Considerăm șirul F definit după formula de mai jos:

Unde:

Bi este un “cuvant” de lungime n-1

J este un caracter din dictionarul limbii studiate

P(Bi,j) reprezinta probabilitatea aparitiei “cuvantului” Bi urmat de litera j

P\_(Bi)(j) reprezinta probabilitatea aparitiei literei j imediat dupa cuvantul Bi (probabilitate conditionata)

Cum orice limba contine “cuvinte” de orice dimensiune, entropia acesteia va fi egala cu:

In cazul in care presupunem veridică existența unei limbi ipotetice în care toate caracterele apar cu exact aceleasi frecvență, entropia acestei limbi va fi unic determinata de lungimea alfabetului acelei limbi si va fi egala cu:

Acest rezultat are două consecințe:

1. Entropia oricarei limbi va fi marginita superior la H\_m
2. O astfel de limba nu poate fi codata pe mai putin de H\_m biti, astfel, codarea binara cu lungime fixa a caracterelor acesteia este cea optima

Astfel, putem concluziona ca entropia limbii romane poate fi limitata la: