

Numérisation du dictionnaire Bailly 1935 (v. 2)

Gérard Gréco, Mark De Wilde, Bernard Maréchal

22 février 2019

Une équipe d'environ **60 bénévoles**, pour la plupart diplômés de lettres classiques.

Un comité central de quelques personnes, pour organiser les travaux, en centraliser et en normaliser les résultats.

Des correspondants essentiellement français, mais aussi plusieurs Belges, un Suisse, un Espagnol et un Camerounais.

Les communications se font par messagerie électronique.

1. **Obtenir le texte jalonné du Bailly 1935, qui est libre de droits.**
 - donc les caractères mêmes du dictionnaire et pas seulement des images de ses pages ;
 - les principaux champs doivent être repérés :
 - les entrées ;
 - les auteurs, les œuvres, les références.
 - le style des caractères doit être préservé pour permettre une reproduction typographique fidèle.

2. Corriger les erreurs manifestes ou relevées dans la littérature.
3. Fournir gratuitement pour un usage non lucratif :
 - un fichier pdf ;
 - une base de données permettant de construire ultérieurement des fichiers html, txt et stardict.
 - Le fichier pdf permet une consultation de l'ouvrage ;
 - le fichier txt permet par exemple une adaptation pour les étudiants aveugles ;
 - le fichier stardict une consultation sur radio-téléphone.

Quelques statistiques :

- **21 691 682 caractères** (hors commandes) [12 099 190 signes pour le Gaffiot 2016,] soit presque deux fois plus que le Gaffiot 2016 !
- **106 315 entrées** + 3 594 entrées secondaires, soit 109 909 entrées au total [Gaffiot 2016 = 83 630] ;
- Par exemple, 21 497 références à l'Illiade qui pointent 11 133 vers sur un total de 15 337.

Les travaux se déroulent en dix étapes :

- 1. Reconnaissance optique des caractères des pages [centralisé].
- 2. Correction par les correspondants par cahiers de 30 pages papier des résultats de la reconnaissance optique [réparti].
- 3. Jalonnement du texte : repérer semi-automatiquement les divers champs du dictionnaire [centralisé].
- 4. Composition typographique du texte pour constituer les épreuves [centralisé].

- 5. **Première correction** des épreuves par les correspondants, par cahiers de 30 pages papier [réparti].
La correction s'effectue par annotation électronique d'un fichier pdf.
Elle porte sur l'*orthographe*, les *références*, la *punctuation* mais aussi le *style des caractères* ce qui permet de valider aussi le jalonnement.
- 6. Report de la première correction sur le fichier central [centralisé].
- 7. Nouveau tirage des épreuves [centralisé] et **seconde correction** par des correspondants **différents** [réparti].

- 8. Report de la seconde correction sur le fichier central [centralisé].
- 9. Revue partielle de l'étymologie et revue complète de la toponymie [réparti et centralisé].
- 10. Corrections en grand, vérification et harmonisation des références, normalisation des abréviations, composition typographique, préparation des fichiers de sortie [centralisé].

En résumé :

- Reconnaissance optique des caractères et correction préalable.
- Deux corrections supplémentaires par des correcteurs différents.
- **Revue de certaines notices étymologiques.**
- Revue complète de la toponymie.
- Corrections en grand et normalisation finales.

Exemple d'épreuve annotée :

The image shows a PDF document titled "/home/gerard/travaux/dictionnaires/bailly/répertoire suivi/epsilon-06_typographie-GG-lecture_BM.pdf" with a toolbar at the top. The document content is a dictionary entry for the Greek word "ἐξουθενέω". The text is annotated with several elements:

- Yellow sticky notes:**
 - Top left: "Autocollant: 28/12/2017 16:01" with a speech bubble icon.
 - Top right: "Autocollant: 28/12/2017 16:10" with a speech bubble icon.
 - Middle right: "Autocollant: 28/12/2017 16:10" with a speech bubble icon.
 - Bottom right: "Autocollant: 28/12/2017 16:10" with a speech bubble icon.
- Red circles:**
 - One around the word "ἐξουθενέω" in the first column.
 - One around the word "ἐξουθενέω" in the second column.
 - One around the word "ἐξουθενέω" in the third column.
 - One around the word "ἐξουθενέω" in the fourth column.
- Red triangle:**
 - Pointing to the word "ἐξουθενέω" in the fifth column.
- Other annotations:**
 - A red box around the text "ἐξουθενέω, depuis que, CALL. Apoll. 6 του. m. sign. SOPH. Ph. 493." in the fifth column.
 - A blue vertical bar on the right side of the page.

The text in the document includes:

ήρια, Luc. Pisc. 33, des mystères : n. irodier, bafouer : τυ...
 1 ; ἱερωσύνην, HDN...
 er la fête d'un dieu.
 ἰσδεῖς, v. ἐξόζω.
 οσιώω-ῶ, dédier, consacrer, P...
 20 || Moy. 1 m. sign. PLAT. Rat. 53
 tourner de soi par une expiation ;
 S. 15, 9 || 3 accomplir les sacrifices
 oires demandés par les dieux,
 M. 586 f.
 οστεῖζω, désosser, d'où ôter le
 , le pépin, la graine, Diosc. 2, 95 ;
 IF. (ORIB. 2, 277 B.-Dar.) (ἐξ,
 v).
 οστρακίζω [ᾶ] 1 bannir par
 isme, HDT. 8, 79 ; AND. 33, 24 ; LYS.
 7 ; PLAT. Gorg. 516 d || 2 p.
 unir, en gén. : ἐκ τοῦ οὐρανοῦ,
 acr. 4, du ciel.
 οστρακισμός, οὐ (ὅ) [ᾶ]
 isement par ostracisme, DS. 11, 87 ;
 Them. 22 (ἐξοστρακίζω).
 ὀστωσις, εως (ῆ) tumeur des os,
 se, GAL. De tumor. præter nat. p.
 ! (ἐξουθενέω) !
 ὅτι... depuis que, CALL. Apoll.
 6 του. m. sign. SOPH. Ph. 493.
 οτρύνω, pousser, exciter
 ent : τινά ἐπί τι, THC. 1, 84, qqn à

ἐξουθενέω-ῶ
 Reg. 2, 30 ; 8, 7 ; 3
 etc. ; NT. 18,
 Jos. (B. I.)
 ἐξουθενέω
 ἐξουθενέω (το)
 SPT. De 21, 7
 c. ἐξουθενέω
 (ἐξουθενέω).
 ἐξουθενήτικ
 faire aucun cas d
 DL. 7,
 ἐξουθενέω
 DL. 7,
 ἐξουθενέω-ῶ, t. ἐξουθενέω, SPT. I
 Reg. 8, 7 ; Sir. 34, 31 (var.).
 ἐξούλης δίκη (ῆ) 1 procès pour
 expulsion illégale (lat. actio unde vi)
 procès intenté par un plaideur qui prétend
 avoir été illégalement dépossédé, DÉM.
 540, 24 || 2 procès pour réclamation de
 propriété (lat. actio rei judicata) procès
 intenté par un plaideur qu'un jugement a
 investi d'un droit de propriété contesté par
 le défendeur qui a saisi le bien et se refuse
 à le restituer, DÉM. 543, 27 ; 1240, 22 ||
 → Gén. sg. seul cas us. d'ord. ; un ex. de
 l'acc. pl. ἐξούλας, AND. 10, 15
 (cf. ἐξειλέω ou ἐξειλλω).
 ἐξουρέω-ῶ : 1 évacuer par les urines,
 ARST. H.A. 6, 23, 2 || 2 uriner, EL. N.A.
 11, 18.
 ἐξούρησις, εως (ῆ) évacuation par
 les urines, AÉT. (ἐξουρέω).

113 ; DH. 4, 2142, le po
 ἡ ταμειυτικῆ ἐξ. DH
 questure ; ἡ τῶν δημ
 1866, le tribunal de la
 θαλάμου, HDN I, 12,
 mœurs (lat. præfectur
 ἐξ. ὄντες, ARST. Rhe
 ἐπ' ἐξουσιῶν, SPT. D
 autorités, les magistra
 ἐξουσία, au sens colle
 eux-mêmes, DH. 11, 3
 Rom. 13, 1 || II p. ext. c
 pouvoir de faire qqe d
 du pouvoir, d'où : 1 ab
 ressources, THC. 4, 39
 828 d || 2 éclat, splend
 (ἐξειμι 1).
 ἐξουσιάζω : 1 avc
 l'autorité, ARST. Eud.
 SPT. 2 Esdr. 7, 24 ; Neh
 Eccl. 19 ; 5, 18, etc. ; 1 A
 NT. Luc. 22, 25, sur qq
 pass. être soumis à un
 Cor. 6, 12 || 2 donner l
 τινος, SPT. Eccl. 6, 5, d
 pouvoir de faire qqe d
 ἐξουσιαστής, οὐ
 maître, SPT. Esai. 9, 6 ;
 Migne (ἐξουσιάζω).

Auxiliis quibus ?

- Parrainage de M. Laurent Lafforgue, mathématicien, médaille Fields 2002, défenseur de l'École libératrice.
- Soutien du Pr Helma Dik de l'université de Chicago (Department of Classics).
- Aucune aide institutionnelle française.
- L'excellent logiciel de reconnaissance optique Abbyy.
- Les logiciels libres : Gnu Linux, Gnu Emacs, Vim indispensables pour traiter de très grands fichiers, Fossil pour l'archivage des versions successives.
- les bibliothèques disponibles sur Internet, en particulier Archive.org.
- La compagnie Monotype Ltd. qui nous a permis d'utiliser gratuitement la belle fonte Greek 90.

Quando ?

- Début des travaux : automne 2016.
- Fin des corrections du texte : fin 2018.
- Parution : ? Tout dépendra des travaux de normalisation en cours.