



Descriptant idiomes: un estudi sobre la freqüència de les lletres

Berta Eixeres

Andrea Gastón

Raquel Torres

4t ESO A



Objectius

Hipòtesis

Disseny de l'experiment

Experiment

Resolució d'hipòtesis

Conclusions

Volem comprovar quines lletres es repeteixen més en cada idioma i verificar si els estereotips de cada llengua es compleixen.

- textos juridico-administratius
 - textos científicotècnics
 - textos humanístics
 - textos periodístics
 - textos literaris
-
- italià
 - català
 - anglès
 - francès
 - castellà
 - alemany
 - portuguès



Objectius


Hipòtesis

Disseny de l'experiment

Experiment


Resolució d'hipòtesis

Conclusions

- 
- El tipus de text influeix en la freqüència de les lletres perquè depenent del text (descriptiu, expositiu, conversacional...) hi ha més verbs, adjectius o noms, i cada categoria gramatical té característiques pròpies que determinen lletres comunes (la “s” del plural d’un nom, la “d” i la “t” dels adjectius provinents de participis...).
 - Es pot distingir un idioma d’un altre segons la freqüència de l’aparició de les lletres.

Hipòtesis basades en els estereotips de cada idioma

- A l’alemany es repetiran més les consonants que a altres idiomes perquè ens sembla molt difícil de pronunciar i en general quan ens costa dir alguna paraula és perquè hi ha poques vocals en relació amb les consonants.
- A l’italià es repetiran més les vocals “i” i “e” perquè al imitar aquest idioma, sempre acabem les paraules amb aquestes dues lletres.

- 
- Al portuguès hi haurà més vocals que a la resta d'idiomes, perquè al escoltar-lo sembla que hi hagi molts hiats i diftongs.
 - A l'anglès, les lletres que més es repetiran són les que formen “ation”, perquè sempre que no et surt una paraula d'aquesta llengua t'inventes una fent-la acabar amb aquesta terminació (per la creença general de que és molt usual).
 - Al català hi ha més “x” que als altres idiomes perquè en els altres no escoltem el so [ʃ].
 - Al francès hi ha més “r” perquè el so /R/ es pronuncia amb molta freqüència, que en aquest idioma és el so de g en posició de r, estereotip de la pronunciació francesa.
 - Al castellà hi ha moltes “y” perquè s'utilitza molt per la conjunció de coordinació.



Objectius

Hipòtesis

Disseny de l'experiment

Experiment

Resolució d'hipòtesis

Conclusions

A l'experiment comptarem les lletres dels textos escollits, que són a, à, á, â, ä, b, c, d, e, è, é, ê, ë, f, g, h, i, ì, í, î, ï, j, k, l, m, n, o, ò, ó, ô, ö, p, q, r, s, t, u, ú, ù, û, ü, v, w, x, y, z, ß, æ, œ, ç, ñ entre tots els idiomes.



Crearem un programa que compti les lletres automàticament.

Dels set idiomes escollits agafarem una mostra de cadascuna de les tipologies textuals.



Objectius

Hipòtesis

Disseny de l'experiment

Experiment

Resolució d'hipòtesis

Conclusions

» `f = open (input_file, "r")`

...

» `cont_l=f.read (). count ("l")`

» `f.seek (0)`

» `cont_L=f.read (). count ("L")`

» `contador_de_l=cont_l+cont_L`

...

» `g = open (output_file, "w").`

» `g.print (contador_de_a + " ;")`

» `g.print (contador_de_b + " ;")`

...



Objectius

Hipòtesis

Disseny de l'experiment

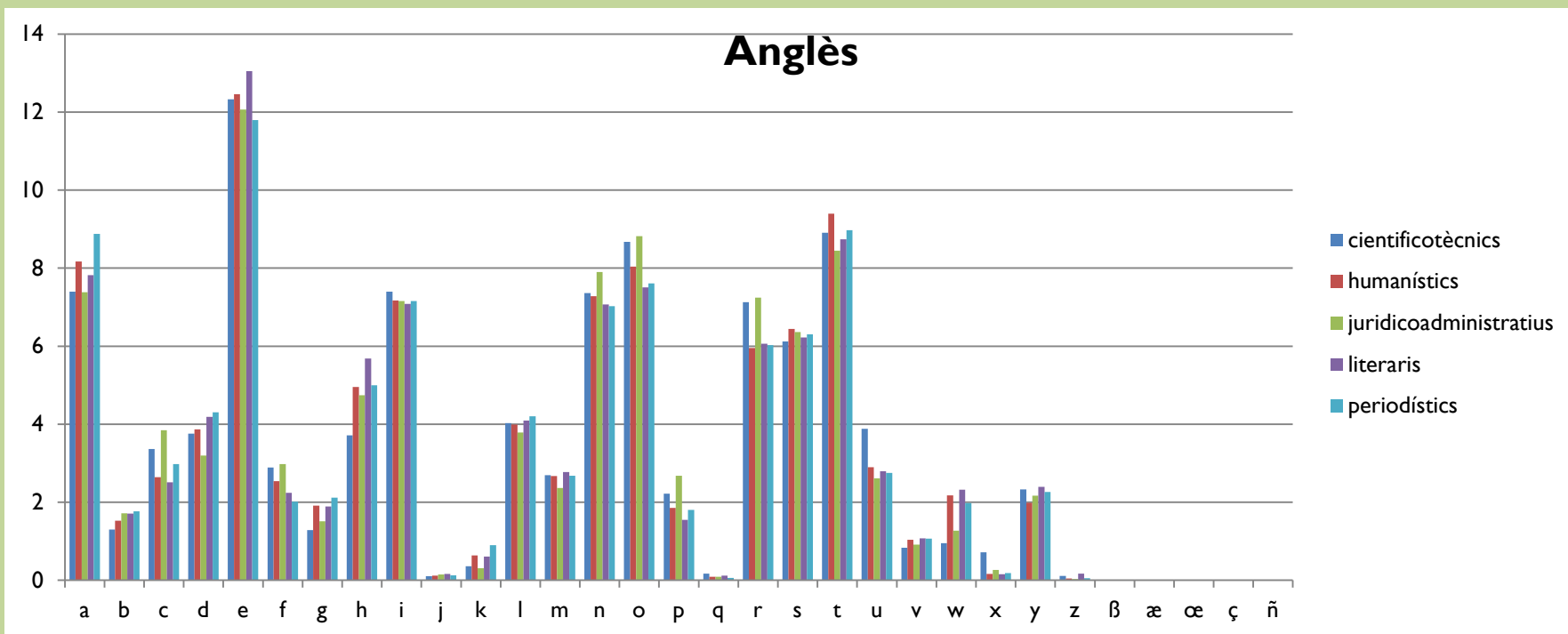
Experiment

Resolució d'hipòtesis

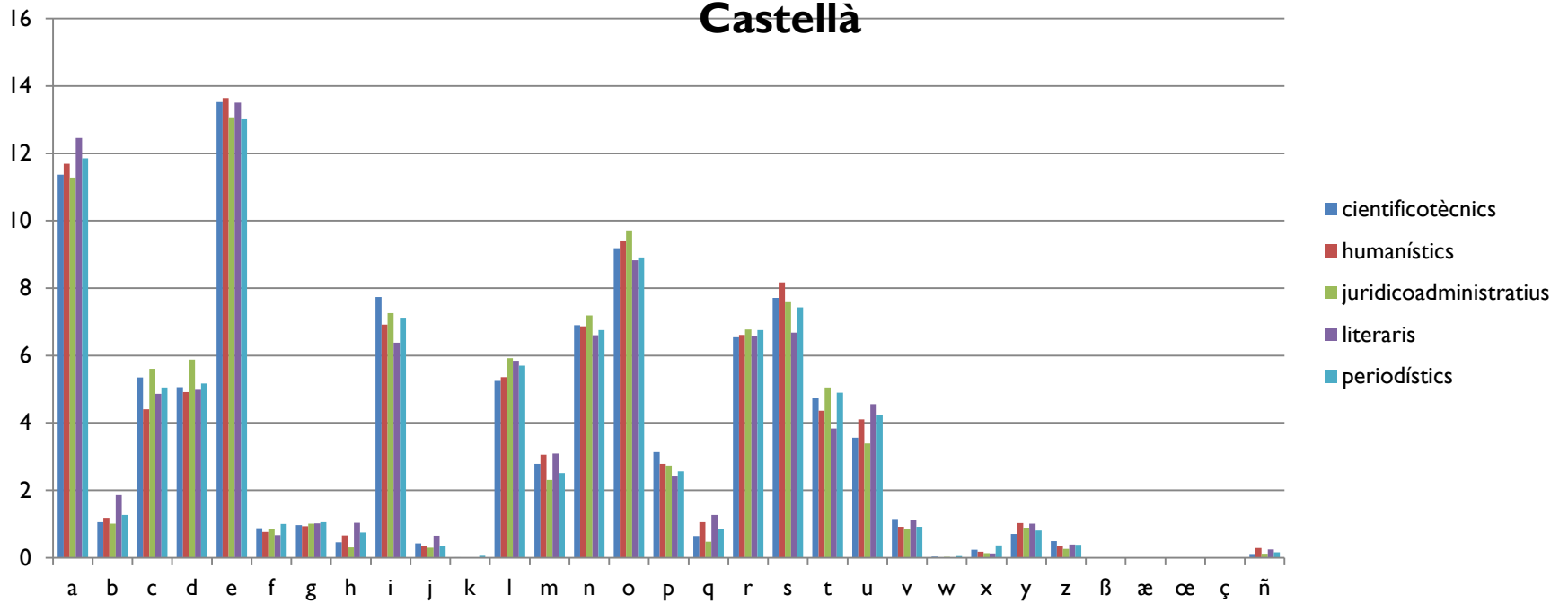
Conclusions

I. El tipus de text influeix en la freqüència de les lletres.

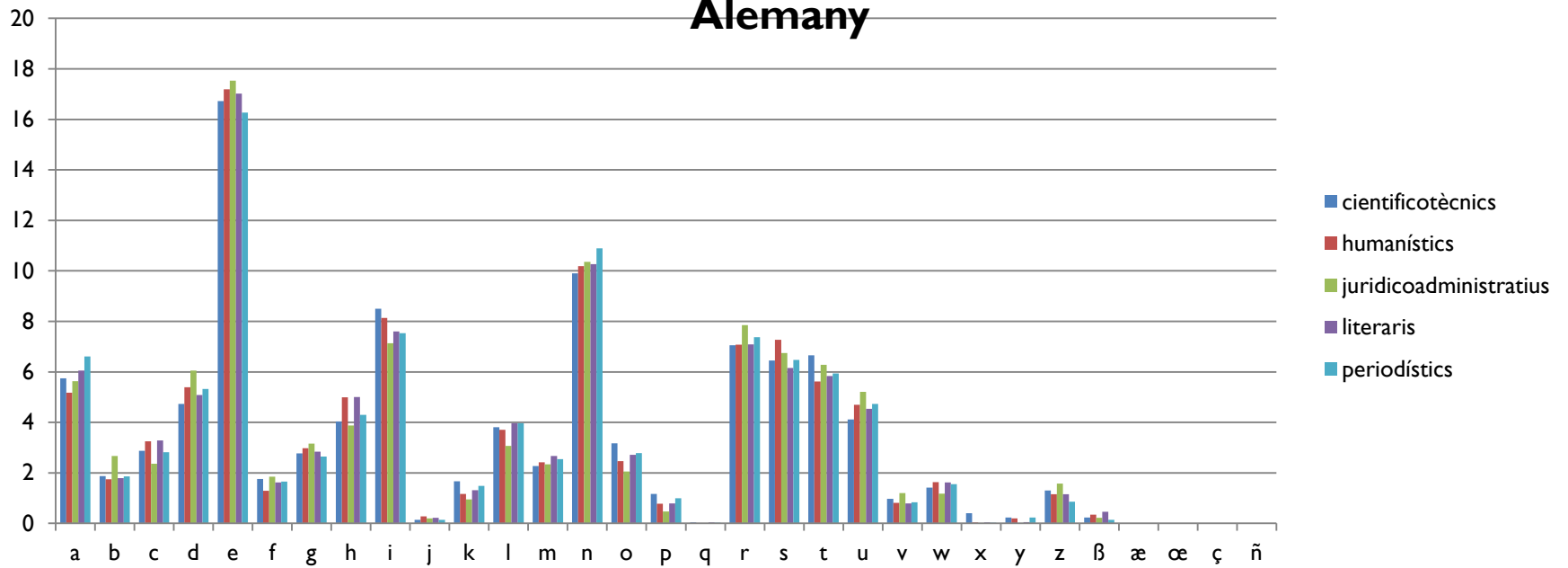
HIPÒTESI ERRÒNIA.



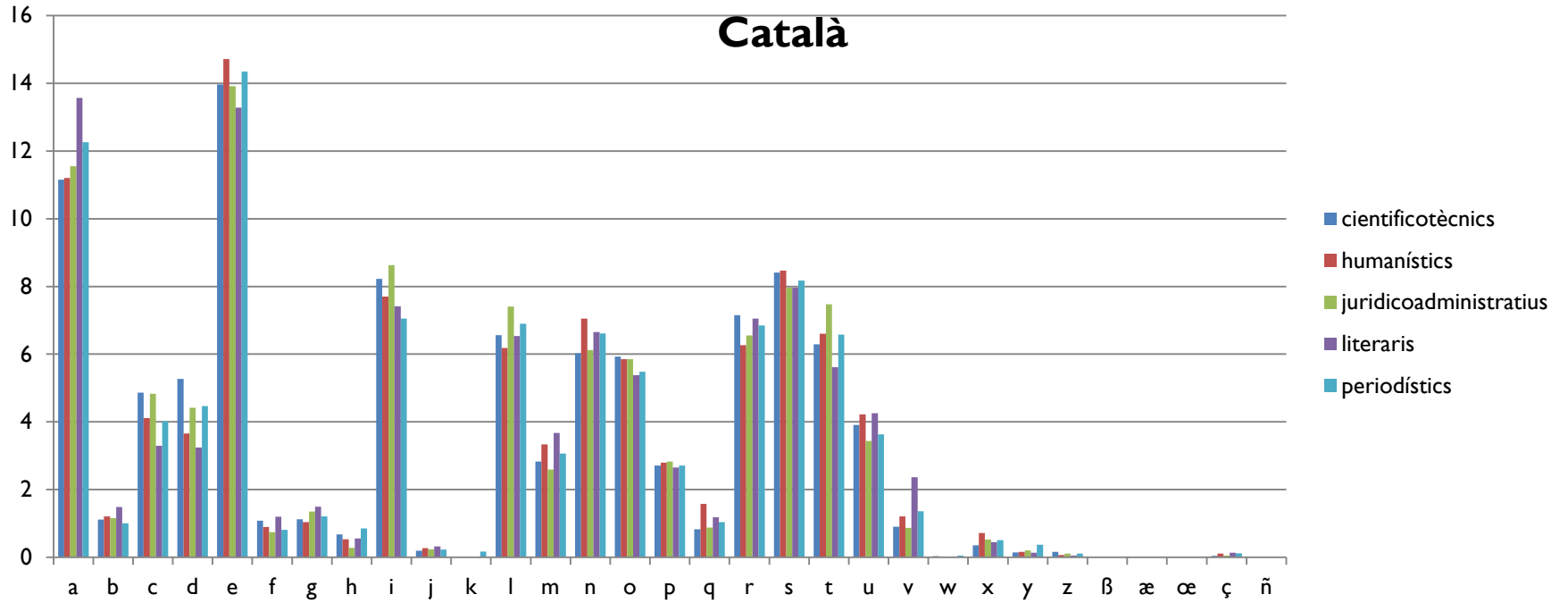
Castellà



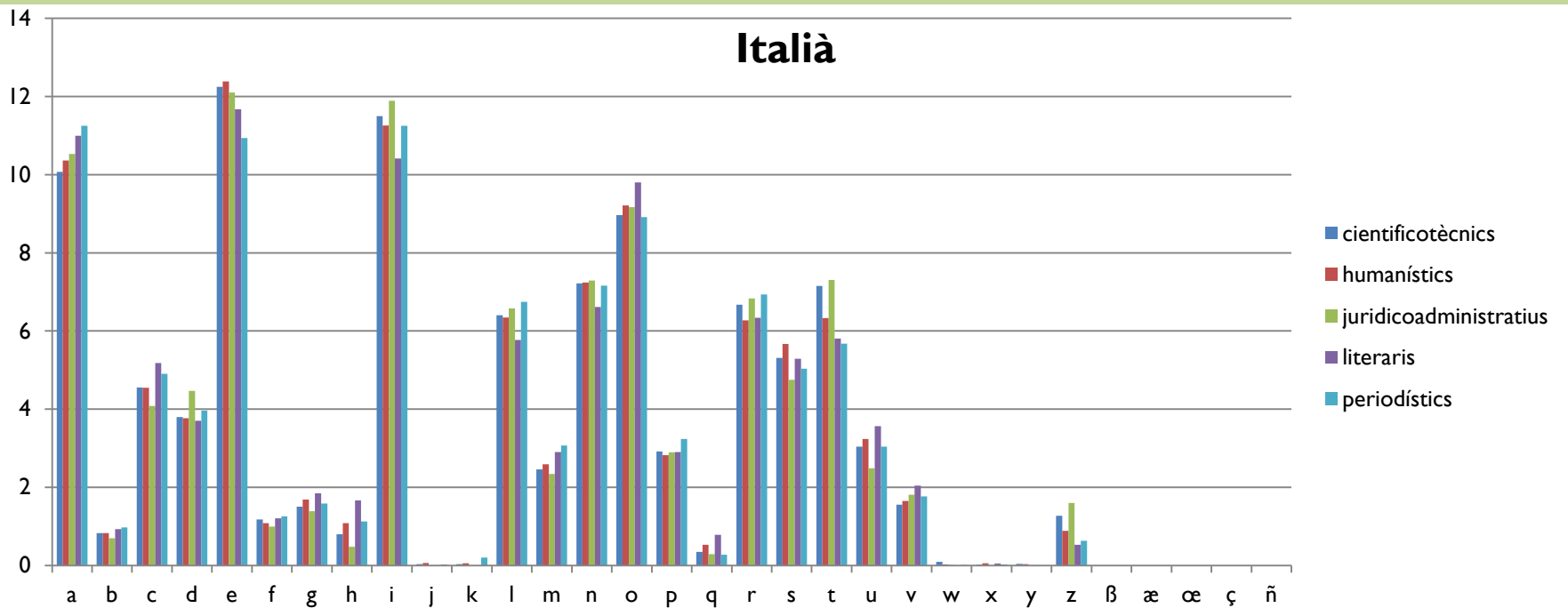
Alemaný



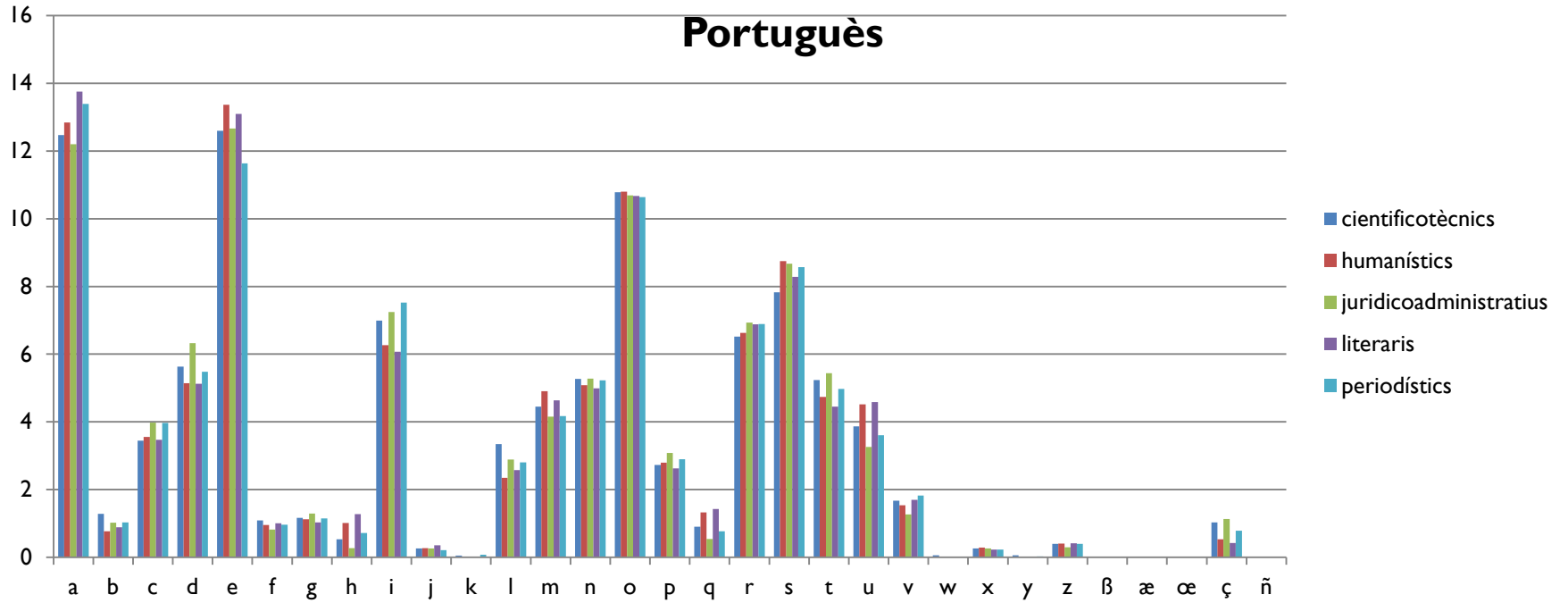
Català



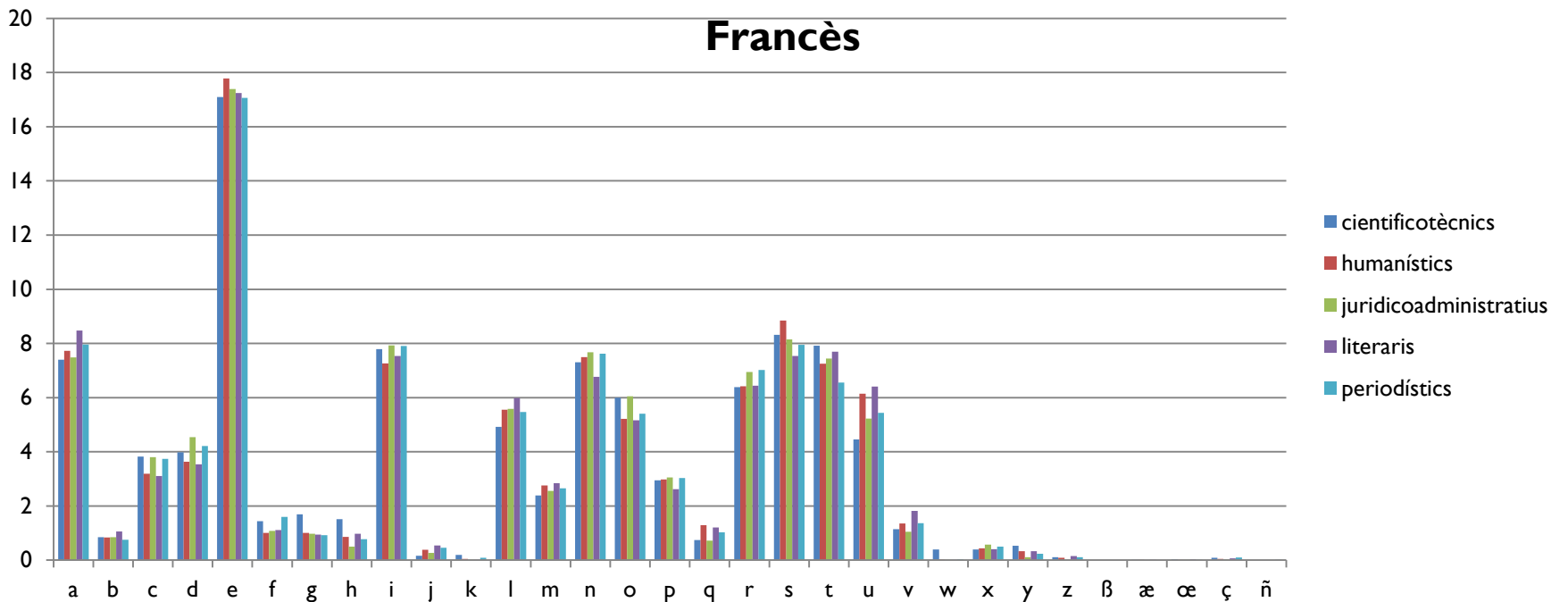
Italià



Portuguès

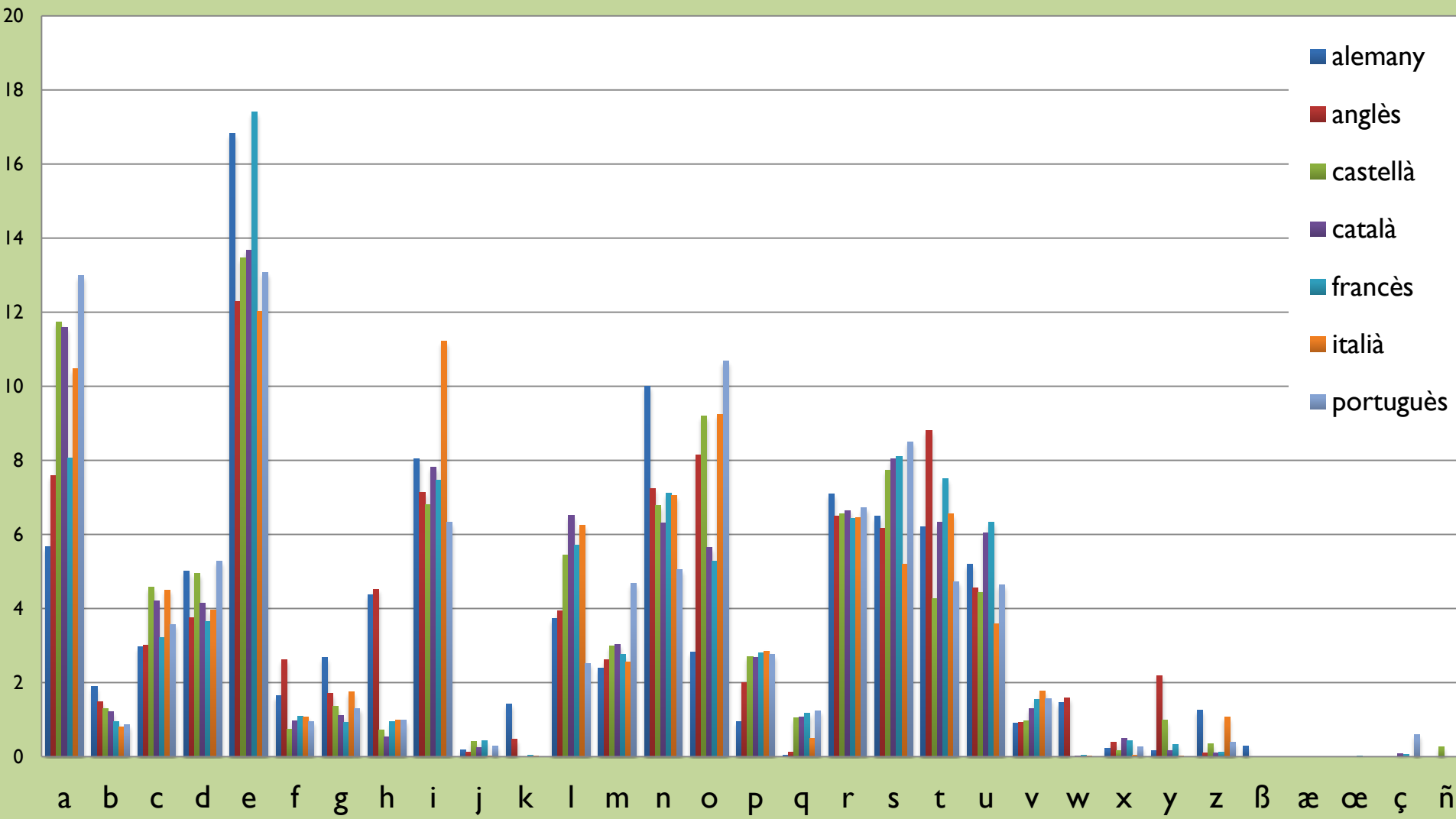


Francès



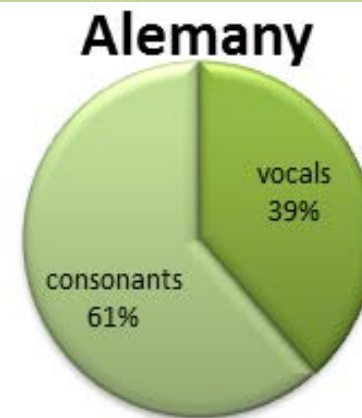
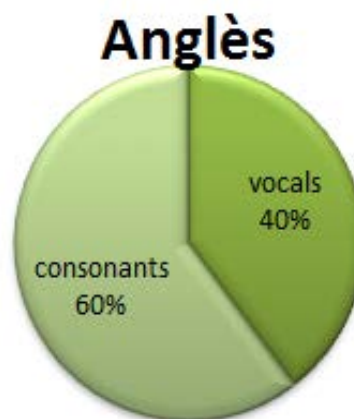
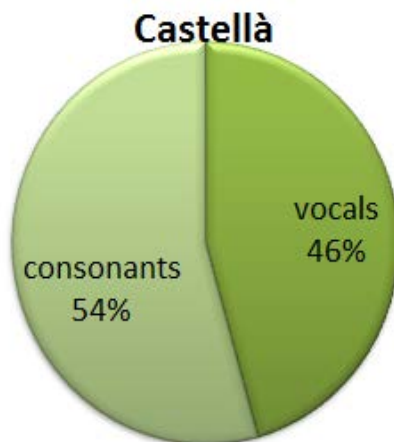
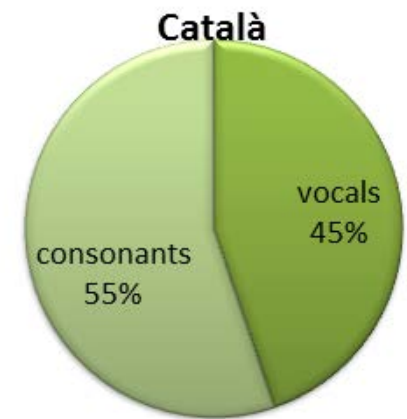
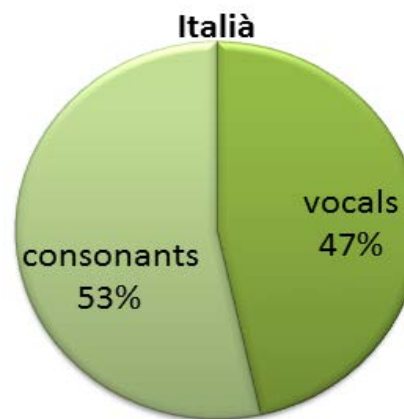
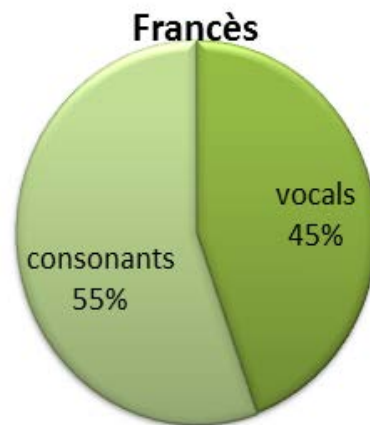
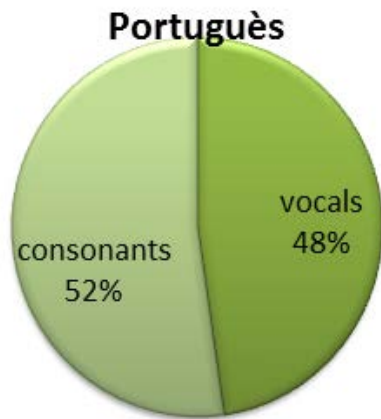
2. Es pot distingir un idioma d'un altre segons la freqüència de l'aparició de les lletres.

HIPÒTESI CERTA.



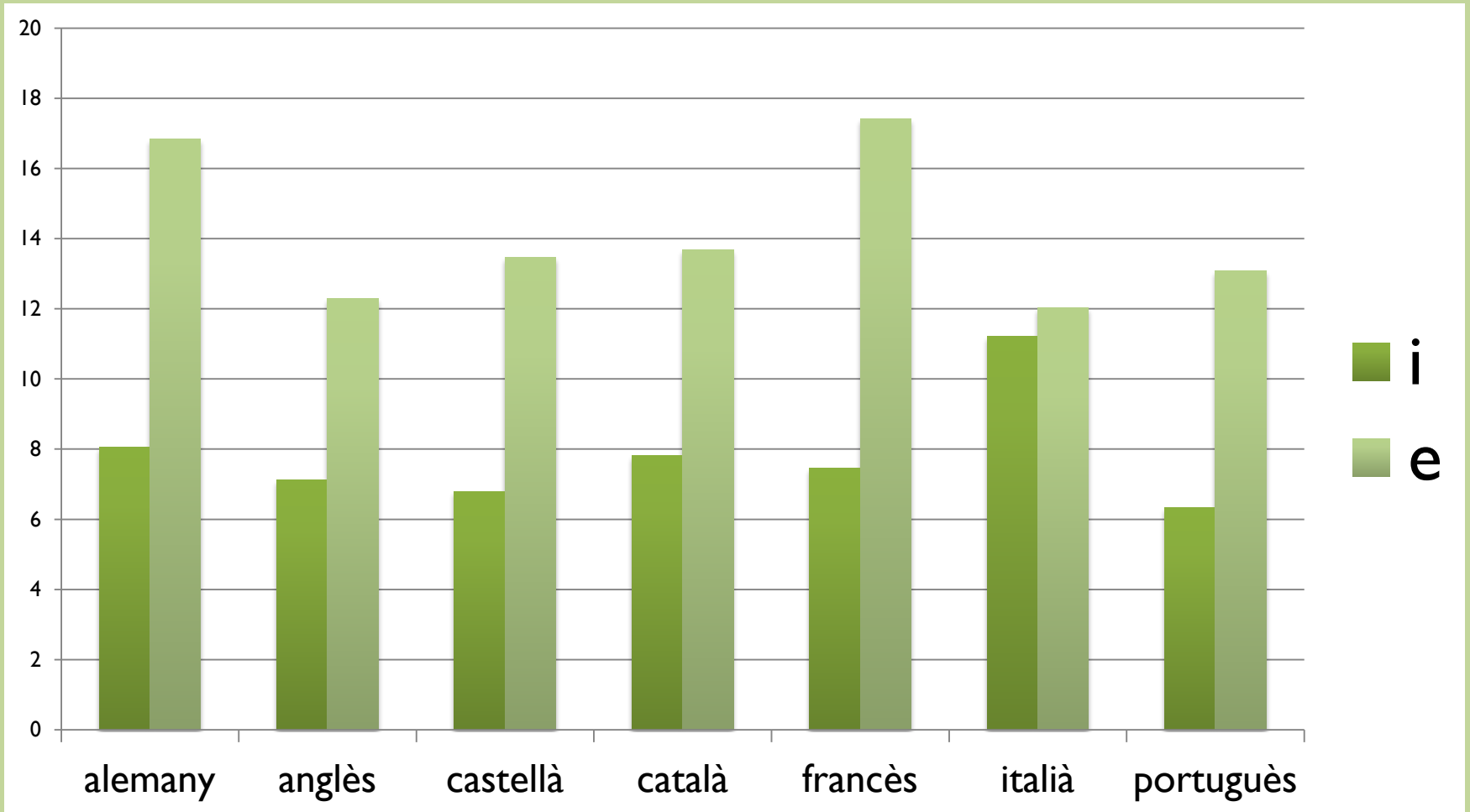
3. A l'alemany es repetiran més les consonants que a altres idiomes.

HIPÒTESI CERTA.



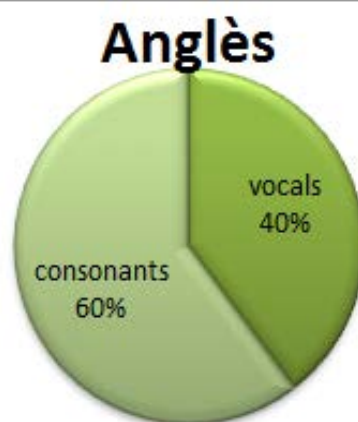
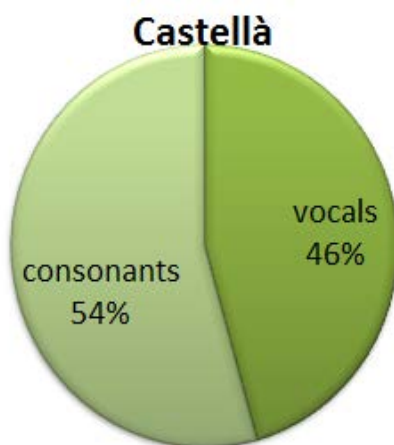
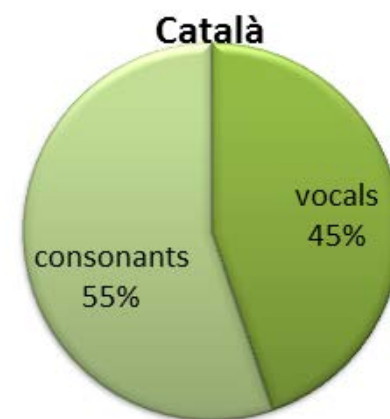
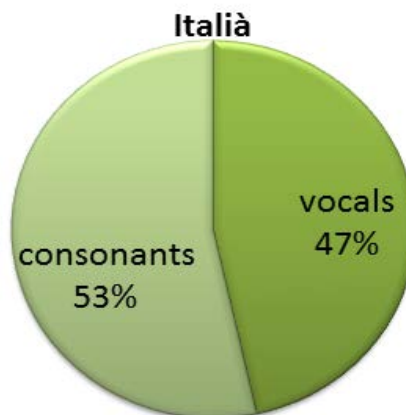
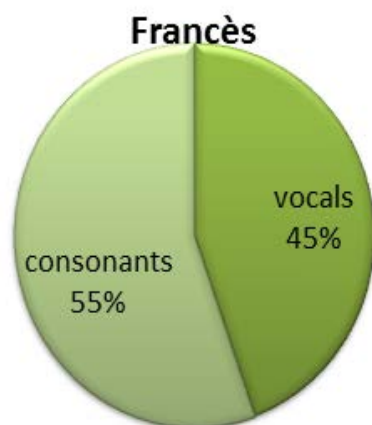
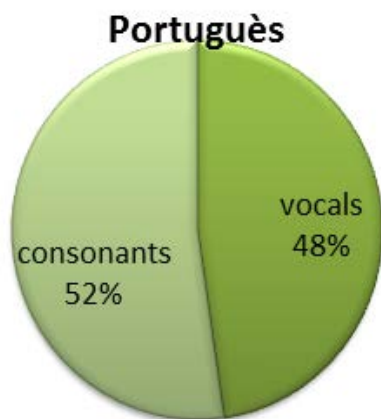
4. A l'italià es repetiran més les vocals “i” i “e”.

HIPÒTESI CERTA I ERRÒNIA.



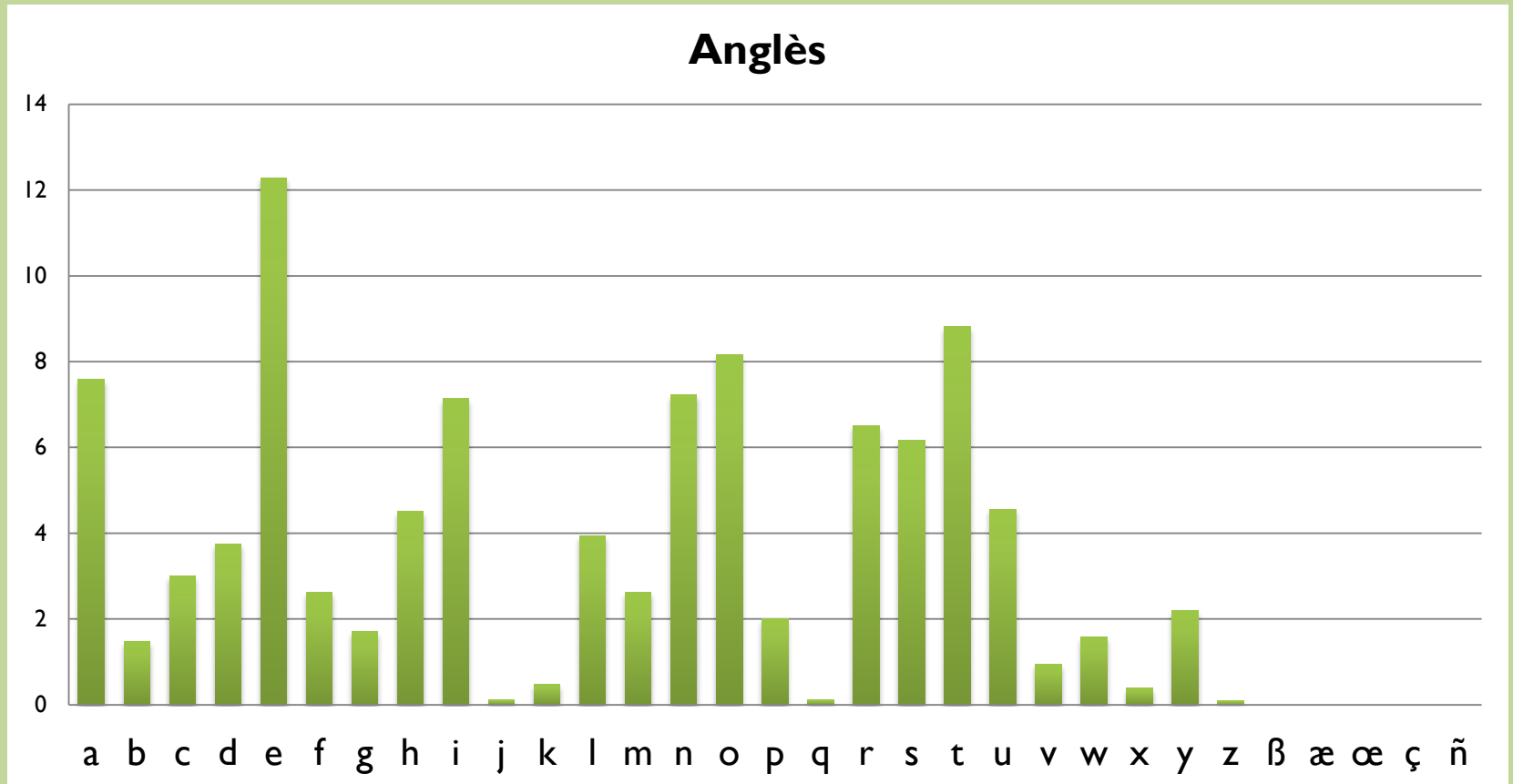
5. Al portuguès hi haurà més vocals que a la resta d'idiomes.

HIPÒTESI CERTA.



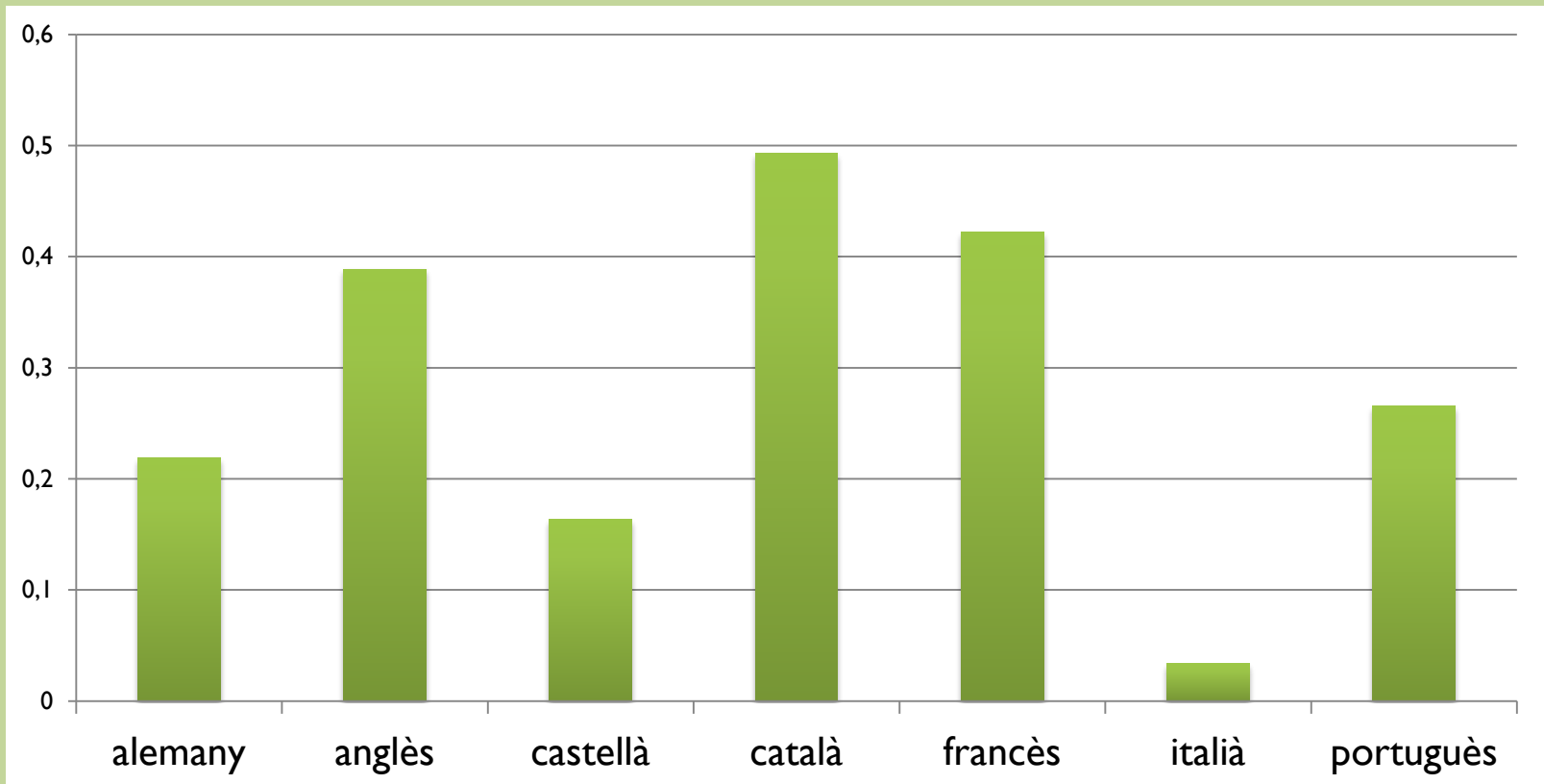
6. A l'anglès, les lletres que més es repetiran són les que formen "ation".

HIPÒTESI CERTA.



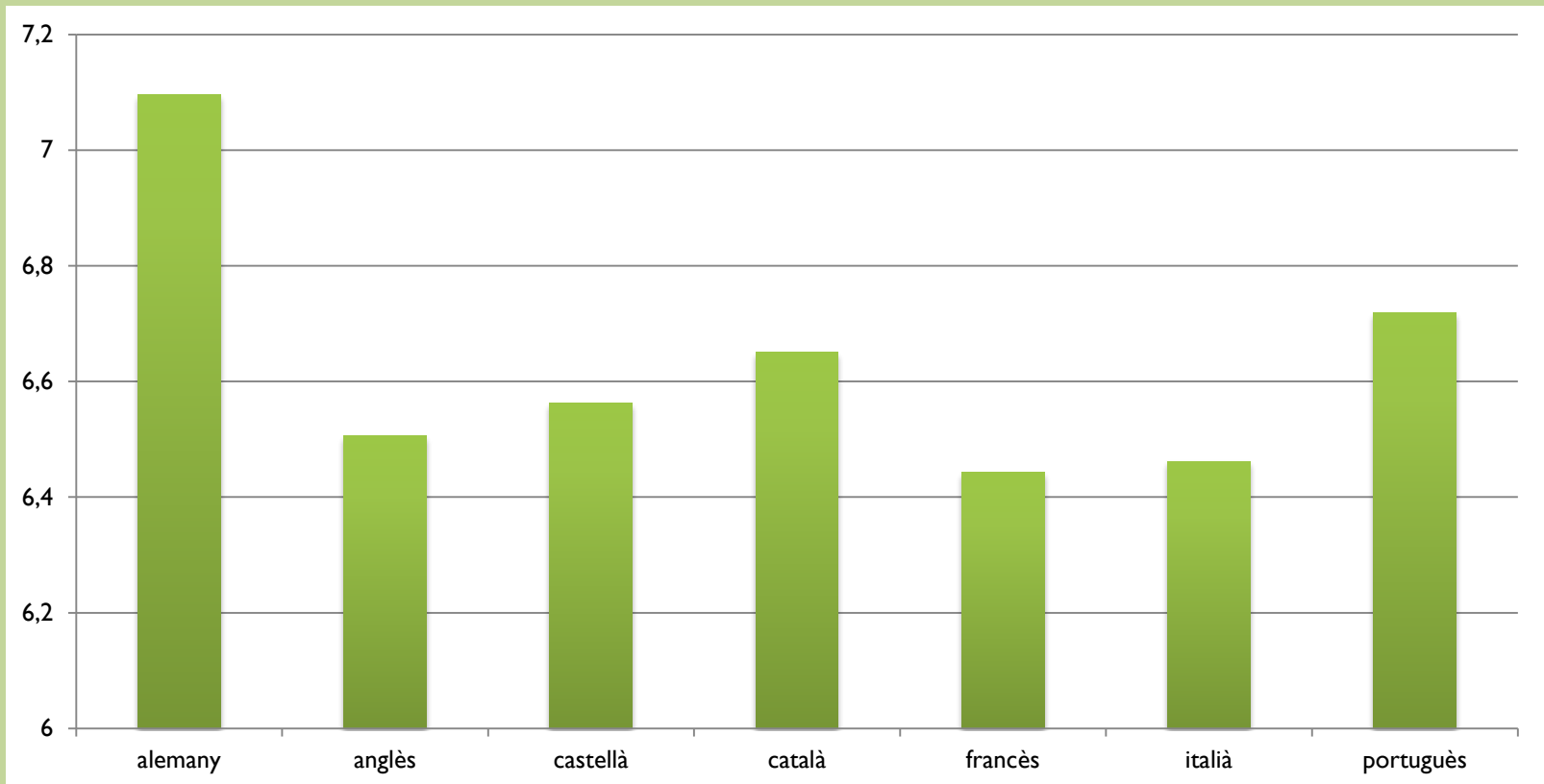
7. Al català hi ha més “x” que als altres idiomes perquè en els altres no escoltem el so [ʃ].

HIPÒTESI **ERRÒNIA.**



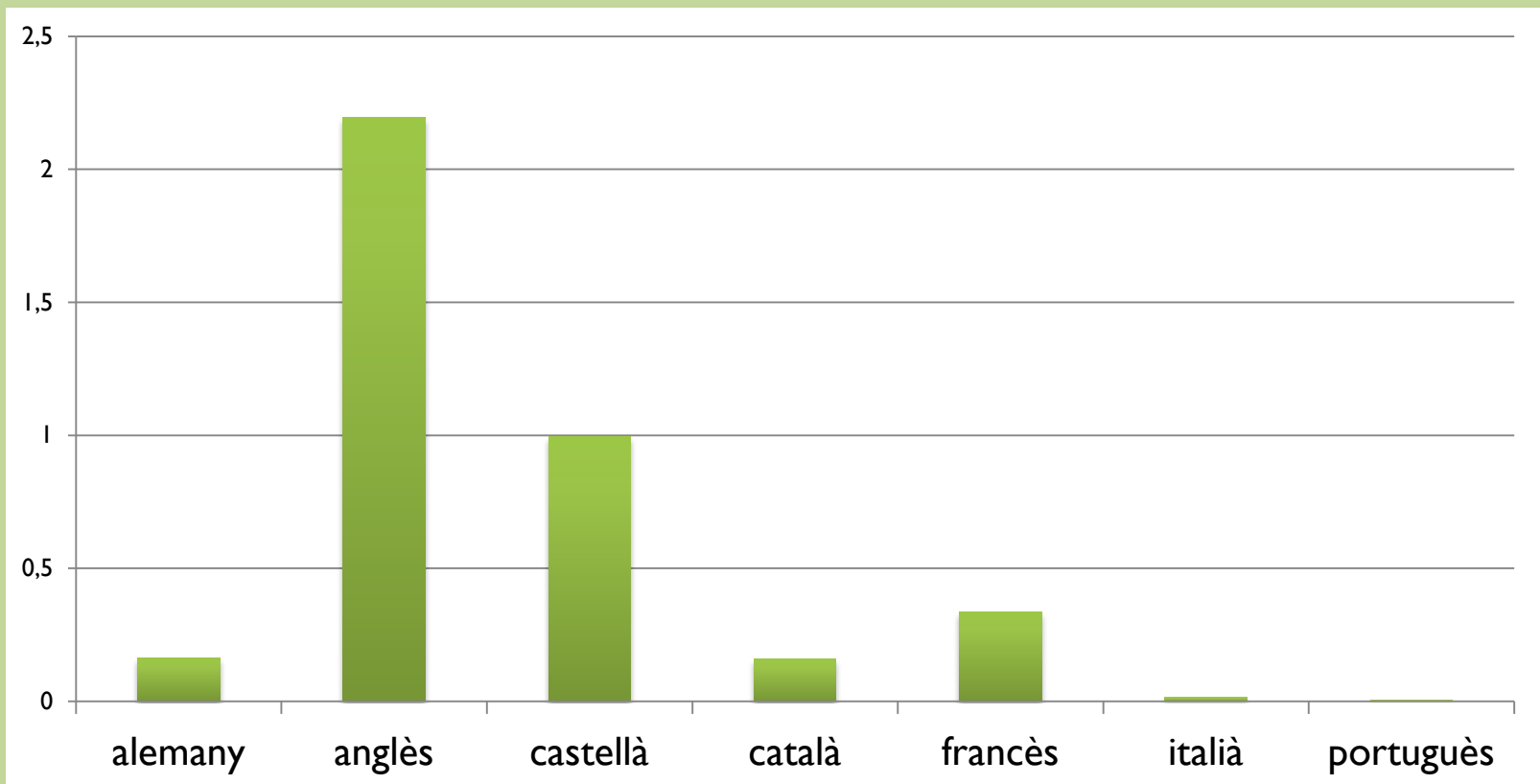
8. Al francès hi ha més “r” perquè el so /R/ es pronuncia amb molta freqüència.

HIPÒTESI ERRÒNIA.



9. Al castellà hi ha moltes “y” perquè s'utilitza molt per la conjunció de coordinació.

HIPÒTESI ERRÒNIA.





Objectius

Hipòtesis

Disseny de l'experiment

Experiment

Resolució d'hipòtesis

Conclusions

Les lletres que més es repeteixen a cada idioma:

- **Català:** e, a, s, i, r, l, t, n, u, o, c, d, m, p, v, b, g, q, f, h, x, j, y, z, ç, w, k
- **Castellà:** e, a, o, s, i, r, n, l, d, c, u, t, m, p, g, b, q, y, v, f, h, j, z, ñ, x, k, w
- **Anglès:** e, t, o, a, n, i, r, s, u, h, l, d, c, f, m, y, p, g, w, b, v, k, x, j, q, z
- **Francès:** e, s, a, t, i, n, r, u, l, o, d, c, p, m, v, q, f, h, b, g, j, x, y, z, ç, w, k, œ, æ
- **Alemanys:** e, n, i, r, s, t, a, u, d, h, l, c, o, g, m, b, f, w, k, z, p, v, ß, x, j, y, q
- **Italià:** e, i, a, o, n, t, r, l, s, c, d, u, p, m, v, g, f, z, h, b, q, x, j, i, w, y
- **Portuguès:** e, a, o, s, r, i, d, n, t, m, u, c, p, l, v, g, q, h, f, b, ç, z, j, x, y, k, w



Gràcies

Gracias

Thank you

Merci

Grazie

Danke

Obrigado