

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-fr.tsv
python3 tsv_to_html_table.py "../PALS/francais_pals/cooc-contextes-fr.tsv"
"../PALS/francais_pals/cooc-contextes-fr.tsv-fr.html"
python3 wordcloud.py »
```

Ces commande ont permis dans un premier temps à partir des du fichier tokenizer des contextes des 50 urls de l'extraire et de calculer les apparitions des mot ".*[oO]rient.*" (ce regex permet de prendre en compte la racine du mot orientalisme et ses variations, peut importe sa place dans une phrase). Le rsultat est un tableau TSV qui est transformé en html grace a la commande suivante permettant de lancer une parition pyhon tsv_to_html.

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

Résultats

Les mots les plus intéressants, selon le tableau des co-occurrences en enlevant les mots peu signifiant 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 a occident, ou encore dans des contextes culturelles et littéraires comme Victor Hugo et ses ecrits inspirés de cet univers , artistes comme Delacroix qui a contribué au mouvement artistique autour de cet imaginaire, ou Edward Said célèbre ecrivains palestino-americain qui a theorisé cette thématique relatives aux études post-coloniale.

Ces résultats confirment que l'orientalisme est le fruits des echanges litteraires, artisitiques et culturelles 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 à contribué a une légitimation d'une domination coloniale. (stéréoptye(17), domination(18), coloniale(23) clichés(11)). Nous pouvons constaté également que l'orientalisme en français fait référence géographiquement surtout au Maghreb(9), au moyen orient(11), et aux ottomans(11)et a 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 faisons avant de commencer ce projet. Ce qui a ete 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.