Utilisation d'un concordancier

Une concordance, dans sa forme la plus simple, est une liste alphabétique des mots d'un texte, donnés avec les contextes dans lesquels ils apparaissent. Le plus souvent on représente les mots sous la forme d'un index KWIC (*KeyWord In Context*), dans lequel chaque mot est centré dans une fenêtre de taille fixe (cf. figure 1).

```
1320      taste it is that such  poor cattle always have in their mouths
948              of sparing the  poor child the inheritance of any part of
778         small property of my  poor father, whom I never saw--so long
1870        desolate, while your  poor heart pined away, weep for it
947                Miss, if the  poor lady had suffered so intensely
1884              the love of my  poor mother hid his torture from me
1615      stockings, and all his  poor tatters of clothes, had, in a long
1577             faded away into a  poor weak stain.  So sunken and
1001           on your way to the  poor wronged gentleman, and, with a
1036          detachment from the  poor young lady, by laying a brawny hand
```

FIG. 1 – Concordance de *poor* dans *A Tale of Two Cities* (Dickens)

Un **concordancier** est donc un programme permettant de réaliser de telles listes sur un corpus de son choix. En général, les concordanciers donnent aussi diverses informations statistiques sur le corpus traité.

Manipulations

1. (a) Lancer le concordancier « Z :/Taln/Concordance/MonoConc/Monopro.exe »[1]
   (b) Charger le fichier `Y:/Amsili/Hugo_ane.txt` (« File|Load Corpus »).
   (c) Essayer les différentes fonctionnalités :
   affichage des concordances (« Concordance|Search/Advanced Search » + « Sort »),
   tris par fréquence des mots, (« Frequency|Corpus frequency data|frequency order/alphabetical order »),
   affichage des collocations (« Frequency|Collocate frequency data »),
   distribution d'un terme dans le corpus (« Display|Distribution »)...
   (d) i. Donner le nombre d'occurences du mot « *Kant* » dans ce texte
   ii. Relever une phrase complète dans laquelle le mot ***âne*** est suivi de l'adjectif ***libre***
   iii. Combien trouve-t-on d'occurence de la forme négative *ne* ?
   Combien trouve-t-on de séquences *ne* un_mot_quelconque *pas* ?
   iv. Comparer la distribution du mot ***âne*** et du mot ***homme***
2. (a) Lancer « Internet Explorer » ou « Netscape », et choisir un texte parmi ceux qui sont proposés sur le site `http://abu.cnam.fr`.
   (b) Sauvegarder l'œuvre choisie en **texte seul**[2]. Indiquez sur le feuille l'œuvre choisie.
   (c) Décharger le corpus précédent (« File|Unload corpus »)
   (d) Charger l'œuvre que vous venez d'enregistrer sur votre ordinateur de travail.
   (e) i. Donner les 10 **mots pleins** les plus fréquents.
   ii. Relever les collocations pertinentes pour 5 de ces mots.
   iii. Donner le nombre d'occurrences (en pourcentage), du mot *en* suivi par un participe présent.

Références

- Le concordancier que nous utiliserons est MonoConc Pro (version 1.0), développé par Michael Barlow (version commerciale, mise à disposition pour l'enseignement).
- Il en existe d'autres téléchargeable gratuitement. Par exemple : `http://vlc.polyu.edu.hk/pub/concapp/concapp.htm`
- À propos de l'utilisation d'un concordancier pour les études littéraires (avec une présentation comparative de plusieurs concordanciers), voir *Using a Concordancer in Literary Studies*, by Maria Rosario Caballero Rodriguez, publié dans *The European English Messenger*, Vol VII/2, pp. 59–62, 1999 (`http://vlc.polyu.edu.hk/Concordance/Review/programa.htm`).

[1] Si les icônes n'apparaissent pas, faire Affichage/Options/Affichage/Afficher les fichiers cachés. Vérifier également que Masquer les extensions pour les types de fichiers connus n'est ***pas*** cochée, pour que les extensions s'affichent.

[2] Cliquer sur **Texte complet** (en bas à droite), puis sur **HTML** pour visualiser le texte au format HTML. Aller ensuite dans le menu Fichier|Enregistrer sous du navigateur, pour enregistrer une version texte (**.txt**).