

Analyse linguistique :

Pour le choix du corpus en français, j'ai décidé de diversifier les sources en sélectionnant des URL issues de domaines littéraires, scientifiques et journalistiques. Cette diversification des sources me permet d'en dégager un sens plus global du terme étudié.

Le mot choisi étant orientalisme, je ne souhaitais pas limiter ma recherche à ce terme précis, mais l'élargir à ses variations lexicales telles que orientaliste, oriental, orient pour capturer une vision plus large et plus complète de ce thème.

Pour pouvoir faire cet analyse nous avons utiliser la partition python cooccurrent.py

```
« python3 cooccurrent.py ../Token/contextes/fr-*.txt --target ".*[oO]rient.*" --match-mode regex  
>../PALS/francais_pals/cooc-contextes-fr.tsv  
python3  
tsv_to_html_table.py  
"../PALS/francais_pals/cooc-contextes-fr.tsv"  
"../PALS/francais_pals/cooc-contextes-fr.html"  
python3 wordcloud.py »
```

Ces commandes ont permis dans un premier temps à partir des fichiers tokenizer des contextes des 50 URLs de l'extraire et de calculer les apparitions du mot ".*[oO]rient.*" (ce regex permet de prendre en compte la racine du mot orientalisme et ses variations, peu importe sa place dans une phrase). Le résultat est un tableau TSV qui est transformé en HTML grâce à la commande suivante permettant de lancer une partition python tsv_to_html.

Et enfin le wordcloud est constitué en retirant les mots stop (mot peu significatif) pour faire l'analyse globale.

Résultats

Les mots les plus intéressants, selon le tableau des co-occurrences en enlevant les mots peu significatifs sont Moyen(210), Occident(136), siècle(128), paris(63), artistes(54), said(49), peinture(49), français(49), Europe(49), femmes(43).

Nous observons que selon les cooccurrences ce terme est souvent cité par opposition à l'Occident, ou encore dans des contextes culturels et littéraires comme Victor Hugo et ses écrits inspirés de cet univers, artistes comme Delacroix qui a contribué au mouvement artistique autour de cet imaginaire, ou Edward Said, célèbre écrivain palestino-américain qui a théorisé cette thématique relatives aux études post-coloniales.

Ces résultats confirment que l'orientalisme est le fruit des échanges littéraires, artistiques et culturels de l'Orient vers l'Occident et d'autres par elle est le reflet d'un imaginaire fantasmé par l'Occident.

Ce courant a entre autres nourri des stéréotypes dans une moindre mesure et a contribué à une légitimation d'une domination coloniale. (stéréotype(17), domination(18), coloniale(23) clichés(11)). Nous pouvons constater également que l'orientalisme en français fait référence géographiquement surtout au Maghreb(9), au moyen orient(11), et aux ottomans(11) et à l'Asie (25).

Conclusion

Nous pouvons conclure qu'en français, l'orientalisme évoque principalement la peintures, Paris, la littérature et que ce terme se construit par opposition à occident. Cela correspond largement à l'idée que nous nous faisions avant de commencer ce projet. Ce qui a été surprenant c'est de voir que ce terme évoque aussi la domination et la colonisation, et qu'en terme de zone géographique cela s'étant de l'Afrique à l'Asie.