



Séance du 13 juin 2015

Initiation au corpus de vieil anglais

Merci à Oxana pour ses suggestions !

Comment se connecter au DICTIONARY OF OLD ENGLISH WEB CORPUS par l'intermédiaire de la Bibliothèque de la Sorbonne ?

À partir de la page d'accueil du site de la Bibliothèque (<http://catalogue.biu.sorbonne.fr/>), il faut cliquer en haut à droite sur *Ressources électroniques*. Sur la page suivante, dans le menu vertical de gauche, cliquer sur *Dictionnaires de langues*. Sur la page suivante, descendre très bas (sous *Autres langues / Anglais*) et cliquer sur le DICTIONARY OF OLD ENGLISH WEB CORPUS. La fenêtre suivante vous réclame vos identifiants (n° de carte de lecteur & date de naissance), dès qu'ils sont saisis vous êtes rebasculé sur la page d'accueil du CORPUS.

Quel type d'informations le DICTIONARY OF OLD ENGLISH WEB CORPUS contient-il ?

Il s'agit d'une saisie informatisée de tous les textes en vieil anglais, qui ont été codés et numérisés à partir des meilleures éditions disponibles à l'époque — donc pas forcément les plus récentes qui soient. Chaque œuvre s'est vu attribuer un numéro, dit « Cameron number », du nom du saint homme qui s'est chargé du travail de classification. Les références en question suivent toutes la même formule : une lettre, puis des nombres, éventuellement séparés par des points (ex. *Beowulf* = A4.1, *Byrhtferth's Manual* = B20.20.2). Vous retrouvez cette référence chaque fois que les œuvres sont mentionnées dans le logiciel et vous pouvez les désigner par cette référence si vous souhaitez restreindre votre recherche à une œuvre. ♦ **Attention** : contrairement au MIDDLE ENGLISH COMPENDIUM, qui est une sélection de citations organisées selon les rubriques du *Middle English Dictionary*, le CORPUS contient tous les textes vieil-anglais, mais n'est pas relié à un dictionnaire (le *Dictionary of Old English* est actuellement achevé pour [A-G]). Ceci a deux conséquences pratiques :

- ♦ Lorsque vous recherchez des occurrences d'un mot ou d'un groupe de lettres, le CORPUS vous permet d'accéder à une liste d'occurrences, sans que le sens ne soit précisé. Ainsi, toutes les occurrences de *þā* vous seront livrées en vrac, quel qu'en soit le sens.
- ♦ Il faut repérer vous-même les différentes variantes orthographiques des termes que vous allez rechercher — un outil automatique est là pour vous y aider.

Comment utiliser le OLD ENGLISH CORPUS ?

À partir de l'écran d'accueil du site, les rubriques permettent de choisir entre divers types de recherches : *Simple Search*, *Boolean Search*, *Proximity Search*, puis *Bibliography Searches*, *Word Wheel* et *Word Count*. NB. *Word Count*, qui indique combien de formes de mots différentes figurent dans un texte donné, tient également compte des mots étrangers, ce qui permet de repérer les textes macaroniques ou glosés.

À quoi correspondent les trois types de recherche possibles ?

Je vous propose de regarder tour à tour les trois premiers types de recherche, en cliquant pour ouvrir les écrans. Il y a tout d'abord des points communs. ♦ À chaque fois, on vous propose de rechercher un ou plusieurs termes que vous tapez vous-même dans une ou plusieurs cases. ♦ À chaque fois, vous pouvez restreindre votre recherche à une catégorie de textes : tout le corpus, seulement la prose, seulement la poésie, seulement les gloses. ♦ À chaque fois, vous pouvez restreindre votre recherche à un ou plusieurs textes (si vous voulez travailler par exemple sur un seul manuscrit). Il faut alors utiliser le « *Cameron number* » ou le « *Short title* ».

♦ **Attention** : La première case porte la mention *Fragmentary*. En faisant dérouler cette case, on voit les autres possibilités : *Begins by* ou *Whole word*. Si vous laissez cette case en l'état, l'ordinateur vous donnera toutes les occurrences de mots qui comportent le groupe que vous avez tapé, ce qui provoque parfois des avalanches de réponses inutiles. Vous pouvez faire l'expérience en cherchant dans la suite de caractère *ece* dans *Beowulf* (dont le *Short title* est « *Beo* ») : on obtient 72 occurrences, la deuxième dans la liste inclut *ece dryhten*, mais d'autres repèrent simplement la suite de lettres *ece* à l'intérieur d'un mot.

♦ Ensuite, tout dépend de ce que vous voulez faire faire à l'ordinateur : rechercher un seul mot (*Simple search*) ou plusieurs.

À quoi sert la Word Wheel ?

La *Word Wheel* permet de voir immédiatement et précisément si une forme donnée est attestée, et en combien d'occurrences. On peut soit taper la forme recherchée dans la case et « l'envoyer » : on accède alors à un menu alphabétique déroulant où le mot est surligné, soit cliquer sur la première lettre de la forme dans l'alphabet proposé, et l'on accède ensuite au même menu déroulant. En cliquant sur le mot, on fait apparaître toutes les occurrences en contexte. *Latin Word Wheel* permet de voir comment un mot latin est traduit en vieil anglais, non seulement dans les gloses mais aussi dans les textes bilingues par ex.

Quelle est la différence entre *Boolean Searches* et *Proximity Searches* ?

♦ Dans le premier cas, l'ordinateur recherche les termes à l'intérieur d'un même « segment ». Ces segments ont été découpés par les compilateurs du corpus au sein des textes, pour pouvoir les indexer plus facilement. En poésie, ils correspondent parfois à des vers, ou à des groupes de vers. ♦ Dans la *Proximity Search*, c'est vous qui précisez de quelle distance maximale vous souhaitez que les termes soient distants. ♦ Vous pouvez par exemple comparer les résultats en recherchant par les deux fonctions possibles les occurrences de la formule *wordum wrixlan* sur l'ensemble du corpus. Je vous conseille de taper seulement **word** et **wrixl** en utilisant la fonction « *Begins by* » pour permettre à des variantes d'apparaître.

Comment rechercher un terme dans une œuvre donnée ?

Il faut préciser soit le « *Short title* » soit le « *Cameron number* ». Attention, dans ce dernier cas il faut absolument taper les lettres en minuscules (pour *Beowulf*: « a4.1 »), et soigner la ponctuation, sinon l'ordinateur se contente de vous asséner « *your query produced no result* » sans vous suggérer que vous avez commis une erreur de syntaxe (c'est énervant) ! Pour trouver le *Cameron number* d'une œuvre, le plus rapide est d'ouvrir *Bibliography Searches* et de taper un des mots du titre ou le nom de l'auteur.

Un outil nécessaire : le *Variant phrase search* disponible sur le site internet du **DICTIONARY OF OLD ENGLISH** <www.doe.utoronto.ca/>

♦ Il s'agit d'un logiciel qui vous permet à partir soit d'un soit de plusieurs termes de rechercher automatiquement les variantes orthographiques de ces termes et de les confronter au CORPUS. Vous pouvez utiliser cet outil sans être connecté par la Sorbonne, il est gratuit.

♦ Pour accéder à cet outil, il vaut mieux quitter le CORPUS (par ex. en ouvrant un nouvel onglet dans votre navigateur), taper l'adresse internet du **DICTIONARY OF OLD ENGLISH** : <www.doe.utoronto.ca/> et ensuite cliquer sur *Research Tools* puis sur *Variant Phrase Search*. Vous voyez ensuite apparaître une case pour taper votre mot/expression. Ensuite, (avec *Performing wildcard*) l'ordinateur va « faire tourner » les voyelles du/des mots, voire, si vous choisissez la deuxième option, les permuter de place avec les consonnes et vous proposer les variantes attestées dans le CORPUS, en précisant pour chacune le nombre d'occurrences.

♦ Pour voir les différences entre les deux fonctionnements, je vous propose de taper **dryhten** et de cliquer la case du haut, puis de comparer avec les résultats en cliquant la case du bas, qui « déplace » également les voyelles. Malheureusement, en cliquant sur les mots, on ne rebascule pas automatiquement dans le CORPUS, il faut faire un copié-collé d'un onglet à l'autre.

♦ Attention, cet outil ne tient aucun compte du sens des mots, il procède par simple proximité graphique. Chaque forme doit être vérifiée en contexte. [En modifiant les lettres de *dryhten* on obtient entre autres la citation suivante : *He se se ðe gedafenlice & endebyrdlice to cymð, hu he ðæron drohtian scyle* (ce verbe signifie « se conduire » : même racine que *drihten*... mais pas lui !)]

♦ L'ordinateur met un certain temps à modifier les expressions à plusieurs termes. Je vous propose de lui demander de rechercher les variantes de **ece dryhten** en cliquant sur la case du bas. On a l'impression qu'il ne se passe rien... Surtout ne touchez plus au clavier ni à la souris, et gardez patience, vous allez bientôt être récompensé par une liste d'occurrences !!! Dans cette liste, vous voyez que toutes les variantes ne sont pas celles que vous « attendiez » : l'ordinateur a perdu de précieux moments à rechercher les co-occurrences de *ic* ou *eac* et *dryhten*, qui ne sont pas pertinentes ici !

Exercices pour explorer le **OLD ENGLISH WEB CORPUS** (les 4 premiers sont inspirés par des suggestions d'André Crépin, les suivants nous sont proposés par Oxana)

N.B. Dans les réponses, les œuvres sont désignées par leur *Short title* : en cas de doute, cliquez sur les références en bleu et vous verrez apparaître la fiche signalétique de l'ouvrage, ainsi que l'environnement de la citation.

1. SIMPLE SEARCH + BIBLIOGRAPHY SEARCH Rechercher les occurrences de **giedd** dans *Wulf and Eadwacer* (il faut commencer par trouver le « *Cameron number* » ou le « *Short title* » du poème, en utilisant la fonction *Bibliography Searches* sur l'écran d'accueil du CORPUS) [Faire la même chose à partir du *DOE*]
2. BOOLEAN SEARCH Affichez à l'écran les occurrences de **fers** (vers, verset), en évitant d'être parasité par les occurrences de *ferscum* (il faut utiliser la fonction « *Not* » à la place de « *And* »).
3. PROXIMITY SEARCH Répondre à la question : la formule **ece dryhten** est-elle utilisée dans les textes en prose ? Qu'en est-il de ses formes déclinées ?
4. DICTIONARY OF OLD ENGLISH RESEARCH TOOLS Quelles sont les variantes orthographiques de **giedd** ? Faire apparaître à l'écran les différentes listes d'occurrences correspondant aux différentes graphies du terme.
5. WORDWHEEL Quelles formes déclinées du mot latin **pestilentia** trouve-t-on dans le corpus vieil-anglais ?
6. WORDWHEEL + RESEARCH TOOLS Trouver les traductions vieil-anglaises du mot latin **pestilentia**.
7. Peut-on identifier le genre des deux principaux mots obtenus qui désignent « la peste » ? Comment ?
8. Peut-on trouver quels autres mots latins sont traduits par **cwyld** et **cwealm** ?