

Humanités Numériques

Cours 8 : exercices Ngram viewer

Loïc Grobol

2021-11-09

Ces questions portent sur le corpus français (French) de Google Books dans le Ngram Viewer :

1. Combien y a-t-il de bigrammes dans la phrase « Le compositeur de cet opéra est un génie. » ?
2. Combien y a-t-il de trigrammes dans la phrase « Marie regarde Pierre avec un télescope » ?
3. Quel est le mot le plus fréquent de la langue française ?
4. À quelle date le mot « internet » devient-il plus fréquent que le mot « minitel » ? Trouvez un second indice de cette évolution technologique.
5. Quel est le terme qui a le plus chuté devant « toujours » en deux cents ans ?
6. Quand « avant-dernier » devient-il plus fréquent que « pénultième » ?
7. Quel est le plus fréquent : « bouffe » (du verbe bouffer) ou « bouffe » (le nom) ?
8. Quelle est la catégorie grammaticale la plus fréquente en français ?
9. Trouve-t-on plus fréquemment un nom ou une conjonction devant « alors » ?
10. Que signifie la requête : `Staline=>*_ADJ`
11. Que signifie la requête : `_ROOT_=>_NOUN_,_ROOT_=>_VERB_`
12. Pourquoi les courbes de « Je peux » et « Je veux » ne sont-elles pas plus similaires ? (Indice : l'explication est linguistique)
13. Quelle est l'évolution des termes « plufieurs », « plusieurs », « prefque » et « presque » entre 1750 et 1830 dans le corpus French 2009 ? Et dans le corpus French ? Pourquoi ?
14. Est-ce que le S long a disparu avant l'esperluette « & » (aux mêmes dates sur le corpus French 2009) ?