

Esther Baiwir / Pascale Renders

Vingt minutes en Utopie : l'ALWi

(conférence présentée le 16 mai 2013 dans le cadre des
Journées liégeoises de Traitement des Sources Galloromanes
de l'Université de Liège)

Date de mise en ligne : 23 décembre 2013

Article à citer comme suit :

Baiwir, Esther & Renders, Pascale (2013). « Vingt minutes en Utopie : l'ALWi », in : Baiwir, Esther & Renders, Pascale (éds), Actes de la 5^e Journée liégeoise de Traitement des Sources Galloromanes (Liège, 16 mai 2013), Liège, Université de Liège, publication électronique (http://lingwa.philo.ulg.ac.be/trasogal/baiwir_renders_2013.pdf), 9 pages.

Vingt minutes en Utopie : l'ALWi

Esther Baiwir et Pascale Renders
Chargées de recherches FNRS

Introduction

L'objectif de l'*Atlas linguistique de la Wallonie* (ALW) est d'éditer les matériaux linguistiques recueillis par Jean Haust et ses successeurs, dans le cadre d'une enquête de grande ampleur menée en Belgique romane entre 1924 et 1959. Éditer consiste ici à donner à voir et à comprendre la matière dialectale. La première étape qui sera exposée ci-dessous est appliquée depuis près d'un siècle par les rédacteurs successifs ; il s'agit du passage des fiches « papier » aux volumes physiques de l'ALW, tels qu'on peut les consulter à travers une lecture linéaire ou des entrées plus ponctuelles (table des matières, index divers). Mais c'est une projection vers l'avenir du projet que nous envisagerons ensuite, qui permettra un jour de passer des volumes traditionnels à une version interactive de l'ALW, l'ALWi (pour *Atlas linguistique de la Wallonie informatisé*).

L'édition « traditionnelle » de l'ALW : quelques jalons

Le parcours menant des fiches manuscrites aux volumes physiques de l'ALW peut être subdivisé en trois phases importantes. Il s'agit d'abord de sélectionner les matériaux pertinents et d'en définir la valeur (a), ensuite de retracer leur histoire (b), enfin de les présenter le plus lisiblement possible (c).

a. La première de ces phases se situe en amont de la conception de l'unité « notice ». Il convient d'abord de découvrir les matériaux, constitués de paquets de fiches manuscrites. Les données sont ensuite minutieusement reportées sur une carte de la Wallonie. Cette étape permet au rédacteur de visualiser l'ensemble des réponses et, dans les cas favorables, de dégager des aires homogènes. C'est tout le « pari » de la géolinguistique : reconstruire l'histoire en étudiant une situation synchronique à travers sa variation géographique. Il convient ensuite de construire le tableau des formes de la notice. Les éventuels « bruits », tels qu'un contexte phrastique envahissant, un gallicisme d'enquête, une imprécision dans les notations, doivent être identifiés et neutralisés.

Dans cette étape du travail, il s'agit de déshabiller et d'isoler nos matériaux, afin qu'ils s'affranchissent de leur statut d'échantillon de parole pour nous dire quelque chose de la langue.

b. Une fois les matériaux vérifiés et classés, il convient d'en retracer l'évolution. Pour ce faire, le dépouillement des divers dictionnaires dialectaux et la consultation des ouvrages de référence tels que les dictionnaires étymologiques fournissent de précieux renseignements. C'est à ce niveau que s'instaure un dialogue avec le *Französisches etymologisches Wörterbuch*

(ou FEW) de Wartburg, afin d'ancrer les données wallonnes dans le paysage linguistique de la Galloromania.¹ Ces démarches aident à la rédaction de l'introduction de la notice et des notes.

Par ailleurs, le grand nombre de types lexicaux présents dans un volume de l'ALW a pour corollaire l'impossibilité de l'exhaustivité; on regrettera sans doute l'absence de certaines étymologies, ou de certaines données livresques, qui auraient pu être mentionnées avec profit - nous y reviendrons.

c. Cette étape achevée, il est possible de définir exactement le schéma de présentation des données. Une introduction est rédigée, qui vise à présenter synthétiquement l'histoire du mot et les formes reprises dans le tableau. Les notes fournissent des informations ponctuelles ; lorsque la notice le permet, le rédacteur confectionne également une carte, reprenant sous forme de symboles tout ou partie des types ou des formes présents dans la notice.

Si cette méthode est celle qui est concrètement appliquée depuis le début de l'entreprise, quelques nouveautés ont vu le jour, qui allègent la rédaction au plan pratique. Ce fut tout d'abord la composition informatique des notices dès le tome 15, puis pour le tome 6, l'informatisation des cartes. Un fond de carte numérisé a été créé et des symboles dessinés sur ordinateur. Cette petite révolution, pour anodine qu'elle puisse paraître, est intéressante, car elle permet aux rédacteurs de visualiser plusieurs versions d'une même carte afin de sélectionner la plus « parlante ».

Dans la logique de cette évolution, une réflexion s'est fait jour² afin d'examiner la faisabilité d'une informatisation globale de l'ALW et ce, à deux niveaux : aide à la rédaction d'abord (mais cet aspect ne sera pas abordé ici) ; aide à la consultation ensuite. Il s'agirait de permettre une consultation en ligne, enrichie par rapport aux versions « papier ».

Vers une édition numérique

Comme l'écrivait Louis Remacle dans l'introduction du premier volume, l'*Atlas linguistique de la Wallonie*, ce sont « des cartes et du texte » (ALW 1 : 17). La formule nous incite à considérer en deux temps l'éventuelle informatisation de l'ALW.

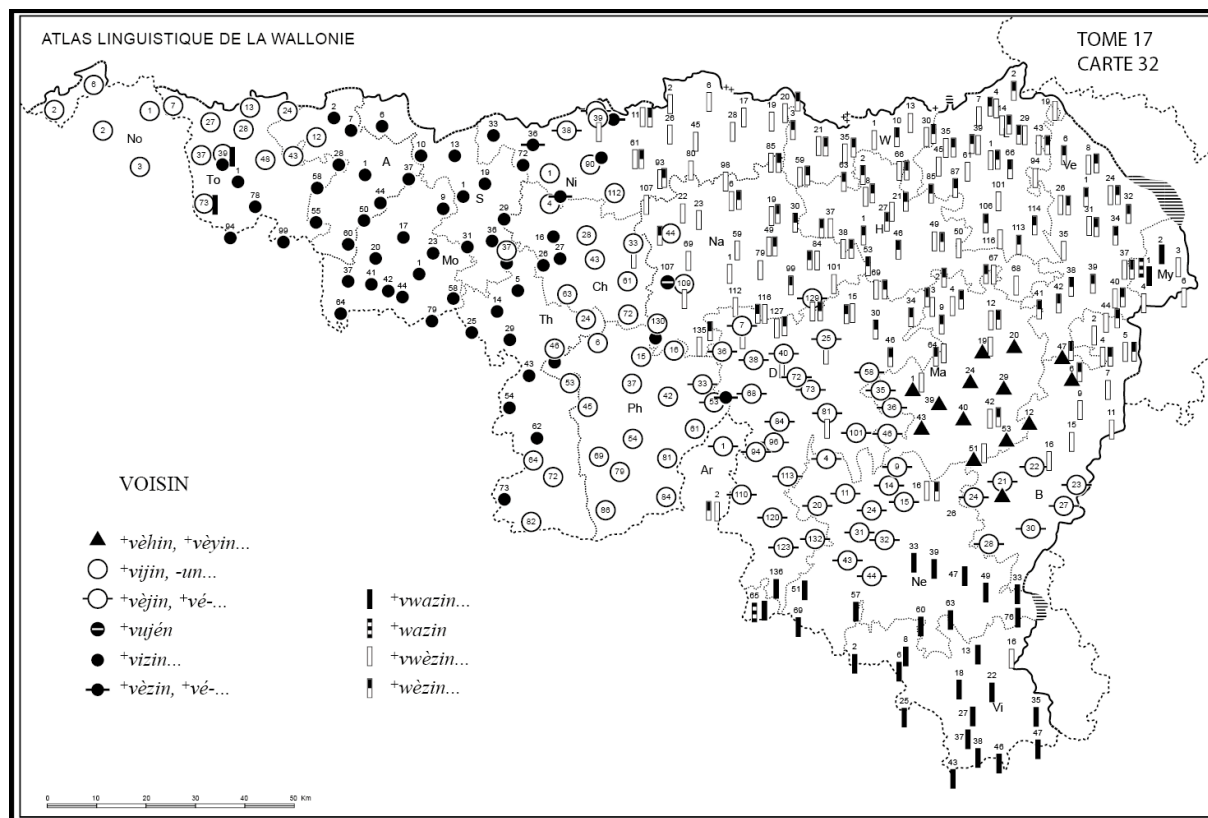
1. Des cartes...

L'apport des outils informatiques au niveau des cartes constituerait principalement en la possibilité d'actualiser, à la demande, des cartes mettant en évidence un phénomène choisi par l'utilisateur. Ce rêve, que font également les responsables de l'informatisation du *Französisches Etymologisches Wörterbuch* ou FEW (voir Renders 2011 : 64-65³), peut être illustré au départ de la carte VOISIN de l'ALW 17. La voici, telle qu'elle se présente dans le volume physique.

¹ V. Boutier 2008 : 309-110.

² Initiée par Nicolas Mazziotta en prolongement d'investigations informatiques dans le cadre du projet *Khartès* (Mazziotta 2008), cette réflexion se poursuit sur la base d'une réflexion théorique de M.-G. Boutier quant aux diverses « relations » régissant les notices de l'ALW (BOUTIER 2008). V. aussi Baiwir / Renders (sous presse).

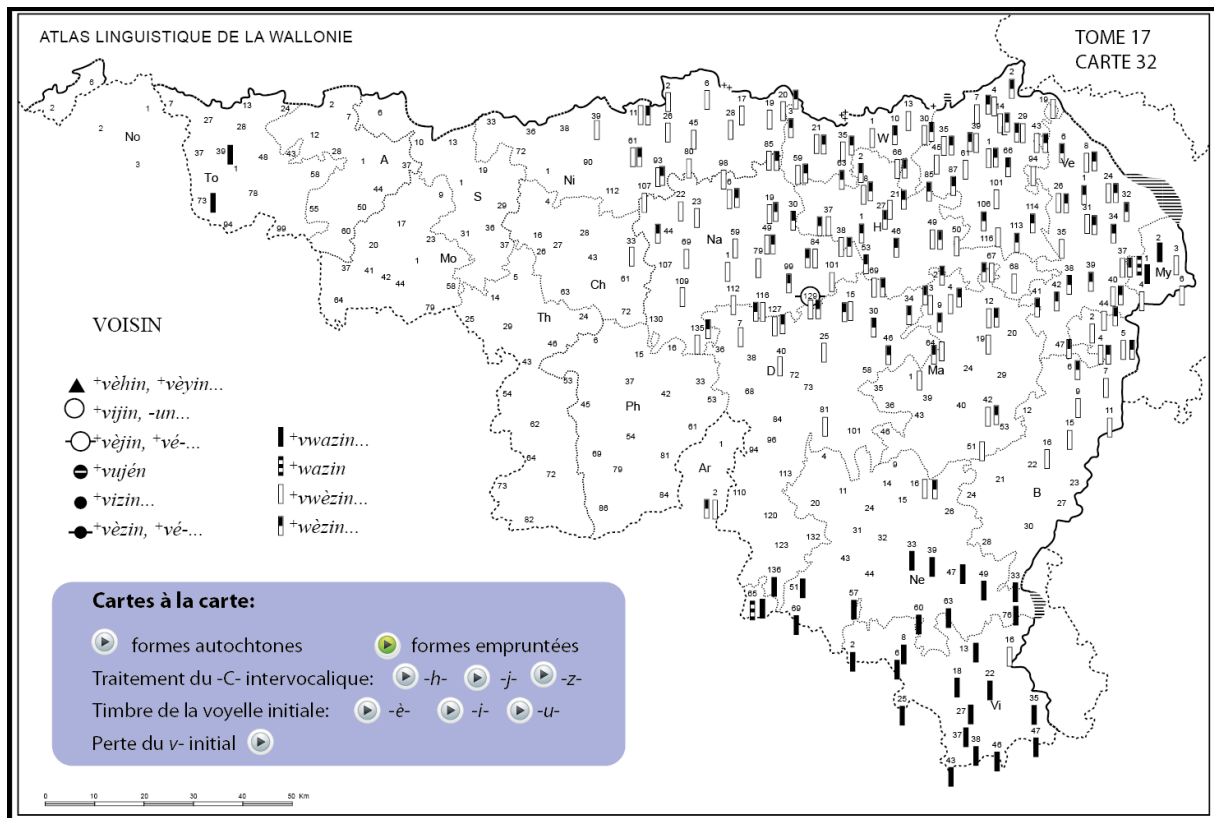
³ Le sujet a également été évoqué en 2009, lors de la première journée liégeoise de *Traitement des sources galloromanes* (Yan Greub/Julia Alletsgruber/Dimitri Chihai, *Aides informatiques à la rédaction et à la lecture du Französisches Etymologisches Wörterbuch (FEW)*, présentation orale, Trasogal I, Liège, 2009).



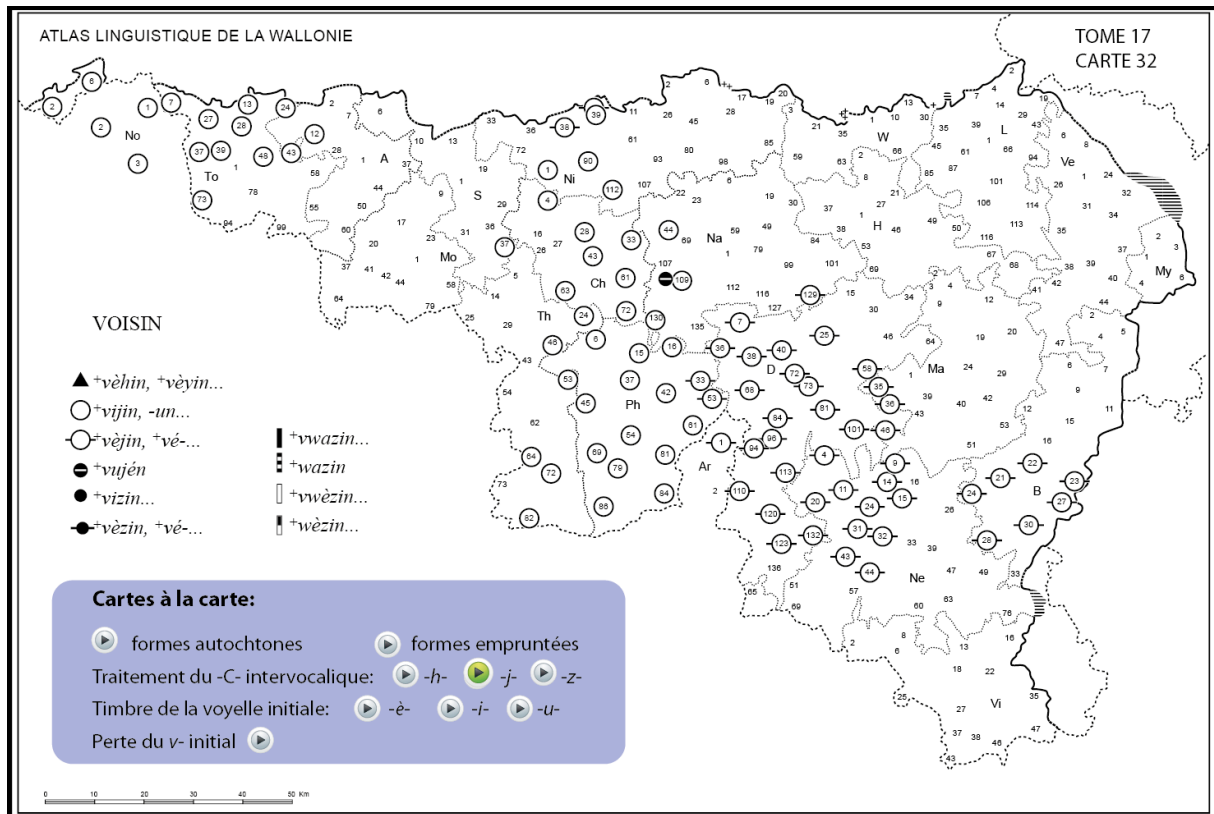
La carte VOISIN de l'ALW 17.

À la lecture du commentaire (voir ALW 17, not. 83), deux enseignements semblent prépondérants : d'une part, le remplacement en de nombreux points des formes autochtones par des formes empruntées à la langue-toit qu'est le français, parfois avec une adaptation phonétique secondaire des lexèmes empruntés ; d'autre part, les isoglosses naturelles des phénomènes phonétiques que sont le traitement du -C- intervocalique et l'aboutissement de la voyelle initiale.

Une carte dynamique pourrait permettre de visualiser ces deux phénomènes. Pour ce faire, deux options nous semblent envisageables. La première, « intelligente », serait de fournir à l'utilisateur différentes options d'affichage en fonction de critères linguistiques. L'intérêt de chacune des cartes ainsi disponible aurait été mesuré et évalué par un linguiste (en l'occurrence, le rédacteur de l'ALWi) et la légende de cette carte interactive reflèterait immédiatement cette analyse. Ainsi, dans notre exemple, il serait pertinent de proposer l'affichage exclusif des formes empruntées, qui seraient largement mises en valeur et faciliteraient l'étude de pôles de diffusion des innovations (projection 1). Ou encore, la zone des aboutissements en -j- du -C- intervocalique latin (projection 2).



Projection 1 : les formes empruntées

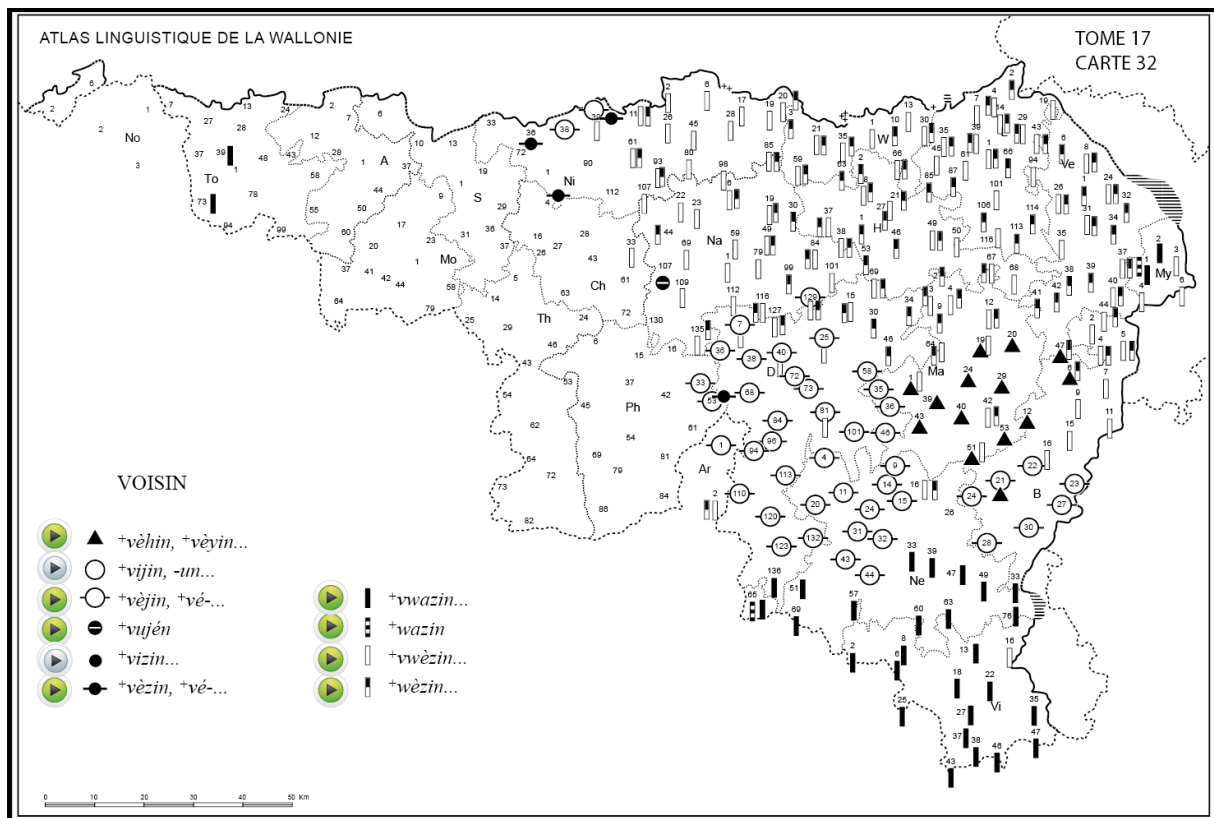


Projection 2 : les formes empruntées

De telles cartes, pour intéressantes qu'elles soient, demandent un travail considérable et conservent les limites que sont l'imagination et l'inventivité de leur concepteur. Un utilisateur qui désirerait étudier un seul type lexical ou seule forme ne pourrait afficher exactement la carte qui lui convient. En outre, si le cas de VOISIN se prête assez bien à l'exercice, d'autres cartes ne permettent pas si facilement d'isoler et de présenter des critères d'affichage lisibles.

En revanche, si cette option était choisie, il serait possible d'aller plus loin et de superposer des données issues de cartes différentes. Par exemple, reproduire sur la carte VOISIN les isoglosses du traitement du -S- de MANSIONE (ALW 1, c. 56 MAISON), opposant des aboutissements en *-h-*, en *-z-* et en *-j-*, permettrait par exemple de voir que d'une part, les isoglosses se superposent et que, d'autre part, celles-ci n'ont pas eu d'impact sur l'extension de l'emprunt.

La seconde option, techniquement plus simple mais scientifiquement moins fondée, consisterait à ajouter un bouton « cliquable » à chaque symbole, que l'on pourrait activer ou désactiver à volonté. Dès lors, l'utilisateur serait seul responsable de la carte qu'il actualise et de la pertinence de celle-ci. Dans l'exemple ci-dessous (projection 3), on a occulté les formes en *vij-* et en *viz-*, ce qui peut être pertinent pour étudier le traitement de la voyelle initiale.



Projection 3 : l'option automatique

Cette solution comporte néanmoins quelque danger, puisque tout est alors possible. Des cartes totalement aberrantes pourraient être générées et mener l'utilisateur à des déductions qui le seraient tout autant. Cet abandon de la responsabilité à l'utilisateur est tout à fait contraire à la philosophie du projet, qu'a louée l'ensemble de la communauté des linguistes et des dialectologues romans. En effet, la caractéristique la plus spécifique de l'ALW est sa volonté d'analyse et d'accompagnement du lecteur dans la compréhension des matériaux.

En revanche, cette option présente l'avantage de la simplicité (pour les concepteurs) et de la liberté (pour les utilisateurs).

2. ... et du texte

Au niveau des textes, l'apport des outils informatiques consisterait en la possibilité (a) d'interroger l'ALW en fonction de critères divers, tels que l'étymologie des lexèmes ou leur localisation ; (b) de faciliter la lecture des notices par la création de liens hypertextuels internes ou externes à l'Atlas.

a. Possibilités d'interrogation

Les informations qui pourraient utilement être identifiées dans l'ALW informatisé et faire l'objet de recherches sont les notions, les étymons, les types et les points d'enquête. À ces quatre types d'information peut être ajoutée la possibilité de recherches « plein texte ».

Envisageons le cas d'un utilisateur qui voudrait rechercher tous les lexèmes wallons ayant un rapport étymologique avec le lexème français « jambe ». La référence (FEW 2, 111 CAMBA), le type (<jambe>) ou la suite de caractères (« jamb ») pourraient servir de voie d'accès aux données. La recherche par référence FEW donnerait les résultats suivants :

- ALW 1, not. et c. 52 JAMBE, qui présente toutes les variantes phonétiques du type <jambe> au sens 'jambe' ;
- ALW 6, not. 80 AMANDE DU NOYAU, n. 3; ALW 6, not. 123 BROU DE LA NOIX, β et n. 57, qui présentent un sens assez surprenant 'lobe ou cuisse de noix' (ayant parfois évolué en 'amande du noyau (d'un fruit)', v. ALW 6, not. 80, *H*) ;
- ALW 15, not. 29, *L* et n. 16 (*s'ingâmbier* 'trébucher') et not. 53 «ENTREFESSION», n. 8 (forme *gambiyons* 'mal aux cuisses, d'avoir marché') ;
- ALW 17, not. 26 BEBE, *J* et n. 12 (type <jambot> signifiant 'bébé') ; ALW 17, not. 33, *D* et n. 7 (collectif <jamboterie> 'marmaille') ; ALW 17, not. 34 GARÇON, *E* et not. 38 gamin, *D*, etc. (<jambot>, signifiant 'enfant') ; ALW 17, not. 99 je les ai fait DEGUERPIR, n. 65 et not. 105 SANS HESITER, SANS TORTILLER, SANS DETOUR, *R'* 12 (<jamb(i)er>).

Ce mode de recherche et de résultats n'est possible qu'à la condition d'informatiser l'ALW via, par exemple, un balisage XML et la création d'une base de données. Une telle informatisation offre de nouveaux parcours de consultation qui sont davantage transversaux, puisque des notices appartenant à des volumes distincts (et donc à des domaines notionnels distincts) sont ainsi atteintes.

Remarquons que la notice 123 du volume 6 (accessible aussi via une recherche plein texte telle que « jamb ») n'apparaîtra pas directement lors de la recherche telle qu'elle est définie ci-dessus, car cette notice ne mentionne pas explicitement la référence au FEW : seule la notice 80 du même volume permet de faire le lien entre la notice 123 et l'étymon CAMBA, via un renvoi interne mentionné dans la notice 80, *H*. Il est donc nécessaire (a) de compléter l'information en ajoutant un « étiquetage FEW » dans les balises, sans pour autant toucher au texte, et (b) de relier les notices entre elles, dans ce cas-ci via les renvois internes. Cette remarque nous conduit au deuxième point, à savoir la possibilité de créer des liens hypertextuels.

b. Création de liens hypertextuels

La création de liens hypertextuels a pour but premier de faciliter la lecture des notices. Il s'agit essentiellement d'animer ce qui pour l'instant est « mort », en permettant de nouveaux parcours de lecture. Des liens peuvent concerner plusieurs types d'information de l'ALW :

- les renvois internes (liens entre les notices de l'ALW) ;
- les renvois vers le FEW informatisé, lorsqu'il sera en ligne ;
- les références bibliographiques, qui peuvent être reliées à leur explicitation et, éventuellement, aux ressources en ligne lorsque ces dernières existent ;
- les points d'enquête, qui peuvent être reliés à leur explicitation, voire à leur localisation sur carte. Dans ce dernier cas, un travail informatique est préalablement nécessaire afin de rétablir l'implicite inhérent au mode de rédaction économique de l'ALW. Ainsi, la séquence « B 11, 12, 15 ; Ne 14, 16 » doit être développée en « B 11, B 12, B 15, Ne 14, Ne 16 » par un algorithme de reconnaissance; c'est seulement à cette étape que la base de données sera à même de restituer les équivalences <B 11 = Limerlé>, <B 12 = Nadrin>, etc.

Conclusion

L'édition traditionnelle se poursuit ; les propositions envisagées dans le cadre d'une édition numérique se situent, quant à elles, en Utopie. Elles requièrent en effet du travail, du temps et des ressources.

Un pas a néanmoins déjà été franchi dans cette direction, via la création d'un site web dédié à l'ALW, que nous avons le plaisir de dévoiler officiellement à l'occasion de cette 5^e journée TraSoGal. Cet espace en ligne (www.alw.philo.ulg.ac.be) propose actuellement une introduction à l'ALW, des outils d'aide à la lecture et quelques notices et cartes, extraites des volumes publiés ou à paraître. Il proposera un jour, nous l'espérons, une interface de consultation donnant accès à une édition numérisée de l'ALW.

Cet espace virtuel pourrait également devenir un lieu de publication, à l'attention d'un public non scientifique. Dans cette perspective de vulgarisation, le travail d'édition prendrait d'autres formes, telles que des cartes et des notices expliquées, choisies pour leur illustration des traits dialectaux propres à chaque domaine linguistique de la Belgique romane. Cette idée, suggérée par les participants à la journée TraSoGal, est séduisante, mais demande également du travail, du temps et des ressources, tout comme l'informatisation et l'édition traditionnelle. Espérons que ces rêves deviennent un jour réalité !

Bibliographie

- ALW = Remacle, Louis, Legros, Élisée *et alii* (1953-...). *Atlas linguistique de la Wallonie*, Liège, Université de Liège, 10 volumes.
- Baiwir, Esther et Renders, Pascale (sous presse). « Les atlas linguistiques sont-ils des corpus ? », *Corpus*.
- Boutier, Marie-Guy (2008). « Cinq relations de base pour traiter la matière géolinguistique : réflexions à partir de l'expérience de l'Atlas linguistique de la Wallonie », *Estudis Romànics* 30, 301-310.
- FEW = von Wartburg, Walther (1922-2002). *Französisches etymologisches Wörterbuch*, Bonn, Tübingen, Bâle, Nancy.
- Renders, Pascale (2011). *Modélisation d'un discours étymologique. Prolégomènes à l'informatisation du Französisches Etymologisches Wörterbuch*. Liège, Université de Liège (thèse de doctorat).