

◆ Un projet ambitieux : l'informatisation du FEW ◆

On connaît l'importance du *Französisches etymologisches Wörterbuch* pour les dialectologues de Wallonie et d'ailleurs. Ce dictionnaire rassemble une masse impressionnante de données dialectales galloromanes, publiées par Walther von Wartburg et ses successeurs depuis 1922 jusqu'à aujourd'hui (cf. www.atilf.fr/few). Les lexèmes, accompagnés de diverses informations (définition, datation, localisation, référence bibliographique etc.) et regroupés sous leur étymon commun, y sont présentés dans des articles structurés, qui offrent une analyse fine de la famille lexicale ainsi définie et de son histoire.

Le FEW constitue donc un ouvrage unique, incontournable et irremplaçable, qui sert de référence pour toute étymologisation de données wallonnes. Dans l'*Atlas linguistique de la Wallonie* par exemple, « le commentaire étymologique est très bref, réduit dans la plupart des cas à un renvoi au FEW et à la mention de l'étymon » (ALW 4, 10). Pour le walloniste, l'apport du FEW est certes sa visée avant tout étymologique (le classement opéré par l'*Atlas* étant d'abord onomasiologique), mais surtout la mise en perspective des données wallonnes dans un ensemble plus vaste (galloroman) :

« La mention du FEW dispense souvent de rappeler tout ce qui a été écrit sur les mots en cause : ce n'est pas dans l'ALW qu'on doit chercher la bibliographie du franco-provençal et du piémontais 'congère' (pas plus que, dans les notes de folklore, on ne s'attarde à dire dans quelles régions en dehors de chez nous un arc-en-ciel moins fort qui double le premier passe pour être l'arc du diable). » (ALW 3, 14.)

Malheureusement, l'ouvrage n'est pas d'un accès aisé. Les 25 volumes qui le constituent sont accessibles uniquement dans des bibliothèques spécialisées. Leur consultation n'est facilitée ni par la présentation condensée des matériaux, ni par la métalangue (une connaissance élémentaire de l'allemand étant obligatoire pour comprendre la plupart des articles), ni par les structures complexes du

dictionnaire ¹, qui nécessitent une initiation préalable. En outre, même l'utilisateur rompu à l'exercice n'est jamais certain de l'exhaustivité d'une recherche, puisque subsiste encore le problème épineux des classements multiples ² et des étymons cachés.

Un projet ambitieux est en route pour remédier à cette situation inconfortable : il s'agit d'informatiser l'ouvrage. Outre l'intérêt que présente un accès direct au dictionnaire sur le net, quatre objectifs sont poursuivis en particulier :

1. **Faciliter la lecture du texte du FEW**, notamment par la résolution des abréviations, l'explicitation du plan des articles longs et la mise en évidence de certaines informations. Il serait par exemple souhaitable de pouvoir repérer aisément des mentions géolinguistiques (« apik. » [= ancien picard], « liég. »), des localisations (« Mons ») ou encore des sigles bibliographiques particuliers (« BSLW », « Haust »).

2. **Permettre des recherches « transversales »**, c'est-à-dire effectuées à travers l'ensemble des articles du dictionnaire. Il s'agit donc non seulement de localiser directement l'endroit (voire les endroits) du FEW où est traité le mot wallon recherché, mais aussi d'obtenir une liste de lexèmes (d'étymons, de sources...) répondant à une requête complexe du type « Quels mots liégeois sont d'origine germanique ? »

3. **Assurer l'exhaustivité d'une recherche** par la récupération des lexèmes et étymons cachés, ainsi que des corrections situées à l'intérieur même des volumes du dictionnaire.

4. **Effectuer la mise à jour** de l'ouvrage. Un grand intérêt de l'informatisation consisterait à intégrer dans le FEW (selon des modalités à définir) les corrections apportées par ailleurs, qu'il s'agisse de données ponctuelles (antédations par exemple) ou de révisions d'une étymologie : « Il est des cas où l'explication de Wartburg est contestée sur la base de comptes rendus ou d'articles de spécialistes du wallon, avant tout d'Él. Legros et de J. Herbillon. » (ALW 4,10.)

L'informatisation du FEW devrait avoir lieu en partie au laboratoire ATILF (CNRS/Nancy Université), où a été informatisé le *Trésor de la langue française*

¹ Cf. Büchi (Eva), *Les Structures du 'Französisches Etymologisches Wörterbuch'. Recherches métalxicographiques et métalxicologiques*, Tübingen, Niemeyer, 1996.

² En ce qui concerne ce problème, voir Jean-Pierre Chambon et al., « Cent cas d'étymologie double dans le FEW », in : *Études de lexicologie, lexicographie et stylistique offertes en hommage à Georges Matoré*, Paris 1987, 165-184, ainsi que la série d'articles publiés dans *TraLiPhi* 27 (1989), 151-179 ; 28 (1990), 25-36 ; 30 (1992), 387-415 ; 32 (1994), 37-68 ; 37 (1999), 251-262.

(TLF)³. La méthode consiste, après avoir acquis le texte sous forme électronique, à y insérer un balisage XML, qui identifiera les divers types d'information (étymon, forme, définition, référence bibliographique, etc.). Ces balises permettront ensuite à un moteur de recherche d'accéder directement aux données demandées par l'utilisateur.

Une étude est en cours à Liège⁴ pour définir le balisage et, surtout, la faisabilité de l'automatisation du processus, les balises ne pouvant être intégrées manuellement dans un dictionnaire de cette ampleur. L'informatisation du FEW pose en pratique un certain nombre d'obstacles, tant au niveau de l'acquisition du texte en format électronique (numérisation et récupération des caractères phonétiques spéciaux) qu'en ce qui concerne la reconnaissance des types d'information par un automate. Des questions importantes – concernant notamment la mise à jour – demandent également réflexion, afin que les solutions adoptées répondent aux exigences de la communauté scientifique.

Dans le cadre de cette étude, un questionnaire a été conçu à l'attention des utilisateurs du FEW, afin de connaître leurs pratiques, leurs difficultés et leurs desiderata concernant l'informatisation. Les lecteurs intéressés peuvent le recevoir en envoyant un courriel à l'adresse suivante :

<pascale.renders@ulg.ac.be>.

Pascale RENDERS
Département de Langues et littératures françaises et romanes
Université de Liège
Place Cockerill, 3-5/40 - B-4000 Liège
Tél. +32/(0)4.366.9659

³ Devenu *Trésor de la langue française informatisé* (TLFi) dans sa version accessible sur la toile, cf. <www.atilf.fr/tlf>.

⁴ Thèse de doctorat effectuée par Pascale Renders, sous la codirection d'Éva Buchi (ATILF/CNRS) et de Marie-Guy Boutier (ULg).