruído e mais culto de que se tem notícia. Simples e humilde, visita com frequência a sua terra natal, onde é recebido e benquisto por seus familiares e conterrâneos. Curioso e sedento de conhecimentos etnolinguísticos, gosta de ter

contato com as pessoas simples, conversar com elas e saber das coisas e fazeres do dia a dia, na cidade e no campo, nas lojas e mercados diversos, nos restaurantes e botecos, na roça, nas indústrias familiares, nos ranchos dos pescadores, enfim com toda gente, de onde busca informações e elementos para suas pesquisas sobre as variedades dialetais, os diferentes modos de falar a língua portuguesa, sua paixão permanente.

Entre as muitas atividades, é o atual coordenador do *Atlas Linguístico-Etnográfico de Portugal e da Galiza*. O ALEPG, iniciado em 1970 por uma equipe dirigida por Luís F. Lindley Cintra, constitui-se de 212 pontos, assim distribuídos: 176 no continente português; 12 na Espanha, na área fronteiriça; sete no Arquipélago da Madeira; e 17 no Arquipélago dos Açores. Como parte desse projeto mais amplo, tem-se o *Atlas Linguístico-Etnográfico dos Açores* (ALEAç), com participação de João Saramago desde 1994. Embora não tenha havido a pretensão de abordar de um modo amplo as realidades linguísticas e etnográficas dos Açores, o ALEAç contempla as especificidades linguísticas nas nove ilhas do arquipélago. Com base em dados recolhidos pelo ALEPG, parte das cartas semântico-lexicais do ALEAç foram elaboradas por João Saramago, sobretudo as relacionadas à criação de gado, à suinicultura, à moagem de cereais, às plantas, à agricultura, às abelhas e à caça, além de algumas outras cartas sobre aspectos morfológicos.

João Saramago é também membro da equipe de dialetólogos e diretor-adjunto do *Atlas Linguistique Roman* (ALiR), que se iniciou em 1987, por iniciativa de Gaston Tuillon e Michel Contini, e tem a sua sede no Centre de Dialectologie da Université Stendhal de Grenoble, na França. Esse projeto está estruturado em 10 comités (português, galego, espanhol, catalão, francês, valão, suíço, italiano, romeno e moldavo) que integram especialistas de 31 universidades ou centros de investigação dos vários países participantes. Portugal

está representado por uma rede de 110 pontos de inquérito, dos quais 10 são no Arquipélago dos Açores e quatro no Arquipélago da Madeira. Os dados dialetais do português foram coletados e sistematizados pelo comitê português vinculado ao Grupo de Estudos de Dialectologia do Centro de Linguística da Universidade de Lisboa, constituído por Luísa Segura (coordenadora), Gabriela Vitorino, Manuela Barros Ferreira, João Saramago, Maria Lobo, Ernestina Carrilho e Celeste Augusto.

Há ainda outros atlas linguísticos e projetos de investigação dialetal que contam com a participação do homenageado neste livro, entre os quais citam-se o *Atlas Linguarum Europae* (ALE), o *Atlas Linguístico do Litoral Português* (ALLP) e projeto *Tesouro do Léxico Patrimonial Galego e Português*, de cujo Comitê Português João António das Pedras Saramago é o coordenador.

Entre muitas outras atividades relacionadas à área de conhecimento em que atua, o homenageado neste livro, organizado por Vanderci de Andrade Aguilera (UEL), Fabiane Cristina Altino (UEL) e Conceição de Maria de Araújo Ramos (UFMA), colaborou com a IBM Portuguesa na “Descrição exaustiva de formas pronominais clíticas hifenadas adequadas a cerca de 5500 verbos”, na “Análise lexicográfica de um ficheiro de vocabulário” com vista à sua correção, nomeadamente através da introdução e suspensão de entradas lexicais e sua classificação gramatical, e na elaboração de um dicionário de sinônimos. Tem atuado em várias Universidades, portuguesas e estrangeiras, em cursos de dialectologia portuguesa; na orientação de trabalhos académicos de alunos de diferentes níveis; e em bancas de dissertações de mestrado e teses de doutorado. É autor ou coautor de vasta obra científica distribuída por livros, artigos e comunicações em eventos científicos em Portugal e em outros países, bem como de alguns trabalhos que versam sobre os Açores. Colaborou na “Enciclopédia Açoriana”, que presentemente pode ser

consultada na Internet, e possui uma obra vastíssima de trabalhos sobre temas diversos.

A descrição de toda a sua obra literária, científica e acadêmica – em livros, artigos, comunicações, participações em obras e traduções e em eventos, além da formação de novos dialetólogos e geolinguistas – revelaria com profundidade e justiça a grandeza do trabalho e a competência de João Saramago no que diz respeito aos estudos dialetais, tanto os de interesse mais amplo sobre a língua portuguesa falada no território continental português quanto os de interesse mais restrito às ilhas do Arquipélago dos Açores, ao Brasil e a outros países.

Os dois volumes deste livro, que tratam dos estudos dialetais brasileiros e portugueses, com as contribuições de mais de duas dezenas de especialistas nessa área de conhecimento, são uma justa e merecida homenagem ao dialetólogo e geolinguista João Saramago, por ocasião do *VI Congresso Internacional de Dialetologia e Sociolinguística* – VI CIDS, realizado em 2022, na cidade de Campo Grande – MS, Brasil.

SUMÁRIO

PREFÁCIO5

APRESENTAÇÃO 11

Felício Wessling Margotti

Róticos na Ilha do Corvo com Base nos Dados do ALEAç:
uma singela homenagem ao dialetólogo e
amigo corvino.....21

Silvia Figueiredo Brandão

Acústica Corvina.....49

Fernando Brissos

Apócope das Vogais Altas [I] e [Ü]:
pegadas açorianas no português falado no Brasil77

Maria do Carmo Sá Teles de Araújo Rolo

Jacyra Andrade Mota

Un Nouveau Rejeton de L'«École Dialectométrique de
Salzbourg»: brève présentation du mode «Beta» de la
dialectométrie de Salzbourg 101

Hans Goebel

Um olhar historiográfico sobre a Geolinguística e
seus reflexos na elaboração do
Atlas Linguístico do Amazonas..... 141

Maria Luiza de Carvalho Cruz-Cardoso

Sobre os Autores 167

APRESENTAÇÃO

Por ocasião do *VI Congresso Internacional de Dialetologia e Sociolinguística*, seus organizadores decidiram homenagear “dois ilustres pesquisadores pela importante contribuição por eles prestada no campo dos estudos dialetológicos e sociolinguísticos”: o português *João António das Pedras Saramago* e a brasileira *Dinah Maria Isensee Callou*. Como parte das merecidas homenagens, as pesquisadoras Vanderci de Andrade Aguilera, Fabiane Cristina Altino e Conceição de Maria de Araújo Ramos, membros do Comitê Nacional do *Projeto Atlas Linguístico do Brasil – ALiB*, organizaram o livro *Estudos Dialetais Brasileiros e Europeus: uma homenagem a João Saramago*.

A presente obra, apresentada em dois volumes, está assim organizada: o volume 1 com cinco capítulos e o volume 2 com onze capítulos escritos por pesquisadores brasileiros e estrangeiros, os quais revelam ao leitor aspectos linguísticos e culturais, percepções, crenças e atitudes das comunidades investigadas, a que se acrescentam questões teórico-metodológicas para a realização de estudos geossociolinguísticos. Cada capítulo reproduz recortes de estudos dialetológicos daqui e de além-mar sobre o português, o galego, o italiano, o sardo, incluindo contatos linguísticos, além de um capítulo sobre software aplicável a estudos dialetométricos geolinguísticos. Trata-se de uma leitura enriquecedora, cheia de revelações em relação às variedades linguísticas e aos processos de

variação e mudança linguística, especialmente no nível fonético e lexical, evidenciando a história e a cultura das línguas estudadas.

O primeiro volume, com cinco capítulos, contempla estudos fonéticos, dialetométricos e uma visão historiográfica da Geolinguística. Em *Róticos na ilha do Corvo com base nos dados do ALEAç: uma singela homenagem ao dialetólogo e amigo corvino*, Brandão, com base no *Atlas Linguístico-Etnográfico dos Açores* analisa as variantes de /R/. Os dados revelam que no Corvo, em contraste com o que ocorre nas demais ilhas dos Açores, predomina a variante [r̄] (a vibrante alveolar) em contexto pré-vocálico, enquanto, em coda final, [r] (o tepe) concorre com o cancelamento.

No texto *Acústica corvina*, Brissos retoma diversos estudos de Saramago que tratam da acústica do português falado nas ilhas do Arquipélago dos Açores, e conclui que: a) o dialeto do Corvo tem uma relação de grande afinidade com o grupo de dialetos centro-meridionais de Portugal, o que se demonstra por meio das vogais /ɛ ɔ o u/ do português padrão, que no Corvo se concretizam respectivamente em [æ ɔ o u-]; b) essa afinidade, porém, não é total, uma vez que o Corvo não deixa de estabelecer correspondências, embora pontualmente e em fenômenos de menor importância, com os dialetos portugueses setentrionais, o que se verifica quanto ao espectro das vogais /e a/, respectivamente [ei a-] no Corvo.

Uma possível relação entre o português brasileiro e o português açoriano é o tema do texto *Apócope das vogais altas [i] e [u]: pegadas açorianas no português falado no Brasil*, escrito por Maria do Carmo Sá Teles de Araújo Rolo e Jacyra de Andrade Mota, com vista a investigar a apócope das vogais átonas finais observadas nos atlas brasileiros e em outras pesquisas, relacionando essas ocorrências à presença açoriana no processo de povoamento do Brasil. Os dados apresentados no artigo reforçam a hipótese de que o fe-

nômeno da apócope observado no português falado no Brasil está associado ao que se observa em Portugal e que foi trazido pelos colonizadores lusos, especialmente açorianos.

Hans Goebel é autor do capítulo *Un nouveau rejeton de l'«école dialectométrique de Salzbourg»: breve presentation du mode «beta» de la Dialectometrie de Salzburg*, por meio do qual apresenta – por meio de um aparato gráfico de três figuras (em preto e branco) e quatorze mapas (em cores) – habilidades analíticas e de visualização de um novo módulo do software dialetométrico VDM (“Visual DialectoMetry”) que foi desenvolvido nos últimos anos no âmbito da “Salzburg Dialectometric School”. Ao final, conclui que novo componente do software VDM se encaixa perfeitamente na lógica metódica de EDMS que, por definição, é múltiplo, demonstrando sua utilidade para a Geolinguística, além de outras áreas de conhecimento.

No capítulo intitulado *Um olhar historiográfico sobre a Geolinguística e seus reflexos na elaboração do Atlas Linguístico do Amazonas*, Maria Luiza de Carvalho Cruz-Cardoso, da UFAM, apresenta uma pesquisa historiográfica sobre o desenvolvimento da Geolinguística no mundo, considerando a classificação dos atlas já publicados e disponíveis para esse tipo de pesquisa no Centro de Linguística da Universidade de Lisboa, sob orientação de João Saramago. O estudo contempla atlas linguísticos representativos dos quatro diferentes tipos apontados por Mario Alinei: (i) regional, (ii) nacional, (iii) de grupo de línguas e (iv) continental. A autora optou, ainda, por incluir uma breve visão do que se realiza numa parte das Américas, e por comentar um atlas de cunho temático para fins de exemplificação, tendo sido focalizados, de uma forma abrangente: (i) o Atlas Linguístico e Etnográfico de Cantábria (ALECant); (ii) o Atlas Linguístico Italiano (ALI) e o Atlas Linguístico do México; (iii) o Atlas Linguistique Roman (ALiR); (iv)

o Atlas Linguarum Europae (ALE). Por último, o Atlas Linguístico do Litoral Português (ALLP). Em síntese, trata dos reflexos desses atlas na elaboração de sua tese, o *Atlas Linguístico do Amazonas*.

O segundo volume, com onze capítulos, é dedicado aos estudos lexicais. Em *Um estudo sobre as variantes lexicais para lanterna registradas pelo ALiB*, as autoras Fabiane Cristina Altino e Vanderci de Andrade Aguilera, com base em dados obtidos pelo projeto *Atlas Linguístico do Brasil – ALiB* nas capitais e nas cinco regiões brasileiras, revelam que, além da predominância da forma *lanterna*, ainda subsistem variantes como *lâmpada*, *farolete*, *pilha*, *flashlight*, *foco*, entre outras. Considerando a distribuição diatópica, a Região Sul mostrou-se mais influente na produção de variantes populares com duas áreas bem delineadas (*farolete* no norte do Paraná e *foco* no norte do Rio Grande do Sul, oeste de Santa Catarina e do Paraná). Merece destaque a distribuição areal de *farolete* em São Paulo e de *lâmpada* na Bahia.

A pesquisadora Rosario Álvarez, do Instituto da Língua Galega da Universidade de Santiago de Compostela, é a autora do texto *Xoaniña, voa voa. Os nomes da coccinella en galego*, com o qual discorre sobre um amplo leque de formas galegas de nomear um pequeno inseto, colorido e brilhante, identificado de modo geral como *joaniña*, cujo nome científico é *Coccinella septempunctata*. Demonstra que a gama de palavras em galego para nomear esse inseto é extensa. Além da forma *coccinella* cultivada, inclui nove formas tradicionais, e estas são apenas algumas das mais comuns: *barrosiña*, *maruxiña*, *papasol*, *papoia*, *reirrei*, *voaniña*, *voíña*, *xoana* e *xoaniña*.

Na sequência deparamo-nos com o texto de Michel Contini, em coautoria com Elisabetta Carpitelli, intitulado *Les designations de la vrillette dans les domaines italo-roman et sarde*, que trata das diferentes formas de nomear um *besouro* em italiano e em sardo,

o qual é geralmente descrito como um pequeno inseto roedor de madeira, mas inclui uma variedade de pequenos insetos fitófagos. Com base em dados do *Atlas Linguístico Italiano – ALI* e do *Atlas Linguístico da Sardenha – AIS*, os autores revelam que o estudo das designações dialetais dos insetos é complexo. O exame dos dados do ALI para *besouro* e do AIS para *mariposa* confirmou amplamente as dificuldades encontradas por outros pesquisadores que abordaram o léxico dialetal relacionado ao mundo animal.

As variantes lexicais para se referir ao inseto que frequenta espelhos d'água, brejos e outras áreas alagadiças é tema do capítulo *O fantástico voo da libélula: um estudo da motivação na criação lexical em designações registradas no ALEAL, ALiB, ALiR E ALEPGI*, texto escrito por Maranúbia Pereira Barbosa Doiron. Embora a unidade lexical padrão para tal inseto seja *libélula*, nos quatro atlas linguísticos investigados pela autora consta um número bastante elevado de variantes, algumas de uso mais geral e outras associadas a certas regiões, tanto no Brasil quanto na área continental de Portugal e nas ilhas do Arquipélago dos Açores. De acordo com os atlas incluídos na análise, o referido inseto contempla relação extensa de lexias, tanto no português brasileiro quanto no português europeu, incluindo o do Açores, e de igual modo em outras línguas românicas.

Com o título de *Os nomes da vacaloura en galego*, Manuel González González, da Universidade de Santiago de Compostela, explora as motivações e crenças relacionadas aos nomes para se referir ao maior escaravelho da Europa. Os machos têm grandes mandíbulas, muito maiores que as das fêmeas, e que servirão como meio de luta contra machos rivais. Seus grandes chifres produziram um profundo impacto na mentalidade popular, inseto a que se atribui poderes quase mágicos, tanto para o mal quanto para o bem. Recebe inúmeros nomes em galego, mas os principais estão relacionados às suas mandíbulas em forma de chifres, que são a base

motivacional dos nomes como *cornuda*, *corneteira*, *escornaboís*, *escornavacas*, *escornacabras*, *vacaloura*, *vacanegra*, *cabraloura*... Outras motivações denominativas importantes têm relação com a dureza da capa que envolve seu corpo, que está na base dos nomes *carroucha*, *carroucho* e *cascudo*, e seus hábitos e comportamento explicam nomes como *furão* ou *rumballón*.

As pesquisadoras Silvana Soares Costa Ribeiro e Aparecida Negri Isquierdo são autoras do artigo *Denominações para “corre-cutia”, “lenço-atrás” e “chicotinho-queimado” na área do falar sulista (nascen-tes, 1953) – dados do ALiB*, com o qual buscam (i) analisar dados referentes à brincadeira infantil conhecida como “corre-cutia”, “lenço-atrás”, “chicotinho-queimado”, documentados pelo Projeto ALiB na área geográfica selecionada (Sul e Centro-Oeste); (ii) reconhecer regiões geográficas marcadas por traços linguísticos peculiares nessas regiões e apontar tendências pontuais em nível nacional, no que se refere às denominações da brincadeira em foco; e (iii) apontar o papel e a importância do léxico para a descrição e delimitação de área dialetais.

O uso das variantes lexicais *abóbora* (de origem portuguesa) e *jerimum* (de origem indígena) e o significado atribuído a esses nomes constituem o fundamento do artigo *O que se vende nas feiras e supermercados paraenses: abóbora ou jerimum?*, elaborado por Marilucia Barros de Oliveira e Luan Costa dos Santos, da Universidade Federal do Pará. Em sites de busca, consta que as formas *jerimum* e *abóbora* convergem para o mesmo nome científico, do gênero *cucurbita*, diferenciando-se apenas pelos tipos e formas. Quando ao uso, por se tratar de um produto que é levado do sul e sudeste para o norte do Brasil, a pesquisa realizada pelos autores revela que a forma *abóbora* (predominante nessas regiões produtoras) está substituindo a forma *jerimum* (antes predominante no norte).

Tendo como referência o fluxo migratório dos Açores para o Maranhão, Conceição de Maria de Araujo Ramos, José de Ribamar

Mendes Bezerra e Theciana Silva Silveira decidiram cotejar, no domínio do léxico os dados recolhidos para o *Atlas Linguístico do Maranhão* (ALiMA) com dados do *Atlas Linguístico- Etnográfico dos Açores* (ALEAç), as designações para os seguintes conceitos: *cria da vaca, cria da ovelha, caminho do gado/no pasto, corno, boi sem chifre, cabra/vaca sem chifre, glândula mamária* (animais) e *cauda* (animais). A análise da base de dados dos dois atlas demonstrou que: “(i) das 14 formas compartilhadas, apenas duas, *bezerro* e *mocho*, ocorreram em toda a rede de pontos de cada um dos atlas; (ii) *caminho do gado/no pasto* foi o conceito que apresentou o maior número de formas denominativas, tanto no ALiMA (17) como no ALEAç (10); (iii) o ALiMA apresentou maior polimorfismo, com exceção da questão concernente ao conceito *corno*; e (iv) além do polimorfismo dialetal, ocorreram casos de polimorfismo individual, sobretudo em relação aos conceitos *caminho do gado/no pasto, boi sem chifre* e *cabra/vaca sem chifre*”.

A variação diatópica do item lexical *cigarro de palha*, que constitui a questão 145 do Questionário Semântico-Lexical (QSL) do projeto ALiB, foi investigada por Abdelhak Razky (UFPA/UnB/CNPq), Eliane Oliveira da Costa (SEDUC) e Regis José da Cunha Guedes (UFRA) e relatada no capítulo *Agrupamentos lexicais do item cigarro de palha nas não capitais do Atlas Linguístico do Brasil (ALiB)*. O conjunto das variantes lexicais cartografadas: *porronca, tabaco* e *cigarro de palha* (Carta L01), *fumo de corda* e *boró* (Carta L02) e *pé duro, pacaia* e *picão* (Carta L03), aponta para quatro dos cinco tipos de agrupamentos estruturantes do contínuo dialetal de agrupamentos linguísticos: Carta L01 - macroagrupamento (1), mesoagrupamento (2); Carta L02 - microagrupamento (1), mesoagrupamento (1); Carta L03 - microagrupamento (1), nanoagrupamento (2).

A descrição parcial dos dados do ALiB obtidos por meio da questão 167 do QSL – “a brincadeira em que as crianças riscam uma

figura no chão, formada por quadrados numerados, jogam uma pedrinha (mímica) e vão pulando com uma perna só” – é o que propõe Valter Pereira Romano no artigo *Contribuições do projeto ALiB para a caracterização de áreas lexicais: o caso da sapata e amarelinha na Região Sul do Brasil*. Foram documentados 160 registros divididos entre 26 formas, com número de ocorrências indicado entre parêntese: *amarelinha* (87), *sapata* (29), *marelinha* (11), *pula-pula* (6), *amarelinho* (4) e *caracol* (3). Com ocorrências únicas, listam-se em ordem alfabética: *amarela*, *amerelinha*, *brincar de queimar*, *caia*, *calha*, *estrelinha*, *jogar as pedrinhas*, *joguinho*, *macaca*, *marelinho*, *pula boneco*, *pula sapato*, *quadra*, *quadrado*, *quadrinho*, *sapato*, *sete pedra*, *tabuada*, *três marias* e *triângulo*. O item lexical *amarelinha* é predominante em toda faixa norte do Paraná, região colonizada principalmente por paulistas. À medida que se adentra ao centro-sul paranaense, por um corredor central do estado, a incidência da variante *amarelinha* vai diminuindo, obtendo baixa produtividade, principalmente, no território gaúcho (20%), espaço em que a forma predominante é *sapata*.

E a presente obra termina com o artigo *Palabras e comidas no Atlas Lingüístico de la Península Ibérica: a parva e o almoço*, de autoria de Xulio Sousa do Instituto da Língua Galega da Universidade de Santiago de Compostela. Nele, o autor explica a mudança de significado da lexias *parva* e *almorzo*, associando-a à mudança de hábitos. De acordo com os materiais do *Atlas Lingüístico da Península Ibérica – ALPI*, cujos primeiros trabalhos de campo foram realizados na década de 1930, a primeira refeição do dia recebia a denominação de *parva* [ˈparβa] na maioria das localidades. E de acordo com o mesmo atlas linguístico, *almorzo* aparece como refeição após a *parva* ou como refeição primeira ou única refeição da manhã. Todavia, por influência da industrialização e urbanização da população, aconteceu a unificação do tempo e a normalização dos horários das refeições. Por conta disso, *almorzo* foi substituído por café da manhã, deixando de

ser uma refeição rudimentar de camponeses para ser uma refeição universal para todas as classes sociais.

Para os amantes da Dialetoлогия e da Sociolinguística, cujo objeto de investigação são as línguas nos diferentes níveis estruturais e lexicais, a variação e a mudança linguística, associadas ao espaço geográfico e a diversas outras dimensões extralinguísticas, com foco nas variedades dialetais, este livro oferece rica e ampla gama de recortes nessa área. Embora nos artigos aqui publicados predominem os estudos no âmbito da lusofonia portuguesa e brasileira, os organizadores não se furtaram de aceitar e incluir estudos sobre outras línguas do mundo românico, o que condiz perfeitamente com o perfil do homenageado. E para concluir essa breve apresentação, tomo a liberdade de retomar uma passagem do artigo de Fernando Brissos, a qual expressa de forma sintética quem é João Saramago e o que ele representa para a Dialetoлогия e a Geolinguística:

Desde a década de 1970 que Saramago pertence aos grupos de trabalho de todos os atlas linguísticos realizados em Portugal, sejam projetos internacionais como o ALE – Atlas Linguarum Europae (CARRILHO, [s.d.]) e o ALiR – Atlas Linguistique Roman (SEGURA, [s.d.]), seja o projeto de atlas linguístico de Portugal, i.e. o ALEPG – Atlas Linguístico-Etnográfico de Portugal e da Galiza (SARAMAGO, 2006); na prática, o autor integra as equipas desses atlas praticamente desde o início dos trabalhos. Essa extensa experiência, a qualidade da produção científica daí resultante e as raras características humanas de Saramago tornam-no uma figura incontornável na geografia linguística portuguesa e românica, sendo colaborador ou consultor de um número significativo de projetos desses domínios, como, nos 25 exemplos mais atuais, a edição digital do ALPI – Atlas Lingüístico de la Península Ibérica (GARCÍA MOUTON, 2009-) e o projeto Dialetopédia (BRISOS, 2022).

Boa leitura!

Felício Wessling Margotti

Universidade Federal de Santa Catarina

UN NOUVEAU REJETON DE L'«ÉCOLE DIALECTOMETRIQUE DE SALZBOURG»: BRÈVE PRÉSENTATION DU MODE «BETA» DE LA DIALECTOMETRIE DE SALZBOURG

Hans Goebel

Salzburg

Hans.Goebel@plus.ac.at

INTRODUCTION

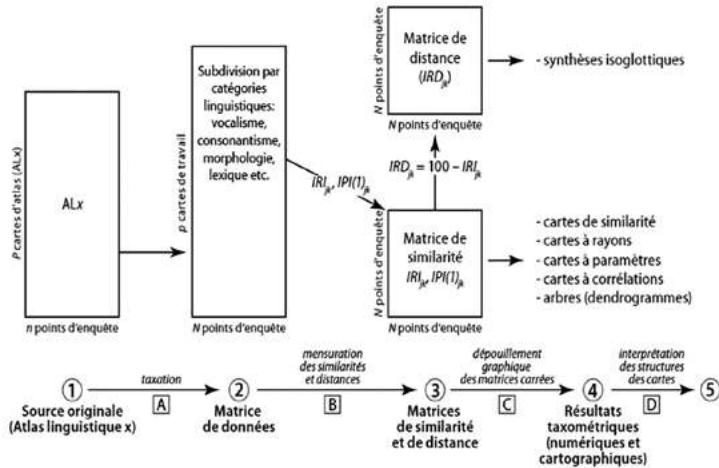
Le but de cet article est de présenter – moyennant un appareil graphique de trois figures (en noir-et-blanc) et quatorze cartes (en couleurs) – les capacités analytiques et visualisatrices d'un nouveau module du logiciel dialectométrique VDM¹ («Visual DialectoMetry») qui avait été développé ces dernières années dans le cadre de l'«École Dialectométrique de Salzburg» (EDMS)².

¹ VDM = Visual DialectoMetry (logiciel dialectométrique créé dans les années 1997-2000 par Edgar Haimerl pour les besoins de l'EDMS).

² Pour une introduction aux méthodes de l'EDMS il y a plusieurs «filères» linguistiques: voir à ce sujet mes travaux suivants: en *allemand*: 1984, en *français*: 1981, 2000 et 2002; en *anglais*: 2009b, 2018b. Comme, de nos jours, le qualificatif de *dialectométrie* (ou de *dialectometry*, *Dialektometrie* etc.) est devenu l'apanage de plusieurs méthodes

Ce nouveau module constitue un élargissement de la chaîne méthodique traditionnelle de l'EDMS telle qu'elle ressort de la Figure 1.

Figure 1: Alignement standard (mode-DM «alpha») des démarches méthodiques de l'«École Dialectométrique de Salzbourg»



Source: Werner Goebel, Vienne.

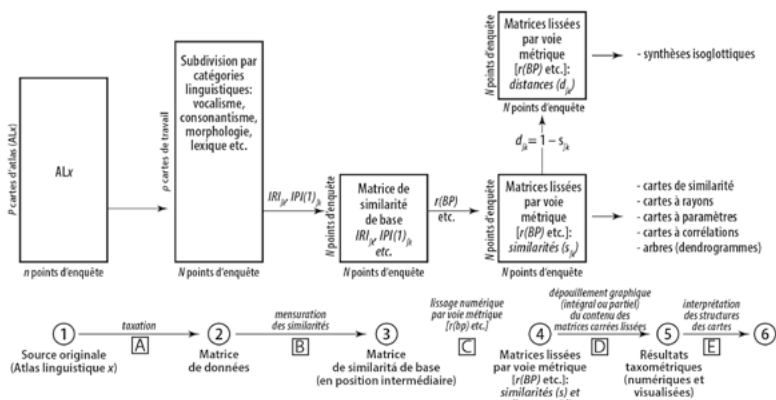
Dans sa configuration classique, cette chaîne méthodique est composée de deux éléments majeurs: de la section *empirique* (qui s'étend de l'*atlas linguistique* [AL] soumis à une analyse-DM [A + 1], jusqu'à la *matrice de données* [A + 2], créée par un processus de *typisation* (ou de "taxation") des données originales de l'AL en question), et de la section *taxométrique* à proprement parler (qui commence par la mesure de la similarité entre les N vecteurs d'attributs de la matrice de données [B + 3] et s'achève par l'accomplissement [C + 4] et l'interprétation de six opérations dialectométriques [D + 5], toutes dûment visualisées.

quantitatives dont le fond commun est en effet très réduit, la mise en vedette de *Salzbourg* pour la caractérisation exclusive des méthodes présentées ici, est très utile, voire indispensable.

Dans cette configuration, la partie *taxométrique* forme un bloc unique centré sur la matrice de *similarité* (*sim*; calculée à l'aide d'un des nombreux indices de similarité de l'EDMS) et sur la matrice de *distance* (*dist*), qui en est dérivée par l'application d'une formule très simple: $dist = 1$ (ou: 100) $- sim$.

Les scores de *similarité* ou de *distance* répertoriés dans les deux matrices émanent donc, à l'instar d'un signal électrique, d'une seule "source" mathématico-statistique avant leur mise en carte circonstanciée.

Figure 2 : Alignement élargi (mode-DM « bêta ») des démarches méthodiques de l' « École Dialectométrique de Salzbourg »



Source: Werner Goebel, Vienne.

Au cours de calculs-DM effectués dans le cadre de recherches scriptométriques³ nous avons découvert l'utilité de l'application supplémentaire d'indices de *corrélation* à proprement parler, parmi

³ Par "scriptométrie" nous entendons l'application des méthodes-DM de l'EDMS aux données (et problèmes) de la scriptologie (médiévale): voir, à titre d'exemple, Dees 1980, nos contributions de 2008, 2016a et 2017 (les deux dernières contributions en coopération avec P. Smečka) et Videsott 2009.

lesquels l'indice $r(\text{BP})^4$, proposé successivement par le physicien français Auguste Bravais (1811-1863) et le mathématicien anglais Karl Pearson (1857-1936), est le plus universel et le mieux connu⁵.

De là est née l'idée de soumettre les scores de similarité stockés dans la matrice de similarité [de base] (voir la position 3 sur les Figures 1 et 2) à un traitement statistique supplémentaire par l'application de l'indice $r(\text{BP})^6$, et de dédoubler ainsi la partie *taxométrique* de notre chaîne-DM traditionnelle: voir la position 4 sur la Figure 2.

Pour bien distinguer les deux traitements statistiques, nous les avons baptisés "mode-DM *alpha*" (ancien) et "mode-DM *béta*" (neuf).

Évidemment, le traitement visualisateur final des scores de *similarité* (ou de *distance*) deux fois filtrés par voie statistique, reste le même.

PLAN DU PRÉSENT ARTICLE

Nous nous proposons de faire voir, à l'aide de quatorze paires de cartes – issues respectivement de l'application des modes-DM *alpha* et *béta* aux six volets taxométriques de l'EDMS – les différences heuristiques qui se creusent entre les deux méthodes.

Les données de base ont été tirées de six volumes de l'ALF qui avaient été dépouillées, peu avant le tournant du millénaire, par une équipe de jeunes romanistes salzbourgeoises⁷, parallèlement

⁴ $r(\text{BP})$ = indice de corrélation de Bravais-Pearson

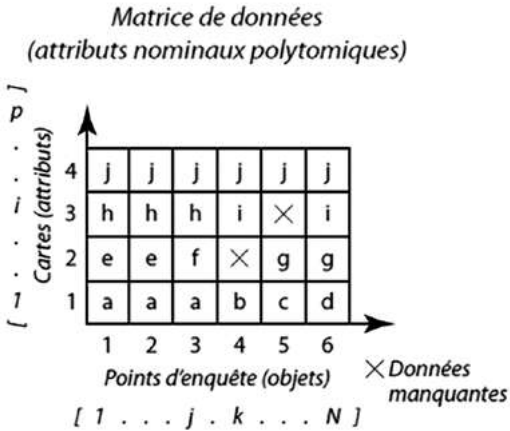
⁵ Pour la formule du $r(\text{BP})$ voir Chandon/Pinson 1981, 65 (→ la formule S4).

⁶ Les scores du $r(\text{BP})$ oscillent entre -1 et +1. Ils signalent la corrélation linéaire entre deux variables métriques.

⁷ En voici leurs noms: Barbara Aigner, Irmgard Dautermann, Hildegund Eder, Susanne Oleinek et Annette Schatzmann. Au cours des deux dernières décennies, notre gratitude

au développement du logiciel VDM par les soins de mon ami Edgar Haimerl.

Figure 3 : Schéma d'une matrice de données tirée par voie taxatoire d'un atlas linguistique



Source: Werner Goebel, Vienne.

PRÉSENTATION DU CORPUS-ALF UTILISÉ

Voir la Fig. 3.

En établissant notre matrice de données (MD⁸) de base nous avons suivi, en dernière analyse, l'exemple de Jules Gillieron à la veille de sa fameuse étude sur l'onomasiologie de l'action de *scier* dans le Sud-Est de la France (GILLIERON, 1905): comme lui nous avons dépouillé les cartes originales (CO⁹) de l'ALF en en classant les données originales par types quitte à les réunir sur une carte

envers elles n'a cessé de croître.

⁸ MD = matrice de données

⁹ CO = carte originale (d'un atlas linguistique)

muette après coup. Notre travail de typisation se distinguait par deux propriétés de celui de Gilliéron: il s'agissait de prendre en compte un nombre aussi grand que possible de CO de l'ALF, et de les analyser ("taxer") à partir de plusieurs points de vue linguistiques: non seulement de celui du *lexique* (comme le faisait jadis Gilliéron), mais aussi de celui de la *phonétique* et de la *morpho-syntaxe*. C'est ainsi que nous avons pu tirer, moyennant l'analyse taxatoire de 626 *cartes originales* (CO) de l'ALF (toutes appartenant à la série A de l'ALF) 1 681 *cartes* "de travail" (CT¹⁰) dont la bigarrure linguistique interne (*polynymie*) est très inégale. Elle va de 2 types (ou "taxats") par CT jusqu'à 91 types/taxats par CT¹¹. Chacun de ces taxats occupe d'ailleurs une aire particulière que nous appelons depuis belle lurette "aire taxatoire" (AT)¹².

Encore quatre remarques explicatives:

- Quant à la qualité intrinsèque des cartes originales (CO) dépouillées: nous avons dû écarter, pour des raisons statistiques, toutes les CO compromises par un surplus (> 5% du réseau intégral de 641 points) de données manquantes.
- Quant aux éléments (→ points) de notre réseau de recherche: nous avons ajouté, aux 638 points d'enquête originaux de l'ALF¹³, trois points factices¹⁴

¹⁰ CT = carte de travail (tirée, moyennant une analyse typisante ou "taxatoire", d'une CO d'un atlas linguistique)

¹¹ Pour une analyse du rôle de la polynymie cf. Goebel 2014.

¹² Les «cartes de travail» (CT) et les "aires taxatoires" (AT) constituent les piliers empiriques centraux de l'EDMS: cf. Goebel / Smečka 2016b.

¹³ Rappelons que le réseau de l'ALF comprend 638 *points d'enquête* et 639 *enquêtes*. Ce petit décalage s'explique par le fait que l'enquêteur Edmond Edmont a fait deux *enquêtes* au P.-ALF 284, Saint-Pol-sur-Ternoise, son village natal.

¹⁴ Voici les correspondances: *français* = P. 999 (situé au cœur de l'Île-de-France), *italien* = P. 998 (accolé à la pointe sud-orientale de la Provence), *catalan* = P. 997 (accolé à la pointe méridionale du Département des Pyrénées Orientales).

qui correspondent aux langues standard du *français*, de *l'italien* et du *catalan*. Il en résulte un total (= N) de 641 points-ALF (ou de polygones).

- Quant au support cartographique de base: une des actions révolutionnaires de Jules Gilliéron (1854-1926) était le recours systématique aux *cartes muettes* (CM¹⁵) que l'imprimerie Georges Protat; Frères (à Mâcon, Saône-et-Loire), chargée de l'impression de l'ALF, avait mis à la disposition des romanistes peu après 1902¹⁶. Il a pratiqué ainsi une espèce de "navette typisante" entre les données originales de l'ALF et leurs métabolismes cartographiques simplifiés. C'est que la géographie linguistique à proprement parler est née par la mise en carte systématique de la variabilité des faits linguistiques sur des vecteurs en papier. Depuis le début du XX^e siècle, le maniement courant de CM est donc de rigueur au sein des études romanes¹⁷. Dans le cadre de l'EDMS la préférence est donnée aux CM *polygonisées* qui, mis à part leur numérisation facile, ont l'avantage de créer un lien cartographique (mieux: sémiotique) logique entre la mise en carte de faits linguistiques par *plages*¹⁸ et par *lignes*¹⁹. Rappelons aussi que la découverte de ce lien est le mérite d'un germaniste allemand (Carl Haag, 1860-1946) qui en

¹⁵ CM = carte muette (pour le dépouillement "taxatoire" des CO d'un atlas linguistique).

¹⁶ C'est Karl Jaberg qui a présenté le premier, en 1906, les nouvelles CM aux romanistes du temps. Voir aussi notre analyse historique de 2018a.

¹⁷ Il y a là une grande différence par rapport aux *germanistes* (et *anglicistes*) dont les atlas linguistiques ne contiennent pas de données brutes (c.-à-d. des transcriptions originales), mais des données préalablement typisées / codées (ou: "taxées") par les éditeurs respectifs.

¹⁸ Terme technique de la cartographie "officielle": cartes "choroplèthes".

¹⁹ La cartographie les appelle cartes "isarithmiques".

avait parlé la première fois en 1898 dans sa thèse sur les dialectes de Souabe (Allemagne du Sud-Ouest).

- Quant à la structure interne de la matrice de données (MD): les superficies (et leurs pourtours) des AT²⁰ sont très changeantes et absolument “imprévisibles”. Leur coïncidence exacte ne se réalise pratiquement jamais, même quand deux AT se réfèrent à la même origine étymologique. Ce phénomène représente le pendant spatial de la fameuse incohérence des “lois” phonétiques qui a fait couler tant d’encre parmi les linguistes à partir du dernier quart du XIX^e siècle (et au-delà). À l’intérieur d’une MD, il en résulte un “désordre” curieux pour lequel nos prédécesseurs ont créé le slogan “Chaque mot a son histoire” (cf. Jaberg 1908 et Christmann 1971), et que nous appelons, depuis plusieurs années, “enchevêtrement particulier” (des AT). Toutes les AT extraites d’un atlas linguistique sont donc enchevêtrées entre elles d’une façon fort complexe ce qui n’empêche nullement leur analyse combinatoire par voie statistique: c’est là le véritable “sel” de la DM en général et de l’EDMS en particulier.

Le tableau suivant est une synopse comparative entre le contenu de la Figure 3 et notre *corpus*-ALF de base:

	Nombre des points d’enquête	Nombre des cartes de travail (CT)	Nombre des aires taxatoires (AT)
Figure 3	6	4	10
Notre <i>corpus</i> -ALF utilisé	641	1 681	28

²⁰ AT = aire taxatoire (aire de diffusion d’un “taxat” ou type linguistique)

PRÉSENTATION DE L'APPENDICE CARTOGRAPHIQUE (CARTES 1-14)

Toutes les 14 cartes multicolores de l'annexe sont bipartites: les clichés de *gauche* sont issus d'un traitement selon le mode-DM *alpha* traditionnel (voir la Fig. 1) alors que les clichés de *droite* se réfèrent au mode-DM *béta* de date récente (voir la Fig. 2): tous les clichés de droite sont donc inédits et partant nettement innovateurs²¹.

Le coloriage des cartes *choroplèthes* et *isarithmiques* de l'annexe a été réalisé à l'aide de deux algorithmes de visualisation (MINMWMAX et MEDMW) qui font partie intégrante du logiciel VDM ("Visual DialectoMetry") et de la philosophie visualisatrice de l'EDMS. Tous les deux réunissent les scores calculés – et stockés ensuite dans une matrice de *similarité* (MS²²) ou de *distance* – en intervalles numériques préalablement définis tout en leur attribuant un coloriage particulier: c'est ainsi qu'aux couleurs *froides* (= en bleu [clair et foncé] et vert) correspondent les scores situés *au-dessous* de la moyenne arithmétique, et aux couleurs *chaudes* (= en rouge, orange et jaune) les scores situés *au-dessus* du même seuil numérique. Ces intervalles disposent soit de *largeurs* numériques égales (→ MINMWMAX) soit de *nombres* égaux *d'éléments* (= points d'enquête ou polygones) (→ MEDMW).

L'agencement des couleurs mises en carte suit la logique de l'arc-en-ciel (ou du spectre solaire).

²¹ Quant à nos travaux-DM antérieurs – réalisés, en mode-DM *alpha*, avec les données de l'ALF – nous ne renvoyons ici qu'à nos contributions, toutes rédigées en langue française, de 2000, 2002, 2003 et 2009a. Les lecteurs intéressés y trouveront un grand nombre d'exemples, tous dûment présentés, expliqués et interprétés. Pour une bibliographie complète de nos travaux-DM, toujours tenue à jour, voir notre site personnel: https://www.sbg.ac.at/rom/people/prof/goebl/dm_publi.htm.

²² MS = matrice de similarité (avec les dimensions N × N).

En règle générale, l'algorithme MINMWMAX engendre des profils choroplèthes plutôt aplanis alors que l'algorithme MEDMW crée des profils plus accidentés. Avec VDM l'on peut facilement sauter entre MINMWMAX et MEDMW tout en variant librement le nombre des intervalles chromatiques entre 2 et 20. Le coloriage algorithmique à la fois multiple et nuancé des scores-DM stockés dans les colonnes des matrices de similarité, de distance ou de corrélation constitue donc un atout majeur du logiciel salzbourgeois VDM.

LES CARTES DE SIMILARITÉ (CS)

Voir les Cartes 1-8²³.

Chacune des CS²⁴ dispose d'un point de référence précis qui sert d'étalon "de comparaison" avec le reste du réseau. Chaque CS comprend donc 640 (= N - 1; N = 641) polygones coloriés alors que le polygone du point de référence reste en blanc. Les points de référence choisis par nous concernent le nord (→ Cartes 1-2) et le centre du domaine d'Oïl (→ Carte 3), le centre du domaine d'Oc (→ Carte 4), le Croissant (→ Cartes 5 et 6) et la partie orientale du domaine francoprovençal (→ Carte 8).

En outre, le contenu et la structure des CS dépendent largement de l'indice de similarité utilisé pour leur calcul. Or, cet indice devrait correspondre – au moins en théorie – à une conception préétablie de la ressemblance linguistique entre les langues. Comme la statistique et la taxométrie offrent une gamme très vaste d'indices de ressemblance qui répondent à des finalités très différentes, il faut que le dialectométricien règle son choix sur les présupposés de la

²³ Sur chacune des cartes choroplèthes, deux des 641 polygones respectifs sont marqués par des hachures blanches. Ce graphisme renvoie soit à la valeur *maximale* (sur un polygone en *rouge*), soit à la valeur *minimale* (sur un polygone en *bleu foncé*) de la distribution de fréquence visualisée.

²⁴ CS = carte de similarité (un des six volets taxométriques de l'EDMS)

géolinguistique. Parmi les indices ainsi adoptés par l'EDMS, l'«Indice Relatif d'Identité» (IRI)²⁵ occupe une place de choix. Il peut être considéré comme indice standard de l'EDMS.

PRÉSENTATION ET INTERPRÉTATION DES CARTES 1 ET 2 (DOMAINE D'OÏL-A)

Le point de référence des deux profils choroplèthes de la Carte 1 se situe à la pointe septentrionale de la Picardie (= p. 297). La structuration beaucoup plus «comprimée» du cliché de droite montre très bien l'effet égalisateur (par un effet de lissage) du mode-DM *béta*. Néanmoins il est possible, par l'application de l'algorithme MEDMW à la place de MINMWMAX, de «décompresser» les grandes plages en *rouge* (→ domaine d'Oïl) et en *bleu foncé* (→ domaine d'Oc) du cliché de droite, et de faire ressortir ainsi la variation géolinguistique sous-jacente: voir la Carte 2, dont les deux moitiés se ressemblent beaucoup malgré d'énormes différences entre leurs bases numériques.

PRÉSENTATION ET INTERPRÉTATION DE LA CARTE 3 (DOMAINE D'OÏL-B)

Le point de référence (= p. 307) se situe en plein domaine d'Oïl. L'effet égalisateur du mode-DM *béta* se manifeste à nouveau. À remarquer, à droite, la mise en vedette (moyennant les intervalles 2-5) des parlers du Croissant et de la plupart des parlers francoprovençaux.

²⁵ La dénomination originale allemande est *Relativer Identitätswert* et l'équivalent anglais *Relative Identity Value*. Pour une explication détaillée de l'IRI cf. Goebel 1981, 2000 et 2002.

PRÉSENTATION ET INTERPRÉTATION DE LA CARTE 4 (DOMAINE D'OC)

Le point de référence 735 se réfère à un parler typiquement languedocien. L'étalement auréolaire des couleurs chaudes (rouge, orange et jaune) sur le profil de gauche (mode-DM *alpha*) reflète très bien la structuration interne du domaine d'Oc. Tout comme sur la carte précédente, le profil de similarité-*béta* (à droite) est caractérisé par deux larges plages monocolores (en *bleu foncé* [→ domaine d'Oil] et en *rouge* [→ domaine d'Oc]), entre lesquelles se glissent deux zones de transition: le Croissant et la marge méridionale du francoprovençal.

PRÉSENTATION ET INTERPRÉTATION DES CARTES 5 ET 6 (CROISSANT)

Le point de référence (p. 704) se trouve dans la Creuze méridionale et, de ce fait, au cœur du Croissant. La configuration "delphino-morphe" du Croissant ressort beaucoup mieux du profil de droite (en mode-DM *béta*), alors que le profil de gauche (en mode-DM *alpha*) souligne, de par la présence des polygones (intermédiaires) en *jaune* tant sur la Manche que sur la Méditerranée, le caractère hybride du locolecte du p. 704.

Dans cette situation, il semble indiqué de jouer sur les modalités visualisatrices du logiciel VDM et d'appliquer non seulement l'algorithme "différenciateur" MEDMW mais aussi un nombre beaucoup plus élevé (= 20) d'intervalles chromatiques: voir les deux profils de la Carte 6.

Aux vingt intervalles numériques des deux profils de la Carte 6 ne correspondent pas autant de teintes empruntées à l'arc-en-ciel; c'est que les classes intermédiaires 4-10 (toutes colorées en *vert*) ainsi que les classes 11-17 (toutes colorées en *jaune*), qui

sont situées respectivement de deux côtés des deux moyennes arithmétiques (à *gauche*: 59,77; à *droite*: + 0,29), ont reçu un coloriage uniforme en *vert* et en *jaune*. Il en résulte un effet optique particulier qui rehausse, par un effet de ricochet visuel, l'éclat des polygones relatifs aux intervalles situés en aval (1-3: en *bleu*) et en amont (18-20: en *rouge*) de l'échelle des scores de similarité.

C'est ainsi que surgit, sur le profil de gauche et autour du point de référence 704, une agglomération "nubi-forme" en rouge, alors que, sur le profil de droite (réalisé en mode-DM *béta*), la silhouette "delphino-morphe" classique du Croissant²⁶ apparaît de nouveau. En comparant les deux profils de la Carte 6, l'on constate, une fois de plus, l'effet niveleur du mode-DM *béta*.

PRÉSENTATION ET INTERPRÉTATION DES CARTES 7 ET 8 (DOMAINE FRANCOPROVENÇAL)

Nous reprenons ici l'expérience des Cartes 1 et 2 desquelles s'est dégagé avec netteté le contraste qui existe entre les visualisations choroplèthes effectuées par l'algorithme MINMWMAX d'un côté, et l'algorithme MEDMW de l'autre. Alors que sur la Carte 7 (établie moyennant MINMWMAX) les deux profils accusent une allure géographique très régulière, les deux profils de la Carte 8 (créés à l'aide de MEDMW) montrent des profils beaucoup plus accidentés et même déroutants; ceci est vrai surtout pour le profil de gauche (produit en mode-DM *alpha*).

UNE CARTE À PARAMÈTRE: LA SYNOPSE DES "COEFFICIENTS D'ASYMÉTRIE DE FISHER" (CAF)

Voir la Carte 9.

²⁶ Voir à ce sujet les cartes en annexe à Goebel 2004.

Un des exploits majeurs de l'EDMS était la découverte que les N distributions de similarités stockées dans une matrice similarité (MS) carrée ($N \times N$) sont porteuses non seulement d'informations linguistiques de type *synchronique*, mais aussi de type *diachronique*. Ces dernières peuvent être repérées surtout à travers les "paramètres typiques" des distributions de similarité tels que le *minimum*, le *maximum*, la *moyenne arithmétique*, l'écart-type ou le *coefficient d'asymétrie de Fisher* (CAF). L'analyse serrée de ces paramètres statistiques permet, entre autres, de mesurer l'intensité de l'"enchevêtrement particulier" des AT.

Or, il est bien connu depuis les recherches de J. Gilliéron (et de beaucoup d'autres) que, dans l'espace, le "changement linguistique" se manifeste par d'interminables "bousculades" – de type expansif ou régressif – parmi les nombreuses AT "en lice". Or, c'est justement le paramètre CAF qui permet de déterminer avec précision les régions où, à l'intérieur du réseau de l'ALF, ces "bousculades" ont été les plus fortes, et celles où ces "télescopages" étaient de nature plus discrète.

Sur la Carte 9 les scores (négatifs) du CAF, coloriés en *bleu foncé* (etc.), renvoient aux régions marquées par d'intenses "bousculades" entre les AT rivales, alors que les scores positifs du CAF, coloriés en *rouge* (etc.) sur la Carte 9, signalent des zones où les AT du lieu menaient une vie beaucoup plus calme.

Dans cette optique, l'architecture du profil de gauche (créé en mode-DM *alpha*) est très éloquente: l'on y discerne deux configurations circulaires en *bleu* (foncé et moyen) qui entourent, d'un côté, le *domaine d'Oil* et, de l'autre, celui du *francprovençal*.

L'encercllement oïlique renvoie au rayonnement bien connu du franc(il)ien à partir de l'Île-de-France caractérisé par d'innombrables effets d'expansion et de refoulements d'AT de toute sorte, alors

que la “fourchette” francoprovençale est due, vers le nord-ouest, à la retraite pluriséculaire de la vieille latinité *lyonnaise* devant l’expansion “tous-azimut” de la latinité *mérovingo-carolingienne*, et, vers le sud (c.-à-d. vers le domaine d’Oc), aux multiples frictions et compromis linguistiques qui se sont produits entre les latinités de *Lugdunum / Lyon* (fondée en 43 av. J.-Chr., c.-à-d. peu de temps après la fameuse conquête de la Gaule par Jules César) et celle de *Colonia Narbo Martius / Narbonne* (fondée déjà en 118 av. J.-Chr., donc au moment de la première mainmise des Romains sur la Gaule).

Aux plages rouges – qui symbolisent des évolutions linguistiques plutôt ralenties – correspondent des provinces occitanes centrales telles que la Gascogne, le cœur du Languedoc et la Provence entre lesquelles se creusent des zones de passage en jaune et orange.

Sur le profil de droite (réalisé dans le mode-DM *béta*), le théâtre des grandes confrontations linguistiques s’est déplacée du pourtour périphérique oriental du domaine d’Oïl vers le milieu de la Gaule, alors que les hauts-lieux du conservatisme linguistique d’Oc ne se sont guère déplacés.

Soulignons encore, sur le profil de droite, l’aplanissement complet du francoprovençal et l’apparition – en rouge – de quelques buttes-témoins conservateurs dans la zone du Croissant.

L’ANALYSE DENDROGRAPHIQUE: LA “CLASSIFICATION HIÉRARCHIQUE ASCENDANTE” (CAH) (MÉTHODE DE WARD)

Voir la Carte 10.

Comme, en ligne générale, l’approche dendrographique fait partie intégrante de la linguistique historique depuis longtemps, on ne s’étonnera point qu’elle réapparaisse aussi en matière de DM.

Du point de vue statistique, la CAH opère avec des algorithmes²⁷ dont le but est de pratiquer, à l'intérieur de $N/2$ ($N - 1$)²⁸ scores de similarités stockés dans une MS, N fusions exclusivement binaires, d'abord entre des éléments (ou lococolectes) isolés et ensuite entre des groupements majeurs créés successivement par voie agglomérative. Ce travail fusionniste débute par les scores les plus petits (c.-à-d. auprès des "feuilles" de l'arbre) pour aboutir, après une "ascension" le long de l'échelle des scores de similarité triés par leur taille, à la génération de la "racine" (ou du "tronc") de l'arbre par la fusion des deux dernières valeurs de cette échelle.

Les ramifications de l'arbre sont toujours binaires; le "branchage" de l'arbre consiste en un certain nombre d'embranchements que nous appelons *dendrèmes*. Pour les propos de la DM, il est de toute première importance de ne jamais considérer isolément l'arbre issu d'une classification de type-CAH, et de toujours l'utiliser pour la projection de ses dendrèmes dans l'espace, c.-à-d. sur la carte de l'espace en question ("spatialisation"), tout en y créant des configurations spatiales typiques, appelés par nous *chorèmes*: voir la Carte 10 où l'identité du coloriage des deux arbres et de leurs spatialisations facilite l'identification des relations qui existent entre les différents *dendrèmes* et les *chorèmes*, leurs équivalents sur la carte.

À noter aussi, du point de vue taxométrique, la différence qui existe entre la variabilité numérique à l'intérieur des dendrèmes (ou chorèmes) [variabilité "*intra-group*"], et celle qui s'établit *entre* les différents dendrèmes/chorèmes [variabilité "*inter-group*"]. La logique des algorithmes utilisés dans le cadre de la CAH veut que

²⁷ L'arbre présenté sur la carte 10 a été construit moyennant l'algorithme-CAH proposé en 1963 par le statisticien américain Joe Ward, Jr: cf. Chandon-Pinson 1981, 123.

²⁸ Toute MS carrée ($N \times N$) consiste en deux moitiés symétriques égales ($N \times N/2$) et une diagonale (N) occupée uniquement par la valeur 100 (ou 0). La somme des scores utilisables est donc de $N \times N/2 - N = N/2 (N - 1)$.

cette variabilité soit minimale au niveau des “feuilles”, et maximale près de la “racine” (ou du “tronc”) de l’arbre. Il existe également une relation directe entre cette variabilité changeante et la “fiabilité” des classifications ainsi établies. C’est que cette dernière diminue avec l’augmentation des deux espèces de variabilité mentionnés ci-dessus. Ceci signifie que les groupements situés près des feuilles sont – au moins en théorie – plus “sûrs” que ceux qui se trouvent près de la racine.

Du point de vue (géo)linguistique il est possible de soumettre les ramifications de l’arbre à une lecture (ou interprétation) diachronique tout en attribuant la plus grande ancienneté aux groupements situés près de la racine. Cette lecture se déroule le long de l’axe horizontal situé au-dessus des deux arbres et se fait en sens inverse par rapport aux nombreuses fusions précédentes. C’est ainsi que la structure des deux arbres de la carte 10 suggère que la première fragmentation du “latin des Gaules” s’est opérée entre les domaines d’Oc et d’Oïl, que l’autonomie classificatoire de l’Est du domaine d’Oïl (en jaune) est plus ancienne que celle du Centre (en rose) et du Sud-ouest (en rouge), et que la position classificatoire du francoprovençal (en orange) est ambiguë: à gauche (en mode-DM *alpha*) celle-ci semble être en retrait (= plus jeune) par rapport au Centre, alors qu’à droite (en mode-DM *béta*) elle suggère une plus grande ancienneté de ce domaine linguistique, augmenté d’ailleurs vers l’Ouest par quelques points du Croissant.

À noter aussi que, sur les deux spatialisations (suivant les modes-DM *alpha* et *béta*), les quatre bifurcations intra-occitanes se trouvent pratiquement à la même hauteur. Ceci laisse conclure à une évolution linguistique uniforme et tranquille du domaine d’Oc, exempté d’expansions et de retraits considérables.

L'ANALYSE INTERPONCTUELLE EN FONCTION DISCRIMINATOIRE ("CARTES ISOGLOTTIQUES")

Voir la Carte 11.

Le terme et le concept de l'*interpoint* ont été créés, à partir de 1949, par l'abbé Théobald Lalanne (1880-1952). Il présuppose l'existence d'un réseau de recherche avec des points d'enquête répartis plus ou moins équitablement dans l'espace, entre lesquels il est possible de pratiquer deux opérations géométriques: d'abord le traçage de *lignes droites* entre les points avoisinants du réseau (triangulation de Delaunay²⁹), et ensuite la construction des verticales (ou: perpendiculaires) respectives quitte à les prolonger jusqu'à leur fusion réciproque.

Dans le premier des deux cas, il en résulte un *treillis triangulaire*, alors que le résultat graphique du second est une *mosaïque polygonale* (polygonisation de Voronoi)³⁰. Les "interpoints" de mémoire lalannienne ont donc deux apparences géolinguistiques et géométriques³¹ qui desservent deux fonctions linguistiques différentes: la *communication* à travers les *côtés de triangle*, la *distanciation* par les *côtés de polygone*.

Suivant ces deux principes géométriques nous avons tiré, des 641 points de notre réseau de recherche, 1 791 interpoints qui, dans la perspective de l'EDMS, correspondent à autant de valeurs soit de *distance* (voir la Carte 11), soit de *similarité* (voir la Carte 12). Leur mise en carte aboutit à deux messages géolinguistiques différents. Alors que la carte 11 signale le *cloisonnement* interponctuel général,

²⁹ Mathématicien russe (B. N. Delone, 1890-1980) dont le nom russe a été francisé après coup en *Delaunay*.

³⁰ Mathématicien russe (G. F. Voronoi, 1868-1908). La paternité de ce principe géométrique très utile, est multiple: voici les noms de deux autres pionniers concurrents: P. G. Dirichlet (1805-1859) et A. H. Thiessen, (1872-1956).

³¹ Cf. à ce sujet notre contribution de 1983.

la Carte 12 renvoie à la *connectivité* interponctuelle. Dans les deux cas, il s'agit de phénomènes très variables avec des implantations spatiales bien définies.

Les deux clichés de la Carte 11 montrent très bien la bipartition de la Galloromania, non seulement par la présence de plusieurs zones de transition, colorées en différentes teintes de bleu et caractérisées par une grande épaisseur des côtés de polygone respectifs, mais aussi par la présence de quelques régions linguistiquement égalisées, et qui sont dotées de polygones entourés de côtés *minces* et *rouges*: il s'agit de l'Île-de-France, du Languedoc et de la partie orientale de la Provence.

Une fois de plus, le message iconique du cliché de droite est (beaucoup) plus clair et simple. À mes yeux, la configuration polygonale en bleu qui sillonne le milieu de la Galloromania a les apparences d'un *dauphin* qui s'élanche de l'est vers l'ouest.

L'ANALYSE INTERPONCTUELLE EN FONCTION COMMUNICATIVE (“CARTES à RAYONS”)

Voir la Carte 12.

Évidemment, la logique iconique des deux clichés de la Carte 12 s'oppose à celle de la Carte 11: ici, l'intensité du coloriage (→ en rouge) et l'épaisseur de la triangulation polygonale signalent d'excellents rapports communicatifs entre les points-ALF limitrophes. À nouveau se dégagent, dans le nord, le Centre, et, dans le sud – bien que dans une moindre mesure –, la Gascogne, une partie du Languedoc et la Provence.

Les zones de transition entre Oc et Oïl sont marquées par une texture triangulaire fine et colorée en bleu. La largeur de

cette “déchirure” entre Oc et Oil³² est plus grande sur le cliché de gauche (produit en mode-DM *alpha*) que sur celui de droite (mode-DM *béta*).

L'ANALYSE CORRÉLATIVE (LA SIMILARITÉ LINGUISTIQUE CORRÉLÉE AVEC LA PROXIMITÉ GÉOGRAPHIQUE)

Voir les Cartes 13 et 14.

La méthode-DM de l'analyse corrélatrice dépend d'une assomption théorique capitale: à savoir qu'il existe, dans les profondeurs de la MD dépouillée, des structures globales issues d'un comportement linguistique particulier des humains face à l'espace. Depuis un certain temps (2005) nous appelons cette activité anthropique “gestion basilectale de l'espace par l'homo loquens”. Évidemment, cette “gestion” peut embrasser toutes les catégories linguistiques ou seulement la phonétique, le lexique ou la morpho-syntaxe. Il est donc légitime d'examiner les liens corrélatifs spatiaux qui existent entre ces différentes espèces *linguistiques* de la gestion de l'espace, et aussi la corrélation qui se manifeste entre la gestion *linguistique* (globale) de l'espace par les humains, et celle qui s'instaure, par voie *euclidienne*, entre les distances (ou proximités) kilométriques des N localités de notre réseau.

La Carte 13 montre, à gauche, la position – générée sans l'ombre de doute par la gestion *glotto-anthropique* de l'espace – du

³² Rappelons dans ce contexte le verdict émis en 1888 (435-436) par Gaston Paris sur la non-existence d'une *muraille* entre le nord et le sud de la France: “Et comment, je le demande, s'expliquerait cette étrange frontière qui de l'ouest à l'est couperait la France en deux en passant par des points absolument fortuits ? Cette muraille imaginaire, la science, aujourd'hui mieux armée, la renverse, et nous apprend qu'il n'y a pas deux Frances, qu'aucune limite réelle ne sépare les Français du nord de ceux du midi, et que d'un bout à l'autre du sol national nos parlers populaires étendent une vaste tapisserie dont les couleurs variées se fondent sur tous les points en nuances insensiblement dégradées”. On doit au même auteur aussi le verdict de la “non-existence des dialectes”. Nous qualifions depuis longtemps cette attitude égalisatrice et anti-classificatoire comme “typophobe” (cf. Goebel 1986).

dialecte de Marcigny (= p. 1 de l'ALF) au sein de la Galloromania, alors que le cliché de gauche, également relatif au p. 1 de l'ALF, repose sur l'application pure et simple du théorème de Pythagore ($a^2 + b^2 = c^2$) sur les coordonnées géographiques (en x et y) des 641 points de notre réseau. La régularité de ce profil n'a pas de quoi surprendre: tout y est régi par les nécessités de la géométrie euclidienne.

Dans ces conditions, il est indiqué de se demander dans quelle manière coïncident les deux espèces de gestion de l'espace, créées soit par *l'homme*, soit par les *impératifs de la géométrie*.

Du point de vue taxométrique, cette comparaison se fait moyennant N (ici: 641) calculs de corrélation (exécutés à l'aide de l'indice de corrélation de Bravais-Pearson $r(\text{BP})$ entre deux séries de valeurs: entre les N colonnes de la matrice de *similarité* (calculées par l'application de l'indice IRI_{jk})³³, et les N colonnes d'une matrice de *proximité* analogue, dans les cellules de laquelle ont été stockées – selon une formule très simple: $d + p = 100$) – les valeurs non pas des *distances* kilométriques (d) calculées, mais de leurs équivalents de *proximité* (p).

De cette comparaison “corrélative résultent 641 scores- $r(\text{BP})$ susceptibles d'être visualisés selon les standards cartographiques de l'EDMS: voir les deux clichés de la Carte 14.

Pour bien comprendre le message de ces deux cartes choroplèthes très clairement structurées il faut connaître la signification des valeurs sur lesquelles elles reposent. Il s'agit de 641 scores de corrélation, toutes oscillant à l'intérieur d'une fourchette numérique située entre -1 (en bleu foncé) et +1 (en rouge), qui signalent le taux de l'interdépendance entre les deux espèces de gestion de l'espace.

³³ IRI_{jk} = Indice Relatif d'Identité (entre les points d'atlas j et k)

Les couleurs *chaudes* (du rouge au jaune) signalent un phénomène que l'on pourrait appeler le "pas cadencé" entre les deux sortes de gestion, alors que les couleurs *froides* (du bleu foncé au vert clair) signalent le contraire: de fortes dissonances (ou: antagonismes) entre ces deux espèces de gestion de l'espace. Or, la structure géographique bipolaire des deux clichés – avec, en rouge, une implantation cunéiforme "expansive" dans le nord, une bande "défensive" le long des Pyrénées dans le sud, et une large ceinture de "dissonance" dans le centre – nous suggère la présence d'antagonismes conflictuels "de longue durée".

Il est en effet très probable que cette structure bipolaire renvoie aux multiples contacts et conflits linguistiques qui se sont déroulés, à travers l'épanouissement et la diminution d'innombrables AT, entre les deux latinités de base de la Galloromania: celle de Narbonne (= Oc), et celle du Centre (= Oïl)³⁴. Une excellente preuve de cette vision "conflictuelle" est la position excentrique (en bleu foncé, avec des hachures blanches superposées³⁵) du p. 635 (Andraut; situé au sud-est de Bordeaux) qui, en tant qu'îlot linguistique d'origine saintongeaise implanté au XV^e siècle en territoire gascon, remplit toutes les prémisses d'une dissonance capitale entre les gestions *linguistique* et *euclidienne* de l'espace.

Dans l'optique de ces deux clichés, la latinité de Lugdunum, héritière directe de la conquête de Jules César (58-51 av. J. Chr.) et souche du francoprovençal, semble avoir été "broyée" entre les "pinces" majeures de Lutèce et de Narbonne.

³⁴ Il s'agit de la latinité "mérovingo-carolingienne", intimement liée à l'essor politique des Francs à partir de 498 (baptême de Clovis) et à l'enracinement de ces derniers au sol de la tribu gauloise des *Parisii*, en Île-de-France.

³⁵ La superposition de hachures blanches renvoie aux valeurs *maximales* et *minimales*: à gauche: le *maximum* (en rouge) est en Picardie, le *minimum* (en bleu foncé) au P. 635; à droite: le *maximum* est dans les Pyrénées orientales (en rouge), alors que le *minimum* (en bleu foncé) se trouve à nouveau au P. 635.

À souligner aussi l'étonnante gradualité des passages finement échelonnés entre les pôles du rouge et du bleu. Ce phénomène est aussi tributaire de l'effet "différenciateur" de l'algorithme MEDMW et du nombre élevé de 10 intervalles chromatiques. La même remarque concerne d'ailleurs aussi les Cartes 11 et 12, dédiées à la présentation des deux versants de la DM interponctuelle.

Évidemment, le champ de telles comparaisons corrélatives est très vaste: surtout les comparaisons entre différentes catégories linguistiques (p. ex. *phonétique vs. lexicale, vocalisme vs. consonantisme* etc.) fournissent d'excellents résultats dont l'intérêt linguistique est considérable (voir GOEBL 2005).

CONCLUSION

Le nouveau volet du logiciel VDM s'intègre parfaitement dans la logique méthodique de l'EDMS qui par définition est multiple. J'espère que les exemples présentés sur les Cartes 1-14 ne laissent pas de doute sur son utilité géolinguistique. Au besoin cet intérêt pourrait s'étendre aussi à d'autres sciences qui, comme la géographie, l'anthropologie ou la sociologie quantitatives, sont engagées dans l'analyse de l'interaction entre l'homme et l'espace.

C'est le bon moment pour souligner, une fois de plus, que le logiciel VDM est à la libre disposition de tous ceux qui en ont besoin et en font la demande³⁶.

RÉFÉRENCES

ALF = GILLIERON, Jules; EDMONT, Edmond. **Atlas linguistique de la France**. 10 v. Paris: Champion, 1902-1910 (réimpression: Bologne: Forni, 1968) [trois fois consultable dans le web: Innsbruck: <https://diglib.uibk.ac.at/urn:nbn:at:at-ubi:2-4568>; Toulouse: <http://symila.univ-tlse2.fr/alf/>; Grenoble: <https://lig-tdcge.imag.fr/cartodialect5/#/>].

³⁶ Voici mon adresse courriel: hans.goebel@plus.ac.at

CHANDON, Jean-Louis; PINSON, Suzanne. **Analyse typologique**. Théories et applications. Paris; New York; Barcelone; Milan: Masson, 1981.

CHRISTMANN, Hans Helmut. Lautgesetze und Wortgeschichte. Zu dem Satz „Jedes Wort hat seine eigene Geschichte“. In: STEMPEL, Wolf-Dieter; COSERIU, Eugenio (eds.). **Sprache und Geschichte**. Festschrift für Harri Meier zum 65. Geburtsta. München: Fink, 1971, p. 119-124.

DEES, Anthonij. **Atlas des formes et des constructions des chartes françaises du 13^e siècle**. (Beihefte der Zeitschrift für romanische Philologie, v. 178). Tübingen: Niemeyer, 1980.

ESTUDIS ROMÀNICS. Barcelona: Institut d'Estudis Catalans, 2006. v. 28. p. 281-298.

GILLIERON, Jules (avec Jean Mongin). **Scier dans la Gaule romane du sud et de l'est**: étude de géographie linguistique. Paris: Champion, 1905. Réimpression de l'original avec une traduction italienne en vue par Lorenzo Massobrio: "*Segare*" nella Gallia romanza meridionale e orientale. Novi Ligure: Grafica editoriale universitaria, 1990.

GOEBL, Hans. Éléments d'analyse dialectométrique (avec application à l'ALS). **Revue de linguistique romane**, n. 45, p. 349-432, 1981.

GOEBL, Hans. Parquet polygonal et treillis triangulaire: les deux versants de la dialectométrie interponctuelle. **Revue de linguistique romane**, n. 47, p. 353-412, 1983.

GOEBL, Hans. **Dialektometrische Studien**. Anhand italoromanischer, rätoromanischer und galloromanischer Sprachmaterialien aus ALS und ALF. (Beihefte der Zeitschrift für romanische Philologie, v. 191-193). Tübingen: Niemeyer, 1984.

GOEBL, Hans. Typophilie und Typophobie. Zu zwei problembeladenen Argumentationstraditionen innerhalb der Questione ladina. In: HOLTUS, Günter; RINGGER, Kurt (eds.). **Raetia antiqua et moderna**. W. Th. ELWERT zum 80. Geburtstag. Tübingen: Niemeyer, 1986. p. 513-536.

GOEBL, Hans. La dialectométrisation de l'ALF: présentation des premiers résultats. **Linguistica**, n. 40, p. 209-236, 2000.

GOEBL, Hans. Analyse dialectométrique des structures de profondeur de l'ALF. **Revue de linguistique romane**, n. 66, p. 5-63, 2002.

GOEBL, Hans. Regards dialectométriques sur les données de l'Atlas linguistique de la France (ALF): relations quantitatives et structures de profondeur. **Estudis Romànics**, XXV, p. 59-121, 2003.

GOEBL, Hans. Das «Croissant» - eine Nachschau im Abstand von 90 Jahren (mit zwei dialektométrisch erstellten Farbkarten). In: NOLL, Volker; THIELE, Sylvia (eds.). **Sprachkontakte in der Romania**. Zum 75. Geburtstag von Gustav Ineichen. Tübingen: Niemeyer, 2004. p. 159-172.

GOEBL, Hans. La dialectométrie corrélative. Un nouvel outil pour l'étude de l'aménagement dialectal de l'espace par l'homme. **Revue de linguistique romane**, n. 69, p. 321-367, 2005 (cartes en couleurs: 356-367).

GOEBL, Hans. Sur le changement macrolinguistique survenu entre 1300 et 1900 dans le domaine d'Oil. Une étude diachronique d'inspiration dialectométrique. **Dialectologia Revista electrònica**, Barcelona, n. 1, p. 3-43, 2008.

GOEBL, Hans. Quelques coups d'œil dialectométriques sur l'Atlas linguistique de la France: structures de surface et structures de profondeur. In: DALBERA-STEFANAGGI, Marie-José; SIMONI-AUREMBOU, Marie-Rose (eds.). **Images de la langue**: représentations spatiales, sémantiques et graphiques. Paris: Editions du CTHS, 2009a. p. 39-60.

GOEBL, Hans. Dialectometry and quantitative mapping. In: LAMELI, Alfred; KEHREIN, Roland; RABANUS, Stefan (eds.). **Language and Space**. An International Handbook of Linguistic Variation, v. 2: Language Mapping (Handbücher der Sprach- und Kommunikationswissenschaft [HSK] 30.2.). Berlin: de Gruyter 2010; 1^{ère} partie: p. 433-457 (texte); seconde partie: p. 2201-2212 (cartes).

GOEBL, Hans. L'impact de la polynymie des cartes d'atlas sur le résultat de calculs dialectométriques. In: POLSKA AKADEMIA UMIEJETNOSCI; INSTYTUT FILOLOGII ROMANSKIEJ UNIWERSYTETU JAGIELLONSKIEGO (ed.): **Linguistique romane et Linguistique indo-européenne**. Mélanges offerts à Witold Mańczak à l'occasion de son 90^e anniversaire. Kraków, Cracovie: Polska Akademia Umiejetnosci. Instytut Filologii Romanskiej Uniwersytetu Jagiellonskiego, 2014. p. 243-260 (avec 10 cartes en couleurs).

GOEBL, Hans. La face cachée de la géographie linguistique. Bref aperçu sur les "cartes muettes" produites pour l'ALF, l'AIS et le FEW. **Revue de linguistique romane**, n. 82, p. 5-63, 2018a (avec 16 cartes en couleurs).

GOEBL, Hans. Dialectometry. In: BOBERG, Charles; NERBONNE, John; WATT, Dominic (eds.). **The Handbook of Dialectology**. Hoboken, New Jersey: John Wiley; Sons, 2018b. p. 123-142 (avec quatre planches en couleurs).

GOEBL, Hans / SMECKA, Pavel. L'interprétation dialectométrique des atlas scripturaux d'Anthonij Dees. **Revue de linguistique romane**, n. 80, p. 321-368, 2016a (avec 50 cartes en couleurs).

GOEBL, Hans; SMECKA, Pavel. The Quantitative Nature of Working Maps (WM) and Taxatorial Areas (TA): A Brief Look at two Basic Units of Salzburg Dialectometry (S-DM). In: KELIH, Emmerich; KNIGHT, Róisín; MACUTEK, Ján; WILSON, Andrew (ed.). **Issues in Quantitative Linguistics 4**. Dedicated to Reinhard Köhler on the occasion of his 65th birthday. Lüdenscheid: RAM-Verlag (Studies in Quantitative Linguistics, v. 23), 2016b. p. 113-127 (avec quatre cartes en couleurs).

GOEBL, Hans; SMECKA, Pavel. Trois regards dialectométriques sur l'aménagement géolinguistique du domaine d'oïl, basés sur une synthèse des données médiévales réunies par Anthonij Dees en 1980 et 1983, et celles de l'ALF. In: KRISTOL, Andres M. (ed.). La mise à l'écrit et ses conséquences. **Actes du troisième colloque «Repenser l'histoire du français»**, Université de Neuchâtel, 5-6 juin 2014. Tübingen: Francke, 2017. p. 15-49 (avec 12 planches en couleurs).

HAAG, Carl (ou: Karl). **Die Mundarten des oberen Neckar- und Donautales** (Schwäbisch-alemannisches Grenzgebiet: Baarmundarten). Reutlingen: Hutzler, 1898.

JABERG, Karl. Zum Atlas linguistique de la France. **Zeitschrift für romanische Philologie**, n. 30, p. 512, 1906.

JABERG, Karl. *Sprachgeographie. Ein Beitrag zum Verständnis des Atlas linguistique de la France*. Aarau: Sauerländer, 1908. Version espagnole: *Geografía lingüística, Ensayo de interpretación del Atlas lingüístico de Francia*. Granada: Universidad de Granada, 1959.

LALANNE, Abbé Théobald. **L'indépendance des aires linguistiques en Gascogne maritime**. Berceau de Saint Vincent de Paul: chez l'auteur, 1949-1952. Réimpression: Labatut: Atelier Histoire Trait d'Union, 2018.

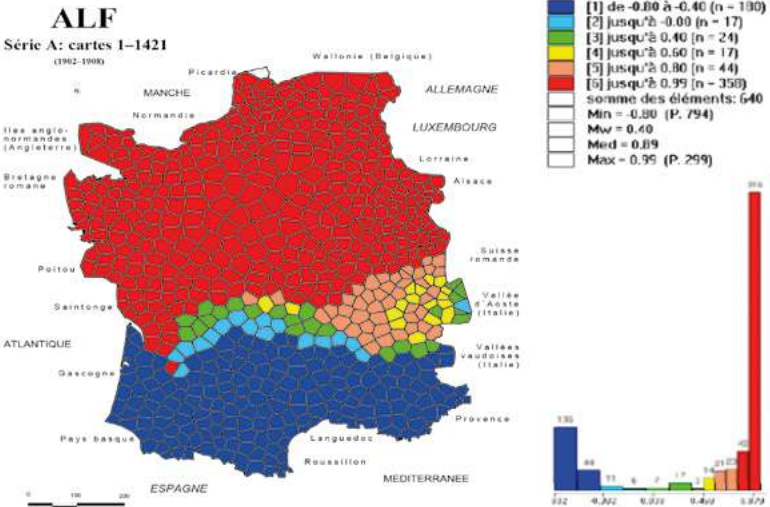
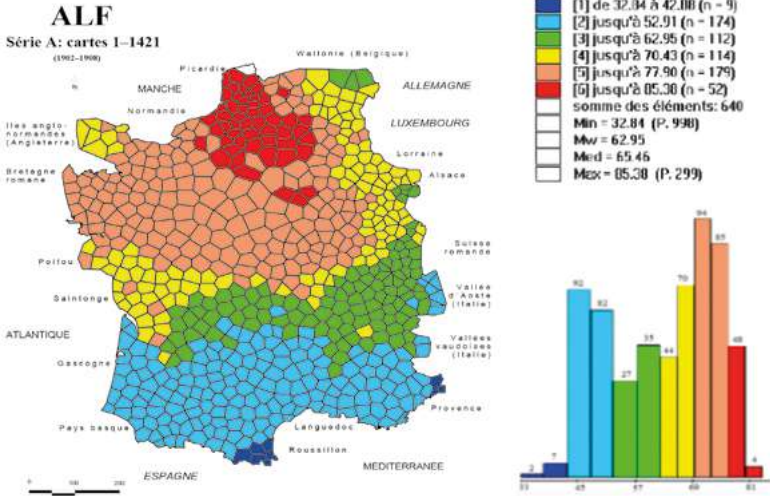
PARIS, Gaston. Les parlers de France. *Revue des patois gallo-romans* 2, 1888, p.161-175. In: PARIS, Gaston. **Mélanges linguistiques**. Latin vulgaire et langues romanes, langue française, notes étymologiques, publiés par Mario Roques. Paris: Champion, 1909. p. 432-448.

VIDESOTT, Paul. **Padania scrittologica**: Analisi scrittologica e scrittometriche di testi in italiano settentrionale antico dalle origini al 1525. (Beihefte zur Zeitschrift für romanische Philologie, v. 343). Tübingen: Niemeyer, 2009.

CARTE 1: DEUX CARTES DE SIMILARITÉ RELATIVES AU P.-ALF 297 (FORT-MARDYCK, NORD)

Corpus: 1 681 CT; **algorithme d'intervallisation :** MINMWMAX
6-tuple

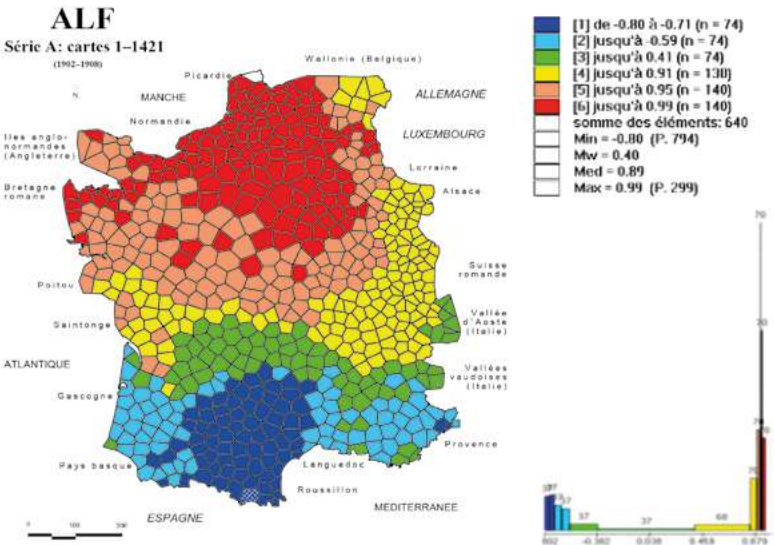
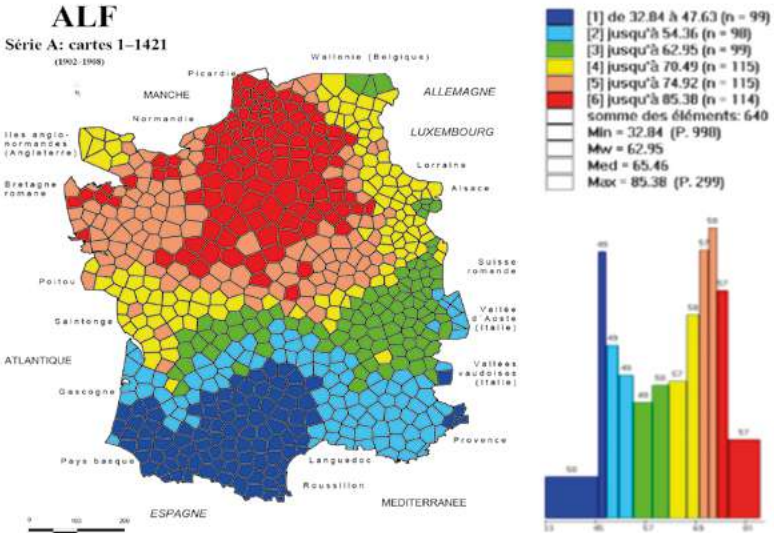
Calcul de similarité: en haut (« alpha »): IRIjk; en bas (« bêta »):
IRIjk + r (BP)



CARTE 2: DEUX CARTES DE SIMILARITÉ RELATIVES AU P.-ALF 297 (FORT-MARDYCK, NORD)

Corpus: 1 681 CT; algorithme d'intervallisation : MEDMW 6-tuple

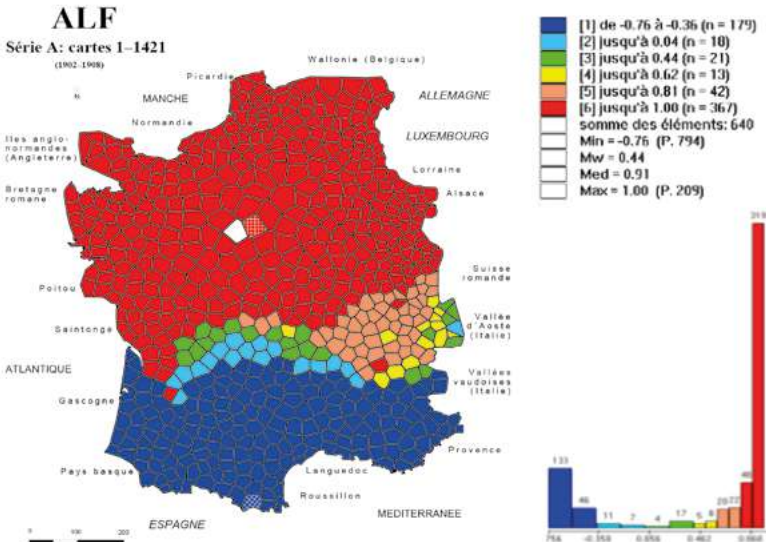
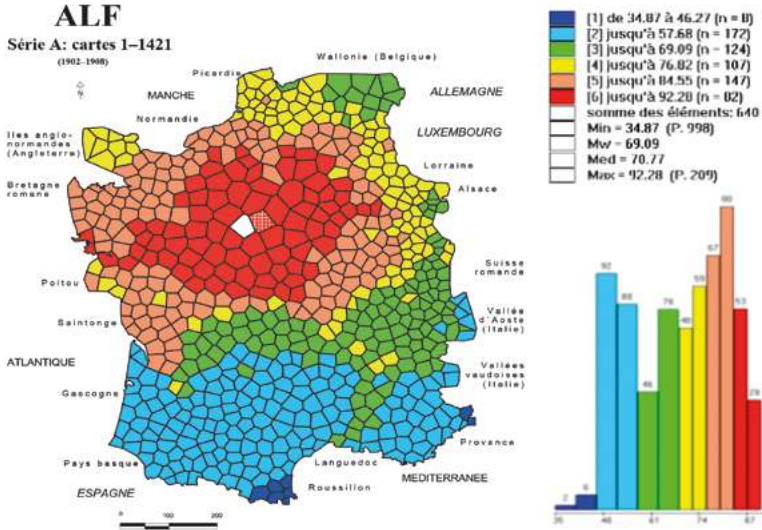
Calcul de similarité: en haut (« alpha »): IRIjk; en bas (« bêta »): IRIjk + r (BP)



CARTE 3: DEUX CARTES DE SIMILARITÉ RELATIVES AU P.-ALF 307 (SAINT-AY, LOIRET)

Corpus: 1 681 CT; algorithme d'intervallisation : MEDMW 6-tuple

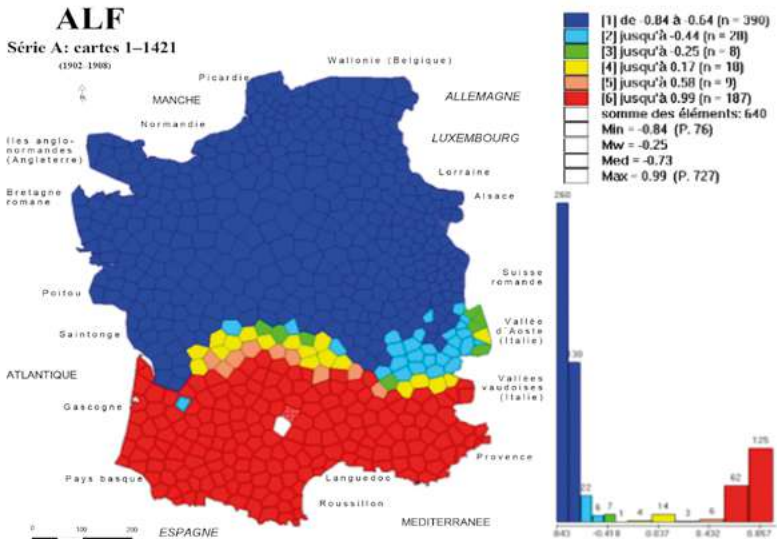
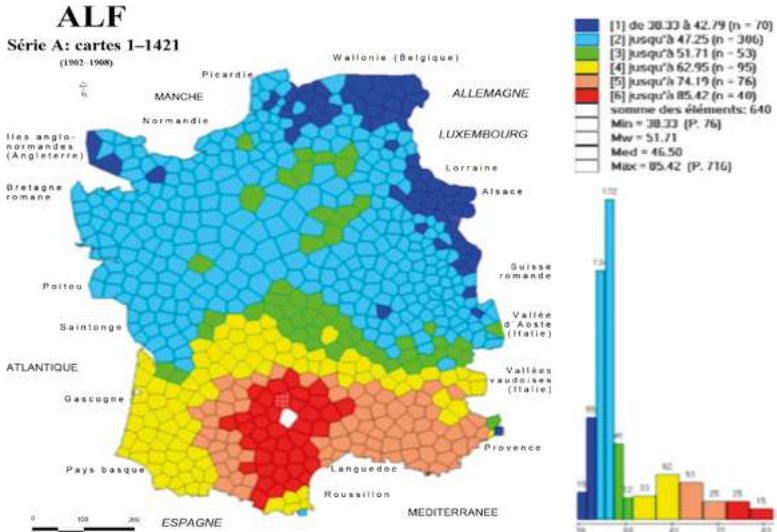
Calcul de similarité: en haut (« alpha »): IRIjk; en bas (« bêta »):
IRIjk + r (BP)



CARTE 4: DEUX CARTES DE SIMILARITÉ RELATIVES AU P.-ALF 735 (CALMONT, AVEYRON)

Corpus: 1 681 CT; algorithme d'intervallisation : MINMWMAX 6-tuple

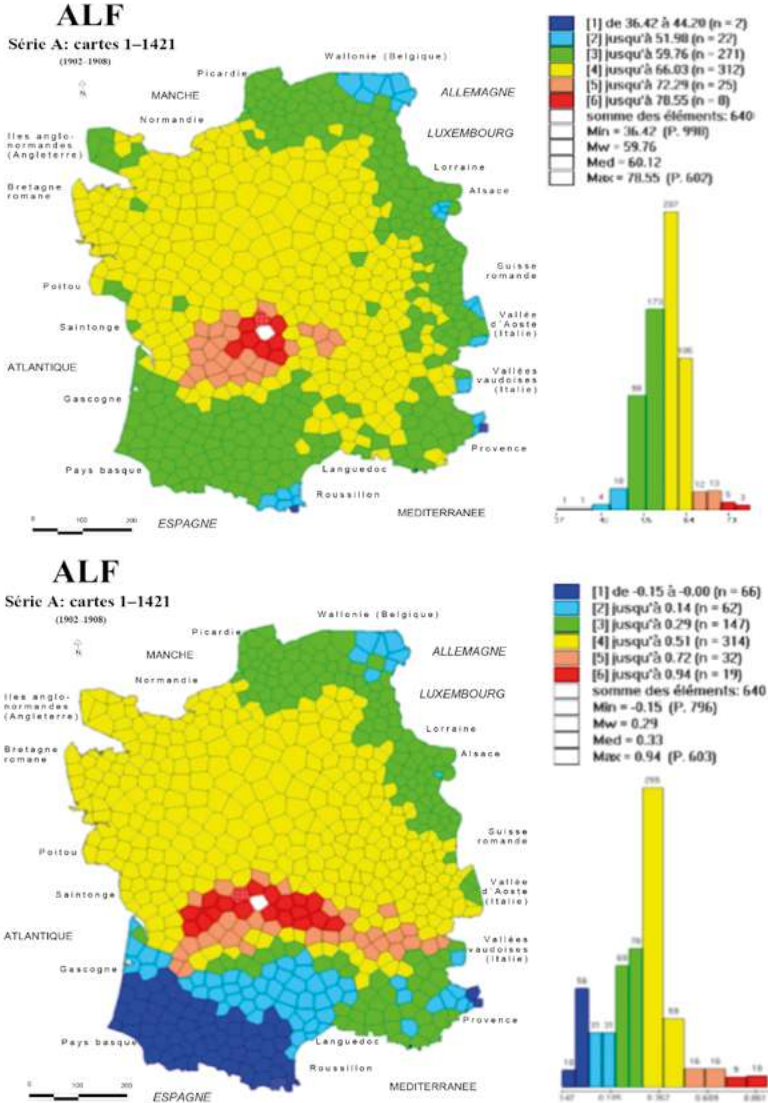
Calcul de similarité: en haut (« alpha »): IRIjk; en bas (« bêta »): IRIjk + r (BP)



CARTE 5: DEUX CARTES DE SIMILARITÉ RELATIVES AU P.-ALF 704 (SAINT-QUENTIN, CREUSE)

Corpus: 1 681 CT; algorithme d'intervallisation : MINMWMAX 6-tuple

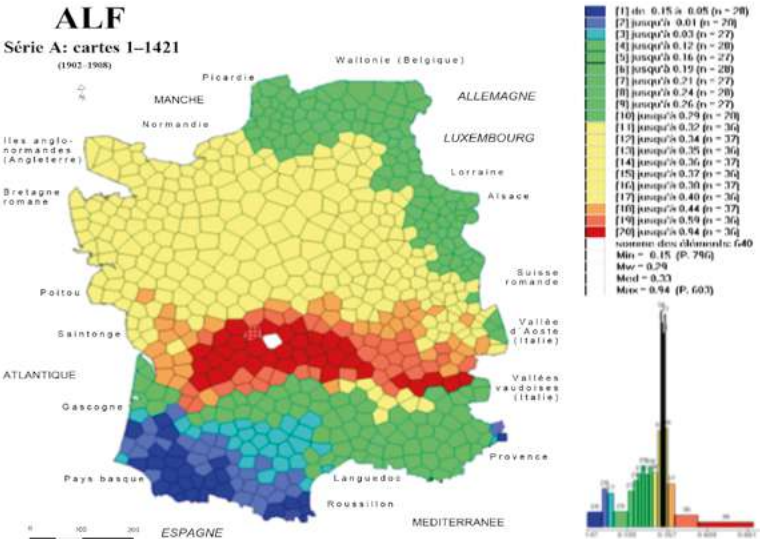
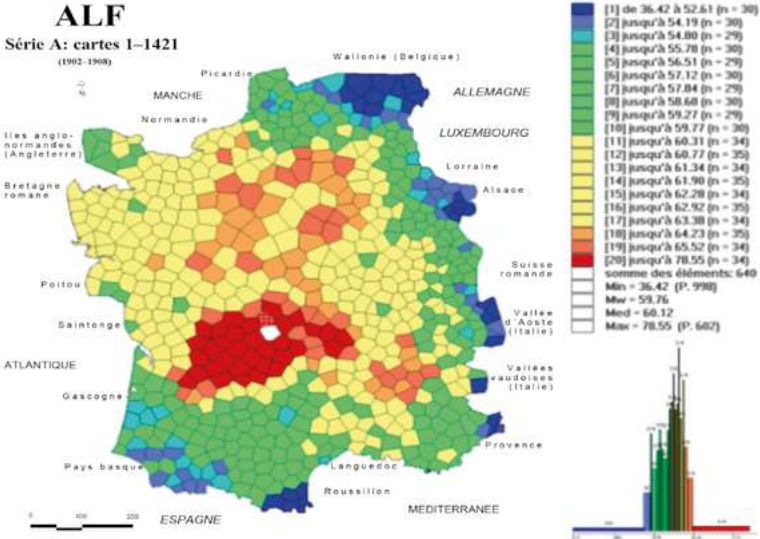
Calcul de similarité: en haut (« alpha »): IRIjk; en bas (« bêta »):
IRIjk + r (BP)



CARTE 6: DEUX CARTES DE SIMILARITÉ RELATIVES AU P-ALF 704 (SAINT-QUENTIN, CREUSE)

Corpus: 1 681 CT; algorithme d'intervallisation : MEDMW 20-tuple

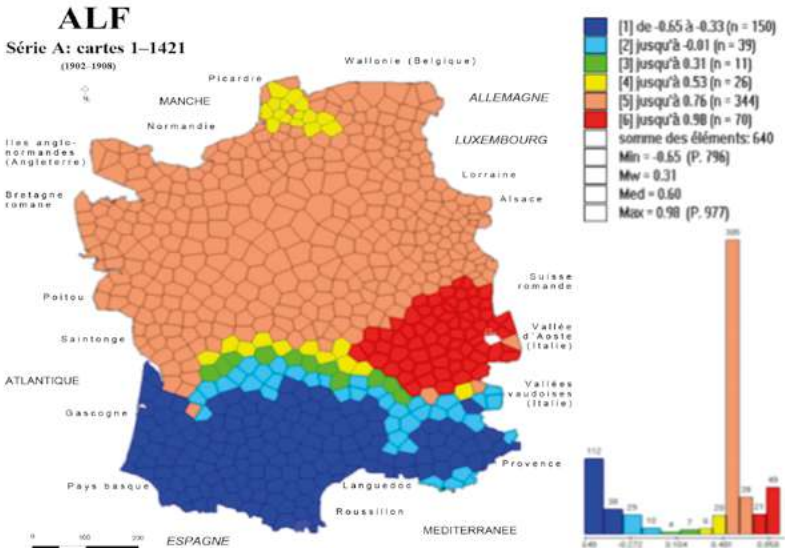
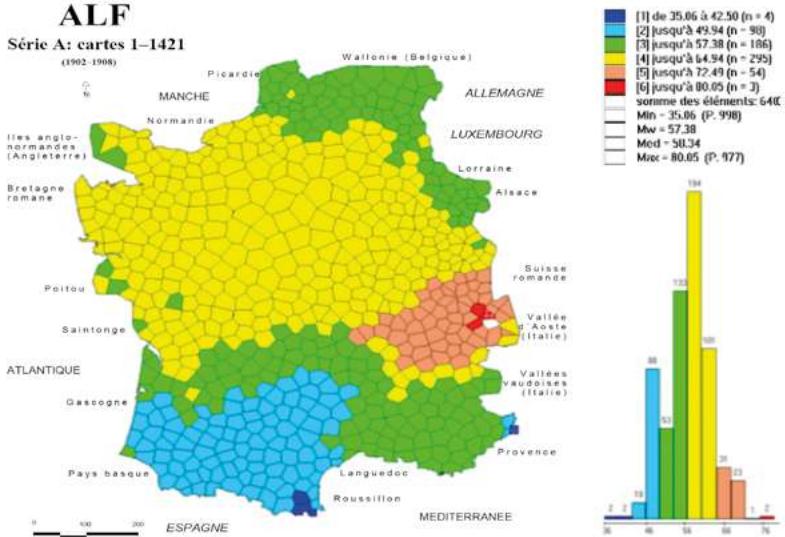
Calcul de similarité: en haut (« alpha »): IRIjk; en bas (« bêta »): IRIjk + r (BP)



CARTE 7: DEUX CARTES DE SIMILARITÉ RELATIVES AU P-ALF 976 (BOURG-SAINT-PIERRE, VALAIS, SUISSE)

Corpus: 1 681 CT; algorithme d'intervallisation : MINMWMAX 6-tuple

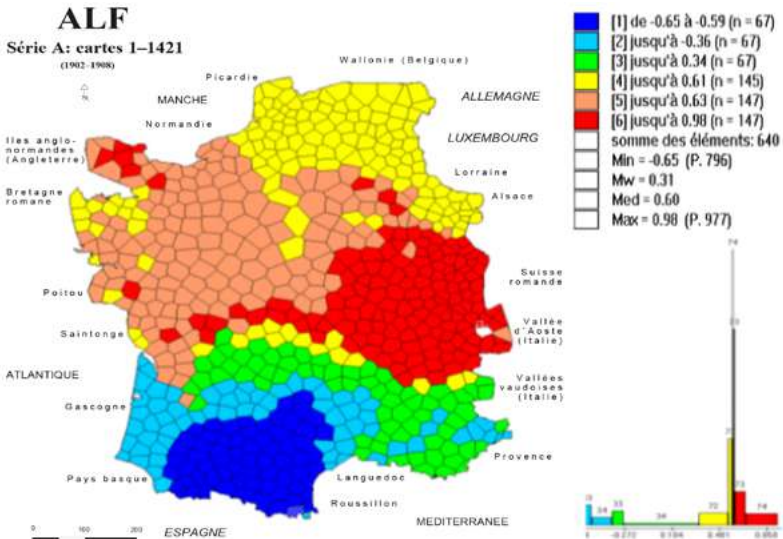
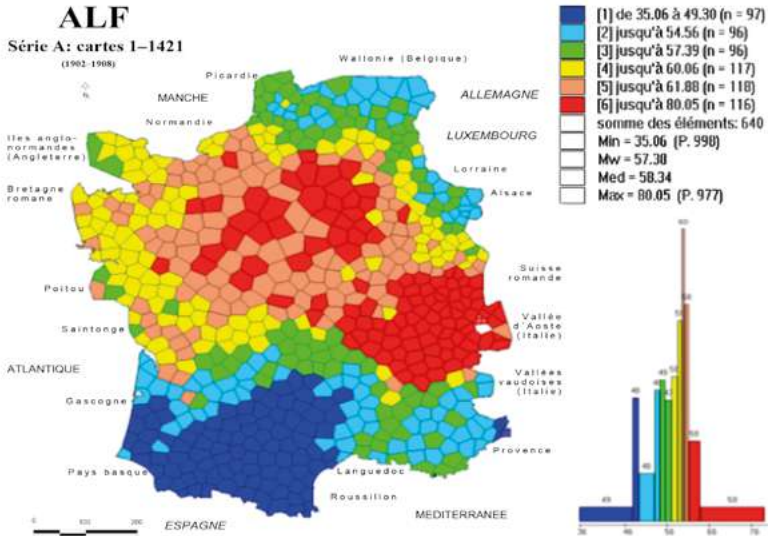
Calcul de similarité: en haut (« alpha »): IRIjk; en bas (« bêta »):
IRIjk + r (BP)



CARTE 8: DEUX CARTES DE SIMILARITÉ RELATIVES AU P.-ALF 976 (BOURG-SAINTE-PIERRE, VALAIS, SUISSE)

Corpus: 1 681 CT; algorithme d'intervallisation : MEDMW 6-tuple

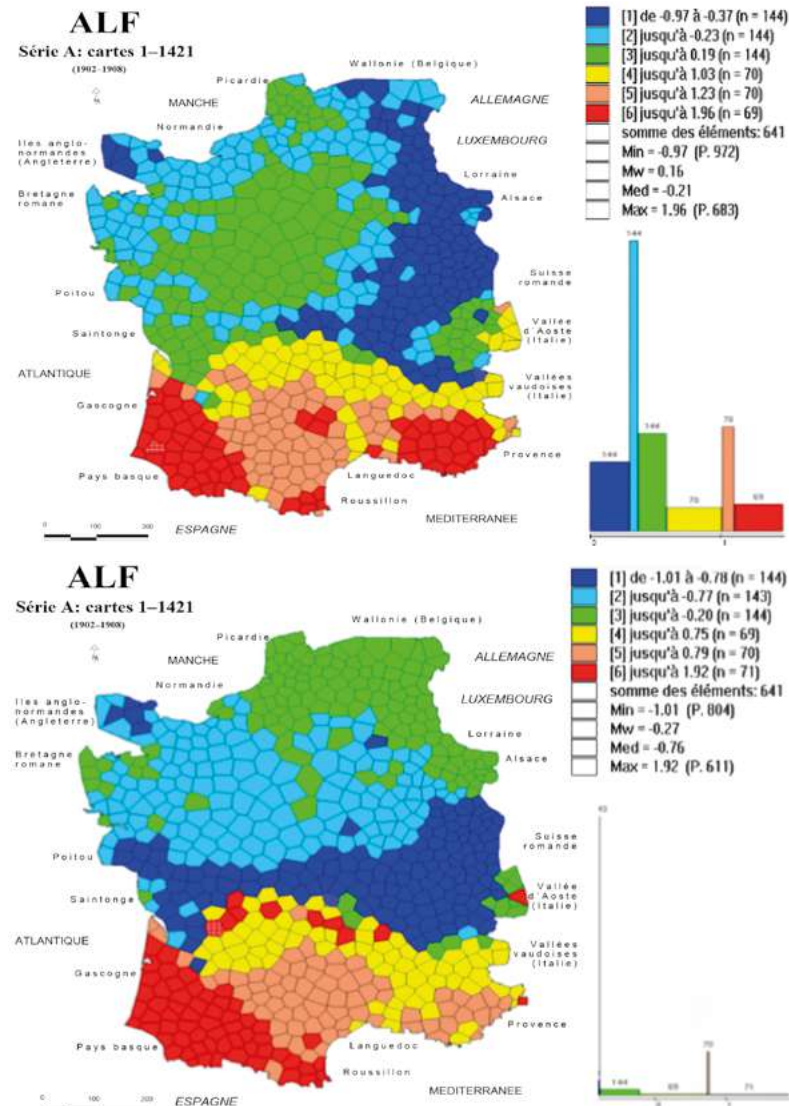
Calcul de similarité: en haut (« alpha »): IRIjk; en bas (« bêta »): IRIjk + r (BP)



CARTE 9: DEUX SYNOPSES DES SCORES DU « COEFFICIENT D'ASYMÉTRIE DE FISHER » (CAF)

Corpus: 1 681 CT; **algorithme d'intervallisation :** MEDMW 6-tuple

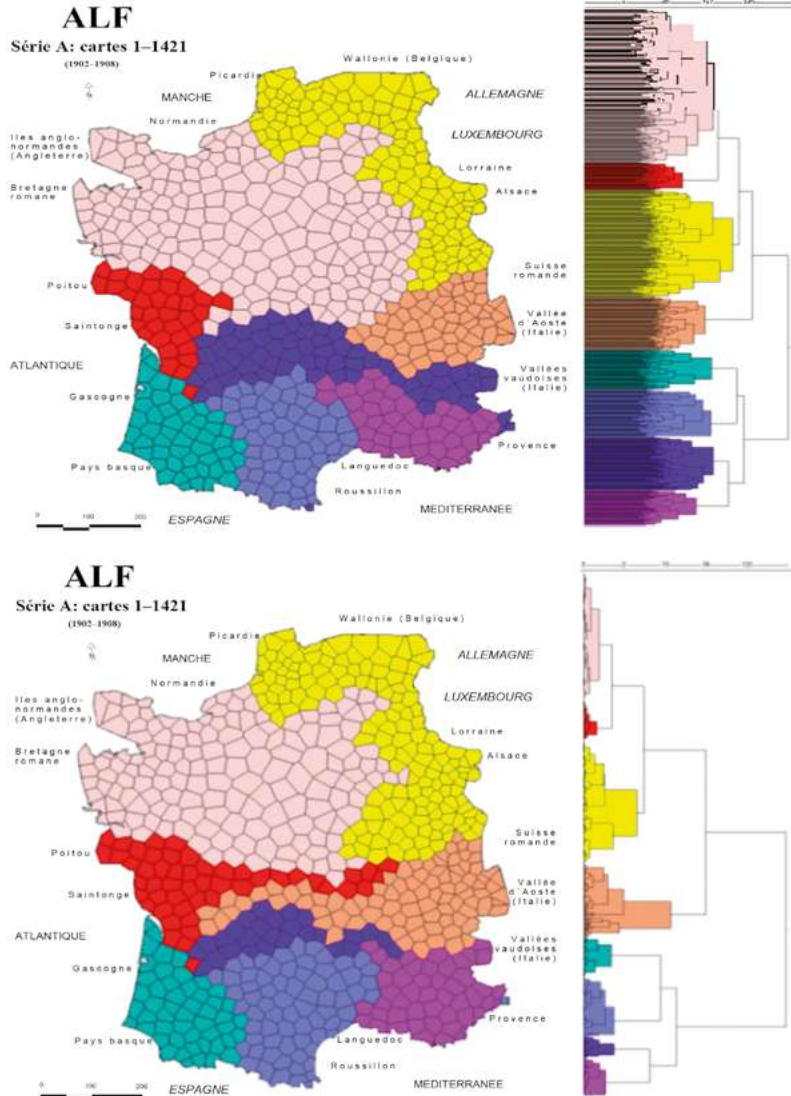
Calcul de similarité : en haut (« alpha »): IRIjk; en bas (« bêta »): IRIjk + r (BP)



CARTE 10: DEUX «CLASSIFICATIONS HIÉRARCHIQUES ASCENDANTES» (CAH) SELON LA MÉTHODE DE JOE WARD JR.

Corpus: 1 681 CT; Nombre des dendrèmes et chorèmes coloriés: 10

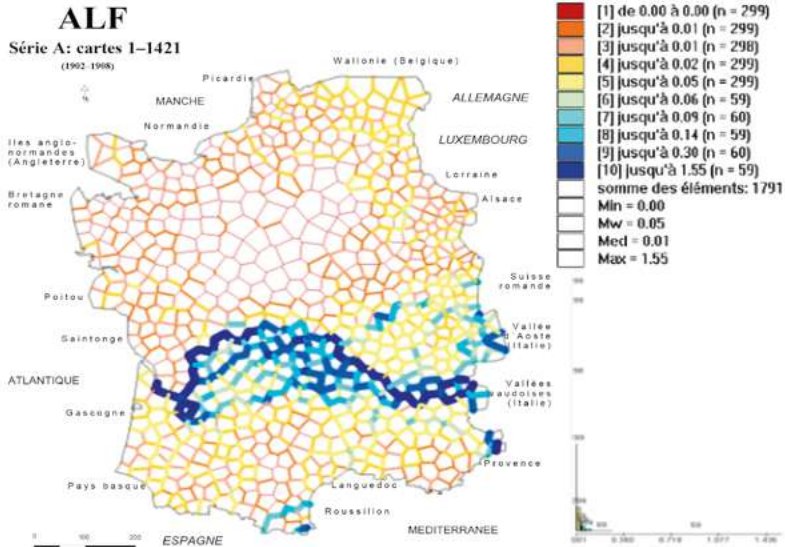
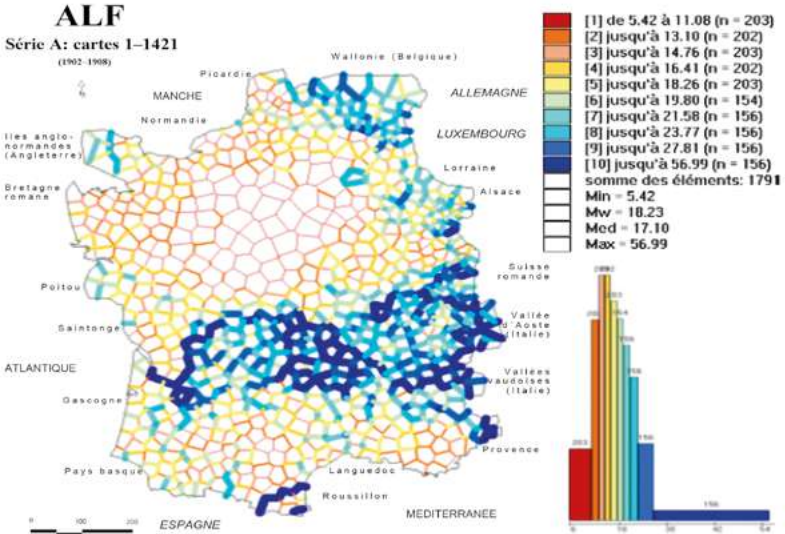
Calcul de similarité: en haut (« alpha »): IRIjk; en bas (« bêta »): IRIjk + r (BP)



**CARTE 11 : DEUX ANALYSES INTERPONCTUELLES EN FONCTION
DISCRIMINATOIRE (→ CARTES ISOGLOTTIQUES)**

Corpus: 1 681 CT; algorithme d'intervallisation : MEDMW 10-tuple

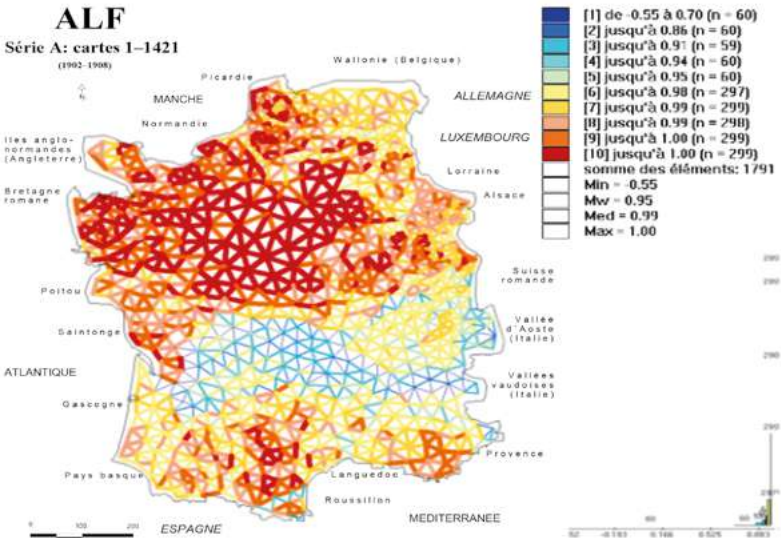
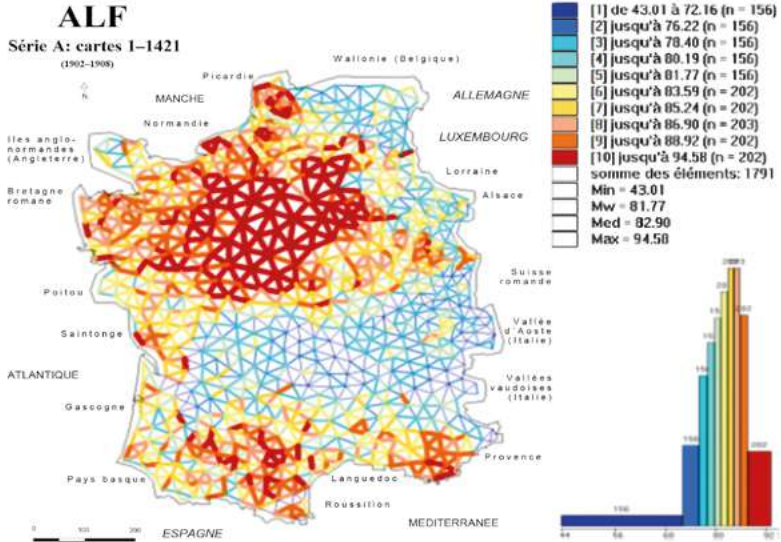
**Calcul de similarité : en haut (« alpha »): IRIjk; en bas (« bêta »):
IRIjk + r (BP)**



CARTE 12: DEUX ANALYSES INTERPONCTUELLES EN FONCTION COMMUNICATIVE (→ CARTES à RAYONS)

Corpus: 1 681 CT; algorithme d'intervallisation : MEDMW 10-tuple

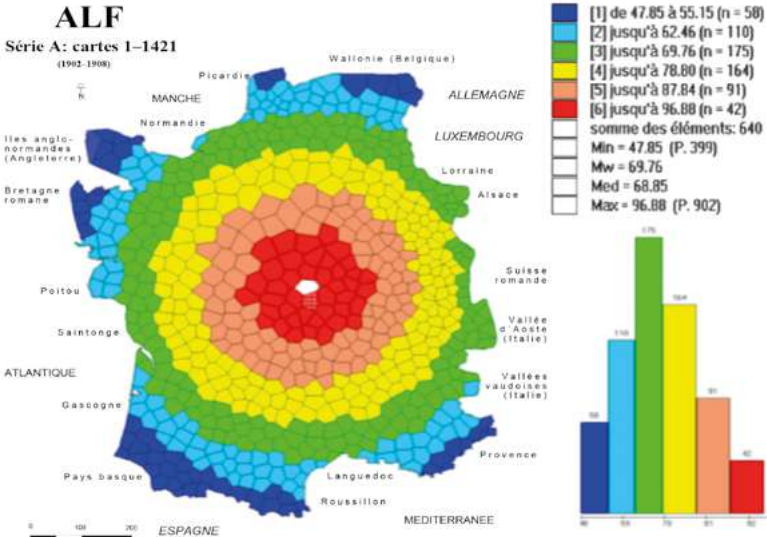
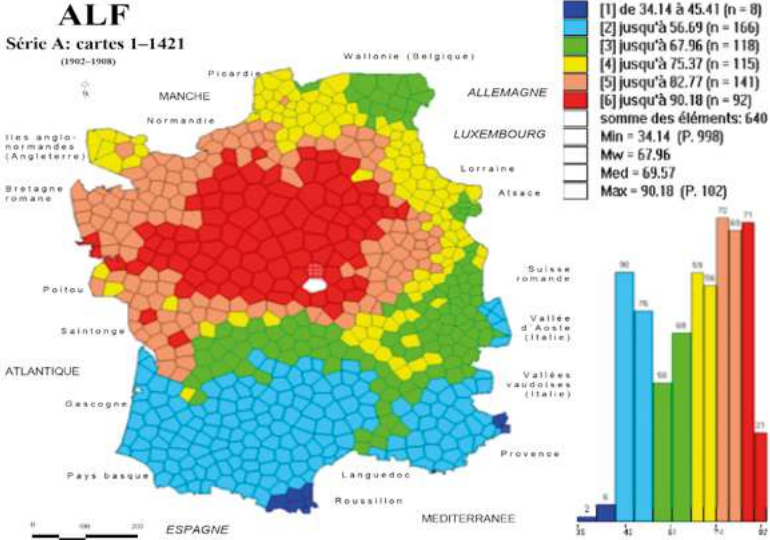
Calcul de similarité: en haut (« alpha »): IRIjk; en bas (« bêta »): IRIjk + r (BP)



CARTE 13: DEUX CARTES DE SIMILARITÉ RELATIVES AU P.-ALF 1 (MARCIGNY, NIÈVRE)

Corpus: 1 681 CT); algorithme d'intervallisation : MINMWMAX 6-tuple

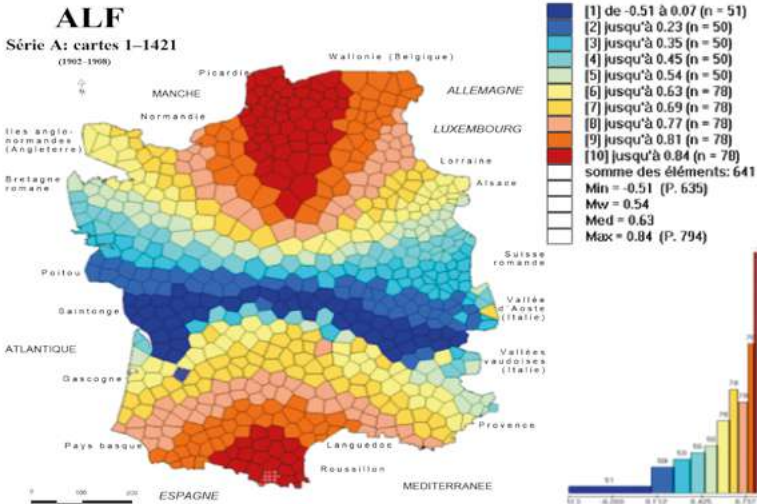
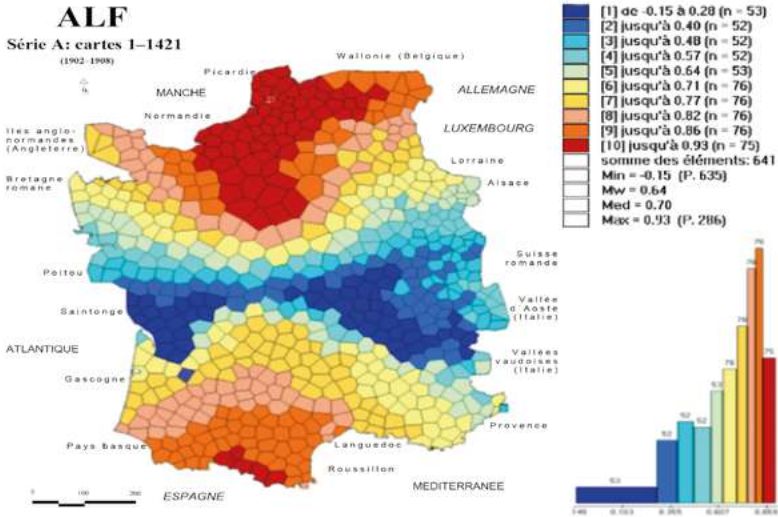
Calcul de similarité/proximité: en haut (« alpha »): IRIjk; en bas: proxPytho (selon le théorème de Pythagore : $a^2 + b^2 = c^2$)



CARTE 14: DEUX CARTES À CORRÉLATIONS (ENTRE LA SIMILARITÉ LINGUISTIQUE ET LA PROXIMITÉ GÉOGRAPHIQUE PROXPYTHO [= 100 - DISTPYTHO])

Corpus (en haut et en bas) : 1 681 CT; algorithme d'intervallisation : MEDMW 10-tuple

Calcul de similarité : en haut (« alpha »): IRIjk; en bas (« bêta »): IRIjk + r (BP)



Sobre os Autores

Felício Wessling Margotti

Doutor em Letras pela Universidade Federal do Rio Grande do Sul (UFRGS), 2004. Professor no Programa de Pós-graduação em Linguística da Universidade Federal de Santa Catarina (UFSC). Diretor Científico do Projeto Atlas Linguístico do Brasil (Regional Santa Catarina) desde 2007.

FERNANDO BRISSOS

Doutorado em Linguística pela Universidade de Lisboa, 2011. Professor na Pós-Graduação na mesma universidade. É membro do Centro de Linguística da Universidade de Lisboa, do Centro de Linguística da Universidade de Zurique, dos Grupos de Pesquisa em Dialectologia e Geolinguística da Universidade de São Paulo e em Sociogeolinguística da Universidade Federal de Uberlândia.

HANS GOEBL

Doutorado em Filologia Românica/Linguística na Universidade de Regensburg (Baviera), 1980. Foi professor na Universidade de Salzburgo (Áustria), chefe do laboratório de pesquisa para dialetometria e do grupo de trabalho de linguística variacional românica na Universidade de Salzburgo. Em 2009 recebeu a Ordem do Mérito da Província do Tirol do Sul e em 2013 o Prêmio Wilhelm Hartel da Academia Austríaca de Ciências.

JACYRA ANDRADE MOTA

Doutora em Letras pela Universidade Federal do Rio de Janeiro (UFRJ), 2002. Participou da primeira equipe de inquiridores do Atlas Prévio dos Falares Baianos – APFB. Professora na Pós-Graduação da Universidade Federal da Bahia (UFBA). Bolsista de Produtividade em Pesquisa do CNPq, Diretora Executiva do Projeto Atlas Linguístico do Brasil (Regional Bahia), 1996-2017 e Diretora-Presidente desde 2018.

MARIA DO CARMO SÁ TELES DE ARAÚJO ROLO

Doutorado em Língua e Cultura pelo Programa de Pós-Graduação em Língua e Cultura da Universidade Federal da Bahia (2016). Professora substituta na Universidade Federal da Bahia e professora efetiva da Secretaria da Educação do Estado da Bahia. Pesquisadora do Atlas Linguístico do Brasil (ALiB).

MARIA LUIZA DE CARVALHO CRUZ-CARDOSO

Doutorado em Letras Vernáculas pela Universidade Federal do Rio de Janeiro – UFRJ, 2004. Professora da Universidade Federal do Amazonas – UFAM, autora do Atlas Linguístico do Amazonas – ALAM.

SILVIA FIGUEIREDO BRANDÃO

Doutora em Letras Vernáculas pela Universidade Federal do Rio de Janeiro (UFRJ), 1988. Professora no Programa de Pós-graduação em Letras Vernáculas na mesma IES. Bolsista de Produtividade em Pesquisa do CNPq. Autora de *A geografia linguística no Brasil* (São Paulo: Ática, 1991) e organizadora de *Duas variedades africanas do Português: aspectos fonético fonológicos e morfossintáticos* (Blucher, 2018).

O HOMENAGEADO



João António das Pedras Saramago

É licenciado em Filologia Românica, atuando como pesquisador e professor do Centro de Linguística da Universidade de Lisboa (CLUL) desde 1978. Com a tese *A ilha do Corvo – alguns aspectos linguísticos*, aprovada com distinção e louvor, a Dialectologia estabeleceu-se em definitivo na vida de João Saramago.

Atualmente atua no Atlas Linguístico do Litoral Português (ALLP), Atlas Linguístico-Etnográfico de Portugal e da Galiza (ALEPG), Atlas Linguistique Roman (ALiR), Atlas Linguarum Europae (ALE), e Tesouro do Léxico Patrimonial Galego e Português (TLPGP), como coordenador ou diretor.

Este livro, que trata dos estudos dialetais brasileiros e portugueses, é uma justa e merecida homenagem ao dialetólogo e geolinguista João Saramago, por ocasião do VI Congresso Internacional de Dialectologia e Sociolinguística – VI CIDS, realizado em 2022, na cidade de Campo Grande – MS, Brasil.

Na realidade, quando se fala de léxico, há que ter em conta o seguinte: (i) muito raramente o léxico tem servido de base para uma tipologia de divisão linguística; (ii) sobretudo, têm sido utilizados traços fonéticos para proceder a divisões e classificações dos dialectos de uma determinada língua.

Ora, é um facto conhecido e aceito pelos estudiosos que fenómenos de índole lexical se diferenciam, na sua essência, dos fenómenos fonéticos pela simples razão de que aqueles devem ser interpretados com base numa análise que envolve factores de índole extralinguística, nomeadamente a história e a etnografia. Esta é a principal razão pela qual, normalmente, os traços fonéticos são escolhidos na delimitação e classificação de dialectos ou de variedades linguísticas.

(SARAMAGO, J. A. das P. Tão longe e tão perto. Tão perto e tão longe. In: ALTINO, F. C. (org.). *Múltiplos olhares sobre a diversidade linguística: uma homenagem à Vanderci de Andrade Aguilera*. Londrina: Midiograf, 2012. p. 147-163).

ISBN 978-65-89995-13-5



9 786589 995135