

ESTHER BAIWIR (ed.), *Les atlas linguistiques galloromans à l'heure numérique: projets et enjeux*, Villeneuve-d'Ascq (Centre d'études médiévales et dialectales de l'Université de Lille 3) 2020, 184 p. (*Bien dire et bien apprendre* 35).

Le volume réunit huit contributions présentées lors de la journée d'étude intitulée «Quel dialogue numérique entre les atlas linguistiques galloromans?» et organisée à l'Université de Lille en septembre 2019. L'avant-propos rédigé par E. Baiwir (3-4) précise les deux objectifs poursuivis lors de la journée, qui s'appliquent *a fortiori* à ce numéro thématique de la revue *Bien dire et bien apprendre*: «faire le point sur les projets numériques galloromans liés aux atlas linguistiques et [...] envisager l'interopérabilité de leurs données» (3). Cette *interopérabilité*, notion transversale à l'ensemble du volume, est garantie «lorsque les données ou les outils de la recherche venant de ressources différentes sont capables de s'intégrer et de fonctionner ensemble» (Colcuc et Mutter, 133). Ceci implique que les atlas soient considérés comme des bases de données rendues accessibles par la numérisation et qui «se réf[èrent] les unes aux autres» (135). Pour atteindre ces objectifs, le dialectologue doit se munir d'*outils* en phase avec l'évolution des modalités de la recherche, ces outils étant vus comme autant de pistes qui «permettront d'*exploiter* et de valoriser ce que [les dialectologues] ont amassé pendant un demi-siècle» (7), si l'on reprend le vœu formulé dans le texte d'ouverture par le regretté Fernand Carton à qui ce numéro est dédié. *Interopérabilité* (ou, de façon plus lâche, *mise en réseau*), *outils* et *exploitation* seront les trois clés de lecture à partir desquelles nous présenterons ce volume thématique.

Les quatre contributions que nous épinglons en premier lieu explorent les possibilités d'une mise en réseau des ressources atlantographiques et lexicographiques existantes, à travers la présentation de projets de numérisation concrets ou à partir de réflexions et d'études de cas.

La contribution de M. Glessgen et M. Sauzet («La trajectoire et l'exploitation lexicale des Nouveaux atlas linguistiques de la France», 9-45) ouvre opportunément le volume en retraçant la genèse et l'histoire des *NALF* dans le panorama des études dialectales de la Galloromania (10-14). Deux cartes récapitulatives (17, 23) et un tableau détaillé en annexe (38) mettent à jour les travaux de J. Séguy¹ et de G. Brun-Trigaud² pour livrer au lecteur une vue d'ensemble de l'état d'avancement, de numérisation et d'indexation des *NALF*, cette dernière opération conditionnant les possibilités d'exploitation (lexicale surtout) de ces immenses bases de données (20-28). Les projets géolinguistiques récents, y compris aux marges de la Galloromania et dans les territoires d'outre-mer, ne sont pas oubliés dans ce vaste état des lieux. Les auteurs concluent en appelant de leurs vœux un «chantier global commun» (29) qui réunirait en une seule interface les données de tous les atlas galloromans. L'atteinte de cet idéal exigerait une coordination à plusieurs niveaux (structures linguistique, informatique et institutionnelle) qui peut s'inspirer des projets existants: le *Thesaurus Occitan (THESOC)* et le projet *Verba Alpina* constituent déjà des exemples fructueux d'une mise en réseau des données dialectales et d'une collaboration entre dialectologues.

Un autre projet de ce type est en cours: l'*Atlas pan-picard informatisé (APPI)*, présenté par E. Baiwir et C. Kaisin (47-59). Cette nouvelle ressource atlantographique pour un domaine linguistique picard débarrassé des «frontières administratives et méthodologiques entre la France et la Belgique» (50) permet déjà la consultation de l'*ALPic* au format image et souhaite rendre interopérables les données de l'*ALPic*, de l'*ALF* et de l'*ALW* (dont les instigatrices de l'*APPI* peuvent désormais profiter de la récente mise en ligne³). Une mise en réseau partielle

¹ SÉGUY, J. 1973: «Les Atlas linguistiques de la France par régions», *Langue française* 18: 65-90.

² BRUN-TRIGAUD, G. 2016: «Vers un renouveau des atlas linguistiques régionaux», *Géolinguistique* 16: 7-20.

³ Depuis 2020 et la publication de ce volume thématique, les dix tomes de l'*Atlas linguistique de la Wallonie* sont disponibles en libre accès sur le site du projet: <https://alw.uliege.be>. Pour un aperçu des étapes de l'informatisation

des trois atlas est déjà opérationnelle et accessible par le biais d'un tableau et de cartes dynamiques. L'*APPI* bénéficie de l'expertise de l'équipe de *VerbaAlpina* (Colcuc et Mutter, 131-146) et utilise avec succès le modèle informatique développé dans le cadre de ce projet, ce qui démontre, si besoin est, les avantages d'une collaboration étroite entre groupes de travail.

P. Brasseur et E. Baiwir (61-70) explorent dans la troisième contribution les pistes d'une numérisation de l'*Atlas linguistique normand*, récemment achevé, et interrogent la possibilité d'une intégration des données de cet atlas régional dans la nomenclature de l'*APPI*, par l'étude sémantique d'un échantillon composé de 205 notions. Les résultats montrent qu'un rapprochement est possible et souhaitable, moyennant un enrichissement des données de l'*ALN* par lemmatisation et étymologisation.

Enfin, M. Robecchi («La place des *Atlas* dans la rétro-conversion du *FEW*», 147-167) examine les possibilités et les apports d'une mise en réseau du *FEW* et des *NALF* (en particulier, les données regroupées dans le *THESOC*) en cours de numérisation. L'auteur analyse les échanges entre le tome 16 du *FEW*, l'*ALLY* et l'*ALMC* et détaille quelques propositions concrètes pour optimiser les renvois entre les différents ouvrages à partir des interfaces de consultation. Dans la même optique, la conclusion de l'article livre une série de *minima etimologica* envisagées comme autant d'efforts indispensables à une coordination des *NALF* et du *FEW*. Si M. Robecchi montre les avantages d'une insertion systématique des *NALF* dans l'immense *thesaurus galloroman*, il souligne également l'importance d'une réflexion à mener sur l'impact de l'informatisation⁴ et sur les limites d'une centralisation des données qui risquerait d'altérer la dimension monographique du *FEW* (164).

Pour rendre accessibles et interopérables les données de la recherche géolinguistique, le dialectologue doit se doter d'outils adaptés à de telles ambitions. Deux contributions se distinguent en ce qu'elles exposent des projets liés au développement d'outils ou présentent des démarches permettant une numérisation intégrée des données de la recherche.

La contribution de G. Brun-Trigaud et al. (71-84) est consacrée au projet *ANR ECLATS*, qui réunit plusieurs équipes de linguistes et d'informaticiens autour d'un but commun: «apporter un outillage logiciel et méthodologique facilitant l'extraction, l'analyse, la visualisation et la diffusion des données contenues dans les atlas linguistiques [...], afin de permettre des recherches novatrices en dialectologie» (72). Quatre applications sont présentées en lien avec des objectifs précis: stockage, accessibilité et annotation des données dialectales de l'*ALF* (CartoDialect); création de nouveaux supports cartographiques (ShinyDialect); exploitation des données dialectales grâce à la génération de cartes interprétatives (ShinyClass) et création de cartes géolinguistiques thématiques par superposition de couches géographiques (DialectoLOD).

L'article de B. Colcuc et C. Mutter («Interopérabilité des données géolinguistiques à l'exemple du projet *VerbaAlpina*», 131-146) montre comment l'évolution des modalités de la recherche (numérisation, accessibilité, outils d'analyse, etc.) amène les chercheurs à mobiliser les quatre principes réunis sous l'acronyme *FAIR* qui orientent «le débat concernant le traitement des données, la communication et la coopération dans l'ère des médias» (132): les données de la recherche doivent être *findable* '(re)trouvables', *accessible* 'accessibles', *interoperable* 'interopérables' et *reusable* 'réutilisables'. Les auteures se focalisent ensuite sur le principe d'interopérabilité, déjà commenté en introduction de ce compte rendu, et exposent la démarche effectuée au sein de l'atlas en ligne *VerbaAlpina* pour structurer les données «de sorte qu'elles puissent être compatibles avec d'autres bases de données, afin de permettre leur interconnexion» (133). Cinq aspects de l'interopérabilité sont détaillés, du traitement structuré

de l'œuvre, cf. BOUTIER, M.-G./RENDERS, P. 2021: «L'*Atlas linguistique de la Wallonie* (ALW): un projet centenaire en libre accès», *Wallonnes* 2/2021: 9-17.

⁴ D'après RENDERS, P. 2009: «Des dangers de l'informatisation d'un document: le cas du *FEW*», *MethIS 2 (Pratiques du document, Méthodes et Interdisciplinarité en Sciences humaines)*: 179-195.

des données à l'accès aux données primaires et aux métadonnées. Ce protocole a déjà été adopté dans le cadre du projet *APPI* et est envisagé pour la numérisation de l'*ALN*: l'appel de Colcuc et Mutter à la constitution d'un stock de données extensible nourri par les projets individuels semble inspirer les dialectologues.

Pour illustrer l'intérêt d'une mise en réseau des projets géolinguistiques, deux contributions montrent plus particulièrement ce que l'interopérabilité et la numérisation peuvent apporter à l'exploitation des matériaux.

P. Boula de Mareüil et al. (85-108) présentent et analysent les données de l'atlas sonore des langues et dialectes de Belgique pour montrer «la richesse du patrimoine linguistique de la Belgique» (103). Directement élaborée dans un format numérique à partir de données obtenues par crowdsourcing, cette nouvelle ressource s'intègre dans un réseau d'enquêtes plus large qui vise à enregistrer et à reproduire la fable d'Ésope «La bise et le soleil» dans plusieurs langues du monde. La contribution présente la méthodologie adoptée pour récolter les données sonores en Belgique romane⁵ avant de consacrer une large section aux analyses lexicales, phonétiques et morphosyntaxiques réalisées à partir des matériaux recueillis. Les auteurs procèdent à une comparaison systématique des données de l'atlas sonore avec celles de l'*ALW*, donnant ainsi une perspective diachronique à l'analyse.

Enfin, la contribution de G. Raimondi («ALEPO et APV: la contribution de l'Italie à l'étude de la *Galloromania peripherica*», 109-130) expose clairement les apports scientifiques d'une prise en compte de la *Galloromania peripherica* italienne, zone de transition au croisement des familles dialectales gallo-romane et gallo-italique, pour l'étude des dynamiques géolinguistiques et de la délimitation des ensembles dialectaux. En rassemblant les données de l'*Atlante Linguistico ed Etnografico del Piemonte Occidentale* (ALEPO⁶) et de l'*Atlas des Patois Valdôtains* (APV), l'auteur réévalue l'opposition entre «centre de diffusion des innovations» et «périphérie conservatrice» à travers l'étude de quelques cas lexicaux (*aune*, *frêne*, *donner*) et phonétique (/a/ atone final dans les continueurs de lat. FOLIA). Si la contribution est moins directement en lien avec la thématique de la numérisation des atlas, elle montre nettement l'intérêt de la mise en réseau des données atlantographiques (ici ALEPO, APV et ALF, parfois AIS) et la nécessité d'un accès numérique direct aux atlas pour exploiter les matériaux.

Dans l'ensemble, le volume répond bien aux deux objectifs de la journée d'étude à l'origine des contributions: il donne au lecteur un état des lieux précis quant à la numérisation des atlas linguistiques galloromans et apporte des pistes méthodologiques concrètes quant aux possibilités d'une intégration des données dialectales au sein de projets communs. L'enjeu de l'accessibilité numérique des atlas et de l'interopérabilité de leurs données est plusieurs fois mis en avant et relève, des mots des contributeurs (Baiwir et Kaisin, 47; Glessgen et Sauzet, 32), de la *survie* des études dialectales parmi les disciplines de la romanistique. Dès lors, on ne peut qu'espérer que ce volume stimulera d'autres collaborations entre dialectologues, en vue d'«une exploitation commune et harmonieuse» (Glessgen et Sauzet, 10) du patrimoine linguistique contenu dans ces objets «magiques» (Baiwir, 4) que sont les atlas.

Léonore Dubru (ULiège)

⁵ Le nombre de participants romans (32) et de points d'enquête a considérablement augmenté depuis la publication de ce volume, si l'on en croit l'interface de consultation présentant la carte interactive.

⁶ Nous actualisons ici l'adresse d'accès au site, renseignée dans la note n°5 de l'article: <https://www.alepo.unito.it>.