

## Définition des homonymes au moyen de l'algorithme SOUNDEX

L'algorithme exact procède comme suit :

1. Supprimer les éventuels 'espace' initiaux
2. Mettre le mot en majuscule
3. Garder la première lettre
4. Conserver la première lettre de la chaîne
5. Supprimer toutes les occurrences des lettres : a, e, h, i, o, u, w, y (à moins que ce ne soit la première lettre du nom)
6. Attribuer une valeur numérique aux lettres restantes de la manière suivante :
  - **Version pour le français :**
    - 1 = B, P
    - 2 = C, K, Q
    - 3 = D, T
    - 4 = L
    - 5 = M, N
    - 6 = R
    - 7 = G, J
    - 8 = X, Z, S
    - 9 = F, V
7. Si deux lettres (ou plus) avec le même nombre sont adjacentes dans le nom d'origine, ou s'il n'y a qu'un h ou un w entre elles, alors on ne retient que la première de ces lettres.
8. Renvoyer les quatre premiers octets complétés par des zéros.

En effectuant cet algorithme, on obtient avec "Robert" et "Rupert" la même chaîne : "R163", tandis que "Rubin" donne "R150".