Quelle est l'expression la plus utilisée en France ?

CC0 Public Domain, via Pixabay

Notre réponse du 23/10/2017

Je n'ai malheureusement pas trouvé de réponse indiscutable à votre question. Le site suivant donne les expressions consultées les plus fréquemment sur le site spécialisé mais ce n'est pas exactement la même chose :

Expressio : Les expressions françaises décortiquées
http://www.expressio.fr/palmares.php
L'expression la plus consultée est « **voir midi à sa porte** », suivie de « **soupe-au-lait** » puis de « **avoir un cœur d'artichaut** »

Si vous vous intéressez aux expressions, vous pourrez lire avec bonheur le livre suivant :
Georges Planelles : **Les 1001 expressions préférées des français**

Le site suivant s'inspire des journalistes et donnent des indications sur les expressions à la mode mais cela veut-il dire pour autant que ce sont les expressions les plus utilisées dans l'absolu ?…

Un topito se consacre aux **Top 15 des expressions « toutes faites » les plus utilisées par la presse**
Qui seraient dans l'ordre :

1. « la cerise sur le gâteau »
2. « dans la cour des grands »
3. « le vent en poupe »

4. « un pavé dans la mare »
5. « caracoler en tête »

Le problème soulevé par votre question est la définition d'un corpus. Si vous considérez la langue parlée, il faudrait pouvoir capter toutes les conversations dans un instant donné afin de pouvoir faire des statistiques.
La question d'un corpus est plus aisée avec les textes écrits. C'est la démarche de l'Atilf **Analyse et Traitement Informatique de la Langue Française** qui a présidé à la constitution du **Trésor de la langue française informatisée.**

Pour en savoir plus sur les corpus sur le site de la Queens University de Kingston :
Les corpus

> *Ni le vocabulaire d'un individu, ni le lexique d'une langue n'est directement accessible dans toute son étendue. Si on veut travailler sur les masses d'unités lexicales, il faut trouver une autre clé d'accès. L'une des plus importantes consiste en l'utilisation de corpus. Un corpus est une collection de productions linguistiques qui se ressemblent par leur origine. Ainsi, au niveau de l'écrit, un roman de Camus formerait un corpus. De même, tous les romans de Camus formeraient un autre corpus, qui engloberait le premier. Et tous les romans français produits au XXe siècle formeraient un corpus même plus gros.*
> *Au niveau oral, une conversation pourrait former un corpus, un ensemble de conversations dans les mêmes circonstances formeraient un autre, et ainsi de suite.*
> *Il existe pour le français plusieurs corpus importants. Au niveau écrit, le Dictionnaire des fréquences du Trésor de la langue française fournit des données sur un corpus important de textes français du XIXe et du XXe siècles. Au total, ce corpus comprend plus de 70,000,000 d'occurrences (32,663,549 occurrences pour le XIXe et 37,653,685 occurrences pour le XXe siècle).*
> *Pour le français parlé, le Français fondamental, basé sur 275*

conversations, comprend 312,135 occurrences divisées entre 7995 formes.
Le nombre d'occurrences d'une forme donnée dans un corpus s'appelle sa fréquence absolue. Pour obtenir la fréquence absolue d'un lexème, on additionne la fréquence de ses différentes formes. Lorsqu'on analyse un corpus selon les fréquences absolues, on constate un certain nombre de relations.

Cordialement,

Eurêkoi – Bibliothèque Publique d'Information