

# ESARTINULOC !

Ne cherchez pas le sens caché de cette suite de lettres. Le mot n'a pas de sens, il n'existe pas et ne figure dans aucun dictionnaire usuel. Il s'agit de l'énoncé des lettres les plus fréquemment utilisées en français contemporain, rangées par fréquence décroissante. Certains ont voulu y voir une anagramme. D'après Georges Perec, grand maître en maniement des mots, seul le terme « ulcérations » serait l'anagramme possible d'ESARTINULOC. Parmi les anagrammes de ces onze lettres on peut trouver la forme conjuguée (« consulterai ») ou une phrase de deux mots (« sire coulant », « ta réclusion », « t'inoculeras », « le courtisan »), voire de trois mots (« rose lac nuit », « un castel roi », « l'ancre souci », « ô lutin sacré », « nul si atroce »), et quelques autres encore. Le même Georges Perec en a fait un poème dont la suite des vers est constituée d'anagrammes composées à partir dudit ESARTINULOC. Pour les amateurs, le voici <sup>1</sup> :



**Dominique Maillard**  
Ingénieur général des mines  
honoraire

LE CUTAS NOIR	L'écu, tas noir
TARECLUSION	Ta réclusion
TINOCULERAS	t'inoculera sa nuit :
ANUITLORSEC	l'or sec,
LASTRECOIUN	l'astre coi. Un roi ?
ROIUNCASTEL	Un castel nul !
NULSIATROCE	Si atroce Souci...
SOUCILANTRE	
ROSETUCALIN	L'ancre rose :
ESUNCORAILT	tu câlines un corail.
OUTLACSIREN	Tout lac, sirène,
EALUSONCRIT	a lu son cri.
UASCONTELIR	Tu as conté l'ire,
EOLUTINSACR	ô lutin sacré.
EARTICULONS	Articulons !

On doit noter que le classement de la fréquence des lettres dépend de l'échantillon retenu. Ainsi *Wikipédia* a-t-il procédé à l'analyse de la fréquence des lettres dans son propre corpus de documents en français, le résultat donne EAISNRTOLUD, sensiblement différent de ESARTINULOC. En particulier le D remplace le C dans le classement et l'ordre de fréquence de certaines lettres change. En effet, la nature, le style et la teneur des messages influent fortement sur le résultat. Par exemple, s'il y a beaucoup de verbes à la deuxième personne du pluriel (le vouvoiement, présent dans beaucoup de dialogues), il y aura significativement plus de Z. S'il s'agit d'un texte technique, le vocabulaire spécifique affecte l'écriture : si l'on parle de chemins de fer, il y aura beaucoup plus de W (wagon) ; s'il s'agit de « petites annonces » on trouvera souvent le symbole Euro (€), qui est absent

<sup>1</sup> La colonne de gauche (majuscules) indique la succession des séquences de onze lettres (ESARTINULOC) en ordre variable). La colonne de droite est le poème proprement dit.

de la plupart des autres documents. De même, l'époque à laquelle le texte a été rédigé modifie la fréquence des lettres : un texte français du dix-huitième siècle ne contiendra pas ou peu de W, car cette lettre était beaucoup moins utilisée qu'aujourd'hui.

Enfin, il est évident que si l'on change de langue, le résultat <sup>2</sup> va refléter les spécificités du langage considéré :

- en anglais, les lettres les plus fréquentes sont, dans l'ordre décroissant ETAONIHSRLD, la voyelle U n'apparaît qu'à la douzième place ;
- en allemand, on observe la suite ENIRSTADHUL, cette fois-ci c'est la voyelle O qui disparaît du classement ;
- sans surprise, l'italien, langue colorée et chantante, place quatre voyelles en tête avec la suite EAIONTRLSCD, U étant relégué en douzième position ;
- l'espagnol, comme le français, affiche toutes les voyelles dans le classement avec un ordre différent EAOSNRILDTU.

Ces langues européennes ont toutefois le point commun d'utiliser en premier lieu la lettre E. Ce qui n'était pas le cas du latin, privilégiant la lettre I (qui tenait lieu aussi de J, il est vrai), la séquence latine était IEAUTSRNOMC.

Derrière l'amusement et la prouesse dans la manipulation des mots, vous seriez tentés de ne voir qu'une distraction gratuite pour retraité désœuvré. Détrompez-vous ! Dans les procédés de cryptage des messages codés, l'analyse des fréquences de lettres revêt une importance cruciale. L'analyse fréquentielle est une méthode de cryptanalyse dont la description la plus ancienne est réalisée par Al-Kindi <sup>3</sup> au IX<sup>e</sup> siècle. A partir d'un texte suffisamment long, il propose de classer les lettres selon leur fréquence d'apparition. On fait ensuite la même chose avec le texte à déchiffrer : « *En classant les symboles par ordre décroissant de leur fréquence d'occurrence, on les remplace par les lettres correspondantes, jusqu'à épuiser tous les symboles du cryptogramme à décrypter* ». Cette méthode est utilisée pour décoder des messages chiffrés par substitution, dont un exemple classique est le « chiffre de César <sup>4</sup> ». Il est possible de casser le chiffre de César par l'analyse fréquentielle ou la recherche de mots probables. Encore faut-il connaître le langage utilisé pour repérer les lettres ou les mots les plus fréquents. C'est pour dérouter les décrypteurs japonais que les américains utilisèrent la langue Navajo au cours de seconde guerre mondiale en 1942. Le code ne fut jamais brisé !

Pour ma part, et sauf demande expresse de votre part, je compte continuer à écrire en clair. ■

<sup>2</sup> Pour assurer la cohérence des dénombrements, j'ai utilisé les résultats de fréquences donnés par le site Internet <fr.sttmedia.com>.

<sup>3</sup> Al-Kindi est né à Bagdad en l'an 801. Il fut médecin mathématicien et linguiste. Dans les archives à Istanbul, on découvrit en 1987 une copie de son traité « Manuscrit sur le déchiffrement des messages cryptographiques ».

<sup>4</sup> Le « chiffre de César » est une méthode de chiffrement simple. Le texte crypté s'obtient en remplaçant chaque lettre du texte original par une lettre à distance fixe, dans l'ordre de l'alphabet.