La place des *Atlas* dans la rétro-conversion du FEW

Cette communication porte sur les interfaces, possibles et souhaitables, entre le *Französisches Etymologisches Wörterbuch*, le « dizionario-tetto[1] » de la lexicographie galloromane, et la numérisation des *Nouveaux Atlas Linguistiques de la France*. Elle repose sur deux points d'appui, l'un épistémologique, l'autre méthodologique. Comme on le sait par les témoignages de Wartburg, sa *darstellung des galloromanischen sprachschatzes* qu'est le FEW n'aurait jamais existé sans l'*Atlas linguistique de la France* de Jules Gilliéron. Le FEW a bien entendu intégré également les données disponibles en son temps de l'ALLo, de l'ALG, de l'ALLy[2] et la refonte de la lettre B intègre pleinement les *Nouveaux atlas*. Du point de vue épistémologique, la fusion entre le FEW et les données des atlas est donc intégrale. D'un point de vue méthodologique, nous nous plaçons dans l'optique de la rétro-conversion numérique

1 — E. Buchi, P. Renders, « Gallo-Romance I: Historical and etymological lexicography », dans *Dictionaries. An International Encyclopedia of Lexicography. Supplementary Volume: Recent Developments with Focus on Electronic and Computational Lexicography*, éd. R. H. Gouws, U. Heid, W. Schweickard, H. E. Wiegand, Berlin/Boston, De Gruyter/Mouton, 2013, p. 653-662, en part. p. 654.

2 — Dorénavant, ALLo = R. Hallig, *Atlas linguistique de la Lozère et des cantons limitrophes du Gard et de l'Ardèche*, 1932-1934 (copie manuscrite en possession du FEW) ; ALG = J. Séguy, *Atlas linguistique et ethnographique de la Gascogne*, Paris, Éditions du CNRS, 1954-1973 (6 vol.) ; ALLy = P. Gardette, *Atlas linguistique et ethnographique du Lyonnais*, Paris, Éditions du CNRS, 1950-1976 (5 vol.) ; ALMC = P. Nauton, *Atlas linguistique et ethnographique du Massif Central*, Paris, Éditions du CNRS, 1957-1963 (4 vol.) ; ALFC = C. Dondaine, *Atlas linguistique et ethnographique de la Franche-Comté*, Paris, Éditions du CNRS, 1972-1991 (4. vol.).

du FEW : celle-ci n'est pas encore achevée, mais elle est en bonne voie, soutenue notamment par l'engagement de Pascale Renders[3].

En reprenant les observations d'Esther Baiwir et de Pascale Renders, il nous a semblé nécessaire de se focaliser sur la question qui « concerne les unités à mettre en relation dans chacune des ressources : s'agit-il du lexème, de l'article, de la notice, de la carte ? En fonction de la granularité des unités ainsi identifiées, diverses modélisations informatiques sont possibles[4]. » Nous nous sommes donc demandé quels pourraient être les problèmes et les empêchements pratiques d'une mise en réseau des atlas numérisés avec le FEW rétro-converti, quelle est l'utilité pour le FEW numérisé d'intégrer de nouvelles données et quel pourrait être l'apport de cette intégration.

Le point de départ de notre travail est, notamment, le volume 16 du FEW, qui contient les étymons germaniques des lettres G-R. En même temps, nous avons intégré à notre réflexion le *Thesaurus occitan* (THESOC), « base de données multimédia qui vise à rassembler toutes les données dialectales recueillies sous forme orale en domaine occitan[5] » et qui fournit à présent la référence la plus vaste pour les *Nouveaux atlas*. Dans cette constellation de travaux en cours de rénovation, restructuration ou encore de conception[6], nous ne pourrons que tracer des pistes de réflexion, parfois encore théoriques. Du côté des atlas, nous nous sommes appuyé pour cette réflexion sur le vol. 5 de l'ALLy de Pierre Gardette et, très brièvement, sur les matériaux inédits qui complètent l'ALMC de Pierre Nauton[7].

---

3 — Cf. P. Renders, *L'informatisation du* Französisches Etymologisches Wörterbuch. *Modélisation d'un discours étymologique*, Strasbourg, ELiPhi, 2015.

4 — E. Baiwir, P. Renders, « Numérisation des données dialectales d'oïl : le projet APPI comme laboratoire », publication en ligne dans le cadre du projet APPI (Atlas pan-picard informatisé, sous la direction d'Esther Baiwir), avril 2019, p. 6 (https://appi.univ-lille.fr/data/medias/baiwirrenders2019).

5 — M. Oliviéri, S. Casagrande, G. Brun-Trigaud et P.-A. Georges, « Le Thesaurus Occitan dans tous ses états », *Revue française de linguistique appliquée*, t. XXII-1, 2017, p. 89-102. en part. p. 89.

6 — Communications privées avec Guylaine Brun-Trigaud du 9 août 2019 et avec Benjamin Husson, IE à l'ATILF de Nancy, du 17 septembre 2019, qui nous ont informé sur l'état des deux projets.

7 — Les deux entreprises atlantographiques que nous avons choisi de traiter pour des raisons pratiques (l'ALLy et l'ALMC, qui ont été intégrées dans le THESOC), ont déjà été comparées par P. Gardette, « Le Lyonnais et le Massif Central d'après les Atlas linguistiques régionaux (à propos de la publication du premier volume de l'*Atlas Linguistique et Ethnographique du Massif Central* de P. Nauton) », *Revue de Linguistique Romane*, t. 21, 1957, p. 209-230 [désormais dans P. Gardette, *Études de géographie linguistique*, éd. B. Horiot, M.-R. Simoni, G. Straka, Strasbourg, Société de Linguistique Romane, 1983, p. 663-684].

1.1. Les liens simples : URL

Le FEW compte actuellement 25 volumes. En dehors des 3 volumes (21-23) qui contiennent les matériaux d'origine inconnue ou incertaine, classés selon le système onomasiologique (le *Begriffssystem*), les matériaux sont classés selon leur entrée étymologique. C'est une des grandes différences avec les *Atlas*, qui sont organisés par concepts. L'autre différence consiste en la présentation des matériaux en diachronie dans le FEW et en synchronie dans les *Atlas*[8].

Pascale Renders a écrit 37 algorithmes qui rendent possible la transformation des articles du FEW à partir d'un format XML avec balisage de base (.ffml, *FEW Font-style Markup Language*) en format XML enrichi (.fsml, *FEW Semantic Markup Language*)[9]. Si le FEW contient une erreur d'écriture, une incohérence dans la structure ou une ambiguïté, l'outil de rétro-conversion s'arrête et demande une correction. Les 37 algorithmes ajoutent 24 balisages sémantiques qui contiennent des informations concernant la macro-structure de l'article (<doc>, <com>, <notes>) et l'identification des caractères des unités lexicales (<etymon>, <geoling>, <biblio>). Ce balisage enrichi est le centre de notre réflexion. Voici un exemple d'« unité minimale de traitement[10] » balisée (FEW 16, 355a s.v. *KRAMP)[11] :

```
<unit>AllierSE. krãpunä „taller (du blé)“ ALLy 54</unit>
<geoling>AllierSE</geoling>
<form>krãpunä</form>
<i>krãpunä</i>
<lb/>
<def>„taller (du blé)“</def>
<precisions>ALLy 54</precisions>
<biblio>ALLy 54</biblio>
```

8 — Déjà N. MAZZIOTTA, « L'informatisation du *Französisches Etymologisches Wörterbuch*. Concepts pour une approche modélisée commune à l'*Atlas Linguistique de la Wallonie* », *Zeitschrift für Romanische Philologie*, t. 127, 2011 p. 36-62, à p. 41 avait rapproché le FEW de l'ALW : « Il peut paraître étrange, à première vue, de rapprocher des projets aussi distincts qu'un dictionnaire étymologique comme le FEW et un atlas linguistique, qui recueille des données synchroniques. En réalité […] il semble que les deux entreprises aient les mêmes caractéristiques fondamentales, bien que le FEW soit, il est vrai, un petit peu plus complexe. »

9 — P. RENDERS, *L'informatisation*, p. 144-145.

10 — E. BUCHI, *Les Structures du* Französisches Etymologisches Wörterbuch. *Recherches métalexicographiques et métalexicologiques*, Tübingen, Niemeyer, 1996, p. 116 et P. RENDERS, *L'informatisation*, p. 55.

11 — URL : https://few-webapp.atilf.fr/consulter/volume/match;volume=16;id=000901;-match=*kramp.

L'article rétro-converti en format .fsml est ensuite transformé en HTML ; la donnée est récupérée en utilisant le moteur Allegro. Celui-ci permet l'interrogation des données à travers le langage Corpus Query Language (CQL), en utilisant une interface REST et par la suite mis à disposition des utilisateurs sur une interface web : https://few-webapp.atilf.fr/. Cette interface permet déjà, entre autres, de mettre en exergue des informations comme les étymons, les formes, les étiquettes géolinguistiques ou bibliographiques, les deux dernières avec des renvois au *Beiheft* 2010, ce qui rend la consultation bien plus rapide et aisée. La recherche, pour l'instant, peut être effectuée sur des étymons, des formes ou en croisant des paramètres comme le volume du FEW, la géolocalisation ou la bibliographie. Il est clair que le souhait majeur est de pouvoir implémenter les fonctionnalités de recherche en vue de recherches croisées de plus en plus complexes. Cette interface ne permet pas, pour l'instant, de signaler des mises à jour ou d'intégrer des liens externes qui soient immédiatement visibles par les utilisateurs[12].

La limite de la configuration actuelle est que les hyperliens via des URL au FEW rétro-converti ne peuvent être liés qu'à l'interface graphique et pas aux articles individuels. Il faudrait donc envisager la possibilité d'un accès direct en utilisant une URI transparente et discriminante qui indique le volume, la page et le numéro de l'article dans la colonne selon la balise <"ici"> ("in-column index"[13]). Nous suivrons la proposition de Renders, Baiwir et Dethier : « our suggestion would be to implement a simple REST API based on HTTP request and using JSON-encoded data[14] ». Le format HTTPS de l'URL porte comme informations le numéro du volume, le numéro de l'article et l'étymon concerné. Prenons l'exemple du lexème *i̯l̃ārạdo* dans l'ALLy 434 "(la branche) est éclatée", qui est à rattacher à la base souabe LANDEL (FEW 16, 443a) :

> https://few-webapp.atilf.fr/consulter/volume/match;volume=16;id=001044;match=landel

12 — Cf. P. RENDERS, *L'informatisation*, p. 12, qui présente les avantages de la rétro-conversion du FEW imprimé en *dictionnaire informatisé*, à savoir : « 1) élargir son accessibilité […], 2) permettre des consultations transversales, grâce à de multiples possibilités de requête […], 3) vérifier l'exhaustivité des résultats obtenus […], 4) et, enfin, faciliter les mises à jour qui s'imposent ».

13 — Cf. E. BAIWIR, « Numérisation des données », p. 6.

14 — P. RENDERS, E. BAIWIR, G. DETHIER, « Automatically Linking Dictionaries of Gallo-Romance Languages Using Etymological Information », dans *Electronic lexicography in the 21st century: linking lexical data in the digital age. Proceedings of the eLex 2015 conference, 11-13 August 2015*, éd. I. KOSEM, M. JAKUBIČEK, J. KALLAS, Herstmonceux Castle, United Kingdom, 2015, p. 452-460, à p. 456.

La version scannée peut être rejointe grâce à cette URL :

https://apps.atilf.fr/lecteurFEW/lire/volume/160/page/443

Avec une URL qui renvoie à l'article rétro-converti, on ne peut pas signaler la page mais uniquement le numéro de l'article (dans ce cas-là le 1044), alors que la limite de la deuxième URL est que dans une page on peut trouver plusieurs articles, comme dans la page suivante 444, qui contient 9 articles (https://apps.atilf.fr/lecteurFEW/lire/volume/160/page/444). Pour éviter toute ambiguïté, la recommandation est très simple mais efficace, puisqu'à l'identifiant URI « FEW 16, 443a » il suffira d'ajouter la vedette « FEW 16, 443a, LANDEL ».

1.2. Apport du THESOC aux articles du FEW

Venons-en à un exemple concret. Comme nous ne connaissons pas encore la forme définitive des fiches lexicales du THESOC, nous cantonnerons nos observations au spécimen de « Fiche réponse par localité : la BERGERONNETTE à Pézenas », dont le lemme est *pastorèla*[15]. L'information étymologique est déjà excellente, car elle renvoie non seulement à REW 6279 – désormais malgré tout dépassé – mais aussi au FEW en indiquant en outre une « formule étymologique » (qui correspond à ce que Martin Glessgen appelle la « structure lexématique »[16]) et qui permet de placer le lexème à l'intérieur d'un article du FEW ; en l'occurrence, le renvoi est à FEW 7, 759a s.v. PASTOR[17]. Dans l'optique d'un renvoi au FEW scanné actuellement disponible, l'hyperlien pourrait être fait par le biais d'une URL de ce type :

https://apps.atilf.fr/lecteurFEW/lire/volume/070/page/759.

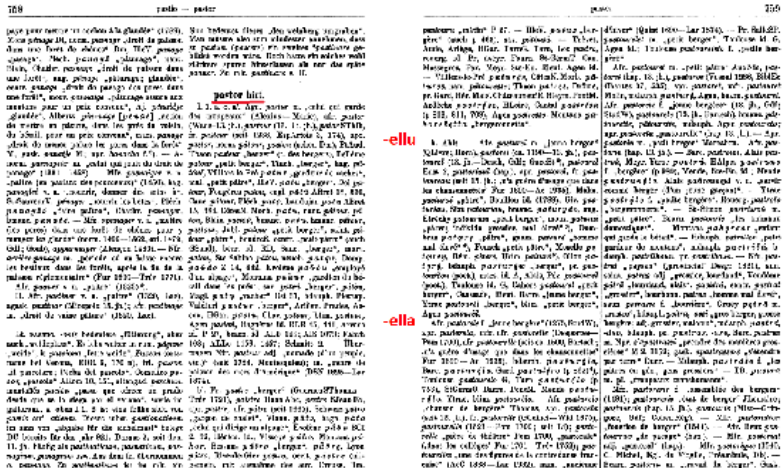
La structure lexématique PASTOR + ELLA permet un lien immédiat vers le bon emplacement dans l'article du FEW qui suit une structure morphologique assez claire : FEW 7, 759a PASTOR § I.1.b.Ablt. classe les dérivés en -ELLU, -ELLA. Si l'on considère l'ensemble du paragraphe, on s'aperçoit que toute la famille des composés en -ELLA se caractérise comme une famille occitane occidentale, avec des représentants éparpillés dans le Valais et la Haute-Savoie. Le

15 — M. OLIVIÉRI, « Le Thesaurus Occitan », p. 91.

16 — Cf. M. GLESSGEN, M. SAUZET, communication orale présentée durant la journée d'étude.

17 — Dans l'article mentionné, le renvoi au FEW est fait de manière classique, avec le renvoi à la page du début de l'article « FEW 7, 758b ». Il est clair qu'un lien plus précis renvoyant à la page concernée au sein de l'article (FEW 7, 759a) serait souhaitable.

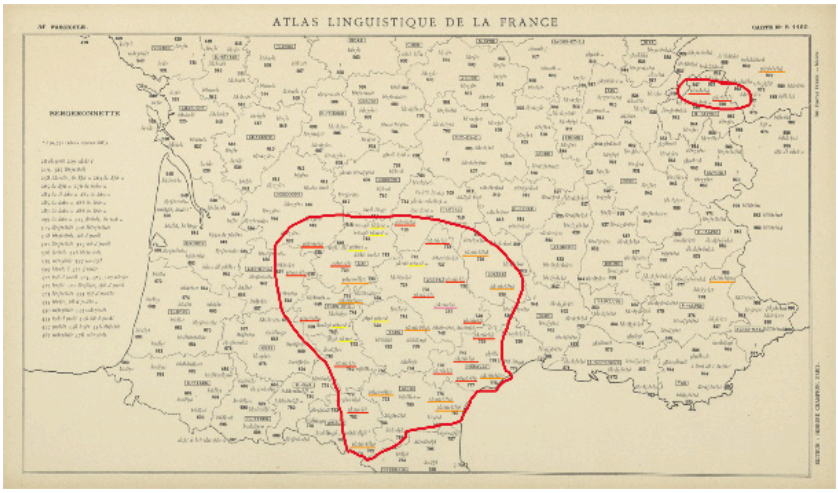
FEW, surtout basé sur l'ALF 1460 et sur l'ALLo 460, permet de décrire une répartition aréale de ce type : « Trèves *p å s t u r é l o* "bergeronnette", Hér. *p a s t u r é l o*, Ariège id., Lavel. (Lavelanet, Ariège, Foix), HGar., TarnG. id., Lot *p o h t o r é l o* (p 720), Tarn, Camarès *pastourèlo*, aveyr., St-GermT., Can. Massegros *p o s t u r é l o*, Cantal *p o h t u r é l o* (p 715). »



Avoir un lien direct vers les données du THESOC, naturellement, permet de cerner la diffusion du type dans l'espace et à une époque quelque peu plus récente, insérant ainsi l'enquête dialectologique synchronique dans l'axe de la diachronie. En transposant les données du THESOC sur la même carte de l'ALF, celle-ci nous montre une aire de diffusion du type <pastorèla> qui se superpose presque parfaitement à celle qui est déjà disponible[18]. De même, si l'on considère aussi le lemme suffixé *pastoreleta*, la répartition aréale est reconfirmée (en effet, les deux types cohabitent dans la même aire). Néanmoins, si l'on élargit l'enquête à l'ensemble de

18 — Le THESOC donne ces attestations (que nous résumons) : « HAlpes (S-Firmin) *pastuR'ela*, Ariège (S-Ibars) *pastur'ɛlɔ*, Aude (8 pt) *pastur'ɛlɔ*, Aveyron (11 pt) *pastur'ɛlɔ / pɔ- / -sθ-*, Cantal (3 pt) *pastur'ɛlɔ / -ɛlɔ*, Gard (4 pt) *pastur'ɛlɔ / pɔ-*, Hérault (16 pt) *pastur'ɛlo / -la / -l* , Lozère (4 pt) *pastur'ɛlɔ / -lo / p -* ». Il faut ajouter, à la base THESOC dans son état actuel, les données de l'ALLOc (X. Ravier, *Atlas linguistique et ethnographique du Languedoc Occidental*, Paris, CNRS, 1978-1993 (4 vol.)). L'ALF enregistrait aussi *pastorèla* dans la Haute-Savoie, p. 947 et 956, mais l'ALJA (vol. 2, liste 94) ne mentionne plus ces attestations, ce qui pourrait indiquer sa disparition en faveur d'autres types.

la famille étymologique de PASTOR pour le sens 'bergeronnette', les NALF fournissent trois attestations du type *pastressa* entre Vaucluse, Bouches-du-Rhone et Alpes-de-Haute-Provence, glissant du sens 'bergère' (attesté dans FEW) à 'bergeronnette'[19]. On signalera, enfin, le lexème *pastręsuna* attesté à Murat, dans le Cantal[20].



Avoir à disposition un programme de cartographie automatique des données d'un article du FEW, ou mieux encore des données intégrées des *NALF*, permettrait une visualisation immédiate des distributions aréales et des éventuelles (voire probables) évolutions en diachronie, décelables grâce à la comparaison des deux ouvrages[21].

2.1. Étymologiser les Atlas

Des questions complémentaires surgissent : suffit-il de faire des renvois à la page/colonne au FEW, ou pouvons-nous exploiter des liens plus précis ? Comment devrait se comporter le FEW ? Quels sont les conseils pour les *Atlas* numérisés afin de se placer dans le cadre du FEW rétro-converti ?

19 — FEW 7, 759a s.v. PASTOR : « Drôme, pr. Gard, Hér. Malz. ChâteauneufR. Blexm. PontM. Ardèche *p a s t r é s o*, HLoire, Cantal *p a s t r é s a* (p. 813, 811, 709) ». Les trois nouvelles attestations du sens sont tirées de ALP 967 (p. 89, 105, 123).

20 — Donnée tirée de ALMC 314 (p. 17).

21 — Un projet de cartographie automatique des données du FEW a été imaginé par Yan Greub mais il est encore inachevé. Cf. P. RENDERS, *L'informatisation*, p. 45, qui met ce projet en lien avec la rétro-conversion du FEW, et M. GLESSGEN, M. SAUZET, communication orale présentée durant la journée d'étude, qui envisagent les « possibilités d'exportation des données sous forme de cartes variables et évolutives ».

Nous nous appuierons surtout sur le vol. 5 de l'ALLy, dans lequel Pierre Gardette reprend chaque carte de l'*Atlas linguistique et ethnographique du lyonnais* et interprète les témoignages en indiquant un classement étymologique sous la forme d'un petit commentaire. Nous avons compté 117 étymons germaniques appartenant au vol. 16 du FEW (G-R)[22]. Les 5 fascicules du vol. 16 ont paru entre 1955 et 1959[23], les 3 volumes de l'ALLy ont paru en 1950, en 1952 et 1955 (alors que le vol. 4, *Exposé méthodologique et tables*, a paru en 1967 et le vol. 5, *Commentaires et index*, en 1976). Wartburg aurait vraisemblablement pu exploiter le travail de Gardette, même si le vol. 3 de l'ALLy est paru en concomitance avec le fascicule 16/1 du FEW. En tout cas, cette chronologie nous donne une base assez sûre pour remarquer que dans le vol. 16 du FEW, l'ALLy est cité exactement dans 12 articles, partiellement dans 32 et non cité dans 73. Le rapport de citation est d'environ 2 : 3, ce qui laisse une lacune sensible dans la dimension « thesaurus » du FEW. Bien entendu, dans un article comme FEW 16, 668a REGINHART, qui donne *renard* en français et qui est emprunté dans les régions occitane et francoprovençale, il ne vaut pas la peine de signaler le renvoi à l'ALLy, qui n'apporterait aucune information supplémentaire[24].

En revanche, la carte 60 de l'ALLy (à l'envers) indique le syntagme *o nor* dans le Nord du Rhône, avec un sens qui est absent de FEW 16, 601b NORÞ (ags.)[25]. De même, pour le souabe LANDEL, la carte 434 ((la branche) est éclatée) de l'ALLy mentionne le signifiant *il̲ãrạdo* (p 36 = Vollore-Montagne, Puy-de-Dôme), qui est à rattacher à FEW 16, 443a : « Alp. *eilandrá* v.a. "déchirer" M, AlpesM. *ẹ l ã n d r á* ALF 1531 p 888 » ; cette nouvelle attestation permet d'élargir l'aire d'attestation du mot vers la Basse Auvergne.

22 — À des fins statistiques, nous signalons 2 bases déonomastiques, 11 bases germaniques, 62 bases ancien bas francique, 1 base burgonde, 1 base gothique, 1 base gothique/burgonde, 1 base longobarde, 1 base souabe, 5 bases ancien-haut-allemandes, 1 base ancien-bas-allemande, 4 bases moyen-haut-allemandes, 2 bases allemandes, 13 bases moyen-néerlandaises, 5 bases néerlandaises, 5 bases ancien-nordique, 1 base anglo-saxonne, 1 base ancien-anglaise. Il serait intéressant de comprendre les trajectoires et l'impact de chaque base sur les langues du domaine francoprovençal et occitan.

23 — Ch. SEIDL, J.-P. CHAUVEAU, Y. GREUB, *Französisches Etymologisches Wörterbuch : eine Darstellung des galloromanischen Sprachschatzes. Beiheft – Complément bibliographique*, Strasbourg, ELiPhi, 2010, p. 409 donne les dates de parution : 16/1 en 1955, 16/2 en 1956, 16/3 en 1957, 16/4 en 1958 et 16/5 en 1959.

24 — Cf. FEW 16,691a « Im occit. ist es zuerst ganz vereinzelt [n. 11: Immerhin findet es sich nicht, wie Arch 134, 102 behauptet wird, nur als eigenname]; erst in späterer zeit setzt es sich hier durch ».

25 — Cette locution se trouve aux points 6 (Saint-Pierre-le-Vieux, Saône-et-Loire), 31 (Saint-Clément-sous-Valsonne, Rhône), 40 (Poleymieux, Rhône). Il s'agit d'un emprunt au fr., cf. ALLy 5, 434 : « Ces expressions anciennes sont remplacées par de plus récentes signifiant "au nord" et "au vent du nord". »

Ce ne sont que deux exemples de formations un peu insolites et qui demandent à être expliquées, mais qui ajoutent des informations précieuses au FEW et à l'histoire de la diffusion des unités lexicales dans l'espace.

Nous avons choisi de traiter certaines formes de l'ALLy qui y ont déjà été étymologisées et qui ont ainsi été intégrées dans le THESOC, afin d'exploiter l'interaction entre ces deux projets et le FEW. Il s'agit notamment des bases : GARÞA (FEW 16, 13b, abfrq.), *GRÛT (FEW 16, 96a, abfrq.), KOLZAAD (FEW 16, 343b, néerl.), *KRAMP (FEW 16, 353b, abfrq.), *KRAPPA (FEW 16,357a, germ.), *MORNÔN (FEW 16,565a, abfrq.).

2.2. Les liens à la structure XML

Nous avons donc cherché à vérifier l'apport au FEW de ces notices étymologisées insérées dans le THESOC et avons voulu savoir comment ces ressources pourraient être liées ensemble, plus d'un point de vue théorique que proprement pratique. Pour l'instant, nous avons retenu trois cas de figure représentatifs[26].

I. Renvoi du FEW à l'ALLy[27]

54, 8 TALLER [THESOC *cramponner* (fr.), renvois à ALLy[28]] (Ø ALMC 977 ; Ø ALG 1107)
ALLy 5, 37 : « 8. *krãpunà* dans l'Allier, à 13, 14, 22, 23 semble dérivé du fr. *crampon* : le blé qui talle forme de nouvelles racines qui le fixent à la terre comme par des crampons. FEW 16, 355a (*KRAMP). »

Emplacement dans le FEW 16, 355a, s.v. *KRAMP : « nfr. *se cramponner* à "s'accrocher fortement (p. ex. à une personne, à un espoir, se dit aussi de la tige d'une plante)" (seit Rich 1759), AllierSE. *k r ã p u n ã* "taller (du blé)" ALLy 54. »

URI : « FEW 16, 355a *KRAMP »
.fsml 16,000901 – lignes : 2261-2268

26 — Nous donnerons, pour chaque typologie, un extrait du fichier .fsml balisé de l'article correspondant, développé par Pascale Renders, afin de montrer comment les données et les balises sont présentées.

27 — À cette catégorie appartient aussi FEW 16, 343b s.v. KOLZAAD (ndl.), qui renvoie déjà à ALLy 76 ; il s'agit d'un emprunt entré dans la langue aux 17e–18e siècles. Dans la perspective d'une base intégrale comme le THESOC, il serait intéressant de voir la diffusion différentielle sur base onomasiologique en format cartographique (notamment avec les dérivés de CRUX).

28 — Localités retenues par le THESOC : Chatel-Montagne (pt 03 ; ALLy 14, Allier), Isserpent (pt 03 ; ALLy 13, Allier), Lachaux (pt 63 ; ALLy 23, PuyD.), Saint-Nicolas-Des-Biefs (pt 03 ; ALLy 22, Allier).

<unit>AllierSE. krãpunä „taller (du blé)“ ALLy 54</unit>
<geoling>AllierSE</geoling>
<form>krãpunä</form>
krãpunä
<lb/>
<def>„taller (du blé)“</def>
<precisions>ALLy 54</precisions>
<biblio>ALLy 54</biblio>

Lignes : 2178-2181

Tags inserted into <p> paragraph, column 355a

<group>Ablt. — Aflandr. apik. cramponchel…
<title>Ablt. —</title>

Commentaire : le vb. connait un sémantisme restreint à l'aire délimitée par ALLy. Le FEW et le THESOC reprennent les données de l'ALLy. Il suffirait de faire un renvoi à l'article du FEW [vol. pag.] ou de renvoyer à l'unité qui contient déjà le renvoi précis à la carte <biblio>ALLy 54</biblio>.

108, 2-3 UNE JOINTÉE [THESOC *grapada*[29]] (ALMC 1021, HLoire Ardèche Lozère ; Ø ALG 483)
ALLy 5, 75 : « B. Deux mots occitans et un hybride se trouvent dans les régions d'influence occitane. […] 2. *grapá (-ạdo)* occupe une aire à l'O et au S, à partir de 36, 48, 60, 67, 70. Ce participe passé du verbe occitan *grapa* "saisir" a pris le sens de "jointée"[30] dans une aire qui s'étend sur une partie du Puy-de-Dôme, la Haute-Loire et quelques localités de la Lozère et de l'Ardèche, *ALMC* 1021. FEW 16, 358a (*KRAPPA)[31]. »

1. grapada
Emplacement dans le FEW 16, 358a, s.v. *KRAPPA : « Vinz. *g r a p á d å* "jointée" RLR 63, 103, Ambert, Thiers, lang. *grapado* "poignée

29 — Localités retenues par le THESOC : Ardoix, Saint-Anthème (pt 63 ; ALLy 57, PuyD), Boulieu, Lalouvesc (pt 07 ; ALLy 74, Ardèche), Saint-Julien-Molhesabate (pt 43 ; ALLy 72, HLoire), Saint-Maurice-En-Gourgois (pt 42 ; ALLy 60, Loire), Saint-Romain-Les-Atheux (pt 42 ; ALLy 67, Loire), Sainte-Sigolène (pt 43 ; ALLy 68, HLoire), Valcivières (pt 63 ; ALLy 47, PuyD), Vanosc (pt 07 ; ALLy 71, Ardèche), Vion, Vollore-Montagne (pt 63 ; ALLy 36, PuyD).

30 — Cf. TLF*i* s.v. *jointée* s.f. « Quantité qui peut être contenue dans le creux formé par les paumes des deux mains placées côte à côte ».

31 — ALLy 5 ajoute : « 3. *grabutó*, qui apparaît à 52, 63, 65, 66, à la jonction des aires de *grapá* et d'*ãmbutó*, est évidemment un hybride de ces deux mots » ; cf. FEW 24, 411a s.v. *AMBOSTA : « LoireSE… *grabuto*… VienneRhN. *grabuto*, VienneRhS. *id.*, *grabota* » et note 6. : « Par croisement avec le type *grapa* (ici 16, 358a) ». Le troisième mot occitan qu'il mentionne est *gafá*, cf. FEW 4, 18b s.v. *GAFFARE.

de grains", LoireSO. Annonay *g r a p á* ALLy 108. » Ajouter les données de ALMC 1021, p 27, 31, 35 (indiquées par le THESOC) qui élargissent l'aire de distribution attestée par le FEW.

URI : « FEW 16, 358a *KRAPPA »
.fsml 16,000912 – ligne 3964 <biblio>ALLy 108</biblio>
Ici le renvoi à l'ALLy est déjà présent (typologie I.).

Commentaire : le lexème est déjà enregistré dans FEW. Les données de l'ALMC, enregistrées par le THESOC, trouvent leur place naturelle avec des renvois à cette unité lexicale et permettent d'élargir l'aire d'attestation.

II. Renvoi de l'ALLy à des unités lexicales du FEW

Dans ce deuxième cas de figure, le FEW contient un lexème qui se retrouve aussi dans l'ALLy mais sans renvoi à l'atlas (en l'occurrence, le sens est le même, alors que la forme peut subir de menues variations phonétiques). Par conséquent, le résultat est souvent l'élargissement de l'aire d'attestation du lexème[32].

421,2 : CROÛTE DE PAIN [Ø THESOC] (Ø ALMC 1122)
(ALLy 5, 304) : « 2. *grwà* dessine une petite aire à 54, 62, 66, 69-71. Ce mot de la région de Saint-Étienne (Veÿ *groua*) et du Forez (Gras *groué*) est aussi lyonnais (P. *grua*) et nord-dauphinois (*DTF* 2709 *grẅé*). On peut le placer sous *GRUT (FEW 16, 97a), à côté de Sainte-Sabine (Côte-d'Or) *gruon* "croûton ou débris de pain". 3. *gruto* à 67, 69 paraît être un hybride de ses deux voisins *grwa* et *kruto*. Mais il faut noter l'alternance de *kr-/gr-* dans cette région. »

Emplacement dans le FEW 16, 97a, s.v. *GRÛT : « Ste-Sabine *gruon* "croûton ou débris de pain". »

URI : « FEW 16, 97a *GRÛT »
.fsml 16,000188 – lignes 1219-1224

<unit>Ste-Sabine gruon „croûton ou débris de pain“</unit>
<geoling>Ste-Sabine</geoling>
<lb/>
<form>gruon</form>

32 — Il y a des cas de figure, comme pour les dérivés de GARÐA (FEW 16, 13b), où les différentes désignations des meules (« tas de gerbes droites », « la meule des gerbes (dans le champ)», « le tas de gerbes (dans la grange)», « la grosse meule longue ») ne semblent pas trouver un correspondant précis dans l'article du FEW ; celui-ci renvoie aux cartes 67-69, mais il resterait à insérer les cartes 65, 70, 77, 92. Dans ce cas-là, un renvoi au § 1.b. de l'article, qui classe le type *gerbier* (< -ARIU) pourrait suffire. À remarquer que le type semble surtout caractéristique de l'aire occitane et francoprovençale.

gruon
<def>„croûton ou débris de pain“</def>

Commentaire : élargissement de l'aire de diffusion du lexème avec le sens "croûte de pain", autrement attesté seulement en bourguignon (Ste-Sabine est située dans le sud de la Bourgogne). Il conviendrait d'insérer un lien vers l'unité lexicale du FEW pour renvoyer aux nouvelles formes enregistrés dans l'ALLy.

1163,4 : LE SON [Ø THESOC] (Ø ALMC 1723, Ø ALG 287)
ALLy 5, 675 : « 4. *gryè$_{œ}$* 47 se rattache à la base germanique *GRÛT (FEW 16, 96) qui a donné l'afr. *gruel* "avoine, orge grossièrement moulue"[33]. »

Emplacement dans le FEW 16, 96b s.v. *GRÛT : « Nfr. *griot* "son servant à faire de l'amidon" Enc 1, 384b, lang. *griôou* "son" S[34], Ambert *g r y œ u* "son de seigle". »

URI : « FEW 16, 96b *GRÛT »
.fsml 16,000188 – lignes 919-942 :

<unit>Nfr. griot „son servant à faire de l'amidon“ Enc 1, 384</unit>
<geoling>Nfr.</geoling>
<form>griot</form>
griot
<def>„son servant à faire de l'amidon“</def>
<lb/>
<precisions>Enc 1, 384</precisions>
<biblio>Enc 1, 384</biblio>

<unit>lang. griôou „son“ S</unit>
<geoling>lang</geoling>
<lb/>
<form>griôou</form>
griôou
<def>„son“</def>
<precisions>S</precisions>
<biblio>S</biblio>

<unit>Ambert gryœ&few-u-boucle; „son de seigle“</unit>
<geoling>Ambert</geoling>
<form>gryœ&few-u-boucle;</form>

33 — Cf. .fsml 16,000188 lignes 742-743 : <group>Ablt. — Afr. gruel „avoine, orge dépouillée de l'enveloppe et grossièrement moulue“ […] <title>Ablt. —</title>

34 — Le sigle S indique : P.-A. BOISSIER DE SAUVAGES, *Dictionnaire languedocien-français, ou choix des mots languedociens les plus difficiles à rendre en français*, Nîmes, 1756. Cf. le commentaire dans Ch. SEIDL, *Complément*, p. 68 : « selon l'auteur ne contient qu'un choix de termes de deux villes, l'une du bas Languedoc, l'autre des Cévennes ».

gryœ&few-u-boucle;
<def>„son de seigle“</def>
<lb/>

Commentaire : l'ALLy confirme l'aire d'attestation du mot relevée par le FEW (Ambert est situé dans le Puy-de-Dôme ; le pt 47 indique Valcivières, canton d'Ambert). Il conviendrait d'insérer la nouvelle attestation de l'ALLy *gryè*$_{œ}$ dans un paragraphe du FEW déjà existant.

III. Données de l'ALLy qui manquent dans le FEW

Le troisième cas de figure est sans doute le plus intéressant. Les atlas peuvent apporter des données absentes dans le FEW et donc exiger l'intégration de nouvelles unités lexicales. On peut imaginer que ce cas de figure puisse être assez fréquent.

102, 9 LA BALLE DU BLÉ [THESOC[35] *balle*, s. *crapa* et *grabeau* (fr.)[36]] (Ø ALMC 1017 ; Ø ALG 286)
ALLy 5, 71-72 : « B. Mots de sens plus général signifiant "poussières, résidus" […] 9. *krapi* (*krapa*$_{e}$) à 61, 66, 69, 70, 73, 75 doit sa terminaison *i* à *buri*, le suffixe -ARIU étant *a*$_{e}$ ou *e*$_{i}$ à 70, 73, 75[37]. Quant à son radical *krap-*, c'est celui de *krạpa* "résidu de la pressée" (ALLy 213). Ce terme a peut-être été attiré vers le sens de "résidu du vannage" par la ressemblance phonique entre *krap-* et *grab-* de *grabó* ».

Emplacement hypothétique dans FEW 16, 359a s.v. *KRAPPA : « Vaux, Couzon *crappa* "grappe de raisin après le pressage" […] Biz. SDT. "grappe dépouillée de son raisin" »

URI : « FEW 16, 357a *KRAPPA »
.fsml 16,000912 – ligne 18 « I 3 Afr. grape18) „rafle de raisin“ (seit 11. jh., REJ 1937, 106; RSt 1, 167) »

Lignes 5187s

Tags inserted into <p> paragraph, column 359b

<pnum>3.</pnum>
<unit>Afr. grape18) „rafle de raisin“

35 — THESOC rattache *grabeau* (fr.) à FEW 2, 1332a CRĪBĒLLUM, mais il faudrait justifier l'ouverture ī > *-a-*.

36 — Localités retenues par le THESOC : Ardoix (pt 07 ; ALLy 73, Ardèche), Boulieu (07 ; ALLy 70, Ardèche), Saint-Sauveur-En-Rue (42 ; ALLy 69, Loire), Vion.

37 — Le point 75 Vion correspond à ALF 827 dans la carte 1452 « balle (de blé, etc.) ».

Les attestations de l'ALLy se rattacheraient aux lignes 5379-5385

<unit>Vaux, Couzon crappa „grappe de raisin après le pressage“ </unit>
<geoling>Vaux</geoling>
<geoling>Couzon</geoling>
<form>crappa</form>
crappa
<lb/>
<def>„grappe de raisin après le pressage“</def>

ou aux lignes 5409-5414

<unit>voir. id., Biz. SDT. „grappe dépouillée de son raisin“ </unit>
<geoling>voir</geoling>
<geoling>Biz</geoling>
<geoling>SDT</geoling>
<imp/>
<def>„grappe dépouillée de son raisin“</def>
<lb/>

Commentaire : le sens manque dans FEW, mais Gardette propose un rapprochement hypothétique avec *KRAPPA, parce que le lexème *grape*, lié aux raisins, est attesté dans la même région, entre l'Isère et l'Ardèche septentrionale. D'un point de vue informatique, il faudrait attacher le lexème de l'ALLy à l'ensemble du paragraphe I.3. du FEW.

512, 5 ROITELET [THESOC *gruda noga*°°[38]] (ALP 956)
ALLy 5, 360 : « Dénominations d'après la petite taille de l'oiseau […] *krou dė nyéi* 75 "coquille de noix" »
cf. FEW 7, 255b s.v. NŪX : « Clerm. *nougóta* f. "roitelet" » et 256b : « Saugues *n u z ē t a* "roitelet" » [Hérault ; HLoire, Le Puy]

Commentaire : d'après le THESOC, la désignation du roitelet liée à l'image de la noix est attestée dans les dialectes du Rhône mérid. (Larajasse 49), Isère occ. (Jardin 63, Clonas 65), Loire (Ste-Croix-en-Jarez 62), HLoire (Saugues), de l'Ardèche sept. (Vion) et de la Drôme sept. (cf. Alixan, Chantemerle, La-Roche-De-Glun, Monteleger dans ALP 956). Les données confirment FEW 7, 255b s.v. NŪX et élargissent l'aire d'attestation du lexème. L'ajout de cette lexie complexe dans les dérivés de NŪX, qui s'insère dans une aire

38 — Localité retenue par le THESOC : Vion (pt 07 ; ALLy 75, Ardèche).

géographique cohérente, enrichit la connaissance des désignations métaphoriques du roitelet créées par expansion sémantique[39].

2.3. L'ALMC et la refonte modulaire du FEW

L'index de l'ALMC rédigé par Nauton, incomplet et sous forme manuscrite, comprend deux fichiers. Le premier, sous forme de fiches lexicales avec renvois non systématiques au FEW, est préparatoire au second fichier qui « présente la liste alphabétique des étymons, chacun suivi de tous les noms patois, y compris toutes leurs variantes, qui y remontent. Il était donc conçu comme un supplément au FEW pour les données des parlers du Massif Central[40]. » Dans un article récent, Wolf propose une réécriture de l'article AUCA, auparavant rédigé par Nauton sur la base de la carte 605, avec le renvoi systématique à FEW 25, ainsi conçue :

> « AUCA (25, 754-775) · *jars 605* : (25, 766b ; + -ARDU) œutsạr, œʊtsạr, ʊkạr, ʊyạr ; (25, 766b ; + -ACEU) aʊkạs, oʊkạs, ʊwtsạs ; (25, 767a ; + -ATTU) oukạ, oʊkạrt · *oie 605* : (25, 755b) ạʊka, ạʊtsa, ọya, wọyo · *oison 605* : (25, 768b ; + -ONE) ʊyʊ, oʊtsụ, ʊwtsụ, aʊtsʊ, ʊtsʊ, œtsụ, eutsụ ; (25, 770a ; + -ITTA) oʊkẹto, aʊkẹto, ʊwtsẹta, ʊyẹto (f.) ; · *cri d'appel des oies 605** : (25, 757b) a l ọyo ; (25, 768b ; + -ONE) ʊwkʊ[41] »

L'article AUCA (FEW 25, 754a-775b) est rédigé par Éva Buchi et André Thibault selon la structure modulaire qui explicite les changements morphologiques et qui lui permettrait d'être lié à la notice de Nauton. La balise tag-affix <affix type="suffix">[42] permettrait, d'un point de vue informatique, de reconnaître la structure explicitée du FEW et de créer un lien avec les données de l'atlas ainsi classée. Par exemple, les suffixés en -ĀCEU sont classés sous le paragraphe 5.a.α. (25, 766b), les suffixés en -ARDU sous 5.a.β. (25, 766b), et ainsi de suite. La nécessité et l'avantage de ce travail d'indexation des matériaux ont été mis en relief par Lothar Wolf :

---

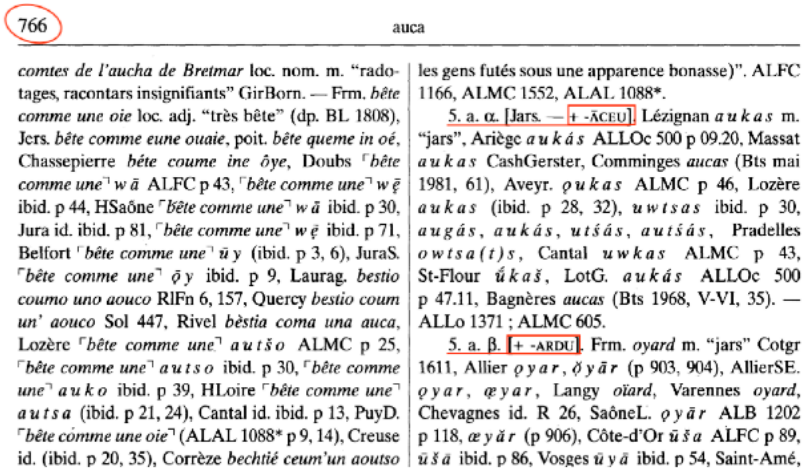
39 — Le *roitelet* est classé parmi les dérivés métaphoriques créés par perception visuelle à partir des « tratti attivi» de dimension (petit), couleur (brun) et forme (ronde), cf. L. STEINER, *I centri di espansione del cambio semantico. Per un'interpretazione cognitiva del* Französisches Etymologisches Wörterbuch, Strasbourg, ELiPhi, 2016, p. 172.

40 — P. NAUTON, L. WOLF, « Le fichier étymologique de l'*Atlas linguistique et ethnographique du Massif Central* de Pierre Nauton informatisé en collaboration avec Walburga Sarcher, Hans-Joachim Töpfer, Hartmut Wolf », p. 1. Nous remercions Guylaine Brun-Trigaud de nous avoir communiqué cette contribution de Wolf, qui semble n'avoir jamais vu le jour.

41 — W. SARCHER, L. WOLF : « Autour de l'indexation électronique de l'ALMC. Identification lexicale et étymologique. État des travaux », dans *Sempre los camps auràn segadas resurgantas. Mélanges offerts à Xavier Ravier*, éd. J.-C. BOUVIER, J. GOURC, F. PIC, Toulouse, CNRS - Université de Toulouse-Le Mirail (Méridiennes), 2003, p. 299-302, p. 300.

42 — P. RENDERS, *L'informatisation*, p. 246-248.

« Ce travail de séparation morphologique des données du fichier est la première tâche à accomplir. Le cas échéant, ce sera un travail pénible sans aucun doute, puisque chaque dérivé devrait être identifié étymologiquement (ce qui ne sera pas toujours possible) et être mis de nouveau dans l'ordinateur, mais comme entrée nouvelle, c'est-à-dire avec l'étymon, le numéro et le titre de la carte de l'atlas, le renvoi précis au FEW, et, bien sûr, les variantes phonétiques[43]. »



766 auca

comtes de l'aucha de Bretmar loc. nom. m. "radotages, racontars insignifiants" GirBorn. — Frm. *bête comme une oie* loc. adj. "très bête" (dp. BL 1808), Jers. *bête comme eune ouaie*, poit. *bête queme in oé*, Chassepierre *bête coume ine ôye*, Doubs ⌜*bête comme une*⌝ *w ā* ALFC p 43, ⌜*bête comme une*⌝ *w ę̄* ibid. p 44, HSaône ⌜*bête comme une*⌝ *w ā* ibid. p 30, Jura id. ibid. p 81, ⌜*bête comme une*⌝ *w ę̄* ibid. p 71, Belfort ⌜*bête comme une*⌝ *ū y* (ibid. p 3, 6), JuraS. ⌜*bête comme une*⌝ *ǭ y* ibid. p 9, Laurag. *bestio coumo uno aouco* RlFn 6, 157, Quercy *bestio coum un' aouco* Sol 447, Rivel *bèstia coma una auca*, Lozère ⌜*bête comme une*⌝ *a u t š o* ALMC p 25, ⌜*bête comme une*⌝ *a u t s o* ibid. p 30, ⌜*bête comme une*⌝ *a u k o* ibid. p 39, HLoire ⌜*bête comme une*⌝ *a u t s a* (ibid. p 21, 24), Cantal id. ibid. p 13, PuyD. ⌜*bête comme une oie*⌝ (ALAL 1088* p 9, 14), Creuse id. (ibid. p 20, 35), Corrèze *bechtié ceum'un aoutso* les gens futés sous une apparence bonasse)". ALFC 1166, ALMC 1552, ALAL 1088*.

5. a. α. [Jars. — + -ĀCEU]. Lézignan *a u k á s* m. "jars", Ariège *a u k á s* ALLOc 500 p 09.20, Massat *a u k a s* CashGerster, Comminges *aucas* (Bts mai 1981, 61), Aveyr. *ǫ u k a s* ALMC p 46, Lozère *a u k a s* (ibid. p 28, 32), *u w t s a s* ibid. p 30, *a u g á s*, *a u k á s*, *u t š á s*, *a u t š á s*, Pradelles *o w t s a (t) s*, Cantal *u w k a s* ALMC p 43, St-Flour *ů k a š*, LotG. *a u k á s* ALLOc 500 p 47.11, Bagnères *aucas* (Bts 1968, V-VI, 35). — ALLo 1371 ; ALMC 605.

5. a. β. [+ -ARDU]. Frm. *oyard* m. "jars" Cotgr 1611, Allier *ǫ y a r*, *ǫ̃ y ā r* (p 903, 904), AllierSE. *ǫ y a r*, *œ y a r*, Langy *oïard*, Varennes *oyard*, Chevagnes id. R 26, SaôneL. *ǫ y ā r* ALB 1202 p 118, *œ y ă r* (p 906), Côte-d'Or *ŭ š a* ALFC p 89, *ŭ š ă* ibid. p 86, Vosges *ŭ y ă* ibid. p 54, Saint-Amé,

La modélisation modulaire des articles de la refonte du FEW a été théorisée et mise en pratique par Anne-Christelle Matthey et Christel Nissille[44]. À Nicolas Mazziotta revient le mérite d'avoir, le premier, évalué la proposition d'une modélisation du FEW rétroconverti selon le modèle de l'*Ontology Web Language*. Il proposait en 2011 une comparaison entre la formalisation OWL de l'ALW, sur la base des « cinq relations de base » théorisées par Marie-Guy Boutier[45], et le FEW, puisqu'à son avis : « les deux entreprises [ont] les mêmes caractéristiques fondamentales, bien que le FEW soit, il est vrai, un petit peu plus complexe[46]. » Sa proposition d'enrichir les fichiers du FEW d'un « formalisme RDF [qui] est plus complet et plus expressif que le formalisme hérité de la représentation XML des documents[47] » représenterait un avantage indéniable

43 — W. Sarcher, « Autour de l'indexation », p. 302.

44 — A.-Ch. Matthey, Ch. Nissille, « L'irruption de l'informatique dans la rédaction du FEW », dans *Actes du 25e Congrès International de Linguistique et de Philologie Romanes (Innsbruck, 3-8 septembre 2007)*, éd. M. Iliescu, H. Siller-Runggaldier, P. Danler, Berlin, De Gruyter, 2010, p. 221-228.

45 — M.-G. Boutier, « Cinq relations de base pour traiter la matière géolinguistique : réflexions à partir de l'*Atlas Linguistique de la Wallonie* », *Estudis Romànics*, t. 30, 2008, p. 301-310.

46 — N. Mazziotta, « L'informatisation », p. 41.

47 — *Ibid.*, p. 60.

pour la mise en réseau du FEW ; à condition que toutes les autres ressources soient formalisées de la même manière. La réécriture de l'article étymologique de Nauton, en suivant le modèle structurel de l'entrée du FEW, permettrait de créer un lien automatique et immédiat entre les deux ressources. En effet, le modèle proposé par Mazziotta peut fonctionner avec les articles de la refonte en B-, qui ont été rédigés de manière « modulaire ». L'avantage de ce type de rédaction est que les matériaux sont préalablement présentés selon une structure qui en explicite les caractères (par exemple, les données morphologiques) et qui permet, comme il le montre bien, une structure hiérarchique explicite[48]. L'avantage de l'ALW est que ses données sont toujours explicitées ; dans le FEW, en vertu du principe de synthèse qui le caractérise, elles sont souvent implicites, et l'organisation structurelle des articles varie en fonction des souhaits des rédacteurs : le .fsml issu de la rétro-conversion n'arrive pas à les expliciter de manière automatique[49]. Nous nous demandons donc si une rétro-conversion du FEW en langage OWL demanderait un effort considérable pour expliciter de façon manuelle toutes les données implicites des articles du FEW. Comme Mazziotta semble l'envisager à la fin de son article, le formalisme OWL devrait être pensé pour pouvoir être appliqué *a posteriori* sur le résultat de la première rétro-conversion.

3. Conclusions : *minima etimologica*

La vocation du FEW d'être « eine darstellung des galloromanischen sprachschatzes » en fait naturellement le lieu idéal pour nourrir son inventaire lexicologique de l'apport des atlas. Le choix de fournir « une présentation […] synthétique des matériaux[50] » et,

48 — *Ibid.*, p. 56 ; Mazziotta le définit comme : « une hiérarchie explicite des critères, qui transforme la "liste" de formes en un ensemble ordonné exprimant leur interprétation. » Déjà N. MAZZIOTTA, P. RENDERS, « Vers un enrichissement raisonné de la rétroconversion du *Französisches Etymologisches Wörterbuch* (FEW) », dans *Proceedings of the XIV Euralex International Congress (Leeuwarden, 6-10 July 2010)*, éd. A. DYKSTRA, T. SCHOONHEIM, Ljouwert, Fryske Akademy, 2010, 1029 avaient préconisé que : « la microstructure rétroconvertie doit être enrichie par une annotation manuelle explicitant : 1/ la nature des regroupements en paragraphes (avec une hiérarchie à plusieurs niveaux) ; 2/ la nature du lien entre les regroupements en paragraphes ; facultativement, 3/ les regroupements non marqués (comme celui qui lie *substantifier* à son déverbal). »

49 — N. MAZZIOTTA, « L'informatisation », p. 56 : « Les principes d'organisation microstructurels du FEW et les relations de hiérarchie-équivalence et d'ordination de l'ALW sont similaires : dans les deux cas, les groupements de formes sont fondés sur une ressemblance qui rapproche des formes similaires en les opposant à toutes les autres. Ces oppositions sont complétées par une hiérarchie explicite des critères, qui transforme la "liste" de formes en un ensemble ordonné exprimant leur interprétation. »

50 — E. BUCHI, *Les Structures*, p. 90. Elle ajoute : « bref, le FEW, s'il est un *thesaurus* de tout le galloroman, n'est pas un *trésor* des patois, contrairement à ce que certains ont

vraisemblablement, les contraintes du *medium* papier dans lequel il a été conçu – il ne pouvait en être autrement – ont obligé Wartburg à des choix et à des sélections[51].

En insérant les données des *NALF* dans le FEW, sa « dimension thesaurus[52] » en serait par conséquent accrue dans sa position de centre collecteur des données lexicologiques galloromanes[53], qui accomplit le rêve d'un « dictionnaire total à la Jorge-Luis Borges ("La bibliothèque de Babel"), inlassablement augmenté, remanié, élagué et reclassé par des machines infaillibles[54] ». En même temps, injecter les données très nombreuses des *NALF*[55] implique une réflexion sur le risque de toucher à la « dimension monographique[56] » de l'œuvre, et donc de « briser la séquentialité du discours et de provoquer la perte d'informations importantes[57] » que seule la lecture intégrale

---

pu penser », puisqu'il « ne se sent pas obligé de citer toutes les formes *in extenso* ».

51 — La dimension quantitative des matériaux tirés des atlas doit prendre en compte la structure du FEW, dont le but reste celui de décrire de manière synthétique les trajectoires principales des parlers galloromans. M. Glessgen, « L'apport des "Inconnus" du FEW à la recherche étymologique », *L'Italia Dialettale*, t. 80, 2019, p. 633-676, p. 647 : « son objectif a été de tracer des lignes claires, dans les évolutions autant phonétiques que sémantiques, pour établir un cadre sûr et fiable aux trajectoires étymologiques galloromanes. »

52 — P. Renders, *L'informatisation*, p. 75.

53 — C'est un peu le même sort que celui du LEI, cf. les considérations de D. Baglioni, *L'etimologia*, Roma, Carocci editore (Le Bussole, 520), 2016, p. 120 : « La sensazione è quindi che il LEI, come del resto in buona parte già il FEW, sia andato sempre più trasformandosi in un *thesaurus* della lingua e dei dialetti, un "dizionario di dizionari" in cui la ricerca degli etimi non è più centrale perché l'etimo, stabilito *a priori* all'inizio del lemma, è meramente funzionale alla sistemazione del materiale. Tuttavia questa caratteristica, lungi dall'essere un difetto, è in realtà il punto di forza dell'opera. »

54 — K. Gebhardt, « L'apport des dialectes d'oïl (surtout entre 1300 et 1600) au lexique de la langue commune (d'après le FEW) », dans *Du mot au texte. Actes du III^e Colloque International sur le Moyen Français*, éd. P. Wunderli, Tübingen, Narr, p. 31-48, à p. 33. La citation est dans P. Renders, *L'informatisation*, p. 11.

55 — Déjà J. Séguy, « Les Atlas linguistiques de la France par régions », *Langue française*, t. 18, 1973 (*Les parlers régionaux*), p. 65-90, p. 87 avait mis en évidence les avantages des *NALF* pour l'étymologie : « Avec les atlas linguistiques l'étymologie a fait un bond en avant. Innombrables sont les monographies, brèves ou de longue haleine, qui procèdent des matériaux des atlas. Certes, les très nombreux dictionnaires et glossaires régionaux ou locaux avaient déjà fourni le corpus massif des parlers gallo-romans, comme on peut s'en rendre compte en lisant le monumental FEW de W. von Wartburg ; mais les atlas apportent des compléments, et, surtout, la méthode géographique possède des vertus incomparables : tel mot, considéré isolément, constitue une énigme étymologique, qui se résout toute seule à la vue des avatars voisins. En matière de morphologie, le contenu des enquêtes dialectologiques paraît souvent absurde, invraisemblable […] ».

56 — Nous reprenons, une fois de plus, cette dénomination de P. Renders, *L'informatisation*, p. 75, qui renvoie à E. Buchi, J.-P. Chambon, « Un des plus beaux monuments des sciences du langage : le FEW de Walther von Wartburg (1910-1940) » dans *Histoire de la langue française 1914–1945*, éd. G. Antoine, R. Martin, Paris, CNRS-Éditions, 1995, p. 935-963, à p. 952 : « L'ouvrage se présente, en fait, comme un ensemble structuré de monographies, dont la forme lexicographique n'est qu'un auxiliaire au service de la "visée globalisante" de Wartburg, qui l'anime et la domine ».

57 — P. Renders, « Des dangers de l'informatisation d'un document : le cas du FEW », *MethIS*, t. 2, 2009 (*Pratiques du document, Méthodes et Interdisciplinarité en*

d'un article et de son commentaire permet, en respectant la conception de Wartburg. Il sera en effet nécessaire de tenir compte de l'ajout de nouveaux sens (cf. *roitelet*), de propositions douteuses (cf. *grwà*), d'élargissements des aires d'attestations (cf. *grapada*). En plus, la dimension synchronique des *Atlas* entrerait dans un dialogue dynamique avec la dimension diachronique du FEW, en permettant ainsi de percevoir les changements des aires de diffusion des mots sur l'axe temporel (cf. *pastorèle*).

Or, la question des modalités de gestion d'une « mise à jour » du FEW est en discussion et nécessite des réflexions plus approfondies, autant d'un point de vue informatique que d'un point de vue théorique. Nous avons relevé, dans cette contribution, comme les possibilités d'hyperliens et de renvois entre FEW et *Atlas* changent au cas par cas, tout en restant assez rationalisables (du moins selon les trois cas de figure que nous avons isolés). Les balises "ici" (in-column index) et le tag-microstructure <p structid=""> (Renders 2015, 275-278) devraient faciliter considérablement des renvois plus fins avec le FEW rétro-converti. Dans le cadre d'un « chantier global commun[58] » les développements concernant le FEW pourront sans doute être harmonisés avec la planification informatique des entreprises de numérisation des *NALF*. Pour cette raison il est important d'envisager des « hyperliens réciproques entre les lexèmes [des Atlas] et le FEW informatisé[59] » qui rendraient plus simples :

1. une homogénéisation et une harmonisation entre les différentes entreprises, tout comme une lemmatisation selon des standards communs ;
2. l'entrée dans la dimension onomasiologique des données du FEW qui, dès le début, a été structuré de manière alphabétique, malgré les souhaits initiaux de Wartburg[60].

L'effort de concertation demandé concerne non seulement les chercheurs dialectologues, mais aussi les informaticiens responsables de concrétiser ces positions théoriques. Du côté du FEW, la

Sciences humaines), p. 179-195, p. 189.

58 — M. Glessgen, M. Sauzet, communication orale présentée durant la journée d'étude.

59 — *Ibid.*.

60 — H. Carles, *Französisches Etymologisches Wörterbuch, Guide*, p. 14. Le fichier onomasiologique, qui contient les fiches onomasiologiques rédigées par Walter Lacher et Marianne Müller, a été partiellement saisi (il couvre les vol. I-VII et le début des vol. XI et XVII) ; il est actuellement disponible à l'adresse : http://stella.atilf.fr/scripts/mep.exe?CRITERE=INDEX_ONOMASIOLOGIQUE;OUVRIR_MENU=MENU_UTILISATION_FEW;ISIS=mep_few.txt;s=s0b2922c8;ISIS=mep_few.txt;s=s0b2922c8;;ISIS=mep_few.txt.

disponibilité à la collaboration est réelle, au moins dans l'idée de son directeur, Yan Greub.

Le linguiste-dialectologue aura surtout le devoir d'interpréter la masse de données, fruit d'un demi-siècle de travail et de dizaines de dialectologues professionnels et engagés[61]. Colette Dondaine a déjà mis en relief l'apport d'un index étymologique de l'ALFC au FEW et à l'ALFC lui-même, puisqu'il « permet d'élucider en diachronie les formes de l'Atlas comtois et de les classer grâce à des références au FEW[62] » :

- révision des matériaux et élimination des éléments inexacts ;
- nouveaux apports ;
- possibilité de situer les types patois dans le temps et dans l'espace grâce au FEW (et au GPSR) ;
- distinction des types héréditaires du latin et des emprunts au français ;
- élimination des "faux inconnus" appartenant à des familles lexicales déjà répertoriées dans le FEW et leur réintégration dans l'architecture de l'œuvre.

Nous conclurons en synthétisant une série de *minima etimologica* mises en avant tout au long de notre contribution et qui pourraient favoriser la coordination des *Atlas* avec le FEW :

1. renvoi à la place précise dans l'article (page, colonne, structure ou unité lexicale selon les exigences) ;
2. informations de morphologie dérivationnelle (composition et dérivation) ;
3. informations sémantiques ;
4. identification du type d'apport au FEW (nouvelle donnée, correction, étymologisation d'un inconnu, etc.) selon des grilles partagées.

Nous sommes bien conscient qu'il ne s'agit là que d'une réflexion partielle et que Pascale Renders ou Nicolas Mazziotta ont déjà entrevu ces différents cas de figure, et bien d'autres. Notre point de vue est celui du rédacteur du FEW et nous réfléchissons donc à partir de ce que le FEW peut apporter à la structuration des matériaux des *NALF*, et ce que ces derniers peuvent apporter au FEW. Nous sommes conscient de l'importance du chantier, que nous souhaiterions commun, tout comme Martin Glessgen

61 — Cf. l'opinion de G. Tuaillon dans l'avertissement du vol. 5 de l'ALLy, p. VIII : « je m'imagine volontiers que, de tous ses travaux, c'est de celui-là, l'interprétation des cartes, que Pierre Gardette a retiré la plus grande satisfaction intellectuelle ».

62 — C. DONDAINE, *Trésor étymologique des mots de la Franche-Comté*, Strasbourg, BiLiRo, 2002 (la citation est à la p. X), les points sont résumés à partir des p. XI-XII).

et Maguelone Sauzet. Mais nous sommes également d'avis que, au fond, la relation entre l'investissement et le rendement est extrêmement positive : grâce à l'excellent travail tant de Wartburg que des auteurs des *NALF*, nous pourrons, en mettant en lien les deux ensembles, les enrichir considérablement et en tirer tout le profit.

Marco ROBECCHI
(Universität Zürich / FEW)[63]

63 — Nous remercions Yan Greub de ses observations sur une première version de cet article.