L'ADN dels escrits de Colom

Lluís de Yzaguirre i Maura
Membre corresponent de la Secció Filològica de l'Institut d'Estudis Catalans
lluis@deyzaguirre.cat

RESUM

Amb aquesta ponència pretenc agrupar les conclusions que diversos treballs dispersos en el temps i en les fonts de publicació m'han permès obtenir sobre els indicis que els escrits en castellà de Cristòfol Colom poden aportar quant a la seva llengua materna. Aquesta recopilació també ha de permetre mostrar la metodologia lexicomètrica emprada i suggerir millores en les transcripcions en què s'ha basat.

1. INTRODUCCIÓ I OBJECTIUS

Les pàgines següents aspiren a oferir als estudiosos de les hipòtesis colombines una versió refosa de treballs anteriors, la majoria dels quals només existeixen en format audiovisual o de presentacions (.ptt). I també a donar eines per a valorar la metodologia emprada o, fins i tot, poder verificar o falsar els meus assoliments. Doncs, situaré les meves aportacions en un itinerari de vint-i-cinc anys, en sintetizaré els resultats, donaré claus per a entendre el procés d'obtenció d'aquests resultats i suggeriré vies per a aprofundir-hi.

2. PRECEDENTS

A les darreries del segle XX, Discovery Channel (DC), un canal de TV via cable de Warner Bross va promoure un documental ficcionat sobre l'origen de Cristòfol Colom que incorporava l'estudi de l'ADN dels ossos suposadament de Colom conservats en indrets controvertits. Quan els resultats promesos no arribaven al ritme habitual en la producció d'aquest tipus de productes, l'empresa que intermediava a l'estat espanyol per a DC (la productora Tráfico de Ideas) va mirar de complementar-ho i es va interessar pels escrits de Colom, tant des d'un punt de vista paleogràfic com estrictament filològic. En aquest context, Tráfico de Ideas va contactar amb el Centre d'Estudis Colombins d'Òmnium Cultural, que tenia de president en Joaquim Arenas i en Francesc Albardaner de vicepresident. Ells es van encarregar, per una banda, d'implicar-hi el doctor Gabriel Roura, de la Universitat de Girona, per tal que fes l'estudi paleogràfic i, per l'altra, em van demanar si podria aportar-hi quelcom que satisfés el requisit de la productora de veure aplicats mètodes moderns a enigmes antics, que era el *leitmotiv* de la sèrie on s'incardinaria el docudrama amb l'objectiu d'aportar llum

sobre l'origen de Colom a partir dels seus escrits.

La lectura dels treballs precedents favorables a una hipòtesi catalana em va fer observar que majoritàriament es basaven en observacions sobre l'ús de mots, de topònims o de construccions sintàctiques suposadament catalanes, i que ningú no havia explorat aspectes fonològics.

Doncs, vaig partir de l'axioma gens qüestionat en el món acadèmic que **el castellà no era la llengua materna de Colom** i vaig plantejar la hipòtesi que l'anàlisi lexicomètrica de les peculiaritats del castellà dels escrits de Colom (concebudes com a interferències de la seva L_1) ens podria aportar indicis sobre la seva llengua materna. Tot aquest plantejament es troba desenvolupat a De Yzaguirre (2006a). he titulat metafòricament aquesta ponència «L'ADN dels escrits de Colom» en consonància amb el fet que el que va motivar que se'ns sol·licités l'estudi filològic fou justament que els resultats sobre l'ADN biològic no avançaven tan prestament com era de desitjar.

3. ASSOLIMENTS

L'estudi que se'm proposà per a DC havia de fer-se sobre la transcripció publicada per la doctora Varela (Varela 1982), amb un conjunt de documents que van del 4 de gener de 1493 fins al 5 de febrer de 1505, un període de més de deu anys en què qualsevol aprenent d'una L_2 té moltes probabilitats de mostrar una evolució de l'error a l'encert, bo i passant sovint per una etapa d'hipercorreccions (com quan hom diu *javia* per *hi havia*).

La divulgació dels resultats a través d'un docudrama va deixar molts estudiosos orfes dels detalls i en els anys successius vaig ser requerit en diverses ocasions a aparcar les meves recerques per a reprendre el tema colombí i divulgar-ne aspectes. Només en un cas podem dir que es tracta d'una publicació seguint els estàndards acadèmics, la resta de materials accessibles són les diapositives de les presentacions.

A continuació repassaré diversos aspectes rellevants, començant per l'esca del pecat, el docudrama de DC, i em basaré en treballs propis on el lector podrà trobar els detalls: De Yzaguirre (2006a, 2006b, 2006c i 2019). Les seves referències són:

- COLUMBUS, Secrets from the Grave, https://www.imdb.com/title/tt0893320/
- De Yzaguirre, Lluís (2006a): Aproximació lexicomètrica a les interferències de base fonològica en els escrits autògrafs de Cristòfor Colom, dins de *Consell interinsular d'Eivissa i Formentera*, "La lengua materna de Cristóbal Colón, Eivissa, 2006, Dipòsit Legal I-283-2006 (annexos 1, 2, 3)
- De Yzaguirre, Lluís (2006b): De Varcelona a Sebilla amb Cristòfor Colom, transparències de la conferència pronunciada en l'acte de commemoració del 500è aniversari del decés de Cristòfor Colom, a Barcelona, el 20 de maig del 2006, organitzat pel Centre d'Estudis Colombins (annex 1)
- De Yzaguirre, Lluís (2006c): Badava Colom?, transparències de la conferència

organitzada pel Rotary Club de Barcelona, al CosmoCaixa de Barcelona, el 27 de juny del 2006, organitzada amb el Centre d'Estudis Colombins

- De Yzaguirre, Lluís (2019): Materials de la conferència presentada a l'Ateneu Barcelonès, el 2019, organitzada pel Centre d'Estudis Colombins

3.1 La llengua materna de Colom no podia ser el ligur o genovès

Vaig triar quatre fenòmens a observar, en què el ligur manté la diferència que el català oriental central ja havia eliminat. Aquests quatre fenòmens són:

- neutralització (confusió) d'o i u àtones,
- neutralització (confusió) d'a i e àtones,
- betacisme (confusió de b i v),
- variacions gràfiques d'essa sorda i sonora ("c, ç, s, ss").

i es manifesten en vacil·lacions en els seus propis escrits i discrepàncies en els usos majoritaris dels seus coetanis que es troben al CORDE (Corpus Diacrónico del Español, de la Real Academia Española). El motiu de triar aquests quatre fenòmens fou que segons autoritats en romanística (Badia i Margarit, Martinet, Lausberg, Alsina, Blecua), aquestes **confusions eren vives al català oriental del S. XV** mentre que **el ligur les desconeixia** radicalment. Cal dir que DC va voler tenir una segona opinió d'un auditor extern i va facilitar el corpus analitzat i les conclusions obtingudes al Pfr. Dr. Dr. Max Pfister, eminent romanista de la Universitat de Saarland, que no va plantejar el més mínim retret a la metodologia emprada.

3.2 Podria ser català oriental central

En cap dels meus treballs he afirmat que Colom fos català, sinó que **les interferències en el seu castellà podrien ser les d'un catalanoparlant.** En concret, vaig afirmar: «*Cal dir que el nostre objectiu inicial (determinar si **la llengua materna de Colom podria haver estat el català) està àmpliament corroborat** (...) però això no exclou que altres llengües que presentin les mateixes oposicions fonològiques que el català en relació als sons estudiats també podrien considerar-se candidates a llengua materna de Colom; (...) **això exclou el protoitalià parlat a Gènova***». També exclou les variants abetacistes del català. I pel principi científic de la Navalla d'Occam, **el català oriental central** resta com el candidat més plausible mentre no apareguin altres hipòtesis compatibles amb les meves troballes.

3.3 Textos autògrafs vs propis

En observar que a les cartes LXIII i LXIV no es detectaven les variables que he observat a tot el corpus i, a més, hi ha altres pautes que es modifiquen des d'aquell moment, em vaig plantejar que aquestes cartes no fossin autògrafes. Totes les variables analitzades presenten una reducció (quasi absoluta) a les dues cartes esmentades.

A tall d'exemple del *forat negre*, reproduïm dos histogrames; primer (Fig. 1), el de la

distribució dels mots amb interferència b/v:

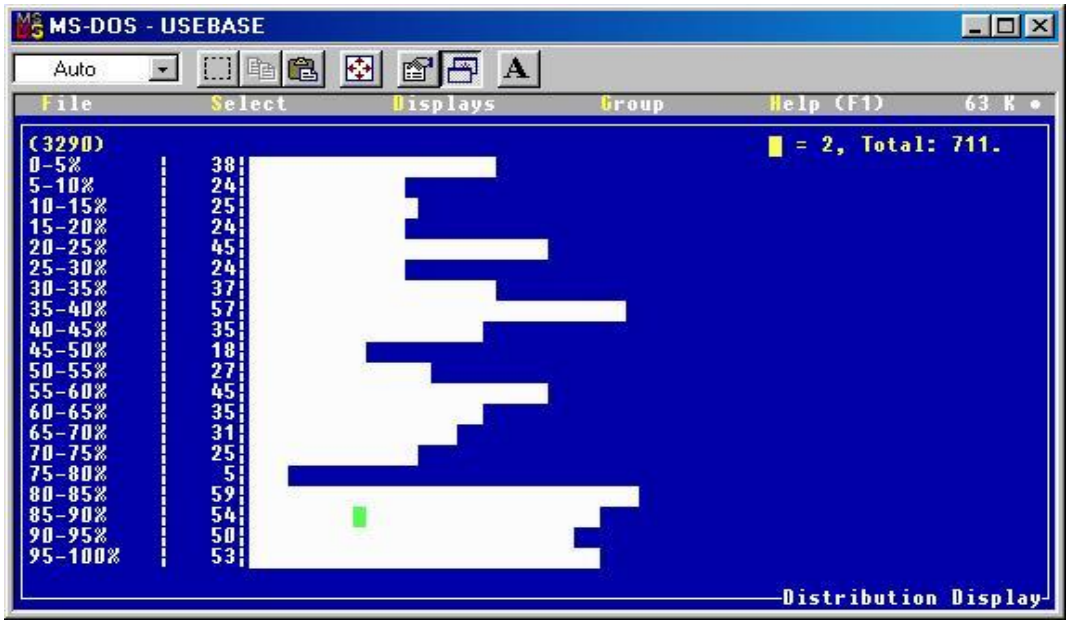


Fig. 1. *Histograma amb la distribució dels mots amb interferència b/v.*

I com a segon exemple (Fig. 2), els mots amb interferència "çe/çi" o "ç=z":

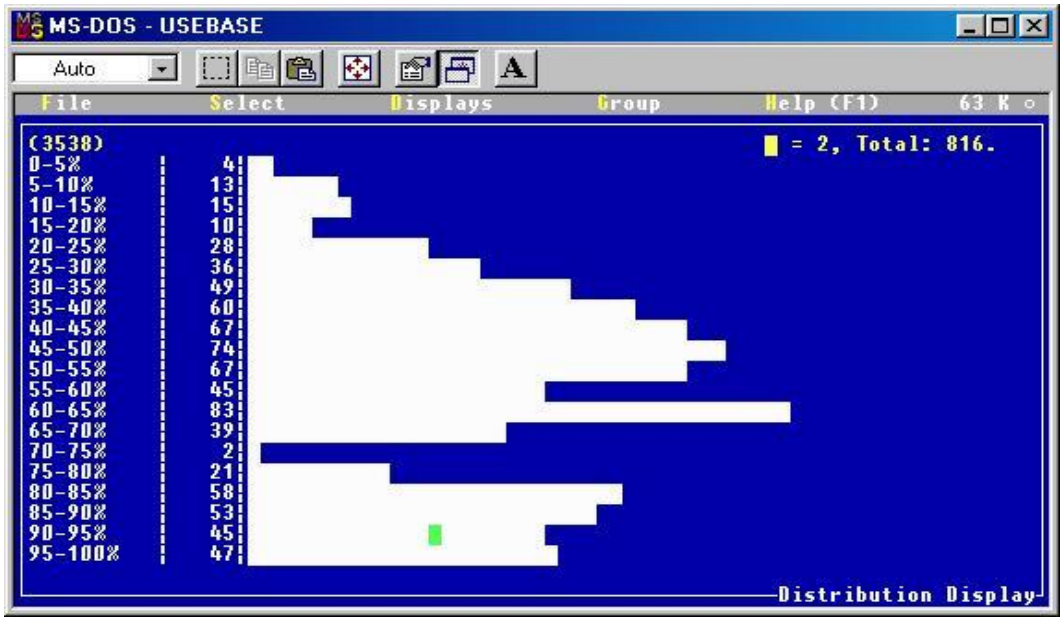


Fig. 2. *Distribució dels mots amb interferència "çe/çi" o "ç=z".*

Vaig demanar al Francesc Albardaner que ho consultés en confiança amb la doctora Varela, la qual es va reafirmar a considerar que són cartes autògrafes, però es va plantejar la possibilitat que haguessin estat fetes a partir d'esborranys revisats o dictades.

En paraules del Sr. Albardaner, *"es produeix en un moment històric decisiu: la seva detenció per Bobadilla i la seva caiguda en desgràcia. Colom fa el tercer viatge de 1498 al 1500 i fa*

el quart viatge del 1502 al 1504. Els textos 63 i 64 són per reclamar els seus drets. A més a més que la raó de la Dra. Varela sigui encertada, que Colom va copiar autogràficament un informe fet per una tercera persona, resulta que Colom perd la confiança, la poca confiança que tenia, en els reis. Per tant està escarmentat i es torna molt cuidadós, recelós i caut en tot el que fa. Doncs, no fóra d'estranyar que a partir del desembre del 1500 els textos colombins tinguin moltes més hipercorrecions que no pas abans de 1500 !"

Em va semblar important poder atribuir l'afirmació a la pròpia doctora Varela, de manera que reprodueixo, amb el seu permís, l'intercanvi de correus que vàrem tenir un cop ella va ser notificada d'aquesta anomalia (que vaig denominar "forat negre", en la mesura que les variables a observar hi deixen de presentar resultats homogenis amb els precedents). Ella no afirma categòricament que siguin textos que Colom va transcriure, però quan fou confrontada l'evidència dels resultats que a mi em suscitaven el dubte de si realment eren cartes escrites per l'Almirall, va sorgir d'ella la hipòtesi que en Colom s'hagués limitat a fer per al seu ús personal, una còpia d'un escrit preparat per algun "leguleyo".

> To: "Consuelo Varela" xxxxx@eehaa.csic.es. Sent: Wednesday, April 05, 2006 11:06am. Subject: Comunicación Ibiza
>
> Estimada Dra. Varela:
>
> por mediación del Sr. Albardaner he recibido respuesta a mis dudas sobre las cartas LXIII i LXIV (edición 92) en el sentido que Ud. ratifica que son de puño y letra de Colón, pero que le parece verosímil que el Almirante actuara aconsejado o que pusiera en limpio un texto revisado por la importancia que tenían estos escritos. (…)
>
> En cualquier caso, mi investigación también revela un cambio en la pauta "orto"gráfica de Colón desde ese momento, cosa que podría dar juego a futuras investigaciones. Una vez tenga su respuesta y cierre el redactado de mi comunicación se la mandaré y sepa de antemano que me encantará cualquier colaboración histórico-lingüística para profundizar en la cuestión.
>
> Reciba el más cordial de los saludos de Lluís de Yzaguirre i Maura (…)

primera resposta

De: "Consuelo Varela" xxxxxx@eehaa.csic.es. Fecha: 7 abril 2006 11:38:38 GMT+02:00 Para: Lluís de Yzaguirre i Maura lluis.deyzaguirre@upf.edu. Asunto: Re: Comunicación Ibiza

Querido amigo. Hasta hoy no he visto el correo, el CICA (centro Informático andaluz) es un servidor pésimo, que o está roto o lo están limpiando. Como le comenté a Francesc, yo tengo la seguridad -dentro de lo que es posible- que esos Memoriales son autógrafos colombinos. Y solo se me ocurre que el almirante estuviera copiando un texto que le había dado algún leguleyo y que, además, no dejara deslizar alguna de las formas que se nos escapan al redactar. Habría que compararlo con la carta que copió de Toscanelli a Martins, que todos dicen que es buena, pero no con la que. según dijo, le escribió Toscanelli a él, que es probable que se la inventara. Siento no poder ayudar mas. Un saludo Consuelo

segon correu

Fecha: 8 abril 2006 19:14:28 GMT+02:00. Para: Lluís de Yzaguirre i Maura lluis.deyzaguirre@upf.edu. Asunto: Re: Comunicación Ibiza

Querido amigo. He tenido estropeado el e-mail de la Escuela. El servidos de la Junta se estropea o lo están arreglando continuamente. Ayer me dio Francesc tu teléfono, pero debí de tomarlo mal porque no logro contactar. Vamos a ver. Yo estoy convencida de que los dos Memoriales escritos por Colón son autógrafos. Para arreglar ese agujero negro **se me ocurre que pueda tratarse de una copia que hizo el almirante de algún documento de sus leguleyos**. De todas formas, aun copiando, es fácil que se produzcan errores, todos tenemos normas que repetimos sin darnos cuenta. Habría que ver, digo yo que soy lega en esa materia, que convendría ver cómo hizo Colón la transcripción de la carta de Toscanelli a Martins, no las otras que parece que se inventó. **Por supuesto que puede citarme.** Me hubiera gustado poder estar en Menorca, pero no se puede estar en todas partes, ´lastima. Un saludo y que disfruten Consuelo

(NB: negretes meves)

Aquesta resposta també ens confirma que les variables triades són significatives. Hem de recordar que no solament hi ha el *forat negre*, sinó que el canvi es detecta, per exemple, en el conjunt del vocabulari amb 'b' (que decreix des d'aquell punt) i amb 'v' (que creix), és a dir que els escrits posteriors intensifiquen les tendències que s'observen al primer 70% del corpus.

4. LA METODOLOGIA LEXICOMÈTRICA

L'aportació del mètode lexicomètric, donant per bona la datació gens dubtosa de tots els escrits analitzats, és que ens permet determinar que els primers escrits discrepen força –pel que fa a les variables observades– del castellà coetani, mentre que en el decurs dels anys van

reduint-se les discrepàncies i fins i tot es detecten hipercorreccions pròpies del procés de presa de consciència de les interferències de la L_1.

He aplicat l'anàlisi de particions, en 20 blocs del mateix nombre de mots, anàlisi que mostra els resultats en histogrames com els de la pàgina 4. Entre els lexicòmetres hi ha diferents escoles en funció de si els mots s'homogeneïtzen gràficament i si es lematitzen:

- en el nostre cas, homogeneïtzar gràficament hauria destruït les diferències que volíem analitzar; hauria tingut sentit p.e. en una recerca semàntica, però no pas en una de fonològica;
- pel que fa a la lematització (substituir cada mot pel seu lema (l'infinitiu en el cas dels verbs, el singular en el cas dels noms...) fou descartada pel calendari atropellat de DC, però a la vista dels resultats no s'ha trobat a faltar.

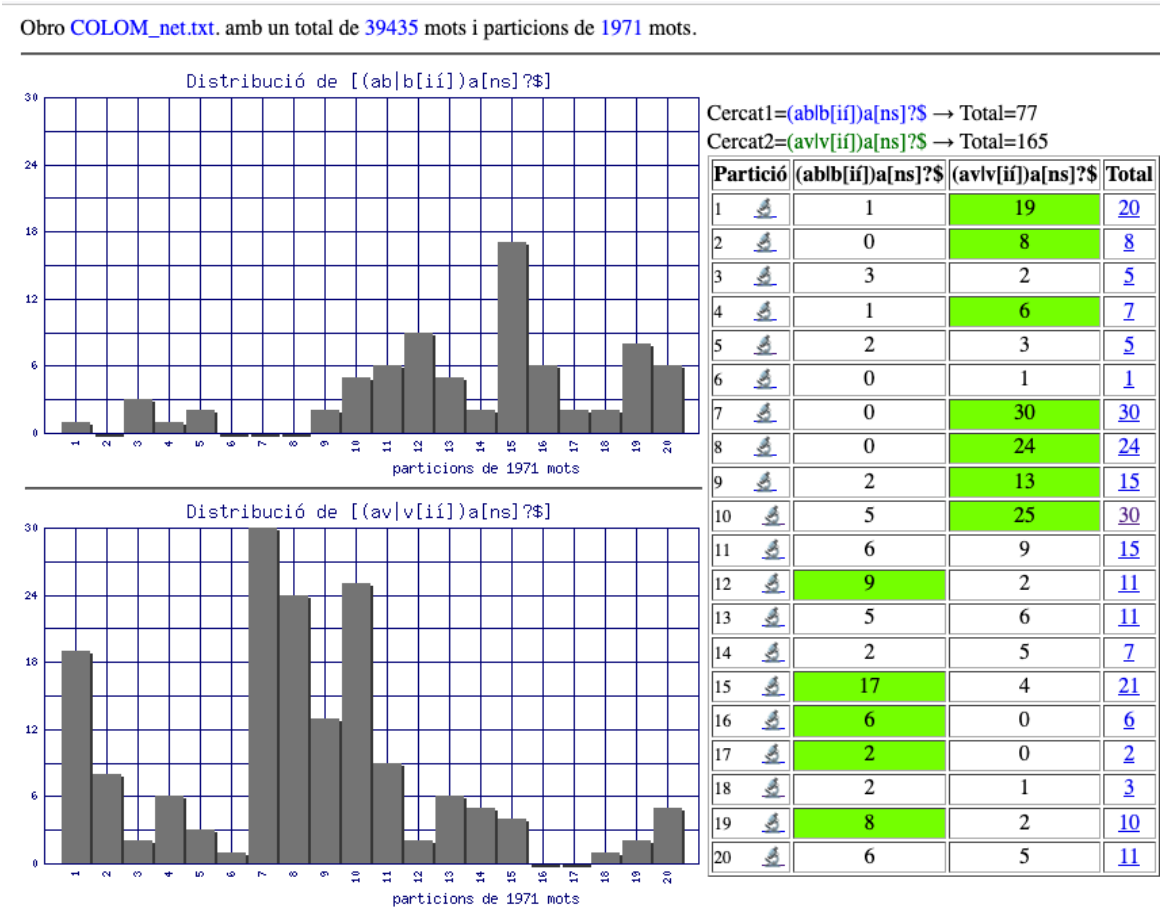
Per a qui en vulgui una presentació sumària, la proposo tot seguit; per a aprofundir-hi, he incorporat un apartat a les referències amb aplicacions a diferents branques de les ciències socials.

> La lexicométrie est l'alliance des sciences du langage, des statistiques et de l'informatique. Elle permet de traiter de vastes ensembles de textes (corpus), d'établir leur vocabulaire, de classer les vocables en fonction de leur fréquence, de leur répartition, de leurs catégories grammaticales. Elle établit les contextes d'emploi d'un vocable et les combinaisons les plus fréquentes dans lesquelles il entre, ce qui permet de déterminer le ou les sens de ce vocable. Elle retrouve les principaux thèmes présents dans un corpus, son genre et son style. Elle segmente ce corpus en fonction des ruptures thématiques ou stylistiques. Pour obtenir ces résultats, des traitements préalables sont nécessaires : balisage des textes, correction et standardisation orthographiques, étiquetage des mots. Le texte peut alors entrer dans une bibliothèque électronique à la disposition des chercheurs.
>
> Cyril Labbé, Dominique Labbé. Lexicométrie: quels outils pour les sciences humaines et sociales?. *Usages de la lexicométrie en sociologie*, Jun 2013, Guyancourt, France. ⟨hal-00834039⟩

Han passat 25 anys des que vaig començar aquestes recerques (un eó, en l'univers informàtic) i ja no es pot disposar de l'eina que empràvem, el paquet TACT (Text Analizing Computer Tools) del **Centre de càlcul per a les humanitats** de la Universitat de Toronto (Canadà); he cregut que calia fornir alguna manera de validar o falsar les recerques que em van permetre assolir les conclusions del conjunt de treballs sintetitzats en aquesta ponència. Per aquest motiu, he desenvolupat una aplicació web que integra les principals funcionalitats del TACT que em van ser tant útils.

Tot seguit mostro un exemple d'interrogació a neoTACT: li he demanat dues distribucions en particions del 5% (20 en total) del corpus colombí (Fig. 3); la primera expressió regular (ab|b[ií])a[ns]?$ ens recupera 77 mots acabats en aba/abas/aban/bia/bias/bian (**amb** i **sense**

accent gràfic, mentre que la segona ens recupera 165 mots amb les mateixes desinències però escrites amb ve baixa. Fins i tot si no teniu cap intenció d'esmerçar estones a desafiar les meves troballes, si heu estat interessats en aquest tema fins aquí, segurament us vindrà de gust de fer-hi provatures, especialment si compreneu el funcionament de les REGEX o expressions regulars.



Partició	(ab|b[ií])a[ns]?$	(av|v[ií])a[ns]?$	Total
1	1	19	20
2	0	8	8
3	3	2	5
4	1	6	7
5	2	3	5
6	0	1	1
7	0	30	30
8	0	24	24
9	2	13	15
10	5	25	30
11	6	9	15
12	9	2	11
13	5	6	11
14	2	5	7
15	17	4	21
16	6	0	6
17	2	0	2
18	2	1	3
19	8	2	10
20	6	5	11

Fig. 3. *Dues distribucions en particions del 5% del corpus colombí.*

5. CONCLUSIONS

D'acord amb el propòsit de partida, he aplegat sintèticament les conclusions i aportacions disperses sobre la nostra aplicació de la metodologia lexicomètrica als escrits autògrafs de Cristòfor Colom i he fornit una manera còmoda d'accedir a tots els treballs anteriors, de manera que quedi clar que la hipòtesi sobre si la llengua materna de Cristòfol Colom podria explicar les **interferències en el seu castellà**, queda **validada per al català oriental central** i **falsada per al ligur o genovès**.

Per a qui vulgui sotmetre aquestes aportacions a una verificació pròpia, he seleccionat unes quantes fonts que mostren l'aplicació de la lexicometria a diversos àmbits de les ciències socials i, finalment, he preparat una aplicació web, neoTACT, que permet emular les eines que ens oferia el programa TACT, amb què vaig tractar els escrits colombins fa decennis.

REFERÈNCIES GENERALS

Alcina Franch, Juan i José Manuel Blecua (1979) *Gramática española*, Ariel, Barcelona,

Badia i Margarit, Antoni M. (1981), *Gramàtica històrica catalana*, Valencia, Tres i Quatre.

Colón, Cristóbal; Varela, Consuelo, editor (1982). T*extos y documentos completos: relaciones de viajes, cartas y memoriales*, Alianza, Madrid, ISBN 84-206-2320-2.

Diversos (2004). *COLUMBUS, Secrets from the Grave*, recuperat el 7 de juliol del 2026 de https://vimeo.com/353669429?fl=pl&fe=cm#t=40m45s.

Lausberg, Heinrich (1976), *Lingüística románica*, Madrid, Editorial Gredos, 2a reimpr. (ed. orig. *Romanische Sprachwissenschaft*, Berlin, Mouton de Gruyter, 1956-1962, 3 vol.).

Martinet, André (1974), *Economía de los cambios fonéticos: tratado de fonología diacrónica*, Madrid, Editorial Gredos.

ACCÉS A LES APORTACIONS DE L'AUTOR

- versions electròniques dels articles, amb alguns annexos: https://latel.upf.edu/terminotica/membres/DE_YZA/publi/colom2006/articlesColom.htm
- aplicació neoTACT: http://latel.upf.edu/cgi-bin/neoTACT.cgi

REFERÈNCIES SOBRE EL MÈTODE LEXICOMÈTRIC

Brunet, Étienne (2011). *Le Vocabulaire de Giraudoux*. Introduction. Céline Poudat. Ce qui compte, Champion, 2011, 978-2-7453-2225-8. ⟨hal-01575144⟩ Recuperat el 8 de juliol del 2026 de

https://hal.science/hal-01575144v1/preview/02_Introduction_th%C3%A8se.pdf#page=2

Diversos autors d'un blog col·lectiu: https://devhist.hypotheses.org/category/a-propos/les-auteurs

- *Utiliser la lexicométrie en histoire (1) : panorama historiographique* Recuperat el 20 de juny del 2026 de https://devhist.hypotheses.org/898.
- *Utiliser la lexicométrie en histoire (2) : les aspects concrets* Recuperat el 20 de juny del 2026 de https://devhist.hypotheses.org/1068.

Labbé, Cyril i Dominique Labbé (2013). *Lexicométrie : quels outils pour les sciences humaines et sociales?* Usages de la lexicométrie en sociologie, Jun 2013, Guyancourt, France. ⟨hal-00834039⟩. Recuperat el 25 de juny del 2026 de https://hal.science/hal-

00834039v1.

Leimdorfer, François et André Salem (1995) *Usages de la lexicométrie en analyse de discours* Cah. Sci. hum. 31 (7) : 131-143 Recuperat el 26 de juny del 2026 de https://horizon.documentation.ird.fr/exl-doc/pleins_textes/pleins_textes_4/sci_hum/41740.pdf