

Analyse statistique des données textuelles

Corpus Street wear

Les Etapes :

- Constitution du corpus, recueil des données
- Nettoyage du texte et encodage. (Lexico3, Hyperbase, Alceste...)
- Lien vers les différents balisages
- Voir aussi les corpus encodés sur la base textopol
- [Textopol.u-pec.fr](http://textopol.u-pec.fr)>base Textopol > Base Dynamique >consulter la base

Les approches exercices sur le corpus "genres" :

Wordle.net

<http://wordle.net/>

Tester un extrait du corpus en le soumettant au site wordle.net.

Exemple comparer la partie "Gothiques" à la partie "skaters" à partir de deux nuages de mots.

Quelles observations ?

TreeCloud

Même procédure

<http://www2.lirmm.fr/~gambette/treecloud/SiteDownloadFra.htm>

Cordial

Quels sont les diagnostics stylistiques, syntaxiques, sémantiques notables ? (comparer deux parties)

Alceste :

Quelles classes sémantico-sémantiques ? Sur l'ensemble du corpus

Tropes :

Quels univers sémantiques ? Quelles catégories de mots?

Lexico3 :

Un outil minimal portant sur la forme graphique.

L'index hiérarchique, les segments répétés, les principales caractéristiques quantitatives du corpus, l'analyse factorielle des correspondances, les cooccurrents spécifiques, les spécificités.

TextObsever:

Extraire des motifs textuels, approfondir le commentaire de l'analyse factorielle des correspondances, projeter des listes thématiques sur la configuration factorielle...

Travail à rendre :

Choisir parmi les outils proposés ci-dessus une méthode permettant de compléter l'approche Lexico3 et ajouter au rapport d'analyse du corpus "genres". Un outil de visualisation, une expérience de catégorisation, un outil d'analyse (Cordial, Tropes, Alceste)

Ressources :

Informations complémentaires disponibles à l'adresse habituelle :
<http://infotext.free.fr/spjp>

(Rubrique genres de discours, TD)

Consulter également le cours « Genre » sur Eprel

Site de référence pour la lexicométrie/textométrie:

<http://texopol.u-pec.fr>

TextObserver:

<http://textopol.u-pec.fr/textObserver>

Tutoriels Alceste sur Eprel et sur Textopol; lien de téléchargement vers les outils logiciels:

textopol.org/secure

- livplasma (<http://www.liveplasma.com/>)
- - Antconc