

# **Els discursos sobre llengües als mitjans de comunicació**

**Proposta d'anàlisi comparativa de  
la premsa catalana i espanyola  
(2015-2021)**

**Dr. Vicent Climent-Ferrando** (coordinador)  
Càtedra UNESCO de Polítiques Lingüístiques per al Multilingüisme  
Departament de Traducció i Ciències del Llenguatge  
Universitat Pompeu Fabra

**Dr. Albert Morales Moreno**  
Estudis d'Arts i Humanitats  
Universitat Oberta de Catalunya

**Recerca finançada per la Càtedra Pompeu Fabra (Universitat Pompeu Fabra)**



<b>Els discursos sobre llengües als mitjans de comunicació</b>	1
Resum	5
Agraïments	6
1. Introducció	7
2. Hipòtesis i objectius	10
3. Marc teòric i metodològic	10
3.1. L'anàlisi del discurs	11
3.2. El discurs periodístic	12
3.3. El marc metodològic	13
3.4. La lingüística de corpus	14
3.4. La lexicometria	15
4. Corpus d'anàlisi	18
4.1. Constitució i explotació	19
4.1.1. Textos pertinents	25
4.1.2. Constitució del corpus	27
4.1.3. Aspectes tècnics i explotació	29
4.1.3.1. Constitució del corpus de titulars	30
4.1.3.2. Constitució del corpus textual	30
4.1.3.3. Altres aspectes tècnics	31
4.2. Explotació estadística del corpus	32
5. Anàlisi	34
5.1. Introducció	34
5.2. Anàlisi dels pics estadístics	34
5.3. Anàlisi del corpus de titulars	37
5.3.1. Contextos d'ús de <i>lleï</i> i <i>ley</i>	41
5.4. Anàlisi del corpus textual	44
5.4.1. Els 25 substantius més freqüents	44
5.4.1.1. Anàlisi dels contextos d'ús de <i>llengua</i> i <i>lengua</i>	46
5.4.1.2. Anàlisi dels contextos d'ús de <i>monolingüisme</i> , <i>bilingüisme</i> , <i>trilingüisme</i> , <i>multilingüisme</i> i <i>plurilingüisme</i>	54
5.4.2. Els 25 adjectius més freqüents	60
5.4.2.1. Anàlisi dels contextos d'ús de <i>lingüístic -a</i>	63

5.4.3. Els 25 sintagmes nominals més freqüents	74
5.4.4. Macroestructures	76
5.4.4.1. Llengua i educació	86
5.4.4.2. Llengua i justícia	90
5.4.4.3. Llengua i política	91
5.4.4.4. Interpretació dels resultats	92
6. Conclusions preliminars	94
6.1. Sobre el corpus d'anàlisi	94
6.2. Sobre l'anàlisi de les dades	97
7. Referències bibliogràfiques	100
8. Annexos	104
Annex 1. Llistats de freqüències	104
Annex 2. Resultats de l'anàlisi de les macroestructures	105
Annex 3. Els 25 noms propis més freqüents	122
Annex 4. Els 25 verbs més freqüents	124
Annex 5. Llistat de taules	126
Annex 6. Llistat de figures	127
Annex 7. Llistat de gràfiques	128

## Resum

Els debats sobre les llengües a Espanya són recurrents als mitjans de comunicació, perquè són objecte de debat polític, sovint en un marc de problematització política i mediàtica. La premsa (coneguda també com *el quart poder*) representa un espai de creació, reproducció i vehiculació d'ideologies sobre les llengües, i canalitza un seguit de posicionaments particulars sobre llengües com el castellà, el català i l'anglès (aquesta darrera, cada cop més present en l'ensenyament i a tots els àmbits socials).

L'objectiu d'aquest treball de recerca és analitzar *què* es diu sobre aquestes llengües a l'Estat espanyol i *com* se'n parla. Per tal de dur a terme aquesta tasca, hem analitzat un corpus d'11.481 titulars i de 286 textos complets de set diaris diferents (quatre en llengua catalana i tres en castellà) durant el període 2015-2021.

Aquesta investigació es proposa analitzar i categoritzar *què* es diu sobre les llengües i *com* es diu, tot identificant les ideologies lingüístiques que s'hi transmeten, com influeixen aquestes ideologies en la conceptualització del multilingüisme i quins mecanismes discursius s'activen per evidenciar-les.

## **Agraïments**

Agraïm la col·laboració de la Dra. Maria Sabaté-Dalmau de la Universitat de Lleida per la seva contribució valuosa a les fases inicial del projecte i de neteja d'un dels set blocs del corpus d'anàlisi del projecte. Per circumstàncies diverses, aquesta investigadora va desvincular-se voluntàriament del projecte de recerca en la seva fase inicial.

Aprofitem l'avinentsa per agrair a la Dra. M. Teresa Cabré Castellví, titular de la Càtedra Pompeu Fabra en el moment de començar aquest projecte, i a la Dra. Rosa Estopà Bagot, successora de l'antiga directora, per fer-nos confiança i haver acceptat de finançar aquesta recerca. Evidentment, també donem les gràcies a tot el personal de suport de la Càtedra (la Dra. Laia Vidal Sabanés i l'Eva Parera Ferrer per totes les gestions realitzades en el marc del projecte.

Finalment, volem agrair les observacions fetes al treball en el marc de la V Jornada "La recerca sociolingüística en l'àmbit de la llengua catalana" (gener de 2022) i en el marc del grup de recerca de discurs GEDIT-UPF (febrer de 2022).

# 1. Introducció

Els debats sobre llengües a Catalunya emergeixen, mediàticament, en un marc de polèmica i de confrontació, el mal anomenat *conflicte lingüístic* (Marimón Llorca, 2021, p. 76) entre el català i el castellà (vegeu, per exemple, Flors-Mas et al., 2022; Woolard, 2016). Aquests debats tracten més aviat la politització del fet lingüístic que no pas el coneixement o l'ús social de les llengües.

En societats amb llengües històricament i políticament minoritzades com la catalana, l'encaix social de la llengua sempre hi ha anat de bracet i ha estat un element cabdal, també discursivament, per als projectes de construcció dels estats-nació (Hambye i Richards, 2012; Heller, 2011). Per aquest motiu, els discursos sobre el multilingüisme se centren sovint en la generació i la circulació de coneixement (Bouzas, 2016; Fairclough, 1989); en concret, se centren sobretot en la instrumentalització dels afers lingüístics per posar-los al servei de l'acció política (Schieffelin, Woolard i Kroskrity, 2012).

En el cas català, aquests discursos altament polititzats han anat *in crescendo* durant la darrera dècada, fins a convertir-se, als voltants del 2020, en un dels temes “tractats amb més visceralitat, irracionalitat i *acientifisme*” (Marimón Llorca, 2021: 81) i, també, amb “anticorrecció lingüística” (Martín Rojo i Elvira Ruiz 2019). Alguns exemples ens poden ajudar a il·lustrar-ne la situació: al 2020, es produïa tot un debat politicomediàtic estatal sobre la *Proposició no de llei sobre la realitat plurilingüe i la igualtat lingüística de l'Estat* (3 de desembre de 2020; PNL, 2020), duta al Congrés dels Diputats el març del 2021.

Els diaris generalistes amb més influència al mercat i nombre de lectors a l'Estat i a Catalunya enfocaven aquest fet de manera molt divergent, seguint línies editorials diametralment oposades que, sovint, sembla que no descriuen els mateixos fets. A tall d'exemple, reproduïm alguns titulars que il·lustren aquestes diferències significatives.

*Entitats en defensa de les llengües cooficials reclamen al Congrés que hi hagi igualtat lingüística a l'Estat (El Punt Avui, 09/03/2021)*

*El Congrés dona suport a la reciprocitat però tanca la porta a la igualtat lingüística a l'Estat (Ara, 12/03/2021)*

*Podemos y nacionalistas piden mismos derechos para las lenguas cooficiales (La Vanguardia, 03/12/2020)*

*El PSOE negocia con Bildu un "encaje jurídico" que permita imponer las lenguas propias en la Administración (El Mundo, 09/03/2021)*

*El PSOE y sus socios se enzarzan en el Congreso por el uso de las lenguas cooficiales (El País, 09/03/2021)*

*Ofensiva de Podemos y Bildu contra el castellano (La Razón, 09/03/2021)*

*El PSOE apoya parte de la polémica propuesta de Bildu y Podemos que quiere hundir el castellano (ABC, 12/03/2021)*

Aquestes notícies apareixen cíclicament a la premsa espanyola. Els exemples següents, d'octubre de 2022 (data de finalització d'aquest informe), presenten nombrosos paral·lelismes amb els contextos anteriors:

*Los socialistas se opusieron a que el Parlamento Europeo envíe a Cataluña una misión para analizar la discriminación lingüística (El Mundo, 22/10/2022)*

*Defensores del bilingüismo piden a la UE que interceda ante la exclusión del castellano en las aulas catalanas (ABC, 12/10/2022)*

*La Generalitat impone el catalán como la Dictadura lo hizo con el castellano, dice Cs (La Voz de Galicia, 12/10/2022)*

Observem que molts diaris, i la política, presenten una tendència cada vegada més habitual: la problematització discursiva de les llengües oficials de l'Estat, més enllà del castellà, vehiculada a través dels mitjans de comunicació, que actuen com a 'amplificadors', divulgadors i 'legitimadors' de certs posicionaments ideològics.

Nombroses veus expertes denuncien aquesta instrumentalització de la llengua amb finalitats politicoeconòmiques. A la sèrie "El futur del català" (elaborada per Vilaweb), diversos sociolingüistes alertaven dels discursos en contra de la llengua catalana. La SOCS organitzava, el 10 de febrer de 2021, la taula rodona *La conflictivització de l'educació lingüística a Catalunya* (SOCS, 2020). L'octubre del 2020, el Comitè de Peticions del Parlament Europeu admetia a tràmit la petició de l'associació *Asamblea por una Escuela Bilingüe* (AEB), entitat que denunciava el sistema d'immersió lingüística a Catalunya davant el Parlament Europeu amb un discurs sobre *imposició lingüística i manca de llibertat i de democràcia a Catalunya*.

Més recentment, el mes d'octubre de 2022, el Comitè de Peticions del Parlament Europeu, a petició de la mateixa associació, l'AEB, enviava un grup d'eurodiputats (una *missió*, en argot europeu) encarregats d'analitzar la situació lingüística a les escoles de Catalunya, missió de la qual la premsa catalana i espanyola se'n feia ressò.

Resulta paradoxal que, simultàniament, els discursos mediàtics empen sovint termes com *multilingüisme* o *plurilingüisme*, als quals assignem valors, formes i conceptualitzacions molt allunyades d'aquesta conflictivització exposada a dalt, sempre positives/positivitzadores en termes lexicosemàntics, com a marcadors ideologitzats d'innovació, modernitat, cosmopolitisme i neutralitat política, usat sobretot per a l'anglès



(Piller i Cho 2013; Sabaté-Dalmau, 2020). Aquests discursos, de fet, acostumen a legitimar polítiques d'incorporació progressiva de l'anglès a l'àmbit educatiu (tant l'obligatori com l'universitari) com a competència necessària per a l'ocupabilitat, la mobilitat i el creixement econòmic. Aquest enfocament utilitarista de les llengües s'emmarca discursivament com un senyal de progrés.

Observem que els diferents discursos que circulen sobre multilingüisme (contradictoris, com hem observat anteriorment) no reflecteixen simplement la natura i el valor objectiu de les llengües, sinó que en creen una representació concreta: n'emfasitzen (o n'amaguen) aspectes, n'inclouen (o n'exclouen) llengües o en vehiculen categoritzacions discursives diferents.

Aquests discursos presenten una dimensió ideològica evident: formen part d'un sistema de creences, valors i normes que legitimen representacions i posicionen actors discursius determinats (en detriment d'altres). És palès, per tant, que són postulats polítics amb conseqüències per a tots els agents (institucionals i cívics) participants de la planificació i la gestió de les polítiques lingüístiques.

Posen de manifest, en definitiva, l'elevat grau d'ideologització dels discursos sobre multilingüisme, que tenen com a finalitat principal legitimar orientacions polítiques a favor o en contra d'unes llengües determinades, o d'una conceptualització concreta del multilingüisme. Creen uns marcs conceptuals i discursius d'inclusió/d'exclusió de llengües, i les categoritzen d'una manera o d'una altra. Repetits constantment i vehiculats sistemàticament a través dels mitjans i les xarxes de comunicació de masses, aquests marcs acaben creant un *habitus* (Bourdieu, 1981), una manera concreta de pensar, definir i conceptualitzar les llengües.

Com veurem amb detall a l'apartat 5, aquesta magnificació i polarització discursiva beu de les nocions metalingüístiques bàsiques dels parlants (Marimón i Llorca, 2021), i es fa palesa en l'àmbit lexicosemàntic, on la terminologia emprada s'agrupa en àrees macrotemàtiques distingides anomenades *topoi* (Bouzas, 2016, p. 157), com són la política, la cultura, el dret, l'educació, etc. Aquests *topoi* s'agrupen en camps semàntics (com ara IMPOSICIÓ, PROBLEMA, CONFLICTE) que donen homogeneïtat de significat al discurs.

És cabdal analitzar els marcs conceptuals, els valors, les normes, les representacions discursives i les creences assignades a les llengües pròpies del nostre entorn (el català i el castellà), però també a l'anglès o a les llengües de la immigració presents en el nostre territori. Només així podrem tenir dades concretes, tant quantitatives com qualitatives, per saber de què parlem quan parlem de *multilingüisme*, amb l'objectiu de desentortolligar les finalitats polítiques que regeix l'activació d'aquests discursos.

## 2. Hipòtesis i objectius

Partim de la base que el *multilingüisme* és una forma de capital simbòlic (Bourdieu, 1991) i un recurs que es distribueix de manera desigual en la societat, el significat del qual es construeix discursivament a partir de contextos polítics, socials o històrics concrets. Aquest treball analitza, per tant, la construcció discursiva sobre el multilingüisme en la premsa més llegida de la nostra societat. Aquesta caracterització ens permetrà conèixer les ideologies lingüístiques associades a cada llengua, definides com la “forma de valorar les llengües”, tot afirmant que una llengua – o una varietat – és millor o pitjor que una altra (Bell, 2014, p. 286), i com a “índex” de les normes que regulen els nostres comportaments sociolingüístics (és a dir, com a vehicle per entendre què està legitimat i què no, en públic, quant a tries, usos i orientacions vers el multilingüisme) (Kroskrity *et al.*, 1998).

Concretament, l'interès nostre és conèixer *què* es diu, *com* es diu i *per què* es diu sobre les llengües seleccionades als mitjans de comunicació de l'Estat espanyol. En concret, caracteritzem set capçaleres de la premsa escrita generalista: quatre en català<sup>1</sup> (*Ara*, *El Punt Avui*, *El Periódico* i *La Vanguardia*) i tres en castellà (*ABC*, *El Mundo* i *El País*).

Amb la investigació del *què* i del *com* es parla de les llengües, el nostre objectiu principal és analitzar i categoritzar les ideologies lingüístiques que s'hi transmeten, com influeixen en la conceptualització del multilingüisme i quins mecanismes s'activen per evidenciar aquestes ideologies (lingüístiques).

Per a fer-ho, a més, partim d'una recollida, una sistematització i un tractament de les dades basats en criteris tant quantitius com qualitius. Pel que fa als quantitius, fem criteris lexicomètrics i analitzem el vocabulari més freqüent. Així, no basem la selecció de les unitats d'anàlisi en criteris subjectius o apreciatius. Les dades qualitatives permeten caracteritzar la dimensió ideològica dels discursos sobre el multilingüisme i proporciona dades quantitatives i concretes sobre la conceptualització mediàtica. La combinació d'ambdues tècniques ens permet fer una radiografia acurada de la realitat i permet conèixer com es conceptualitzen les llengües als diaris catalans i espanyols.

## 3. Marc teòric i metodològic

Aquest apartat recull els fonaments teòrics i metodològics del treball per analitzar la conceptualització discursiva del multilingüisme als mitjans seleccionats. El treball està

---

<sup>1</sup> Som conscients que tant *El Periódico* com *La Vanguardia* es publiquen cada dia en edició bilingüe català i castellà i que s'empren eines de traducció automàtica (juntament amb la postedició corresponent). Tot i que el fet que una versió en sigui l'original i l'altra la traducció pugui tenir un impacte en la selecció lèxica de cada versió, ens interessa analitzar-ne el lèxic d'una de les dues per tal de saber *què* diuen i en quins termes ho fan quan parlen de llengua i de multilingüisme.

enfocat des de l'anàlisi del discurs (apartat 3.1), entès com a estudi sistemàtic de les relacions dialèctiques entre els textos i el seu context social; és a dir, entre els textos, les condicions socials en què es produeixen i les intencionalitats implícites o explícites.

Centrem el treball en el discurs, vehiculat a través de la premsa escrita, la qual gaudeix d'una gran capacitat per influir en l'opinió pública: el discurs periodístic (apartat 3.2).

Posteriorment, presentem el marc metodològic del treball, per tal de posar en relació l'anàlisi del microtext amb l'anàlisi del context en què circulen els discursos (apartat 3.3). Aquesta metodologia, a més, s'enriqueix de les tècniques pròpies de la lingüística de corpus (apartat 3.4) per a identificar els textos rellevants que ens permeten constituir i explotar un corpus textual de referència dels set diaris seleccionats sobre llengua i multilingüisme. Per últim, apliquem els principis teòrics i metodològics de la lexicometria (apartat 3.5) per a identificar les tendències quantitatives del vocabulari. Aquest tractament estadístic del corpus ens permet seleccionar les unitats d'anàlisi a partir de criteris estadístics i basar l'estudi qualitatiu posterior en les paraules (i els contextos d'ús) més representatius dels diaris analitzats.

### 3.1. L'anàlisi del discurs

Aquesta recerca parteix del concepte de *discurs* entès com la forma de representar i valorar el món, tot seguint la definició de Fairclough (2003, p. 124). Segons Van Dijk (1997), el concepte de discurs està format per les tres dimensions següents:

- 1) La paraula (el text)
- 2) La comunicació de creences, opinions o ideologies (la cognició social)
- 3) La interacció social

Per tant, el món (i també les llengües) poden representar-se de formes diferents a través del discurs. Els discursos tenen unes característiques distintives: són compartits pels grups socials i, alhora, són mutables i immutables. Sovint, són combinacions de múltiples discursos. Poden aparèixer en forma de declaracions estructurades, sostingudes en el temps i també a través d'opinions aparentment desestructurades, com per exemple els discursos de la ciutadania en la seva vida diària (Fairclough, 2003, p. 124-126; van Dijk, 1998, p.8).

Per denominar les diverses formes de representar i valorar el món (i, per tant, les llengües), la literatura en ciències socials ha emprat sovint conceptes diversos. Un dels termes utilitzats sovint és el de *marc* (o *framing*), noció que fa referència a les estructures cognitives que guien la percepció i la representació de la realitat (Martínez de Luna, 2013a, p. 24).

En relació amb l'anàlisi de la conceptualització discursiva de les llengües als mitjans de comunicació, un dels conceptes més pertinents és la *ideologia lingüística*, entesa com el conjunt d'idees i creences de les persones sobre les llengües. Aquestes es manifesten a

través de l'ús (o del no ús), o de les actituds i prejudicis de les llengües mateixes. Tal com afirma Bell, la ideologia lingüística és la forma de valorar les llengües, tot afirmant que una llengua (o una varietat lingüística) és millor que una altra (2014, p. 286).

Els discursos i la societat en què es produeixen estan intrínsecament associats: les relacions entre els discursos formen part de les relacions interpersonals. Alhora, les persones els utilitzen en les seves relacions interpersonals com a mitjà per a col·laborar, competir o sotmetre-les (Fairclough, 2003, p. 124).

En definitiva, l'objectiu final dels discursos és ser efectius socialment. Bourdieu el denominava l'*eficàcia performativa*: establir una visió concreta i una divisió del món social (1982, p. 90). Per això, l'anàlisi del discurs és clau: emprant la llengua d'una determinada manera (o, *tout court*, no utilitzant-la) o amagant-ne alguns elements en detriment d'altres, comuniquem una forma d'interpretar i d'entendre el món, tot cercant la primacia en la interacció social. En definitiva, aspiren a crear i a canviar la realitat social (Martínez de Luna, 2013, p. 17).

Atès el caràcter aplicat d'aquesta recerca, no ens proposem analitzar des d'un vessant teòric els diferents corrents de l'anàlisi del discurs ni els principis teòrics dels marcs discursius que predominen als diaris analitzats. Per al nostre treball, és important entendre que l'objectiu de les ideologies, els marcs i els discursos no és reflectir la realitat de forma veraç, sinó dirigir i orientar les creences cap a uns postulats determinants. Tal com afirma van Dijk, "el criteri de validesa de les ideologies no és la veritat sinó l'eficàcia social. Siguin o no verídiques, les ideologies controlen el que els mateixos grups consideren normalment que són creences vertaderes" (van Dijk, 1997, p. 55). Quan la realitat i les creences no coincideixen, prevalen les creences i no pas la realitat. Segons Lakoff, "la gent pensa mitjançant marcs. La veritat, per ser encaixada, ha d'encaixar en els marcs de la gent. Si els fets no encaixen en un marc determinat, el marc es manté i els fets reboten [...] s'ignoren els fets i el marc continua funcionant" (2004, p. 39).

En aquest context, resulta d'especial rellevància conèixer i analitzar els discursos sobre les llengües seleccionades a Espanya que vehiculen els set diaris generalistes escrits, identificar-hi els marcs discursius activats i les intencionalitats que hi són subjacents.

### **3.2. El discurs periodístic**

Com apuntàvem, els mitjans de comunicació tenen un paper cabdal com a altaveu i com a creadors de discurs. Els mitjans transmeten diferents punts de vista, estableixen categories que ens permeten interpretar la realitat i permeten construir o perpetuar actituds i prejudicis.

En general, per tant, l'anàlisi dels mitjans de comunicació i en concret de la premsa escrita, ens permet identificar-hi els marcs mentals activats i, a més, permet constatar el manteniment o el canvi de determinades estructures socials.

Així, la premsa fa d'altaveu (o, directament, els crea) de determinats prejudicis, actituds i (pre)concepcions sobre la diversitat lingüística, que la societat acaba encaixant. D'una banda, amb una conceptualització del multilingüisme i la preservació de llengües com a recurs 'additiu' que suposa un avantatge cognitiu per a la ciutadania. De l'altra, com a barrera 'sostractiva' que impedeix una gestió efectiva, correcta i 'economitzada' de les llengües en contacte en una societat. L'anàlisi de la premsa espanyola i catalana ens permet identificar les pràctiques discursives que funcionen ideològicament amb relació a la representació discursiva de les llengües que se'n deriva.

Les anàlisis sociolingüístiques de la representació de pràctiques multilingües als mitjans de comunicació daten de finals dels anys 1990 (Jaworska i Themistocleous, 2017, p. 59). Des de llavors, els investigadors coincideixen a afirmar que els efectes dels discursos mediàtics són cumulatiu i es basen en estratègies de repetició sistematitzada (denominada *recursivitat* o *iterativitat*) de determinats missatges en diferents mitjans. Ho aconseguen emprant *topoi* o camps semàntics entextualitzats, acompanyats de la repetició de la iconització o l'omissió completa (*erasure*) de certs aspectes socioeconòmics referents a la 'llengua' pertanyents a altres camps semàntics (Camargo-Fernández, 2021).

Malgrat aquesta unanimitat en la importància de la lexicosemàntica per a abordar la creació i circulació de discursos sobre el multilingüisme, és relativament nou l'estudi de la premsa amb la perspectiva que la concep com un espai de gestació, reproducció i circulació de postulats ideològics. A més, tenen un enfocament clar en els recursos discursius que instrumentalitzen el que hauria de ser una llengua, el seu estatus, expansió i defensa tot fent èmfasi en el *topos* PROBLEMA o CONFLICTE LINGÜÍSTIC (Marimón Llorca, 2021, p. 81).

### 3.3. El marc metodològic

Aquest treball pren com a punt de partida metodològic l'hermenèutica de Thompson (1984). L'aproximació *thompsoniana* posa en relació l'anàlisi del microtext amb l'anàlisi del context en què circulen els discursos, entesos com a fruit, reflex i mirall dels macro processos polítics, econòmics, demogràfics i culturals de la societat en què es produeixen, en un període de temps i en un espai situat i concret.

Aquest enfocament requereix un procés d'anàlisi metodològica basada en tres aspectes:

1. L'anàlisi sociohistòrica: s'endinsa en el context polític, social i econòmic en què es produeixen els discursos;
2. L'anàlisi textual del discurs: beu de la lingüística de corpus, analitza el *topoi*, les cadenes de raonament discursives i els elements retòrics; i, per últim,

3. L'anàlisi interpretativa: lliga la segona fase i permet que emergeixin les ideologies subjacents als discursos sobre llengües als diaris seleccionats.

Argüim que el marc de Thompson és útil per al nostre estudi, ja que ens permet tenir en compte el dinamisme i la variació contínua en la manera com la llengua s'instrumentalitza a través dels mitjans de comunicació. A més, contribueix a identificar de quina manera aquesta relació pren forma i evoluciona constantment cap a uns postulats o d'altres, segons la llengua de què tracta. És a dir, ens permet analitzar uns textos en el marc d'un context històric, social, econòmic i polític reals contextualitzat; per últim, ens permet copsar i resseguir les veus d'oposició, silenci i contestació mobilitzades dins d'uns marcs dominants concrets.

Existeix un grup de recerca amb un projecte similar al que desenvolupem. Des del 2008, Francisco M. Carriscondo Esquivel desenvolupa i encapçala (juntament amb estudiants de la Facultat de Ciències de la Comunicació i Filosofia i Lletres de la Universitat de Màlaga) el projecte "Análisis de noticias sobre divulgación lingüística, las lenguas de España y sus variedades (Proyecto Lengua y Prensa)"<sup>2</sup>.

Aquest projecte buida mitjans de comunicació en castellà per a identificar-hi notícies sobre divulgació lingüística, les llengües d'Espanya i llurs varietats. Segons les dades quantitatives de l'apartat "Hemeroteca", la premsa escrita és la categoria més freqüent, amb escreix (vegeu la Taula 1).

Categories	Quantitat
Blogs	374
Canals	314
Documents	45
Pàgines web	590
Plataformes virtuals	929
Podcasts	218

Categories	Quantitat
Premsa escrita	21.468
Ràdio	543
Revistes	120
Televisió	238
Vídeos	500

Taula 1. Dades quantitatives del "Proyecto Lengua y Prensa" (Universitat de Màlaga)

Tot i que el corpus de notícies es continua alimentant, no ens consta que les dades del període analitzat en aquesta recerca (2015-2021) s'hagin explotat amb un enfocament discursiu i lexicomètric com el d'aquest treball.

<sup>2</sup> <https://www.lenguayprensa.uma.es/index.php> [Data de consulta: 22/11/2022].

### 3.4. La lingüística de corpus

La lingüística de corpus (LC) és una branca fonamental de la Lingüística Aplicada actual, tot i que es comença a desenvolupar durant la segona meitat del segle XX. L'adoptem com a punt de partida perquè, com afirma Stefanowitsch, “*whatever aspect of language we study and from whatever perspective we do so, if we are doing so scientifically, observation and experimentation should have a role to play*” (2020, p. 1).

Fins als anys vuitanta, aquesta disciplina passa desapercebuda, però creix l'interès gràcies a factors com el desenvolupament de la informàtica a gran escala, l'abaratiment progressiu i la millora de prestacions dels equips, i l'accés generalitzat d'institucions i de la població a aquestes tecnologies. Molts autors consideren el 1961 l'any de constitució del Brown Corpus, el naixement de la lingüística de corpus moderna.

Els primers treballs de lingüística de corpus, però, són previs. El 1897, el lingüista Friedrich Wilhelm Kaeding va analitzar un corpus d'aproximadament 11 milions de paraules per calcular la freqüència absoluta de cada lletra i les seqüències més habituals en alemany. Altres pioners en la lingüística de corpus foren F. Boas, Z. Harris o C. Fries.

L'*Enciclopèdia Catalana* la defineix com la “branca de la lingüística que es basa en l'estudi dels fenòmens lingüístics a partir de grans col·leccions de textos en format màquina, anomenats corpus” (GEC, 2022). La definició més acceptada actualment en l'àmbit acadèmic és la següent:

*Corpus linguistics is the investigation of linguistic research questions that have been framed in terms of the conditional distribution of linguistic phenomena in a linguistic corpus.*

(Stefanowitsch, 2020, p. 56)

En aquest context, adoptem la definició de *corpus* següent:

*A collection of samples of language use with the following properties:*

- *The instances of language use contained in it are authentic;*
- *The collection is representative of the language or language variety under investigation;*
- *The collection is large.*

(Stefanowitsch, 2020, p. 22-23)

Segons Meyer, la LC ha afegit un gran valor a la recerca en Lingüística:

*Corpus linguistics has been one of the more exciting methodological developments in linguistics since the Chomskyan revolution of the 1950s. It reflects changing attitudes among many linguists as to what constitutes an adequate “empirical” study of language, and it has drawn upon recent developments in technology to make feasible the kinds of empirical analyses of language that corpus linguists wish to undertake.*

(Meyer, 2002, p. 141)

De fet, es considera que les teories generativistes de Chomsky són un punt d'inflexió en el desenvolupament d'aquesta branca. Aquest autor hi planteja “unes idees que van influir molt en la visió del llenguatge i que han estat objecte de debat en altres camps, com la psicologia humana” (Castellà, 1992, p. 22). Tot i això, Chomsky no acceptava el corpus lingüístic com una eina d'anàlisi lingüística, ja que considerava que calia estudiar la competència (i no l'actuació lingüística), perquè la tasca de la gramàtica fos “construir un model formal explicatiu de les regles, principis, etc. que constitueixen la competència lingüística dels parlants” (Castellà, 1992, p. 103).

A diferència del generativisme, tanmateix, la LC prioritza el component descriptiu per sobre de l'explicatiu:

*Corpus linguists see complexity and variation as inherent in language, and in their discussions of language, they place a very high priority on descriptive adequacy, not explanatory adequacy. Consequently, corpus linguists are very skeptical of the highly abstract and decontextualized discussions of language [...], because such discussions are too far removed from actual language use.*

(Meyer, 2002, p. 3)

Aquesta recerca parteix d'una exploració lexicomètrica dels titulars i articles de premsa sobre aspectes lingüístics i es basa en la LC perquè integra una metodologia d'anàlisi empírica i sistemàtica que permet realitzar una tria de les unitats d'estudi basada en criteris quantitius (i, per tant, més objectius que no pas si partíssim d'una anàlisi aleatòria i subjectiva de les unitats).

La LC, per últim, ens permet construir un recull de notícies de premsa sobre qüestions lingüístiques, el qual podem analitzar des de l'aproximació teòrica i pràctica de l'anàlisi del discurs. Hi apliquem, a més, una metodologia de detecció de les unitats d'anàlisi basada en criteris estadístics, gràcies a l'exploració lexicomètrica del corpus de notícies en català i castellà del període 2015-2021.

### **3.4. La lexicometria**

La lexicometria és una branca de la lexicologia que analitza els textos partint de la freqüència d'ocurrència dels mots que l'integren. El *Vocabulari de la lexicometria* la defineix com la “disciplina que permet descriure quantitativament els textos constitutius d'un corpus a partir de la seva segmentació en formes gràfiques. Consta d'un conjunt de mètodes que permeten fer reorganitzacions formals de la seqüència textual i anàlisis estadístiques de vocabulari” (Salem, Cabré i Romeu, 1990, p. 26).



Aquesta branca també s'anomena estadística lingüística (Guiraud, 1960), estadística lèxica (o lingüística quantitativa, Muller, 1974), estadística textual (Lebart i Salem, 1994), anàlisi de dades en lingüística (Benzécri, 1980) o lexicologia quantitativa (Niklas-Salminen, 1997, p. 35). Tot i la variació, nosaltres fem lexicometria per la seva precisió, com argumentava M. Teresa Cabré l'any 1977 (“Preferimos el término lexicometría a estadística textual porque, como muy bien justifican, es más preciso en cuanto al objeto y más amplio en cuanto al método”) (Cabré Castellví, 1977, p. 197)

Aquesta aproximació defineix la noció de *mot*, *paraula* o *forma* des d'un punt de vista formal, és a dir, com “le groupement de lettres, séparé, à gauche et à droite, par un blanc des autres éléments du texte” (Niklas-Salminen, 1997, p. 13-14). Els pioners dels estudis lexicomètrics són Wagner (1949) i Guiraud (1953) i, posteriorment, De Bazin (1967) i Muller (1967). Nombrosos autors (entre d'altres, Guilhaumou, 2002, p. 14) afirmen que la branca, la denominació mateixa i els orígens metodològics de la lexicometria neixen al Laboratoire de lexicologie politique de l'ENS de Saint-Cloud, a França. El primer treball exhaustiu realitzat des d'aquesta aproximació és l'anàlisi dels pamflets del maig del 68 francès, recollit a *Des tracts en mai 68* (Demonet et al., 1975).

Segons el volum *La lexicometria*, aquesta metodologia “permet superar el simple recompte de mots i establir, a partir del tractament sistemàtic i automatitzat de corpus de textos, les complexes relacions estadístiques que es produeixen entre les unitats lèxiques” (Centre de Lexicometria 1993, 1). Això permet una selecció del vocabulari que no es basa en criteris atribuïbles a l'investigador/a –com la tria subjectiva de paraules clau d'un àmbit o d'una temàtica–, sinó que, a partir d'un estudi inicial basat en càlculs matemàtics, es detecten les unitats que resulten més significatives estadísticament. La seva objectivitat comença, doncs, amb la selecció i l'anàlisi:

*Pour la lexicométrie, l'objectivité commence quand les projections tendent à faire place à des formalisations. On ne fixe d'avance ni mots-témoins, ni grilles, ni schèmes, ni hiérarchies. On prend tout (exhaustivité des relevés), de la même manière (uniformité du dépouillement) et selon un critère unique (invariance d'une unité de base anonyme). Cette «neutralité» n'est pas un postulat innocent: elle signifie que chaque occurrence est également significative et que seules entrent en ligne de compte les valorisations quantitatives ou formelles qui sont le fait du texte.*

(Demonet, et al., 1975, p. 20-21)

Aquesta metodologia compta ja amb una tradició madura i consolidada. Podem esmentar els estudis de Benoit (1982), Salem (1993), Bigey (2007) o Savoy (2014), entre d'altres. La tradició en recerca lexicomètrica té dues grans aproximacions: la formalista i la lematitzadora. Fonamentalment, la diferència rau en la unitat d'anàlisi: el primer grup considera que cal estudiar les ocurrencies de cada *forma*. Parteix d'un enfocament semasiològic, ja que “privilégie les signes pour aller vers le sens-cette liste de formes

constitue seulement un matériau de départ” (Martinez, 2003, p. 19). Alguns autors que l’han emprat son Pierre Lafon, André Salem o Maurice Tournier.

El segon grup parteix de l’agrupament de totes les formes d’un paradigma o flexió: llurs recerques parteixen del *lemma*. Investigadors com Étienne Brunet o Charles Muller l’han emprat, adduint que “la lemmatisation, malgré ses imperfections, a le mérite de masquer l’une de ces réalités, donc d’isoler l’autre et d’en permettre la mesure” (Muller, 1984, p. XI).

Tradicionalment la tria d’una escola era cabdal, perquè les eines d’anàlisi lexicomètrica n’empraven l’una o l’altra. Ara, tanmateix, programes com *#LancsBox*, permeten explotar les dades a partir d’ambdues unitats.

Les quatre darreres dècades s’han desenvolupat nombroses recerques i eines lexicomètriques, sobretot en l’àmbit francòfon. Inicialment, la universitat catalana rep amb entusiasme la lexicometria. Poc després que se’n publiquin els primers estudis lexicomètrics a França, també n’apareixen a casa nostra, com el dels vocabularis dels debats parlamentaris de la Constitució espanyola de 1932 (Cabré 1977), el dels programes i manifestos electorals d’aquella dècada (Anglada 1978) o el dels pamflets partidistes publicats en clandestinitat a Espanya durant el període 1965-1974 (Lleal 1980). Darrerament, també l’han emprat autors com Fernández Moreno (1998), Grasset (2011) i Morales (2015, 2018, 2021).

## 4. Corpus d'anàlisi

El corpus d'aquest treball està format per set diaris, editats a Madrid i a Catalunya: *La Vanguardia*, *El Periódico*, *l'Ara* i *El Punt Avui*, *ABC*, *El Mundo* i *El País*.

Premsa en català	Premsa en castellà
ARA 	ABC 
PUN 	MUN 
PER 	PAI 
VAN 	–

Taula 2. Diaris analitzats

La tria d'aquests set diaris es justifica pels criteris que exposem tot seguit:

1. El criteri quantitatiu, basat en el nombre total de lectors dels diaris. D'una banda, *La Vanguardia*, *El Periódico*, *l'Ara* i *El Punt Avui* són les capçaleres generalistes en català més llegides a Catalunya. De l'altra, *El País*, *ABC* i *El Mundo*, són les més llegides en castellà a l'Estat espanyol. A l'apartat 4.1 exposem el nombre de lectors i els percentatges corresponents a cada diari.
2. El criteri ideològic. Hem seguit els postulats de Reig (2011) per garantir una tria que representa diferents línies editorials i tot l'arc ideològic. D'un costat, *l'Ara* (creat el 2010) i *El Punt Avui* (fusió de dos diaris en català i creat el 2011) tenen una marcada línia nacionalista catalana i d'esquerres. *La Vanguardia*, amb més de 130 d'història i associat tradicionalment al centre-dreta català, s'alinea ideològicament a la desapareguda coalició de Convergència i Unió (el seu partit actual hereu seria *Junts x Catalunya*). *El País* ha estat vinculat tradicionalment a l'esquerra mediàtica estatal i, segons Reig, ha tendit a reforçar els posicionaments socialistes. Val a dir, però, que aquests posicionaments han anat canviant, segons la direcció del diari. En la mateixa

línia socialista estatal trobem *El Periódico*. En darrer lloc, els diaris més marcadament conservadors en clau estatal són *El Mundo* i *ABC*.

El doble filtre (nombre de lectors i espectre ideològic) ens ha permès tenir una mostra rica pel que fa a abast i representativitat ideològica.

#### 4.1. Constitució i explotació

El nostre objectiu, com ja hem assenyalat, és identificar de quina manera es representen discursivament les llengües seleccionades (i, en especial, la llengua catalana) als discursos mediàtics,. De les opcions disponibles, hem optat per mitjans de premsa escrita.

El projecte endegat, però, ha estat molt ambiciós. Només la constitució del corpus ha superat les nostres expectatives inicials quant a abast i dimensions. Per tal de garantir l'especificitat de l'anàlisi, l'objectivitat i la sistematicitat en la recollida de les dades i la reproductibilitat de la constitució del corpus, hem emprat la base de dades *Factiva*.

*Dow Jones Factiva*<sup>3</sup> és una plataforma que ofereix més de 32.000 fonts internacionals de premsa i actualitat econòmica i financera, disponibles en 28 llengües (entre d'altres, l'anglès, l'àrab, el danès, l'espanyol, el tailandès i el català). La Taula 3 resumeix les fonts disponibles en català i castellà:

Tipus de font	Àmbit català	Àmbit estatal	Total
<b>Blog</b>	–	27	27
<b>Lloc web</b>	22	142	164
<b>Publicació</b>	24	223	247
Total	46	392	438

Taula 3. Fonts disponibles en català i castellà d'àmbit espanyol (*Factiva*)

En català inclou 46 fonts d'àmbits diferents (local, regional, autonòmic o estatal), com *El 9 nou*, *e-notícies*, *L'Econòmic*, *Naciódigital.cat*, *El Periòdic d'Andorra*, *Regió 7*, *SEGRE.com*, *Tot Sant Cugat* o *VilaWeb*, entre d'altres. En castellà, hi ha 1.382 fonts, relatives a àmbits i països hispanòfons diferents, com Argentina, Mèxic o Veneçuela. N'hi ha 392 de l'àmbit espanyol. Se n'inclouen, entre d'altres, fonts com *20 Minutos*, *Andalucía*

<sup>3</sup> <https://www.dowjones.com/professional/factiva/>

*Económica, Cinco Días, El Correo de Burgos, El Día de Valladolid, Diario de Sevilla, Gaceta.es, Nuevatribuna.es, La Tribuna de Ciudad Real, La Voz de Cádiz o Vozpópuli.* Aquest projecte es va començar a gestar durant la primera meitat de l'any 2021. Per aquest motiu, la tria de fonts s'ha basat en els resultats del rànquing de premsa generalista del Baròmetre d'Opinió Política (BOP) del 2020 (3a onada)<sup>4</sup>. Segons aquesta classificació, els diaris generalistes d'àmbit català més llegits són *La Vanguardia, Ara, El Periódico* i *El Punt Avui* (vegeu la Taula 4).

Diari	Freqüències	%
<b>La Vanguardia (català)</b>	126	20,4 %
<i>La Vanguardia</i> (castellà)*	79	12,8 %
<i>El Periódico</i> (castellà)*	72	11,7 %
<b>Ara</b>	68	11,0 %
<b>El Periódico (català)</b>	39	6,3 %
<i>El País</i>	35	5,7 %
<b>El Punt Avui</b>	25	4,1 %
<i>Vilaweb</i>	20	3,2 %
<i>Diari de Tarragona</i>	13	2,1 %
<i>Diari Segre</i>	10	1,6 %

Taula 4. Els 10 diaris més llegits en català (Font: BOP – Dades 2020 (3a onada))

Pel que fa al castellà, hem adoptat com a referència l'*Estudio General de Medios* del 2020 (3a onada) per identificar-ne els més llegits. Segons aquesta classificació, els diaris generalistes més llegits en castellà són *El País, El Mundo, La Vanguardia, ABC* i *El Periódico* (vegeu la Taula 5).

Diari	Lectors diaris (000)	%
<i>Marca</i>	1.355	18,1%
<b>El País</b>	951	12,7%
<i>As</i>	631	8,4%
<b>El Mundo</b>	622	8,3%

<sup>4</sup> <https://upceo.ceo.gencat.cat/wsceop/7808/Taules%20estad%C3%ADstiques%20-985.pdf> [Consulta: 09/08/2022].

<i>La Vanguardia*</i>	482	6,5%
<b>ABC</b>	433	5,8%
<i>La Voz de Galicia</i>	418	5,6%
<i>Mundo Deportivo</i>	299	4,0%
<i>El Periódico*</i>	287	3,8%
<i>El Correo</i>	281	3,8%

Taula 5. Els 10 diaris més llegits en castellà (Font: EGM – Dades 2020 (3a onada))

Atesa la situació política dels darrers anys, l'equip de treball es proposava d'analitzar des d'una òptica diacrònica de quina manera es representen les diferents llengües seleccionades als diaris esmentats durant el període 01/09/2015-30/09/2021. Com ja hem assenyalat durant la presentació del marc teòric i metodològic de l'estudi, accedim a les notícies pertinents mitjançant la cerca de paraules clau. El cercador de *Factiva* permet fer-hi cerques amb operadors booleans combinats.

L'expressió *booleana* combinada permet que el motor de cerca trobi tots articles en català que emprin les paraules *anglès*<sup>5</sup>, *castellà*, *espanyol*, *català* i *valencià* combinades amb formes que al·ludeixen a la llengua (*idioma*, *llengua*, *lingüístic*, *bilingüisme*<sup>6</sup>, *multilingüisme* i *plurilingüisme*). Vam replicar les mateixes cerques en castellà (*inglés*, *castellano*, *catalán*, *valenciano* i *español*, combinades amb *lingüístico/s*, *idioma/s*, *lengua/s*, *bilingüismo*, *multilingüismo* y *plurilingüismo*). La Taula 6 presenta la sintaxi dels operadors booleans emprats per a executar les cerques.

<b>CA</b>	<i>(angles* OR castella* OR catala* OR valencia* OR espanyol*) AND (linguistic* OR idiom* OR llengua* OR bilingu* OR multilingu* OR plurilingu*)</i>
<b>ES</b>	<i>(ingles* OR castellan* OR catalan* OR valencian* OR espanol*) AND (linguistic* OR idiom* OR lengua* OR bilingu* OR multilingu* OR plurilingu*)</i>

Taula 6. Operadors booleans emprats per a obtenir les dades (*Factiva*)

Des del primer moment érem conscients que la tria de les unitats lèxiques, atès llur caràcter freqüent i polisèmic, implicaria haver de filtrar els resultats manualment per tal

<sup>5</sup> Naturalment, les cerques combinades inclouen totes les formes de cada paradigma. Per exemple, la cerca amb l'operador *anglès\** inclou les formes *anglès*, *anglesa*, *anglesos* i *angleses*.

<sup>6</sup> Naturalment, les cerques combinades de *bilingüisme*, *multilingüisme* i *plurilingüisme* n'inclouen també l'adjectiu corresponent (*bilingüe/s*, *multilingüe/s*, *plurilingüe/s*...).

de seleccionar-ne només aquells textos pertinents que tractessin sobre el fet lingüístic amb relació a les llengües seleccionades. Tanmateix, aquesta cerca combinada en garantia la sistematicitat, l'exhaustivitat i la compleció de les dades de partida per a l'estudi.

Les llicències acadèmiques de *Factiva* emprades no permeten una exportació automàtica dels resultats, la qual n'inclogui les metadades i els textos dels articles. *Factiva* permet visualitzar els llistats dels resultats obtinguts de vint en vint (vegeu la Figura 1).

1. **La España plural, 500 años después Aragón Galicia Castilla y León Madrid País Vasco Andalucía Comunidad Valenciana Castilla-La Mancha...**  
ABC, 13 October 2015, 4808 words, (Spanish)  
Aragón no ha tenido necesidad de «fabricar» su pasado para ver reconocida su condición de nacionalidad histórica, reconocida expresamente en la primera frase del preámbulo de su Estatuto de Autonomía. Ese carácter de nacionalidad histórica, ...  
1 duplicate article(s) identified
2. **ANÁLISIS: En el sistema idiomático mundial**  
ABC, 13 November 2015, 369 words, (Spanish)  
Las cifras internacionales de la lengua española son reflejo de una comunidad vital, así como de una evidente pujanza cultural y educativa. Sin duda, las comunidades que disponen de un idioma internacional como primera o segunda lengua son ...
3. **ANÁLISIS: En el sistema idiomático mundial**  
ABC, 13 November 2015, 369 words, (Spanish)  
Las cifras internacionales de la lengua española son reflejo de una comunidad vital, así como de una evidente pujanza cultural y educativa. Sin duda, las comunidades que disponen de un idioma internacional como primera o segunda lengua son ...
4. **El tripartito evaluará el «nivel» de valenciano de 2.500 funcionarios**  
ABC, 8 December 2015, 571 words, (Spanish)  
La Concejalía de Presidencia del Ayuntamiento de Alicante, que dirige Natxo Bellido (Compromís), realizará en breve una encuesta entre los más de 2.000 funcionarios municipales con el objetivo de «conocer con exactitud el nivel actual de ...
5. **«Cataluña y el Estado han politizado el uso del catalán»; El profesor de la UPO Juan Jiménez-Salcedo ha coordinado un libro sobre las...**  
ABC Sevilla, 16 January 2016, 580 words, (Spanish)  
Juan Jiménez-Salcedo y Jorge Cagiao, profesores de las universidades Pablo de Olavide (UPO) y de Tours (Francia), respectivamente, han coordinado a una decena de autores que han analizado desde la perspectiva de la Sociolingüística, el ...
6. **Puig prima el valenciano pese a que solo lo habla el 36 por ciento**  
ABC, 14 February 2016, 1339 words, (Spanish)  
El Gobierno que preside el socialista Ximo Puig premiará con bonificaciones fiscales a las empresas (especialmente las pymes) que primen el uso del valenciano en la rotulación de sus negocios y en sus páginas web. Una política que ya ha ...
7. **La Junta lleva el segundo idioma al cole sin cumplir su plan de centros bilingües; El próximo curso se impartirá en Primaria otra lengua...**  
ABC Sevilla, 21 April 2016, 731 words, (Spanish)  
LA Consejería de Educación implantará de cara al próximo curso 2016-2017 el segundo idioma extranjero en 3º y 5º de los casi 2.500 centros públicos de Educación Primaria de Andalucía, con el objetivo de fomentar el conocimiento de otras ...  
1 duplicate article(s) identified
8. **Catorce localidades acogen los 38 cursos de verano de la UCLM**  
ABC, 9 May 2016, 847 words, (Spanish)  
La Universidad de Castilla-La Mancha (UCLM) ha programado 38 actividades para los Cursos de Verano 2016. Una edición, la vigesimonovena, que busca favorecer el encuentro de especialistas de distintas materias y disciplinas con el objetivo ...
9. **LA ENSEÑANZA CONCERTADA VUELVE AL DEBATE; El consejero de Educación, Ángel Felpeto, culpó al PP de querer provocar «polémicas artificiales»**  
ABC, 3 June 2016, 1003 words, (Spanish)  
El Grupo Parlamentario del Partido Popular en las Cortes Regionales presentó ayer un debate general sobre la enseñanza concertada, con la intención de que el Gobierno regional «aclare» si va a cumplirse el acuerdo firmado con Podemos para ...
10. **«Catalán y castellano deben convivir sin violencia»; Sociedad Civil Catalana defiende el bilingüismo en la semana de la Diada**  
ABC, 9 September 2016, 503 words, (Spanish)  
La sala del Paraninfo de la Universidad de Barcelona se quedó pequeña ayer para acoger a las más de 400 personas que asistieron a un acto en favor del bilingüismo organizado por la entidad antiindependentista Sociedad Civil Catalana (SCC).  
1 duplicate article(s) identified

Figura 1. Interficie (*Factiva*)

Com s'observa, les pàgines de resultats de *Factiva* inclouen les dades i metadades següents:

- Titular
- Font
- Data
- Càmput de paraules de l'article
- Llengua de l'article
- Fragment del text

En total, els set diaris analitzats inclouen 46.960 articles que contenen, al títol o al cos del text, les paraules seleccionades combinades amb els operadors booleans emprats. Aquests articles es distribueixen per llengua i diari de la manera següent (vegeu la Taula 7).

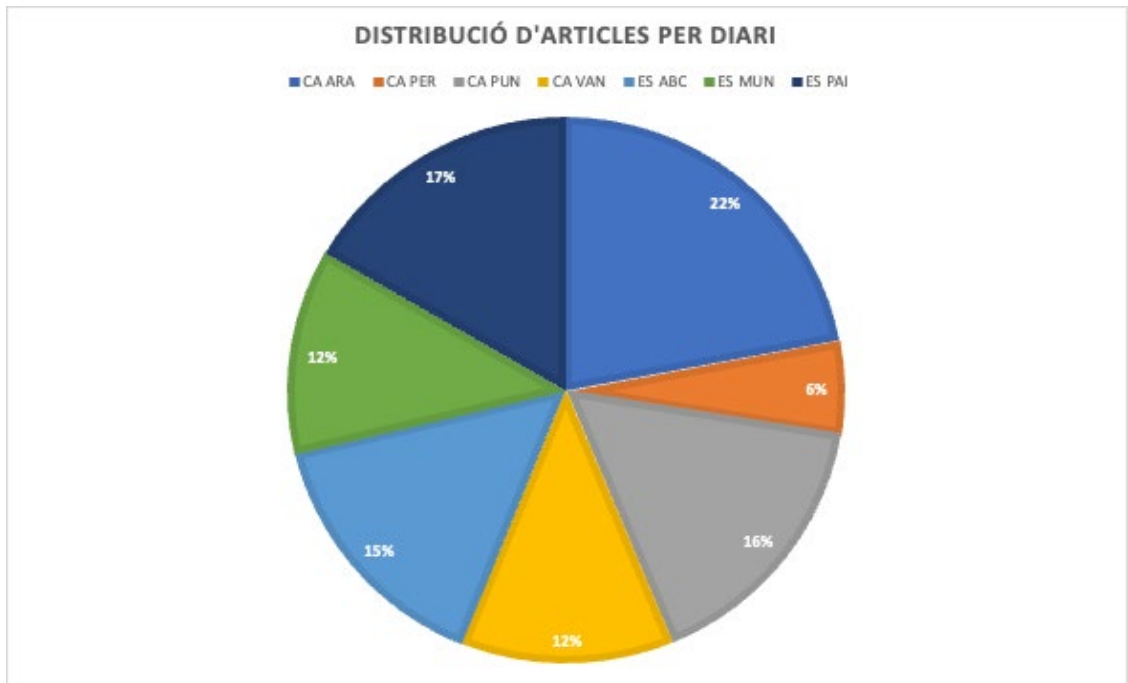
Diari	CA				Total CA	ES			Total ES	Total
	ARA	PER	PUN	VAN		ABC	MUN	PAI		
<b>2015</b>	686	108	533	425	1.752	532	321	372	1.225	2.977
<b>2016</b>	1.886	210	1.560	1.042	4.698	1.055	846	1.289	3.190	7.888
<b>2017</b>	1.766	255	1.816	805	4.642	1.087	1.092	1.466	3.645	8.287
<b>2018</b>	1.938	225	1.240	952	4.355	1.291	1.030	1.419	3.740	8.095
<b>2019</b>	1.915	674	990	873	4.452	1.309	1.037	1.331	3.677	8.129
<b>2020</b>	1.456	580	812	1.023	3.871	1.020	836	1.044	2.900	6.771
<b>2021</b>	700	537	655	684	2.576	837	526	874	2.237	4.813
<b>Total</b>	<b>10.347</b>	<b>2.589</b>	<b>7.606</b>	<b>5.804</b>	<b>26.346</b>	<b>7.131</b>	<b>5.688</b>	<b>7.795</b>	<b>20.614</b>	<b>46.960</b>

**Taula 7. Distribució dels articles obtinguts amb els operadors booleans (*Factiva*)**

Aquestes dades globals ja ens permeten constatar el fet que la premsa en català parla molt més de qüestions lingüístiques que la premsa en castellà: els diaris en català seleccionats inclouen 5.732 articles més (en termes relatius, un 21,7%) sobre llengua que els diaris en castellà (vegeu la Gràfica 1).

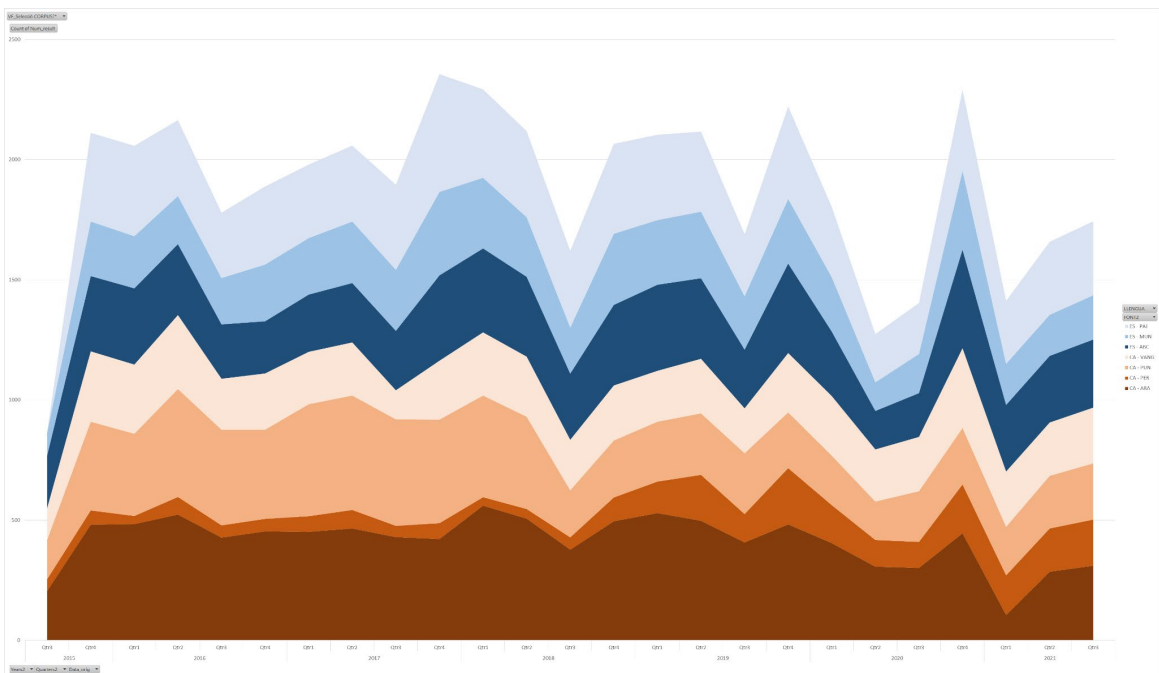
Globalment, a més, constatem que en termes absoluts, el 52,3% dels articles (en total, 24.551) corresponen al període 2017-2019. Cal assenyalar també que el diari més en parla és l'*Ara* (vegeu la Gràfica 1): el 22% dels articles recuperats (10.347 textos), correspon a aquesta font, i se situa cinc punts percentuals per sobre de la segona font més productiva (*El País*, amb 7.795 articles). Els que inclouen menys notícies de llengua són *La Vanguardia* (5.804 articles), *El Mundo* (5.688 articles) i, especialment, *El Periódico* (6%).





**Gràfica 1. Distribució d'articles per diari (dades sense filtrar)**

Cronològicament, aquests textos es distribueixen com es representa a la Gràfica 2.



**Gràfica 2. Distribució cronològica d'articles per diari (dades sense filtrar)**

#### 4.1.1. Textos pertinents

El percentatge de soroll dels resultats sense duplicats era elevadíssim, atesa la naturalesa polisèmica i el caràcter general d'algunes paraules clau de les cerques a *Factiva*. Pràcticament, un de cada quatre articles no era pertinent, tal com il·lustren els exemples següents.

*Sagarra, Montellà i Millàs, a les proves de selectivitat de setembre (Ara, 03/09/2015)*

*El canal cultural ARTE incorpora l'anglès i el castellà (Ara, 18/11/2015)*

*La Mansana de la Discòrdia és avui una de les grans icones patrimonials de... (El Punt Avui Catalunya, 17/01/2016)*

*Isabel-Clara Simó, Premi d'Honor de les Lletres Catalanes (El Periódico de Catalunya, 17/01/2017)*

*Booking contractarà un centenar de persones per reforçar-se a Barcelona (La Vanguardia, 26/01/2019)*

*Pérez-Reverte no vol canvis lingüístics a la Constitució (La Vanguardia, 13/07/2018)*

*El idioma como herramienta ideológica (ABC, 11/12/2019)*

*El giro en el lenguaje sobre Venezuela (El País, 07/02/2020)*

*El Rey: «Los españoles somos solo copropietarios de nuestro idioma» (ABC, 20/10/2015)*

*Trippier: «Mi prioridad ahora es el idioma» (El País, 19/07/2019)*

Per tal d'identificar els articles pertinents per a la nostra anàlisi, per tant, l'equip va revisar manualment el llistat íntegre d'articles de cada diari i va indicar-ne els que, en efecte, tracten dels fets lingüístics ja explicitats. Per a garantir la validesa de les dades, cadascun dels set llistats es van sotmetre a la doble verificació per part de tots dos investigadors de l'equip.

Així doncs, un cop revisades i verificades manualment, es van identificar 11.481 articles pertinents per a la nostra anàlisi (vegeu la Taula 8).

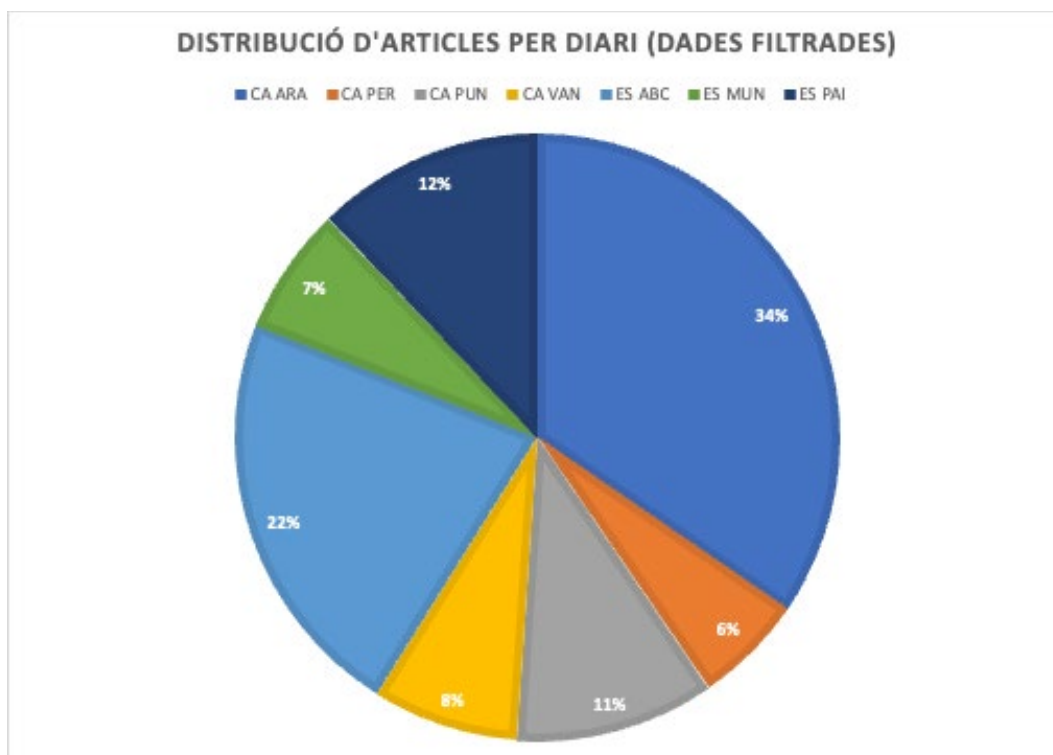
	CA				Total CA	ES			Total ES	Total
Diari	ARA	PER	PUN	VAN		ABC	MUN	PAI		
<b>2015</b>	270	37	55	103	<b>465</b>	186	26	56	<b>268</b>	<b>733</b>
<b>2016</b>	741	61	304	193	<b>1.299</b>	319	61	206	<b>586</b>	<b>1.885</b>
<b>2017</b>	588	57	249	142	<b>1.036</b>	341	95	273	<b>709</b>	<b>1.745</b>
<b>2018</b>	682	83	200	180	<b>1.145</b>	459	174	316	<b>949</b>	<b>2.094</b>

<b>2019</b>	805	201	159	115	<b>1.280</b>	436	156	232	<b>824</b>	<b>2.104</b>
<b>2020</b>	578	136	128	117	<b>959</b>	361	183	189	<b>733</b>	<b>1.692</b>
<b>2021</b>	305	100	126	47	<b>578</b>	442	85	123	<b>650</b>	<b>1.228</b>
<b>Total</b>	<b>3.969</b>	<b>675</b>	<b>1.221</b>	<b>897</b>	<b>6.762</b>	<b>2.544</b>	<b>780</b>	<b>1.395</b>	<b>4.719</b>	<b>11.481</b>

Taula 8. Distribució dels articles pertinents per a l'anàlisi

De nou, observem la mateixa tendència que a les dades en brut (vegeu la Taula 7): la premsa en català inclou més articles sobre qüestions lingüístiques que no pas la premsa en castellà (6.762 i 4.719 articles, respectivament; aquests 2.043 textos suplementaris identificats en català representen un 30,2% més). De mitjana, en termes absoluts, constatem que cada any de la sèrie analitzada es publiquen 1.640 articles sobre fets lingüístics. Tanmateix, el període 2016-2020 se situa per sobre d'aquesta mitjana, en especial els anys 2019 (2.104 textos) i 2018 (2.094 articles).

En termes relatius (vegeu la Gràfica 3), constatem tendències similars a les de l'apartat anterior: el diari *Ara* aporta el 34% d'articles del corpus (3.969 textos), seguit de l'*ABC* (22% del corpus, amb 2.544 unitats) i *El País* (1.395 articles, el 12% de textos del corpus).



Gràfica 3. Distribució d'articles per diari (dades filtrades)

Constatem, a més, que *El Periódico* és la font que menys articles sobre llengua publica: n'inclou 675 (el 6% del corpus) i, de fet, el 75% d'aquests articles correspon al període 2019-2021 (entre 2015 i 2018 es van publicar, només, 238 articles sobre llengua).

#### 4.1.2. Constitució del corpus

Com ja hem avançat a l'apartat anterior, aquest projecte es proposa d'analitzar tots els articles pertinents. Segons el còmput de paraules que consta per a cada article a la base de dades *Factiva*, la mida total del corpus està al voltant dels set milions de paraules (vegeu la Taula 9).

	CA		ES		Total (CA & ES)	
	<u>Articles</u>	<u>Paraules</u>	<u>Articles</u>	<u>Paraules</u>	<u>Articles</u>	<u>Paraules</u>
<b>2015</b>	<b>465</b>	291.114	<b>268</b>	159.600	<b>733</b>	<b>450.714</b>
<b>2016</b>	<b>1.299</b>	796.963	<b>586</b>	294.349	<b>1.885</b>	<b>1.091.312</b>
<b>2017</b>	<b>1.036</b>	643.950	<b>709</b>	396.495	<b>1.745</b>	<b>1.040.445</b>
<b>2018</b>	<b>1.145</b>	721.913	<b>949</b>	509.977	<b>2.094</b>	<b>1.231.890</b>
<b>2019</b>	<b>1.280</b>	845.911	<b>824</b>	455.817	<b>2.104</b>	<b>1.301.728</b>
<b>2020</b>	<b>959</b>	664.340	<b>733</b>	431.010	<b>1.692</b>	<b>1.095.350</b>
<b>2021</b>	<b>578</b>	394.877	<b>650</b>	398.203	<b>1.228</b>	<b>793.080</b>
<b>Total</b>	<b>6.762</b>	4.359.068	<b>4.719</b>	2.645.451	<b>11.481</b>	<b>7.004.519</b>

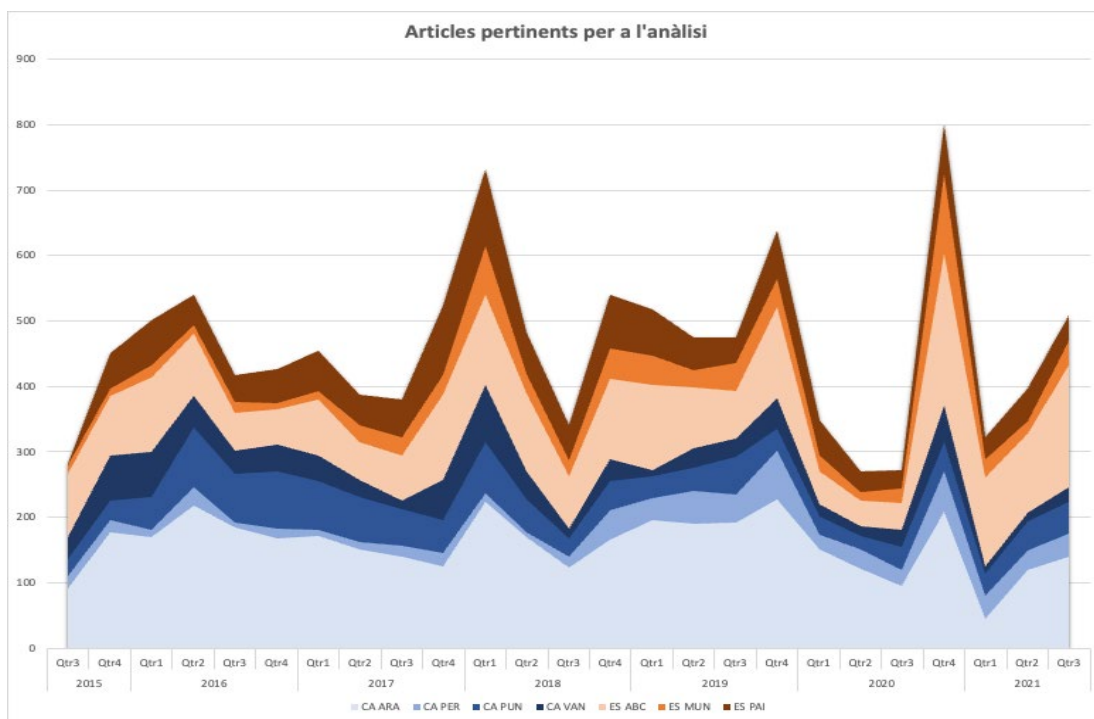
Taula 9. Documents i còmput de paraules dels articles pertinents per a l'anàlisi

Com es constata a la Taula 10, els cinc anys que més n'inclouen (pel que fa a còmput total de paraules) són, en ordre decreixent, el 2019, el 2018, el 2020, el 2016 i el 2017.

	Total CA & ES	
	Articles	Paraules
<b><u>2019</u></b>	2.104	1.301.728
<b><u>2018</u></b>	2.094	1.231.890
<b><u>2020</u></b>	1.692	1.095.350
<b><u>2016</u></b>	1.885	1.091.312
<b><u>2017</u></b>	1.745	1.040.445
<b>2021</b>	1.228	793.080
<b>2015</b>	733	450.714

**Taula 10. Articles filtrats (còmput de paraules decreixent)**

La representació de les dades de la Taula 10 permet identificar visualment en quins trimestres hi ha hagut pics de notícies sobre la temàtica analitzada (vegeu la Gràfica 4).



**Gràfica 4. Distribució cronològica d'articles per diari (dades filtrades)**

Els cinc pics més alts són: el quart trimestre de 2020 (801 articles), el primer trimestre de 2018 (731), el quart trimestre de 2019 (638 articles), i el segon trimestre de 2016 i el quart trimestre de 2018 (540 articles, cadascun).

Ateses les limitacions que ja hem assenyalat, aquesta investigació analitza un corpus representatiu i estratificat que representa el 2,5% del total del corpus textual: inclou 286 textos triats aleatòriament (en total, contenen 168.756 paraules). Per a les dades quantitatives completes per llengua i diari, vegeu la Taula 11.

	CA								Total CA		ES						Total ES		Total CA & ES	
	ARA		PER		PUN		VAN				ABC		MUN		PAI					
	A	P	A	P	A	P	A	P	A	P	A	P	A	P	A	P	A	P		
2016	38	24.922	3	1.333	15	6.016	9	3.703	65	35.974	16	7.283	3	1.678	11	5.591	30	14.552	95	50.526
2018	34	20.621	5	3.230	10	7.106	9	6.689	58	37.646	23	13.494	9	7.273	15	9.027	47	29.794	105	67.440
2020	30	19.754	7	3.842	7	4.384	5	3.399	49	31.379	18	8.821	9	4.890	10	5.700	37	19.411	86	50.790
Total	102	65.297	15	8.405	32	17.506	23	13.791	172	104.999	57	29.598	21	13.841	36	20.318	114	63.757	286	168.756

Taula 11. Dades quantitatives del corpus de l'anàlisi qualitativa

#### 4.1.3. Aspectes tècnics i explotació

Com ja hem assenyalat, la llicència acadèmica institucional de *Factiva* emprada no permet l'exportació massiva automàtica de resultats. El sistema permet exportar, de vint en vint, els articles en què apareixen les paraules clau de la cerca en format de text enriquit (RTF) o en format de document portàtil (PDF) (vegeu les fletxes de la Figura 2).

The screenshot shows the Factiva search results page. At the top, there are navigation tabs: Home, Search, Alerts, Newsletters, Companies/Markets. Below that is a search bar with filters for language (English or Catalan), date (09/01/2015 to 09/30/2021), and source (El Mundo (Spain, Spanish L...)). The main area displays a list of search results, each with a checkbox and a brief description. Two red arrows point to the 'RTF' and 'PDF' export options in the top right of the results list. The sidebar on the left shows various filters like Date, Companies, Sources, and Subjects.

Figura 2. Interfície de resultats i opcions d'exportació (*Factiva*)

Tanmateix, la interfície d'aquesta plataforma no ens ha permès d'exportar la taula de metadades (llengua, data de publicació, diari...) de resultats. Per aquest motiu, hem hagut de dur a terme manualment les operacions de copiar i enganxar manualment cadascuna de les pàgines de resultats (vint per pantalla, com la de la Figura 2) en un fitxer de text (.TXT).

Posteriorment, hem aplicat expressions regulars amb l'editor *TextMate* (versió 2.0.23) per a estructurar les dades en format de text separat per tabuladors (.TSV) i, posteriorment, poder assignar un codi identificador<sup>7</sup> a cada text i classificar-los amb *Microsoft® Excel for Mac* (versió 16.67) per tal de destriar els articles pertinents dels no pertinents per a la nostra anàlisi.

#### 4.1.3.1. Constitució del corpus de titulars

L'estudi lexicomètric dels titulars dels 11.481 articles del nostre corpus d'anàlisi s'ha de fer per llengua, perquè les eines d'estadística textual treballen amb corpus monolingües. Per tant, s'han agrupat els titulars corresponents als 6.762 textos en català en un arxiu de text (.TXT) i els titulars dels 4.719 articles en castellà en un altre.

#### 4.1.3.2. Constitució del corpus textual

Posteriorment, per a la constitució del corpus analitzat qualitativament, integrat per 286 articles dels anys 2016, 2018 i 2020, hem hagut de tornar a la base de dades *Factiva* i exportar-ne el contingut complet de l'article a un arxiu .TXT. S'hi han hagut d'eliminar manualment les parts relatives a les metadades, com en aquest exemple del diari *Ara*, en què el color vermell ratllat indica els fragments eliminats (vegeu la Figura 3) i el verd el text que s'ha mantingut.

---

<sup>7</sup> Els codis dels noms dels arxius segueixen una sintaxi senzilla i intuïtiva que inclou informació sobre la llengua (CA/ES), l'any de publicació (2015-2021), la font (ARA, LV, PER, PUN, ABC, MUN, PAI) i l'identificador numèric que n'indica l'ordre de més antic a més recent (000001, 000002...). Per exemple, l'arxiu *ES\_2020\_ABC\_01738.txt* correspon a l'article "[Brotes de xenofobia en el Parlamento de Cataluña](#)", publicat en castellà al diari *ABC* l'any 2020 (12/02/2020) i identificat amb el número 1.738 a la sèrie d'articles del corpus.

~~ara llegim~~

Un tast de català; Una col·laboració que pot donar molts fruits

~~albert pla nualart~~

~~390 words~~

~~2-January-2016~~

~~Diari Ara~~

~~ARA~~

~~AraPrimera~~

~~Catalan~~

~~Copyright 2016. Edicio de Premsa Periodica ARA S.L.~~

Com que m'agrada encetar l'any amb optimisme avui us vull parlar del potencial que veig en un espai virtual, l'Acadèmia Oberta (AO), que fa possible un diàleg enriquidor i productiu entre els responsables lingüístics del mitjans (i d'altres àmbits) i els acadèmics de l'IEC. L'AO va néixer aquest juliol amb el propòsit de documentar i debatre la distància que separa la norma dels mitjans (els seus llibres d'estil) de la més estrictament normativa (la de l'IEC). I de fer-ho des del respecte mutu i procurant que s'escurci.

Sovint he defensat que és bo que aquesta distància existeixi. La norma dels mèdia, en una llengua viva, ha d'anar un pas endavant respecte a l'acadèmica. Entre altres coses, perquè la primera és el millor camp de proves per actualitzar la segona. Però no és bo que la distància s'estanqui o creixi: desconcerta els usuaris i fa perdre prestigi a la llengua.

Per això convé tant un diàleg en què uns i altres revisem posicions per reforçar el consens normatiu. I el reforcem de dues maneres: perquè uns abandonem les transgressions més injustificades, i l'IEC acull en la seva norma les més fonamentades.

El dia de Nadal el coordinador de l'AO, Oriol Camps, ens va fer arribar dues cartes firmades per Joan Martí i Castell, president de la comissió de lexicografia de la Secció Filològica, en què ens feia saber que arran del debat mantingut a l'AO sobre la bondat del verb disseccionar i del substantiu subratllat totes dues paraules serien introduïdes al DIEC en una pròxima actualització del diccionari.

És una manera de normalitzar la llengua; de fer que el que ja era català sigui també normatiu, i que, amb el DIEC a la mà, "El forense pugui disseccionar [i no pas dissecar, com fins ara] un cadàver" i "En un text el subratllat [i no pas la subratllada] pugui ser de qui l'ha escrit".

Em direu que són dos mots que gairebé tothom feia servir ignorant que fossin incorrectes. Però és un primer pas: ja hi haurà temps d'abordar qüestions més complexes que sí que creen maldecaps. Esperem que l'AO ho faci aquest 2016. Bon any!

~~Edicio de Premsa Periodica Ara, S.L.~~

~~Document ARA0000020160105ec120008t~~

**Figura 3. Exemple de neteja d'un text (Ara)**

#### 4.1.3.3. Altres aspectes tècnics

Malgrat la gran dedicació que ha suposat la importació i la preparació de les dades, un cop les tenim en format de text (.TXT), no hem hagut de dur a terme altres operacions tècniques. En general, la qualitat de les dades del corpus és bona.

Les dades d'*El Periódico*, tanmateix, inclouen, de manera molt puntual, errors de codificació relacionats amb símbols ortogràfics com l'apòstrof (vegeu la Figura 4).



### La llei d'educació blindarà la immersió lingüística a Catalunya

La futura llei **deducació** que s'està cuinant al Congrés eliminarà la referència expressa que diu que el castellà ha de ser la llengua vehicular en **l'ensenyament**, fruit **dun** acord entre el PSOE, Unides Podem i ERC que busca blindar el sistema **dimmersió** lingüística a Catalunya. Fonts republicanes expliquen a EL PERIÓDICO que han pactat una esmena amb els partits que sostenen el Govern per canviar el redactat de la norma que, una vegada entri en vigor, obligarà els alumnes i les alumnes a conèixer perfectament tant el català com el castellà, una cosa que ja passa en **l'actualitat**.

Aquest acord suposa un nou gest **dacostament** de socialistes i podemistes amb ERC en vigílies de la primera votació dels Pressupostos del 2021. Els republicans encara no han confirmat si presentaran una esmena a la totalitat a aquests comptes, una cosa per a la qual tenen de termini fins divendres. Serà el dia abans, dijous, quan la iniciativa sobre la immersió lingüística es voti en la ponència del Congrés que està analitzant les esmenes a la llei educativa.

[...]

“La llei d'educació blindarà la immersió lingüística a Catalunya”,  
03/11/2020, *El Periódico de Catalunya*

Figura 4. Exemple d'un text amb errors de codificació (*El Periódico*)

Aquests errors esporàdics plantegen, naturalment, reptes per a la correcció automàtica. Per a la recerca actual no s'han pogut esmenar els quinze textos (8.405 paraules), de manera que els recomptes quantitativs d'*El Periódico* podrien presentar alguns errors.

## 4.2. Explotació estadística del corpus

El paquet *#LancsBox*, desenvolupat per la Lancaster University, inclou un programari de nova generació amb eines d'anàlisi de dades lingüístiques i de corpus, que permet explotar corpus existents i propis, visualitzar-ne els resultats i analitzar dades en qualsevol llengua.

El programa funciona amb tots els sistemes operatius (Windows, Mac, Linux) i etiqueta morfosintàcticament textos en una vintena de llengües, entre les quals el català, el castellà, l'italià o el suec. Entre d'altres, l'arquitectura informàtica d'aquesta eina es basteix a partir d'Apache Tika, Gluegen, Groovy, JOGL, minlog, QuestDB, RSyntaxTextArea, smallseg o TreeTagger.

Un cop recollides les dades i convertides en fitxer de text (.TXT) s'han importat, tractat i analitzat els arxius amb el programari *#LancsBox* (versió 6.0.0). Algunes de les funcions que més hem emprat han estat les taules de freqüències (absolutes i relatives) del vocabulari, l'anàlisi de coocurrències lèxiques i l'eina de detecció de concordances de les paraules clau en context (KWIC).

D'una banda, s'han analitzat separatament els dos arxius de text que inclouen els titulars dels articles en català (6.762 titulars) i en castellà (4.719 titulars). Aquesta primera anàlisi estadística permet identificar, gràcies als índexs estadístics (relatius i absoluts) d'ús del vocabulari, les paraules més emprades als titulars de premsa sobre qüestions lingüístiques. És, per tant, una primera aproximació als temes més importants quan la premsa tracta les qüestions lingüístiques.

De l'altra, s'han importat, en dues anàlisis separades els 286 articles seleccionats per a l'anàlisi qualitativa, també en funció de la llengua de publicació: en català n'hem processat 172 (104.999 paraules) i en castellà 114 (63.757 paraules). Les freqüències relatives i absolutes, i altres funcions de *#LancsBox* (com l'eina de detecció d'*n-grams*, de concordances i de coocurrències lèxiques) també ha permès d'analitzar-los qualitativament.

## 5. Anàlisi

### 5.1. Introducció

Com hem assenyalat a l'apartat 3, aquest treball s'emmarca metodològicament en l'hermenèutica de Thompson (1984), que posa en relació l'anàlisi del microtext amb l'anàlisi del context en què circulen els discursos. Un dels elements clau, per tant, en tota anàlisi del discurs és l'anàlisi de les circumstàncies contextuais (enteses com els esdeveniments polítics, socials i econòmics) en què es produeixen els discursos. Les línies que segueixen presenten, en conseqüència, aquesta anàlisi succinta a l'apartat 5.2. Posteriorment, als apartats 5.3 i 5.4, s'analitzen quantitativament i qualitativament el corpus de titulars i el d'articles de premsa, les macroestructures i les temàtiques principals a través de les quals s'emmarquen els discursos mediàtics sobre llengües.

### 5.2. Anàlisi dels pics estadístics

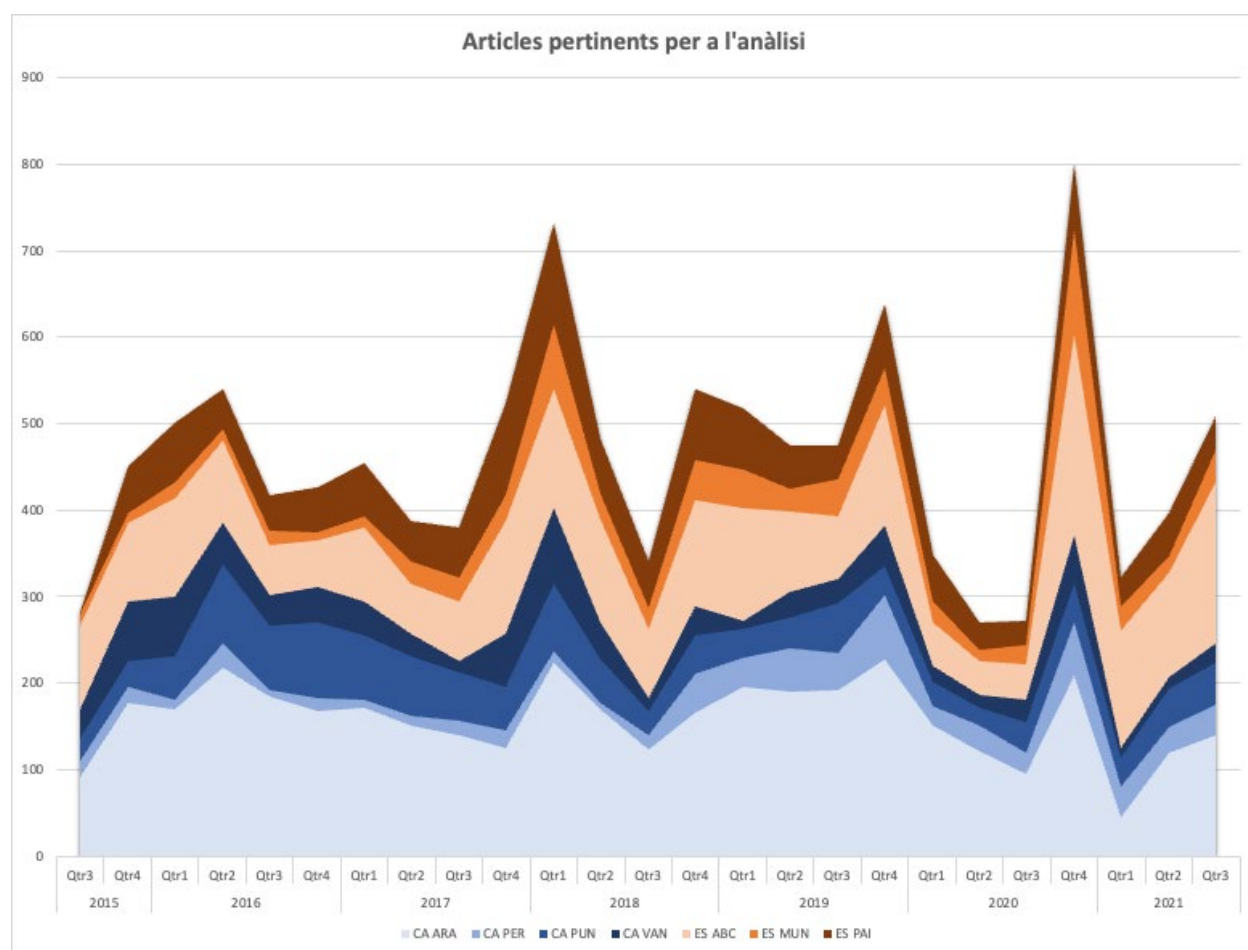
Aquesta recerca analitza els articles publicats entre setembre de 2015 i setembre de 2021. Aquesta forquilla forma part d'una franja temporal més extensa (l'anomenat *procés sobiranista* a Catalunya) que arrenca, entre d'altres, de les demandes de més autogovern a Catalunya. En bona part, aquestes demandes creixen per les aspiracions fallides en la reforma de l'Estatut d'autonomia de Catalunya (2006) i la sentència posterior STC 31/2010 amb què el Tribunal Constitucional es pronuncia contra preceptes relacionats amb aspectes clau de l'autogovern (com la llengua catalana o les tipologies competencials), els invalida i els declara inconstitucionals. Aquest fet, juntament amb un context de gran crisi econòmica, comporta un viratge social i polític d'una gran part de la població a favor de la independència a partir del 2012.

Els mitjans de comunicació, com a altaveus cognitius entre el poder social (propaganda, política, lleis) i el públic, s'han encarregat de difondre i legitimar el cos de creences d'ideologies confrontades en aquest procés (Gifreu, 2018, p. 53). En aquest context, la llengua ha ocupat un paper destacat, com ja s'ha indicat a l'apartat 3, i s'han normalitzat actituds lingüístiques en conflicte (Calafat, 2021).

Tant els mitjans de comunicació estatal(iste)s com els sobiranistes han emmarcat la realitat amb uns discursos avaluatius al voltant del *procés*, amb un vocabulari amb significació pragmàtica oposada. Aquest marc contextual ha donat com a resultat un enfrontament discursiu entre dues esferes mediàtiques amb capacitat perlocutiva (és a dir, que provoca un efecte en la societat) i una polarització ideològica que ha anat *in crescendo*.

Com hem esmentat anteriorment, el nombre d'articles trobats ha superat amb escreix les xifres inicialment previstes: 11.481 articles identificats que parlen sobre llengua i multilingüisme. Hem recollit, sistematitzat i tractat tot aquest material, tot i que l'anàlisi de corpus d'informació tan gran supera l'abast d'aquest estudi. Per això, hem identificat quatre moments específics del període analitzat en què la llengua ha tingut un ressò mediàtic més intens.

Aquests moments constitueixen els pics de la Gràfica 4 (la tornem a reproduir, tot i que ja apareix a l'apartat 4.1.1) per tal d'analitzar el context en què es produeixen els pics. En conclusió, analitzem la fase 1 de les tres fases analítiques que estableix Thompson (1984).



Gràfica 4. Distribució cronològica d'articles per diari (dades filtrades)

El **primer pic** correspon als mesos de març (finals), abril i maig de 2016. Aquest enfocament mediàtic sobre el tema de les llengües sorgeix arran de la publicació del *Manifest per un veritable procés de normalització lingüística a la Catalunya*

*independent* (també conegut com a *Manifest Koiné*), en què un grup de més de 200 intel·lectuals signaven un document sobre l'oficialitat única del català en una hipotètica república catalana futura i alertaven del procés de substitució lingüística en què es trobava el català (a favor del castellà). Hi reclamaven una normalització per al català que passava per una oficialitat única i plena del català (i l'aranès), i per un reconeixement del castellà, sense rang d'oficialitat.

Les reaccions van ser immediates, sobretot als diaris estatal(iste)s, amb un rebuig frontal i amb reaccions des de diversos àmbits (polítics, acadèmics i periodístics) tot destacant l'arraconament del castellà i construint un marc mental al voltant de les nocions d'*imposició*, *radicalització*, *discriminació*, *aniquilació* i *apartheid lingüístic*.

Com a reacció, de l'altra banda, trobem els defensors dels postulats del manifest, que alerten del mite de la societat bilingüe equilibrada, tot argumentant que en qualsevol societat oficialment bilingüe, la llengua amb més poder econòmic, polític i social acaba imposant-se a l'altra.

El **segon pic** correspon al quart trimestre de 2017 i a principis de 2018. És el segon pic més alt de tota la sèrie cronològica analitzada. Aquest pic cal emmarcar-lo en tres fets polítics, socials i amb repercussions econòmiques molt rellevants.

D'una banda, l'1 d'octubre de 2017 se celebrava el referèndum per consultar a la ciutadania si volia una Catalunya independent i que l'Estat va declarar il·legal. Tot seguit, l'Estat aplicava l'article 155 CE, el qual s'aplica quan una comunitat autònoma "no complia les obligacions que la Constitució o altres lleis li imposen, o actuava de forma que atemptés greument contra l'interès general d'Espanya". *De iure* i *de facto*, per tant, es destituïa el govern de la Generalitat i les institucions catalanes passaven a dependre del Govern de l'Estat. En conseqüència, el president de la Generalitat s'exiliava a Bèlgica, juntament amb diversos consellers, i els membres del Govern que romanien a Catalunya eren empresonats. El 21 de desembre de 2017 se celebraven noves eleccions.

En aquest context es produeix una activitat discursiva molt remarcable en què la llengua forma part central dels debats, sobretot dels mitjans estatal(ist)es, tot creant un marc que vincula aquest conflicte polític a un conflicte sociolingüístic. Cal remarcar les referències constants al model lingüístic de l'educació a Catalunya (tal com es constata a les diverses seccions de l'anàlisi de l'apartat 5.4) i se li atribueixen els valors de *conflicte*, *radicalització* i *separatisme*.

El **tercer pic** presenta, en línies generals, dues tendències oposades. D'una banda, les celebracions el novembre de 2019 de les reunions d'alt nivell de la *Asociación de Academias Españolas de la Lengua*, presidida pels reis d'Espanya. Com a resultat,

incrementava el nombre de referències a la llengua castellana, tot atribuint-li uns valors de *fortalesa, unió, cohesió, agermanament entre pobles i d'expansió numèrica*.

Aquests discursos són presents només en una part de la premsa escrita en castellà. En canvi, els mateixos diaris continuen situant la llengua catalana en un marc de conflictivització permanent en àmbits com l'escola, la universitat i la societat en general.

El **quart pic** s'emmarca exclusivament en l'àmbit educatiu, i concretament en l'aprovació de la nova llei educativa a finals de 2020, la Llei orgànica 3/2020, de 29 de desembre, per la qual es modifica la Llei orgànica 2/2006, de 3 de maig, d'Educació (anomenada també *LOMLOE* o *Llei Celáa*). La llei, a petició d'Esquerra Republicana de Catalunya i amb el suport de *Podemos*, no incorporava de manera explícita que el castellà ha de ser llengua vehicular en l'educació obligatòria a Catalunya. Això provocava un cataclisme mediàtic, fet que es manifesta en el pic més elevat de tota la sèrie cronològica.

Declaracions polítiques, manifestacions ciutadanes (la majoria de les quals, fora de territori català) articles d'opinió d'autors reeixits com Mario Vargas Llosa contraris al que anomenen, un cop més, l'aniquilació del castellà a Catalunya. A aquest fet cal afegir la sentència del Tribunal Superior de Justícia de Catalunya, del 16 de desembre de 2020, en què per primera vegada per a tot el sistema educatiu català, s'establia que el percentatge de castellà com a llengua vehicular no pot ser inferior al 25%.

Aquella decisió judicial trencava el sistema de conjunció lingüística dels darrers gairebé 40 anys. Les reaccions eren immediates per part de tots els sectors. D'una banda, a través d'un discurs emmarcat en termes bel·licistes contraris cap a la llengua catalana (*atacs, imposicions, blindatges*) i de l'altre, amb un marc de *seny, ordre, restauració i drets lingüístics*. La mateixa realitat, emmarcada de manera completament oposada.

Aquesta ciclogènesi política i jurídica sobre la llengua en l'àmbit educatiu ha estat una constant durant tot el període analitzat i ha marcat, de manera transversal i recurrent, els debats (altament polititzats i no basats en dades sociolingüístiques) als mitjans de comunicació escrits de l'Estat i de Catalunya.

### 5.3. Anàlisi del corpus de titulars

Hem constituït el primer subcorpus amb els titulars dels 11.481 articles pertinents. La seva distribució per llengua i les seves característiques lexicomètriques principals consten a la Taula 12.

Llengua	Titulars	Tokens	Types	Lemes
CA	6.762	72.538	13.061	15.062

<b>ES</b>	4.719	59.812	9.902	8.185
<b>TOTAL</b>	11.481	132.350	— <sup>8</sup>	—

Taula 12. Dades lexicomètriques del corpus de titulars (#LancsBox)

Com ja s'ha constatat (vegeu apartat 4.1.1), la premsa en català dedica un 30,2% més d'articles que la premsa en castellà a qüestions lingüístiques en el període analitzat. Ateses les diferències en les dimensions d'ambdós blocs, prenem com a referència la freqüència relativa d'ús del vocabulari. La Taula 13 presenta els 25 substantius més freqüents en cada llengua.

Premsa en CA			Premsa en ES		
Substantiu	Freq_Abs	Freq_Rel	Substantiu	Freq_Abs	Freq_Rel
<i>català</i>	832	114,7	<i>lengua</i>	433	72,4
<i>llengua</i>	417	57,5	<i>castellano</i>	409	68,4
<i>castellà</i>	213	29,4	<i>catalán</i>	408	68,2
<i>any</i>	174	24,0	<i>español</i>	377	63,0
<i>lleï</i>	144	19,9	<i>ley</i>	208	34,8
<i>immersió</i>	112	15,4	<i>educación</i>	167	27,9
<i>escola</i>	95	13,1	<i>idioma</i>	160	26,8
<i>llengua</i>	91	12,5	<i>año</i>	133	22,2
<i>govern</i>	73	10,1	<i>inglés</i>	108	18,1
<i>país</i>	64	8,8	<i>inmersión</i>	101	16,9
<i>cultura</i>	64	8,8	<i>escuela</i>	99	16,6
<i>premi</i>	62	8,5	<i>alumno</i>	98	16,4
<i>model</i>	58	8,0	<i>euskera</i>	91	15,2

<sup>8</sup> La suma de les paraules diferents (*types*) i dels lemes no és necessàriament la suma de les quantitats en català i en castellà, ja que pot haver-hi encavalcaments entre llengües.

<i>día</i>	57	<b>7,9</b>	<i>libro</i>	80	<b>13,4</b>
<i>alumne</i>	56	<b>7,7</b>	<i>aula</i>	79	<b>13,2</b>
<i>frase</i>	56	<b>7,7</b>	<i>país</i>	75	<b>12,5</b>
<i>món</i>	55	<b>7,6</b>	<i>colegio</i>	72	<b>12,0</b>
<i>llibre</i>	52	<b>7,2</b>	<i>autor</i>	72	<b>12,0</b>
<i>directora</i>	47	<b>6,5</b>	<i>uso</i>	72	<b>12,0</b>
<i>president</i>	44	<b>6,1</b>	<i>centro</i>	68	<b>11,4</b>
<i>campanya</i>	44	<b>6,1</b>	<i>valenciano</i>	66	<b>11,0</b>
<i>debat</i>	43	<b>5,9</b>	<i>profesor</i>	64	<b>10,7</b>
<i>tast</i>	42	<b>5,8</b>	<i>niño</i>	61	<b>10,2</b>
<i>literatura</i>	41	<b>5,7</b>	<i>comunidad</i>	58	<b>9,7</b>
<i>selectivitat</i>	41	<b>5,7</b>	<i>cultura</i>	58	<b>9,7</b>

Taula 13. Els 25 substantius més freqüents (corpus de titulars)

Com és natural, la franja alta de les freqüències<sup>9</sup> presenta un grau de coincidència elevat: en aquest cas, quatre de les cinc paraules més freqüents són les mateixes en ambdós llistats, tot i que es constaten diferències amb relació a llur freqüència d'aparició (vegeu la Taula 14).

	Freq_Rel (CA)	Freq_Rel (ES)
<i>català</i>	<u>114,7</u>	68,2
<i>llengua</i>	57,5	<u>72,4</u>
<i>castellà</i>	29,4	<u>68,4</u>
<i>lleï</i>	19,9	<u>34,8</u>

Taula 14. Comparació de freqüències relatives (#LancsBox)

<sup>9</sup> L'apartat 5.4.1 inclou una anàlisi de les concordançes d'ús, entre d'altres, d'unitats d'aquest top 5 (com *llengua* i l'adjectiu *lingüístic -a*).



Com és natural, la premsa catalana tendeix a incloure més titulars sobre notícies amb relació a la llengua catalana que no pas de castellà, i a la inversa passa igual. Per això, les freqüències relatives són significativament més altes en cada cas quan es refereixen a la llengua: a la premsa en català, el substantiu *català* (amb la resta d'unitats del paradigma lèxic) s'empra amb una freqüència relativa de 114,7 (versus 68,2 a la premsa en castellà); a la premsa en castellà, es fa referència al *castellà* amb una freqüència de 68,4 (versus 29,4 a la premsa en català). Constatem, per tant, que la premsa catalana dedica nombrosos titulars al català i que en parla amb una freqüència relativa que se situa molt per sobre que la de totes les paraules del llistat d'unitats més freqüents dels titulars en castellà.

Entre els cinc substantius més utilitzats a la premsa en llengua catalana hi apareix *any*. Un cop revisats qualitativament els resultats (vegeu la Figura 5), constatem que són molt recurrents les referències temporals per a presentar tendències, diferències i comparacions estadístiques interanuals (com l'article "El castellà a selectivitat creix per segon any seguit", *Ara Balears*, 02/07/2016).

	Node	
les audiències onze anys seguits Un altre	any	Tres proves de foc per al català
La tossuda deriva extremista de Bauzá Aquest	any	arribaran a Catalunya 1.900 'menes' més que
refugiats; Aquestes unitats s'havien d'haver convertit aquest	any	en serveis de suport...lingüístic 'Loreak' enlla els
dies que vindran' triomfa al FIC-CAT Aquest	any	tampoc: els alumnes de 4t d'ESO no
punts mínims en les competències numèriques Aquest	any	tampoc: els de 4t d'ESO no se'n
marca el curs «És difícil adaptar-se cada	any	a canvis» Societat Civil acusa Mas de
sector... L'APUNT: Llegiu enJoan-Lluís Lluís Com cada	any	el portal web Ésadir, coincidint amb la
nipons ja són considerats adults, i cada	any	els ajuntaments organitzen una... Piolin vs. Piuet:
vida quotidiana Vicent Sanchis: "Improviseu massa, cada	any	cal negociar de zero" TV3, un futur
Més de 30.000 estudiants es presenten cada	any	a Catalunya a les proves d'accés a
més inclòs"; 45.000 nouvinguts adults segueixen cada	any	els cursos del Consorci per a la
Fonsi és la punta de l'iceberg d'un	any	en què el reggaeton i els... Piratejat
de maig i 11 de juny d'aquest	any	fins a 3.108 persones es van... El
a president El 21 de març d'aquest	any	un ballari heterodox, Antonio el Tekila, guanyava...
places per intentar dignificar els estudis l...	Any	nou, català nou; TV-3 comença l'1 de
perd presència entre els màsters; Durant l'últim	any	han caigut un 60% les sol·licituds per
afavorir la seva... Recordo que al primer	any	de carrera, quan arribes a la Universitat
El llibre en català creix per segon	any	consecutiu La crisi s'instal·la al llibre infantil
El castellà a selectivitat creix per segon	any	seguit; Un 78,7% dels alumnes contestà en
en castellà; Augmenta aquest curs, per segon	any	consecutiu, el nombre... La literatura torna al
de la facturació, que creix per tercer	any...	Els estrangers de Barcelona també es queixen
pantalla del 3,2% El Consell ajorna un	any	la implantació de la llei de plurilingüisme
la política espanyola; La qüestió catalana Un	any	mogut Tornada a l'escola d'incògnit El gallec
MIR: El Govern vol que duri un	any	i s'implanti primer a secundària a partir
572 milions, 5 més que fa un	any	L'Institut Cervantes preveu que a mitjans de
de la civada Discriminar de franc Un	any	fora, el programa més demanat L'eurodiputat popular
que girar la truita és fàcil! Un	Any	per redibuixar Pompeu Fabra Per a que
el cine independent"; Sorprès Fa gairebé un	any	que es va estrenar la primera pel·lícula...
me gusta Joana" L'experiència La iniciativa: un	any	més. Que corri la llengua! Els centres
les mans que fan un llibre Un	any	sense solucions per als immigrants tancats a
els discursos de l'atenuu, amb l'ara Un	any	que comença ahir a Palma reconeixements: Les
sabers" a Can Oleo Manel Mari, un	any	després La Policia Local de Burjassot identifica
quan s'instal·lin nous centres de 'menes' Un	any	per redescobrir la literatura de Teresa Pàmies
Immersió a Letònia; històries del món Un	any	sense diacrítics; lletra petita Que el 155

Figura 5. Concordances del lema *any* (#LancsBox)

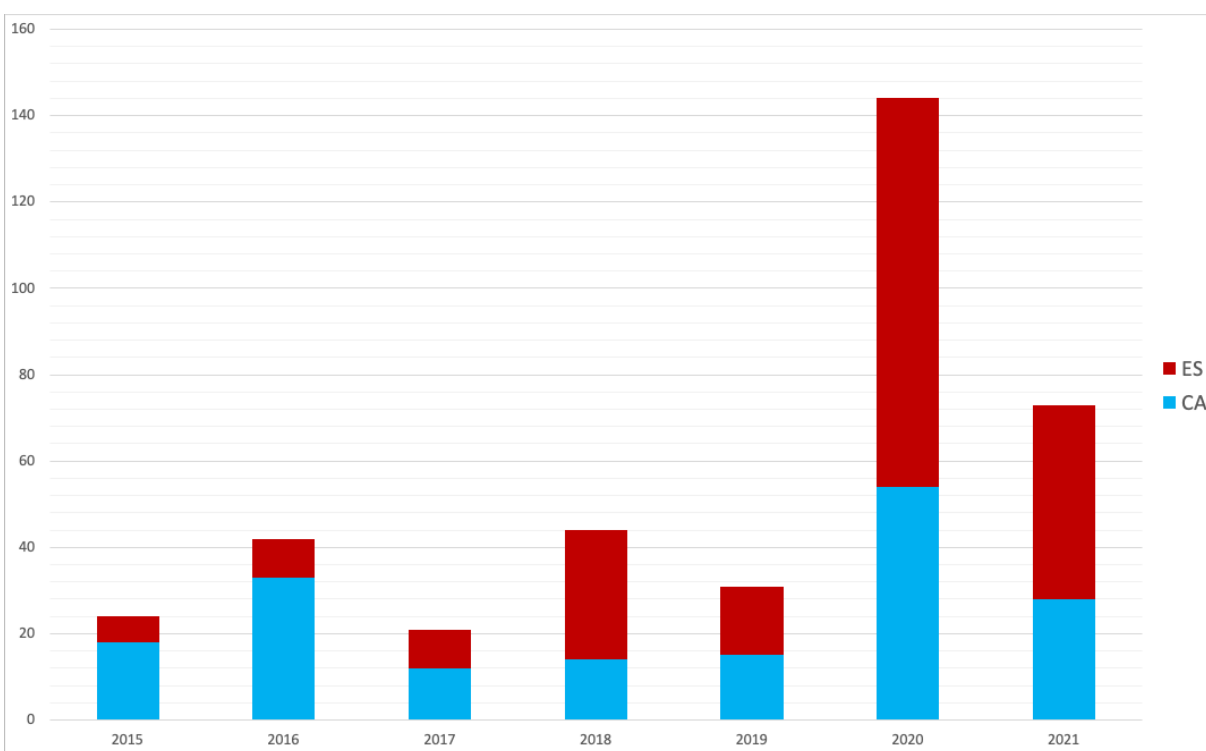
En ambdues llengües, els llistats de paraules més utilitzades contenen unitats relacionades amb el món educatiu: en català, hi identifiquem *immersió*, *escola*, *model [lingüístic]*, *alumne* o *selectivitat*; en castellà, s'hi identifiquen *educación*, *enseñanza*, *inmersión*, *escuela*, *alumno*, *centro [(público) bilingüe, de educación, de primaria o educativo, entre d'altres]* o *profesor*. Es constata, per tant, que l'ensenyament és un tema prioritari amb relació al tractament discursiu dels aspectes lingüístics a la premsa analitzada.

S'aborden també altres qüestions que hi estan relacionades, com l'àmbit cultural (*literatura, cultura, premi, llibre...*) o l'institucional i polític (*govern, país, directora, president...*) en el cas dels titulars en català, i en castellà amb relació a l'encaix lingüístic de les llengües (*idioma, lengua, catalán, castellano, español, euskera, uso, comunidad [valenciana]...*), paraules de l'àmbit jurídic i institucional (*ley, Estado, derecho, país*) i d'altres com *año, presidente, vez, parte, partido, mundo y ciudadano*.

### 5.3.1. Contextos d'ús de *lleí* i *ley*

Hem constatat que, en ambdues llengües, la paraula *lleí* apareix entre les cinc més freqüents del llistat. S'observa, en conseqüència, que tant la premsa en català com en castellà emergeix sovint l'aspecte legislatiu quan hi tracten aspectes lingüístics.

Els titulars s'hi refereixen, sobretot, els anys 2016 (la premsa catalana, especialment), el 2018, el 2020 i el 2021 (vegeu la Gràfica 5).



Gràfica 5. Distribució cronològica de les concordances de *lleí* i *ley*

Entre d'altres, el pic de l'any 2016 conté múltiples titulars relacionats amb la commemoració dels trenta anys de l'aprovació de la Llei 3/1986, de 29 d'abril, de normalització lingüística a les Illes Balears. L'especial "30 anys de la Llei" del diari *Ara Balears* del 29 d'abril de 2016 inclou articles com "Els fills de la normalització lingüística", "La Llei que va néixer fruit del consens polític" o "Un milenar de cors per festejar el català". Tal com afirma Joan Miralles (un dels tècnics lingüístics que va participar en la redacció

de la Llei), se'n commemora l'efemèride perquè "el català era una llengua que en aquells moments estava «molt malalta» i per aquest motiu "no es podien col·locar les dues llengües oficials al mateix nivell". Això suposa dotar el català de més presència a tots els àmbits de la societat".

L'any 2018 hi ha, sobretot, titulars de la premsa en castellà. Un repàs breu dels titulars d'aquell any permet identificar que la premsa madrilenya (en especial, l'ABC i *El Mundo*) esperona el govern de Rajoy a aprofitar l'aplicació de l'article 155 CE per a canviar el model lingüístic de l'escola catalana. Posteriorment, però, constata la impossibilitat de fer-ho atès el pronunciament del Tribunal Constitucional a la STC 1377/2014 que anul·lava la possibilitat de la disposició addicional 38 de la LOMCE de finançar amb 6.000 euros l'escolarització en castellà en centres privats, tal com s'il·lustra en els exemples següents:

*El Gobierno garantizará el castellano como vehicular en la escuela catalana (ABC, 17/02/2018)*

*"La ley educativa bien vale el 155", El Mundo, 16/02/2018*

*El Gobierno renuncia a aplicar el 155 para garantizar el castellano en la preinscripción escolar en Cataluña, (El Mundo, 04/04/2018)*

*El Gobierno temía que el TC tumbara su medida sobre el castellano en Cataluña (ABC, 04/04/2018)*

L'increment notable d'articles sobre aspectes relacionats amb la legislació lingüística de, sobretot, l'any 2020 (i, en menor mesura, de 2021), com ja s'ha assenyalat, correspon a l'aprovació i l'aplicació de l'anomenada *Llei Celaá* (o *LOMLOE*), com ja s'ha constatat amb relació al quart pic (vegeu l'apartat 5.2).

Pel que fa a les coocurrències de les paraules *lleï* i *ley* en els titulars d'ambdós blocs, constatem algunes diferències (vegeu les Figures 6 i 7, i la Taula 15).

## Coocurrències de *lleï*

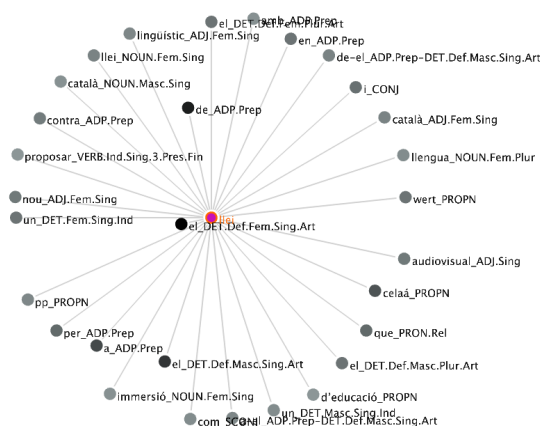


Figura 6. Coocurrents de *lleï* (#LancsBox)

## Coocurrències de *ley*



Figura 7. Coocurrents de *ley* (#LancsBox)

La taula següent presenta els coocurrents lèxics de les representacions anteriors (sense les paraules gramaticals ni les etiquetes *part-of-speech*):

	Premsa en CA	Premsa en ES
<b>Substantius</b>	<i>català, editorial, lingüístic, oficial, propi</i>	<i>constitucional, educativo, español, lingüístico, nuevo</i>
<b>Adjectius</b>	<i>any, campanya, castellà, català, catalunya, cultura, fabra, jurista, justícia, keep, llengua, apunt, país, plataforma, pp, ruta, signe, tribuna, unitat, valenciana</i>	<i>castellano, catalán, cataluña, celaá, educación, el, español, gobierno, la, lengua, ley, pp, uso, wert</i>
<b>Verbs</b>	<i>demanar, fer, parlar, tenir</i>	<i>aprobar, cumplir, pedir, ser</i>

Taula 15. Coocurrències lèxiques de *lleï* i *ley* (#LancsBox)

Es constata, per exemple, com alguns coocurrents de la premsa en català (com *oficial* o *propi*) fan al·lusió a la terminologia oficial de la legislació lingüística (*llengua pròpia*, *llengua oficial*). En castellà emergeixen referències a la llei educativa (*educativo*) i a l'adjectiu *constitucional* (hi ha nombroses referències al Tribunal Constitucional i, a més, s'empra en sintagmes com *defensa constitucional*, *proyecto constitucional* o *reforma constitucional*).

Els substantius amb què coocorren les paraules *lleï* i *ley* denoten també diferències en la manera d'abordar la qüestió lingüística. És palesa la importància del component jurídic

de la regulació lingüística en els titulars en català (alguns dels cocorrents són *jurista* o *justícia*)

És significatiu també identificar que, quan es parla de llei i de llengua, entre els verbs més usats en català apareixen *demanar* o *parlar* i, en castellà, *aprobar* o *cumplir*<sup>10</sup>. És especialment dramàtic el to d'alguns:

“¿Qué nos queda si no se cumple la ley? Una familia con ayudas de la 'Ley Wert':  
"Los niños son rehenes del sistema catalán” (El Mundo, 26/02/2018)

## 5.4. Anàlisi del corpus textual

L'exploració lexicomètrica del corpus parteix del llistat de freqüències (absolutes i relatives) obtingudes amb l'eina #LancsBox. L'apartat 4.1.2 inclou les característiques lexicomètriques generals del corpus (vegeu la Taula 11). Tot seguit, analitzem els 25 substantius, adjectius i sintagmes nominals més freqüents en els dos subcorpus<sup>11</sup>.

### 5.4.1. Els 25 substantius més freqüents

L'anàlisi lexicomètrica del corpus textual presenta resultats molt similars als obtinguts a l'anàlisi del corpus de titulars (vegeu l'apartat 5.3). Constatem, tal com s'observa a la Taula 16, que predomina el vocabulari relacionat amb el món educatiu (entre d'altres, *escola*, *model [lingüístic]*, *immersió [lingüística]*, *educación* o *alumno*). Ja s'ha constatat en nombrosos apartats d'aquesta recerca que és un tema prioritari en tractar la qüestió lingüística.

	Lemma	Freq_abs	Freq_rel	Lemma	Freq_abs	Freq_rel
1	<i>llengua</i>	575	59,2	<i>lengua</i>	367	58,1
2	<i>català</i>	518	53,3	<i>castellano</i>	204	32,3
3	<i>castellà</i>	245	25,2	<i>catalán</i>	202	32,0
4	<i>any</i>	182	18,7	<i>español</i>	162	25,6
5	<i>escola</i>	129	13,3	<i>año</i>	141	22,3
6	<i>manera</i>	93	9,6	<i>ley</i>	127	20,1

<sup>10</sup> Vegeu l'Annex 4 per a més informació.

<sup>11</sup> A l'Annex 3 també podeu trobar el llistat dels 25 noms propis més emprats a cada subcorpus.

	Lemma	Freq_abs	Freq_rel	Lemma	Freq_abs	Freq_rel
7	<i>model</i>	81	8,3	<i>comunidad</i>	108	17,1
8	<i>cas</i>	81	8,3	<i>educación</i>	103	16,3
9	<i>president</i>	77	7,9	<i>idioma</i>	99	15,7
10	<i>govern</i>	74	7,6	<i>estado</i>	97	15,4
11	<i>país</i>	74	7,6	<i>centro</i>	89	14,1
12	<i>cosa</i>	73	7,5	<i>enseñanza</i>	78	12,3
13	<i>lleí</i>	70	7,2	<i>presidente</i>	73	11,6
14	<i>fet</i>	65	6,7	<i>alumno</i>	73	11,6
15	<i>persona</i>	63	6,5	<i>escuela</i>	72	11,4
16	<i>món</i>	57	5,9	<i>vez</i>	66	10,4
17	<i>centre</i>	57	5,9	<i>profesor</i>	66	10,4
18	<i>part</i>	53	5,5	<i>parte</i>	63	10,0
19	<i>immersió</i>	53	5,5	<i>inmersión</i>	61	9,7
20	<i>gent</i>	52	5,4	<i>derecho</i>	61	9,7
21	<i>projecte</i>	51	5,2	<i>partido</i>	56	8,9
22	<i>alumne</i>	50	5,1	<i>euskera</i>	56	8,9
23	<i>día</i>	49	5,0	<i>país</i>	56	8,9
24	<i>temps</i>	48	53,3	<i>mundo</i>	56	8,9
25	<i>casa</i>	48	59,2	<i>ciudadano</i>	55	8,7

Taula 16. Lemes més freqüents en català i castellà (corpus textual)

Pel que fa al català, constatem que hi ha algunes diferències en la freqüència relativa d'algunes unitats entre el corpus de titulars i el corpus textual. Així, per exemple, *lleí, cultura, premi, frase, món, llibre, directora, president, campanya, debat* o *tast* s'empren amb una freqüència més alta al corpus de titulars.

És natural que la tria lèxica dels titulars –i, per tant, les posicions més altes del rànquing de paraules– siguin substantius més concrets i nodes conceptuals més importants. Aquesta prioritització de la tria lèxica dels titulars explica que, al corpus textual, constin entre les paraules més freqüents unitats més generals com *manera, cas, cosa, fet, persona, món, centre, part, gent, temps* o *casa*.

En el cas de la premsa en castellà, constatem que ambdós llistats presenten també un grau de coincidència molt elevat. Hi emergeix el lèxic relacionat amb les llengües de l'Estat (*castellano, catalán, valenciano, euskera...*) i, com ja s'ha assenyalat, paraules relatives a l'àmbit educatiu: *educación, inmersión, escuela, alumno, libro, aula, centro, profesor, niño...*

Resulta interessant indicar que *comunidad* és una paraula molt emprada en aquest bloc, no només perquè s'empra en el sintagma *Comunidad Valenciana*, sinó perquè la premsa en castellà acostuma a recordar que les competències en matèria d'ensenyament corresponent a les comunitats autònomes, com il·lustra el context següent:

*Como no podía ser de otra manera, teniendo en cuenta las hipotecas con las que subsiste el Gobierno socialista, la ley de Educación que prepara Isabel Celaá supone una concesión en toda regla a Podemos y los independentistas. Tal como consta en el documento de 17 páginas que la ministra de Educación ha hecho llegar al sector, la norma pone trabas a la expansión de la enseñanza concertada, rebaja el peso de la Religión y **devuelve a las comunidades la competencia en lenguas oficiales.** (El Mundo, 10/11/2018)*

Atès l'interès d'alguns substantius, hem analitzat qualitativament els usos de substantius com *llengua, monolingüisme, bilingüisme, trilingüisme, multilingüisme* i *plurilingüisme*. Per a fer-ho, hem cercat els contextos d'ús de cadascuna d'aquestes paraules, hem identificat el sintagma en què s'empren i els hem classificat i analitzat, tal com es presenta als dos apartats següents.

#### 5.4.1.1. Anàlisi dels contextos d'ús de *llengua* i *lengua*

Hem analitzat les concordances de *llengua* del corpus textual. En total, com s'observa a la Taula 17, l'anàlisi quantitativa de #LancsBox que hi ha 942 ocurrències de *llengua* (575) i *lengua* (367). Hem exportat els 942 contextos d'ús i els hem classificat.

Per a aquesta recerca, només ens hem fixat en els usos com a nom comú, motiu pel qual el total de la Taula 17 és 891 (i no pas 942): 558 corresponen als contextos de la premsa en català i 333 als de la premsa en castellà.

Sintagmes en què s'empra <i>llengua</i> (nom comú)	Premsa en CA	Premsa en ES	Total
~ ø [sense adjectiu]	131	83	214
~ catalana / ~ [catalana]	167	36	203
~ vehicular	31	40	71
~ oficial	24	30	54
~ pròpia	34	13	47
~ castellana / ~ espanyola / ~ [castellana] / ~ [espanyola]	16	21	37
~ cooficial	11	21	32
*possessiu + llengua	20	6	26
~ estrangera / ~ [estrangera]	11	14	25
~ comuna	7	10	17
~ materna	11	5	16
~ primera / primera ~	10	2	12
~ habitual	9	1	10
~ principal	5	-	5
~ universal	-	4	4
segona ~	2	2	4
~ mare	3	-	3
~ diferent	3	-	3
~ d'origen	3	-	3



Sintagmes en què s'empra <i>llengua</i> (nom comú)	Premsa en CA	Premsa en ES	Total
~ <i>basca</i>	-	2	2
~ <i>àrab</i>	-	2	2
~ <i>de Cervantes</i>	-	2	2
~ <i>d'escolarització</i>	1	1	2
~ <i>de l'Estat</i>	2	-	2
~ <i>minoritària</i>	2	-	2
~ <i>familiar</i>	-	2	2
~ <i>minoritzada</i>	2	-	2
~ <i>romànica</i>	1	1	2
~ <i>de l'aula</i>	2	-	2
<i>unitat de la ~</i>	2	-	2
~ <i>oral</i>	2	-	2
~ <i>de Catalunya</i>	1	1	2
~ <i>destacada</i>	-	2	2
~ <i>d'ús habitual</i>	2	-	2
~ <i>franca</i>	1	1	2
~ <i>de les noves generacions</i>	1	-	1
~ <i>paterna</i>	1	-	1
~ <i>nova</i>	1	-	1
~ <i>de signes</i>	-	1	1
~ <i>residual</i>	-	1	1

Sintagmes en què s'empra llengua (nom comú)	Premsa en CA	Premsa en ES	Total
<i>~ de signes catalana</i>	1	-	1
<i>~ de referència</i>	1	-	1
<i>~ del país</i>	-	1	1
<i>~ de l'emigració</i>	-	1	1
<i>~ del territori</i>	1	-	1
<i>~ de l'Imperi</i>	-	1	1
<i>~ dels altres</i>	1	-	1
<i>~ territorial</i>	-	1	1
<i>~ d'imposició</i>	-	1	1
<i>altra ~</i>	-	1	1
<i>~ d'oficialitat</i>	-	1	1
<i>~ normal</i>	-	1	1
<i>~ dinàmica</i>	1	-	1
<i>~ obligatòria</i>	-	1	1
<i>~ dominant</i>	1	-	1
<i>~ original</i>	1	-	1
<i>~ escombrada, menystinguda, substituïda, censurada, folkloritzada, castigada, prohibida</i>	1	-	1
<i>~ preeminent</i>	1	-	1
<i>~ escrita</i>	1	-	1
<i>~ proscriu</i>	1	-	1
<i>~ estàndard</i>	1	-	1

Sintagmes en què s'empra llengua (nom comú)	Premsa en CA	Premsa en ES	Total
~ de la Justícia	1	-	1
~ civil	1	-	1
~ de la televisió	1	-	1
~ exclusiva	-	1	1
~ vernacle	1	-	1
~ exògena	1	-	1
~ de Mercè Rodoreda	1	-	1
~ curricular	1	-	1
~ de Shakespeare	1	-	1
~ forta	1	-	1
~ nouvinguda	1	-	1
~ anglesa	1	-	1
~ d'estrangers	-	1	1
~ germana	1	-	1
~ òbvia	-	1	1
~ d'ensenyament	1	-	1
~ de l'escola	1	-	1
~ hebrea	-	1	1
~ parlada a Espanya	-	1	1
~ imperial	1	-	1
~ per defecte	-	1	1

Sintagmes en què s'empra llengua (nom comú)	Premsa en CA	Premsa en ES	Total
<i>~ important</i>	1	-	1
<i>~ d'estudi obligatori</i>	-	1	1
<i>~ imposada</i>	-	1	1
<i>~ de la immigració</i>	-	1	1
<i>~ inicial d'atenció</i>	1	-	1
<i>~ que ens uneix</i>	-	1	1
<i>~ institucional</i>	1	-	1
<i>~ respectable</i>	-	1	1
<i>~ local</i>	1	-	1
<i>~ russa</i>	1	-	1
<i>~ majoritària</i>	-	1	1
<i>~ única d'ús diari</i>	-	1	1
<i>~ autonòmica</i>	-	1	1
<i>~ valenciana</i>	-	1	1
<i>~ de cohesió i vehicular</i>	1	-	1
<i>~ veïna</i>	-	1	1
<i>~ de dominació</i>	-	1	1
<i>~ viva</i>	1	-	1
<i>~ de l'alumnat</i>	1	-	1
<i>~ de Lorca, Neruda o García Márquez</i>	-	1	1
<i>~ molt diferent</i>	-	1	1

Sintagmes en què s'empra <i>llengua</i> (nom comú)	Premsa en CA	Premsa en ES	Total
~ de Miguel de Cervantes	1	-	1
~ moribunda	1	-	1
~ de segona	1	-	1
~ morta	-	1	1
~ no marcada	1	-	1
<b>Total</b>	<b>558</b>	<b>333</b>	<b>891</b>

Taula 17. Sintagmes en què es documenta *llengua* (#LancsBox)

D'acord amb la Taula 17, constatem que la premsa en llengua catalana es refereix sobretot a la *llengua catalana* (és el sintagma més emprat: 167 contextos), així com a la *llengua* sense adjectiu (és a dir, en abstracte, de manera genèrica o com a el·lipsi d'un sintagma anterior) (131 contextos).

Entre els usos més freqüents identifiquem que també s'utilitza a sintagmes de la terminologia que regeix la política lingüística, com *llengua pròpia*, *llengua vehicular*, *llengua oficial* i *llengua cooficial*. Observem que en el llistat de freqüències d'aquests sintagmes a la premsa castellana es fan servir aproximadament els mateixos termes.

El llistat de freqüències també ens permet constatar que hi ha tota una sèrie d'unitats que només es documenten únicament en un dels dos blocs (vegeu la Taula 18):

Sintagmes documentats únicament a la premsa en CA	Sintagmes documentats únicament a la premsa en ES
~ principal, ~ mare, ~ d'origen, ~ diferent, segona ~, ~ de l'Estat, ~ oral, ~ minoritària, ~ d'ús habitual, ~ minoritzada, ~ de l'aula, unitat de la ~, ~ paterna, ~ moribunda, ~ de la Justícia, ~ dels altres, ~ civil, ~ de l'alumnat, ~ de signes catalana, ~ dinàmica, ~ de les noves generacions, ~ dominant, ~ nouvinguda, ~ escombrada, menystinguda, substituïda, censurada, folkloritzada, castigada, prohibida, ~ de referència, ~ escrita, ~ de segona, ~ estàndard	~ universal, ~ familiar, ~ àrab, ~ basca, ~ de Cervantes, ~ destacada, segona ~, ~ obligatòria, ~ de signes, ~ per defecte, ~ del país, ~ normal, ~ de dominació, ~ de la immigració, ~ d'estrangers, ~ exclusiva, ~ d'imposició, ~ de l'emigració, ~ de l'Imperi, ~ òbvia, ~ residual, ~ parlada a Espanya, ~ respectable, ~ de Lorca, Neruda o García Márquez, ~ territorial, ~ que ens uneix, ~ valenciana, ~ veïna, ~ imposada, ~ única d'ús diari, ~ majoritària, ~ d'estudi obligatori, ~ autonòmica, altra ~, ~ molt diferent, ~ d'oficialitat, ~ morta, ~ hebrea

Taula 18. Sintagmes de *llengua* documentats només en CA o ES

La lectura manual dels 891 contextos d'ús també ens ha permès identificar-hi de quines llengües parlen. La Taula 19 en presenta els resultats.

Llengua/Llengües a què fa referència <sup>12</sup>	Prensa en CA	Prensa en ES	Total
<b>CA</b>	300	60	<b>360</b>
<b>Referència abstracta/genèrica</b>	132	75	<b>207</b>
<b>ES</b>	46	122	<b>168</b>
<b>Altres llengües</b>	26	19	<b>45</b>
<b>Llengües cooficials / oficials de l'Estat</b>	25	12	<b>37</b>
<b>CA/ES</b>	13	18	<b>31</b>
<b>VA</b>	5	9	<b>14</b>
<b>EU</b>	1	9	<b>10</b>
<b>EU/ES</b>	1	2	<b>3</b>
<b>EN</b>	3	-	<b>3</b>
<b>OC</b>	-	2	<b>2</b>
<b>LSC</b>	1	1	<b>2</b>
<b>FR</b>	2	-	<b>2</b>
<b>VA/ES</b>	-	1	<b>1</b>
<b>FR/ES</b>	-	1	<b>1</b>
<b>CA/ES/EN</b>	1	-	<b>1</b>
<b>GA</b>	-	1	<b>1</b>
<b>EN/ES/CA</b>	1	-	<b>1</b>
<b>FR/CA</b>	1	-	<b>1</b>

<sup>12</sup> Per tal d'optimitzar les taules, hem emprat la llista de codis ISO 639-1 per a identificar les llengües.

Llengua/Llengües a què fa referència <sup>12</sup>	Prensa en CA	Prensa en ES	Total
CA/LSC/OC	-	1	1
<b>Total</b>	<b>558</b>	<b>333</b>	<b>891</b>

Taula 19. Llengües a què fan referència les concordances

Constatem que la premsa en llengua catalana fa referència, sobretot, a la llengua catalana quan tracta qüestions lingüístiques (300 dels 558 contextos analitzats) i, en segon lloc (132 contextos), a la llengua de manera abstracta o genèrica, sense especificar quina, seguit del castellà (46 concordances). La premsa en castellà presenta, com és previsible, una tendència a la inversa: hi predominen les referències a la llengua castellana (122 contextos), seguides dels usos de llengua en abstracte (o de forma genèrica) (75) i de les referències al català (60).

Les dades sí que ens permeten constatar, però, que amb relació al valencià o l'èuscar, la premsa en castellà inclou, en termes relatius, un percentatge de notícies més elevat que la de llengua catalana. Per tant, la premsa en castellà tendeix a diferenciar català i valencià, motiu pel qual aquesta segona variant hi està més representada (ja s'ha constatat a la Taula 19 que, en el corpus de premsa en castellà, es documenta entre les unitats més freqüents).

#### 5.4.1.2. Anàlisi dels contextos d'ús de *monolingüisme*, *bilingüisme*, *trilingüisme*, *multilingüisme* i *plurilingüisme*

Aquest apartat presenta l'anàlisi de les concordances de *monolingüisme*, *bilingüisme*, *trilingüisme*, *multilingüisme* i *plurilingüisme*. La Taula 20 presenta els contextos en què es documenten.

	Prensa en CA	Prensa en ES	Total
<b>bilingüisme</b>	<b>28</b>	<b>69</b>	<b>97</b>
bilingüisme (el)	10	25	35
centre/escola/institut (privat/públic) bilingüe	2	9	11
(ser) bilingüe	4	6	10
model/programa bilingüe	-	8	8
ensenyament bilingüe	1	4	5

	<b>Prensa en CA</b>	<b>Prensa en ES</b>	<b>Total</b>
secció bilingüe	-	3	3
societat bilingüe	1	2	3
realitat bilingüe	-	2	2
en [versió] bilingüe	2	-	2
bilingüització (la)	-	2	2
competència bilingüe	1	-	1
sistema escolar bilingüe	1	-	1
diari bilingüe	1	-	1
escriptor bilingüe	1	-	1
bilingüització forçada de la població	-	1	1
causa bilingüe	-	1	1
revista bilingüe	-	1	1
bilingüisme integral	-	1	1
sèrie bilingüe	-	1	1
país bilingüe	1	-	1
sistema espanyol bilingüe	1	-	1
ple domini bilingüe	1	-	1
Catalunya bilingüe	-	1	1
procés de bilingüització	1	-	1
matèria bilingüe	-	1	1
ideologia bilingüe	-	1	1
<b>monolingüisme</b>	<b>7</b>	<b>8</b>	<b>15</b>



	<b>Prensa en CA</b>	<b>Prensa en ES</b>	<b>Total</b>
monolingüisme (el)	3	-	3
model (educatiu/d'ensenyament) monolingüe	-	3	3
Catalunya monolingüe	-	2	2
(ser) monolingüe	2	-	2
sistema monolingüe actual	-	1	1
manifest monolingüe	-	1	1
confort monolingüe.	1	-	1
política lingüística monolingüe	1	-	1
règim monolingüe	-	1	1
<b>multilingüisme/plurilingüisme</b>	<b>25</b>	<b>6</b>	<b>31</b>
multilingüisme/plurilingüisme (el)	7	3	10
competència plurilingüe	3	-	3
ensenyament multilingüe/plurilingüe	2	-	2
societat multilingüe/plurilingüe	1	1	2
decret de/llei de plurilingüisme	2	-	2
model plurilingüe	2	-	2
escola multilingüe	1	-	1
condició plurilingüe	1	-	1
(ser) plurilingüístic -a	1	-	1
centre bilingüe i multilingüe	-	1	1
cos de professors plurilingüe	1	-	1
població plurilingüe	1	-	1

	Prensa en CA	Prensa en ES	Total
formació plurilingüe	1	-	1
quota multilingüe exòtica	-	1	1
mestre plurilingüe	1	-	1
realitat plurilingüe	1	-	1
<b>trilingüisme</b>	<b>3</b>	<b>7</b>	<b>10</b>
trilingüisme (el)	1	2	3
ensenyament trilingüe	1	1	2
decret de trilingüisme	-	1	1
sistema escolar trilingüe	1	-	1
trilingüisme peculiar	-	1	1
(ser) trilingüe	-	1	1
model educatiu trilingüe	-	1	1
<b>Total</b>	<b>63</b>	<b>90</b>	<b>153</b>

**Taula 20. Contextos d'ús de monolingüisme, bilingüisme, trilingüisme, multilingüisme i plurilingüisme**

Cal assenyalar, en primer lloc, que, malgrat ser un subcorpus més petit, la premsa en castellà inclou moltes més ocurrencies de *bilingüisme* (la més freqüent de les cinc nocions analitzades): acumula 69 ocurrencies d'aquesta paraula i la premsa en català en conté 28.

Es constata que apareix, sobretot, a contextos que tracten sobre l'ensenyament bilingüe, amb termes com *bilingüisme (el)*, *centre/escola/institut (privat/públic) bilingüe*, *(ser) bilingüe*, *model/programa bilingüe*, *ensenyament bilingüe* o *secció bilingüe*. La premsa en català, però, predomina en les referències al *multilingüisme/plurilingüisme*: 25 contextos i només apareix a 6 de la premsa en castellà.

A més, és interessant identificar quines llengües emergeixen quan s'hi apel·la. Quan revisem qualitativament els resultats i identifiquem de quines llengües tracten els

contextos d'ús, constatem (vegeu la Taula 21) que, sobretot, es refereix al bilingüisme català-castellà (54 contextos) i anglès-castellà (34 contextos).

	Prensa en CA	Prensa en ES	Total
<b>bilingüisme</b>	<b>28</b>	<b>69</b>	<b>97</b>
CA/ES	19	35	54
EN/ES	4	30	34
Referència abstracta/genèrica	2	2	4
CA/FR	2	-	2
ES/EU	-	2	2
ES/FR	1	-	1
<b>monolingüisme</b>	<b>7</b>	<b>8</b>	<b>15</b>
CA	2	8	10
FR	2	-	2
ES	2	-	2
Referència abstracta/genèrica	1	-	1
<b>multilingüisme/plurilingüisme</b>	<b>25</b>	<b>6</b>	<b>31</b>
Referència abstracta/genèrica	23	4	27
EN/ES	1	1	2
EN/ES/CA/FR	1	-	1
CA/ES	-	1	1
<b>trilingüisme</b>	<b>3</b>	<b>7</b>	<b>10</b>
CA/ES	-	3	3
Cooficials + EN/ES	2	-	2

	Premsa en CA	Premsa en ES	Total
EN/ES/CA	1	1	2
EN/FR/AR	-	2	2
ES/ES/GA	-	1	1
<b>Total</b>	<b>63</b>	<b>90</b>	<b>153</b>

**Taula 21. Llengües a què fan referència monolingüisme, bilingüisme, trilingüisme, multilingüisme i plurilingüisme**

Aquestes diferències s'expliquen perquè, quan la premsa en castellà apel·la als programes bilingües castellà-anglès, en general se'n remarquen els beneficis i els avantatges de gaudir d'aquesta oportunitat, tal com s'afirma als contextos següents:

*La Consejería de Educación incorporará 19 nuevas secciones bilingües en centros docentes de Castilla y León para el próximo curso, con lo que la red dispondrá de un total de 598 secciones bilingües, distribuidas en 489 colegios e institutos, que impartirán algunas asignaturas en un idioma extranjero. Además, la Junta está desarrollando una evaluación del modelo de bilingüismo con el objetivo de consensuar uno nuevo. El inglés ha sido el idioma elegido por todos los centros y, en cuanto a las etapas educativas, las 19 son de Secundaria. (ABC, 26/08/2016)*

*Los 1.148 alumnos procedentes de los 32 centros del programa bilingüe han obtenido 0,78 puntos más que la media de la Comunidad en la fase general de la PAU celebrada el pasado junio: 7,28 frente a 6,50. Esta diferencia se acentúa en la prueba específica de Inglés, donde los bilingües obtienen un 7,58 frente al 6,54 de la media regional, es decir, 1,04 puntos más. La mejora de resultados también se detecta en otras materias evaluadas en la PAU (Lengua, Historia de España, Historia de la Filosofía, Matemáticas, Biología, Ciencias de la Tierra y Medioambientales, Física y Química), según el informe de la Comunidad. (El País, 22/12/2016)*

L'altre context habitual a la premsa en castellà és el que fa referència al bilingüisme castellà-català. En aquests casos, a diferència del cas anterior, el bilingüisme ha estat, tal com s'expressa a l'article següent, un instrument del nacionalisme català de manipulació en l'ensenyament (segons l'*Asamblea por una Escuela Bilingüe*) i de marginació de la llengua castellana (per això, existeixen associacions que fan de bastions del bilingüisme com *Asamblea por una Escuela Bilingüe* o *Hablamos Español*, esmentades als articles següents):

*Para la Asamblea por una Escuela Bilingüe, entidad minoritaria pero de creciente actividad, "el nacionalismo lleva 30 años manipulando la educación". "Nunca ha habido consenso, sino sometimiento", zanja Ana Losada. (El País, 13/05/2018)*

*Cargó de razones y argumentos a la Cataluña bilingüe, a esa Cataluña real que lleva décadas batallando por "una enseñanza también en español", lema de la concentración convocada ya antes*

de que el pasado jueves el Tribunal Superior de Justicia de Cataluña rebatiera los falaces argumentos de la Generalitat para arrinconar al castellano. (El Mundo, 21/12/2020)

*Hablamos Español defiende que es imprescindible «la elección de lengua vehicular en la enseñanza», «el bilingüismo en la administración» y «no primar el uso de una lengua oficial a la competencia profesional». Apuesta, además, por «la apertura de la cultura a los creadores en ambas lenguas, y que los topónimos prohibidos en español vuelvan a ser oficiales». (ABC, 17/09/2018)*

#### 5.4.2. Els 25 adjectius més freqüents

L'anàlisi lexicomètrica de #LancsBox permet identificar automàticament els adjectius més emprats a cadascun dels dos subcorpus (vegeu la Taula 22). El programari lematitza totes les formes d'un paradigma (*català, catalana, catalans, catalanes*), les agrupa i en dona un llistat amb les freqüències de totes les formes flexionades sumades.

	Lemma	Freq_abs	Freq_rel		Lemma	Freq_abs	Freq_rel
1	<i>català</i>	305	31,4		<i>lingüístico</i>	175	29,2
2	<i>lingüístic</i>	205	21,1		<i>catalán</i>	169	28,2
3	<i>nou</i>	139	14,3		<i>nuevo</i>	157	26,2
4	<i>espanyol</i>	135	13,9		<i>español</i>	124	20,7
5	<i>primer</i>	125	12,9		<i>educativo</i>	80	13,4
6	<i>gran</i>	93	9,6		<i>solo</i>	66	11,0
7	<i>públic</i>	86	8,9		<i>bilingüe</i>	63	10,5
8	<i>polític</i>	84	8,6		<i>público</i>	55	9,2
9	<i>bo</i>	82	8,4		<i>político</i>	50	8,4
10	<i>propí</i>	80	8,2		<i>oficial</i>	49	8,2
11	<i>educatiu</i>	64	6,6		<i>vasco</i>	45	7,5
12	<i>social</i>	57	5,9		<i>abierto</i>	43	7,2
13	<i>oficial</i>	54	5,6		<i>baleár</i>	42	7,0
14	<i>cultural</i>	44	4,5		<i>valenciano</i>	41	6,8

	Lemma	Freq_abs	Freq_rel		Lemma	Freq_abs	Freq_rel
15	<i>important</i>	40	4,1		<i>junto</i>	40	6,7
16	<i>passat</i>	40	4,1		<i>escolar</i>	40	6,7
17	<i>general</i>	36	3,7		<i>último</i>	38	6,3
18	<i>segon</i>	36	3,7		<i>gran grande</i>	73	12,2
19	<i>millor</i>	35	3,6		<i>bueno</i>	35	5,8
20	<i>internacional</i>	35	3,6		<i>nuestro</i>	35	5,8
21	<i>principal</i>	33	3,4		<i>constitucional</i>	34	5,7
22	<i>vehicular</i>	27	2,8		<i>común</i>	33	5,5
23	<i>habitual</i>	27	2,8		<i> europeo</i>	31	5,2
24	<i>escolar</i>	24	2,5		<i>cultural</i>	29	4,8
25	<i>fàcil</i>	22	2,3		<i>propio</i>	27	4,5

Taula 22. Els 25 adjectius més freqüents en català i en castellà (#LancsBox)

Constatem que ambdós llistats comparteixen nombrosos adjectius, com *català, lingüístic, nou, espanyol, gran, públic, polític, bo, propi, educatiu, oficial* i *escolar* (en castellà *lingüístico, catalán, nuevo, español, educativo, público, político, oficial, escolar, gran|grande, bueno* i *propio*). Veiem que, com és previsible, els adjectius *català, lingüístic* i *espanyol* encapçalen els llistats.

A més, hi apareix *nou/nuevo*. En català, s'empra sovint al sintagma *nou model lingüístic* i en contextos que, sociolingüísticament, són importants per a la transmissió lingüística en català, com *noves generacions, nous parlants, nous aprenents* o *nous espectadors*. En castellà, s'empra sobretot en contextos que fan referència a canvis normatius, com *nueva ley educativa/de educación, nueva normativa* o *nueva reforma educativa*, sobretot amb relació a l'anomenada *llei Celáa*.

D'una banda, les unitats que apareixen a les primeres 25 posicions en català que no consten al llistat homòleg en castellà són: *primer, social, cultural, important, passat, general, segon, millor, internacional, principal, vehicular, habitual* i *fàcil*.

La premsa en català narra sovint el/la *primer/a X*, tal com es veu a la Figura 8, i altres referències al *passat* (adjectiu emprat, sobretot, amb *curs* i *any*):

Index	File	Left	Node	Right
1	CA_2016_AR	a l'acadèmica. Entre altres coses, perquè la	primera	és el millor camp de proves per
2	CA_2016_AR	ignorant que fossin incorrectes. Però és un	primer	pas: ja hi haurà temps d'abordar qüestions
3	CA_2016_AR	l'Estatut d'autonomia de 1983 va representar la	primera	porta des de l'any 1908 a Menorca
4	CA_2016_AR	llengües oficials. Tres anys més tard del	primer	text estatutari, l'any 1986, s'aprovà la Llei
5	CA_2016_AR	que he exposat, queda clar que el	primer	és restituir i deixar sense efecte les
6	CA_2016_AR	tots els públics El LADAT crea la	primera	sèrie d'animació sobre el mallorquí més universal
7	CA_2016_AR	viatger i tantes altres coses, tindrà la	primera	sèrie d'animació sobre la seva vida i
8	CA_2016_AR	Ramon Llull recordarà i ens contarà en	primera	persona sobre els aspectes més cabdals de
9	CA_2016_AR	Algèria i Tunísia. Així transcorren els sis	primers	capítols de la sèrie. En el darrer,
10	CA_2016_AR	institució que es va convertir en la	primera	universitat de l'estat espanyol que va incorporar
11	CA_2016_AR	la UIB al món empresarial. El seu	primer	projecte va ser educatiu, i va oferir
12	CA_2016_AR	al Parlament Europeu; Cinc eurodiputats porten per	primer	cop un informe sobre greuges a la
13	CA_2016_AR	coincideix amb la setmana en què, per	primera	vegada, els greuges al català es portaran
14	CA_2016_AR	va assegurar que no era la	primera	vegada que denunciava un pacient per parlar
15	CA_2016_AR	dos anys i mig, el matí del	primer	dia a Santiago de Xile. Era el
16	CA_2016_AR	una nova vida a Sud-amèrica. Durant els	primers	dies al Canadà, l'Olvia preguntava sobretot pel
17	CA_2016_AR	a Toronto per viure una aventura. "Les	primeres	setmanes ho va viure una mica angaiada
18	CA_2016_AR	ho va passar pitjor va ser el	primer.	Tenia sis anys quan van marxar a
19	CA_2016_AR	i ell començava la primària. Durant les	primeres	setmanes a l'escola ningú l'entenia, però després
20	CA_2016_AR	per tant, màxima predisposició a l'aprenentatge. Els	primers	dies a Santiago de Xile, el Leo
21	CA_2016_AR	l'aprenentatge de l'anglès va ser natural. Els	primers	dies a la llar d'infants no parlava,
22	CA_2016_AR	Berlin. "Els nens a casa durant el	primer	any parlaven català però ara ja es
23	CA_2016_AR	"Jo no ho sabia i a la	primera	fiesta vaig convidar-ne 15!", diu la Mar.
24	CA_2016_AR	en els països preferits pels catalans, els	primers	llocs els ocupen França, Alemanya, l'Argentina i
25	CA_2016_AR	Huffington Post es vanta d'haver estat el	primer	a filtrar els fóruns digitals. No volien
26	CA_2016_AR	cada vegada menys pròpia Situem-nos en la	primera	reunió d'una AMPA en una escola pública
27	CA_2016_AR	no va ser l'única sorpresa. Tristament, els	primers	sorpresos foren els presents a la Sala,
28	CA_2016_AR	d'aquest personatge es va conèixer des del	primer	segon. L'any 2011 havia concedit fins i
29	CA_2016_AR	Els punts de conflicte són diversos. En	primer	lloc –i el menys conflictiu perquè, segons
30	CA_2016_AR	ha certes consideracions que caldria avançar. La	primera	consideració és que, encara que hi estiguin
31	CA_2016_AR	gent ha decidit dues coses fonamentals. La	primera,	que vol ser mestressa del seu propi
32	CA_2016_AR	quan el gènere espanyol va instaurar per	primer	cop aquests exàmens, als quals s'han d'enfrontar
33	CA_2016_AR	l'Estat". Desembre del 2014 Se celebra la	primera	prova d'accés, només en castellà i a
34	CA_2016_AR	de parla catalana, i aquest és el	primer	reconeixement que ens han donat. No és
35	CA_2016_AR	paraula impresa. Com hem pogut comprovar en	primera	persona, aquest que deixam enrere ha estat
36	CA_2016_LV	El viatge del Monarca a Lisboa, el	primer	des de les eleccions del 20-D, va
37	CA_2016_LV	el protocol, ja que, segons ell, en	primer	lloc s'han de contestar les preguntes en
38	CA_2016_PU	de Franco amb dos grans èxits. El	primer,	aconseguir que amb molt poc recorregut aquest
39	CA_2016_PU	d'ESO han estat avaluats aquest curs, per	primer	cop, en les competències científicotecnològiques, un àmbit
40	CA_2016_PU	dels resultats, ja que es tracta del	primer	test. El departament està estudiant traslladar aquesta

Figura 8. Contextos d'ús de *primer -a* (#LancsBox)

Pel que fa a l'adjectiu *social*, s'empra especialment en els sintagmes *cohesió social* i *xarxes socials*, àmbits clau, respectivament, per a la supervivència i la projecció de futur de la llengua catalana.

A més, com ja hem assenyalat anteriorment, entre els adjectius més usats consten termes de la sociolingüística (*[llengua] vehicular*, *[llengua] habitual*, *[llengua] principal*) i de l'organització administrativa (*[direcció] general*, *[secretaria] general*).

De l'altra, els adjectius que consten al top 25 en castellà que no consten entre les 25 més freqüents en català són: *solo*, *bilingüe*, *vasco*, *abierto*, *balear*, *valenciano*, *junto*, *último*, *nuestro*, *constitucional*, *común*,  *europeo* i *cultural*. Es constata l'ús elevat d'adjectius que fan referència a zones d'Espanya amb llengües cooficials (*vasco*, *balear*, *valenciano*) i els contextos que, com ja s'ha assenyalat prèviament, esmenten els programes bilingües anglès-castellà (*enseñanza bilingüe*, *programa bilingüe*, *instituto bilingüe*, *centro público bilingüe* i *sección bilingüe*).

En castellà, l'adjectiu *constitucional* apareix amb una freqüència d'ús elevada. Com s'ha assenyalat, s'empra en 34 contextos (vegeu la Figura 9) i, per tant, té una freqüència relativa de 5,7 (a la premsa en català només s'ha documentat a 4 contextos, amb una freqüència relativa de 0,41).

Index	File	Left	▲ Node	Right
2	ES_2016_MU	que ponen trabas, por ejemplo, al derecho	constitucional	de poder escolarizar en castellano. Y las
3	ES_2016_PA	Un proyecto	constitucional	castalán excluye la oficialidad del castellano Una
6	ES_2018_ABI	generacional, pero les une el mismo espíritu	constitucional	y la misma condena ante los lazos
11	ES_2018_ABI	los socialistas, con las reglas del juego	constitucional	y la obscena desobediencia a las resoluciones
31	ES_2018_PA	siglos en Cataluña y que "el régimen	constitucional	de 1978 ha reafirmado la continuidad de
5	ES_2018_ABI	Lomce. Además, varias sentencias firmes del Tribunal	Constitucional	(TC) y del Tribunal Supremo (TS) reconocen
8	ES_2018_ABI	y una cierta capacidad crítica, el Tribunal	Constitucional	(TC) admite una mayor libertad del profesor
10	ES_2018_ABI	senda abierta por la sentencia del Tribunal	Constitucional	sobre el Estatuto de Cataluña de 2006
12	ES_2018_MU	fuertes La anulación por parte del Tribunal	Constitucional	del procedimiento para escolarizar en Cataluña a
13	ES_2018_MU	elementos de la sentencia muy relevantes. El	Constitucional	avala la nueva distribución competencial prevista en
14	ES_2018_MU	Carlos Vidal Prado es catedrático de Derecho	Constitucional	de la UNED.
17	ES_2018_MU	enseñanza de la lengua cooficial" El Tribunal	Constitucional	ha desestimado un conflicto de competencias planteado
19	ES_2018_MU	Estado "imponga aspectos de estricta ejecución". El	Constitucional	no le da la razón en ninguna
20	ES_2018_MU	norma acorde con la sentencia del Tribunal	Constitucional	que anuló el mecanismo para garantizar en
22	ES_2018_PA	en la lengua que elijan El Tribunal	Constitucional	ha avalado el precepto del Código de
23	ES_2018_PA	Los diputado recurrieron otros artículos que el	Constitucional	no aborda al no haber pormenorizado en
24	ES_2018_PA	la normativa lingüística aplicable. No obstante, el	Constitucional	recuerda que ese reconocimiento no puede suponer
25	ES_2018_PA	pueden recibirlos en castellano. Por ello, el	Constitucional	no ve inconstitucionales ninguno de los artículos
26	ES_2018_PA	de materias. Una sentencia de 2010 del	Constitucional	sobre el Estatut estableció que el castellano
27	ES_2018_PA	la Generalitat. Por ejemplo la sentencia del	Constitucional	sobre el artículo 35 del Estatut: "Nada
30	ES_2018_PA	se arrastra desde la sentencia del Tribunal	Constitucional	de 1994, ha llevado a 40 familias
32	ES_2020_ABI	judicial. Las tres formaciones acudirán al Tribunal	Constitucional	en cuanto el proyecto de ley esté
18	ES_2018_MU	no puede sostenerse, conforme a la doctrina	constitucional,	como es que la comunidad autónoma es
29	ES_2018_PA	2.300. La ley acabó en el Tribunal	Constitucional,	que en 1994 "avaló lo hecho hasta
4	ES_2016_PA	como se prevé ahora en el proyecto	constitucional.	Los promotores admiten que el documento está
9	ES_2018_ABI	educativo en esta comunidad dentro del marco	constitucional.	Según la propuesta, con el actual reparto
1	ES_2016_ABI	desacato reiterado a las sentencias del Tribunal	Constitucional.	Tantas han sido las denuncias como nulos
28	ES_2018_PA	del Tribunal Superior de Cataluña y del	Constitucional.	Estas obligan a que el castellano sea
33	ES_2020_ABI	recurrirá la ley educativa ante el Tribunal	Constitucional.	«En España tenemos una gran materia prima,
21	ES_2018_MU	la observancia de los principios y normas	constitucionales	aplicables y demás normas básicas que desarrollan

Figura 9. Contextos d'ús de *constitucional* (#LancsBox)

La majoria de concordances correspon a l'any 2018, sobretot en el sintagma *Tribunal Constitucional*. Això demostra que la judicialització de l'afer català, sobretot arran dels fets d'octubre de 2017, també emergeix discursivament quan es parla de la qüestió lingüística.

És rellevant, a més, veure de quina manera diaris com *El Mundo* apel·len als drets lingüístics de la ciutadania amb relació a l'estatus del castellà a Catalunya:

*Un proyecto constitucional catalán excluye la oficialidad del castellano (El Mundo, 14/05/2016)*

*Y cabría tomárselo a broma si no fuera porque la realidad es que las autoridades de Cataluña hace mucho tiempo que ponen trabas, por ejemplo, al derecho constitucional de poder escolarizar en castellano. Y las políticas de la Generalitat para imponer el catalán como lengua exclusiva en ámbitos oficiales ya denigran el idioma que comparten con todos los españoles. (El Mundo, 02/04/2016)*

#### 5.4.2.1. Anàlisi dels contextos d'ús de *lingüístic -a*

Evidentment, un dels adjectius clau és *lingüístic -a*. Per aquest motiu, n'hem revisat les concordances a la premsa en català i en castellà (vegeu la Taula 23).

Sintagmes amb <i>lingüístic -a</i> (i adjectius derivats)	Premsa en CA	Premsa en ES	Total
<i>immersió lingüística</i>	42	30	72
<b>*Noms propis institucionals (CPNL, DGPL, SGPL...)</b>	21	6	27
<i>model lingüístic</i>	15	6	21
<i>política lingüística</i>	10	10	20



Sintagmes amb <i>lingüístic</i> -a (i adjectius derivats)	Premsa en CA	Premsa en ES	Total
<i>projecte lingüístic</i>	10	10	20
<i>drets lingüístics</i>	9	2	11
<i>normalització lingüística</i>	7	2	9
<i>ús lingüístic</i>	6	3	9
<i>parella lingüística</i>	7	-	7
<b>*Noms propis legislació (LNL, LPL)</b>	<b>6</b>	<b>1</b>	<b>7</b>
<b>*Noms - Altres noms propis</b>	<b>6</b>	<b>-</b>	<b>6</b>
<i>realitat lingüística/sociolingüística</i>	5	1	6
<i>àmbit lingüístic</i>	1	4	5
<i>diversitat lingüística</i>	5	-	5
<i>riquesa lingüística</i>	3	1	4
<i>emergència lingüística</i>	4	-	4
<i>matèria no lingüística</i>	2	2	4
<i>casella lingüística</i>	-	4	4
<i>pluralitat lingüística</i>	4	-	4
<i>situació sociolingüística</i>	3	-	3
<i>frontera lingüística</i>	2	1	3
<i>tema lingüístic</i>	1	2	3
<i>requisit lingüístic</i>	3	-	3
<i>competència lingüística</i>	2	1	3
<i>domini lingüístic</i>	3	-	3
<i>comunitat lingüística</i>	3	-	3

<b>Sintagmes amb <i>lingüístic</i> -a (i adjectius derivats)</b>	<b>Premsa en CA</b>	<b>Premsa en ES</b>	<b>Total</b>
<i>discriminació lingüística</i>	1	2	3
<i>inversió lingüística</i>	-	2	2
<i>tertúlia lingüística</i>	2	-	2
<i>opció lingüística</i>	2	-	2
<i>disponibilitat lingüística</i>	-	2	2
<i>patrimoni lingüístic</i>	1	1	2
<i>qüestió lingüística</i>	-	2	2
<i>serveis lingüístics</i>	1	1	2
<i>raó lingüística</i>	2	-	2
<i>conjunció lingüística</i>	1	1	2
<i>regulació lingüística</i>	1	1	2
<i>títolació lingüística</i>	1	1	2
<i>legislació lingüística/en matèria lingüística</i>	1	1	2
<i>imposició lingüística</i>	-	2	2
<i>responsabilitat lingüística/responsable lingüístic</i>	2	-	2
<i>política monolingüística</i>	-	2	2
<i>identitat lingüística (i cultural)</i>	2	-	2
<i>normativa lingüística</i>	1	1	2
<i>modalitat lingüística</i>	-	2	2
<i>dimensió lingüística</i>	1	-	1
<i>vents lingüístics</i>	1	-	1
<i>submersió lingüística</i>	1	-	1

<b>Sintagmes amb <i>lingüístic -a</i> (i adjectius derivats)</b>	<b>Premsa en CA</b>	<b>Premsa en ES</b>	<b>Total</b>
<i>exclusió lingüística</i>	-	1	1
<i>proposta lingüística</i>	-	1	1
<i>exigència en matèria lingüística</i>	1	-	1
<i>sensibilització lingüística</i>	1	-	1
<i>fet lingüístic</i>	1	-	1
<i>dret d'opció lingüística</i>	1	-	1
<i>formació lingüística</i>	1	-	1
<i>procés de substitució lingüística</i>	1	-	1
<i>capacitació lingüística</i>	1	-	1
<i>recursos estilístics i lingüístics</i>	1	-	1
<i>gestió lingüística</i>	1	-	1
<i>ressorgir culturalment i lingüísticament</i>	1	-	1
<i>hàbit lingüístic</i>	1	-	1
<i>sistema lingüístic</i>	-	1	1
<i>homogeneïtat lingüística</i>	-	1	1
<i>terreny lingüístic</i>	1	-	1
<i>aposta lingüística</i>	-	1	1
<i>uniformitat lingüística</i>	1	-	1
<i>ideologia lingüística</i>	1	-	1
<i>enquesta d'usos lingüístics</i>	1	-	1
<i>idioma lengüicida</i>	-	1	1
<i>decret lingüístic</i>	1	-	1

<b>Sintagmes amb <i>lingüístic</i> -a (i adjectius derivats)</b>	<b>Premsa en CA</b>	<b>Premsa en ES</b>	<b>Total</b>
<i>catàleg de drets i deures lingüístics</i>	1	-	1
<i>díctadura lingüística</i>	-	1	1
<i>imperialisme lingüístic</i>	-	1	1
<i>reglament lingüístic</i>	1	-	1
<i>colonització lingüística</i>	-	1	1
<i>barbaritat lingüística</i>	-	1	1
<i>influència lingüística</i>	-	1	1
<i>segregació lingüística</i>	-	1	1
<i>influir lingüísticament</i>	1	-	1
<i>servei de normalització lingüística</i>	1	-	1
<i>inseguretat lingüística</i>	1	-	1
<i>societats lingüísticament heterogènies</i>	1	-	1
<i>integrisme lingüístic</i>	-	1	1
<i>supremacisme lingüístic</i>	-	1	1
<i>atenció lingüística</i>	1	-	1
<i>territori lingüístic</i>	1	-	1
<i>lamentació lingüística</i>	1	-	1
<i>treball lingüístic</i>	1	-	1
<i>composició lingüística</i>	-	1	1
<i>caixa de pandora lingüística</i>	-	1	1
<i>litigi polític lingüístic</i>	-	1	1
<i>eminència lingüística</i>	-	1	1

Sintagmes amb <i>lingüístic</i> -a (i adjectius derivats)	Premsa en CA	Premsa en ES	Total
<i>llibertat lingüística</i>	-	1	1
<i>espai lingüístic</i>	1	-	1
<i>llibre lingüístic</i>	1	-	1
<i>professora de lingüística</i>	-	1	1
<i>mal lingüístic</i>	1	-	1
<i>projecte nacional, lingüístic i cultural alternatiu a l'espanyol</i>	1	-	1
<i>autoestima lingüística</i>	1	-	1
<i>qualitat lingüística</i>	1	-	1
<i>Matrix lingüístic</i>	1	-	1
<i>diferència lingüística</i>	1	-	1
<i>actitud lingüística</i>	1	-	1
<i>règim lingüístic</i>	1	-	1
<i>model de coexistència lingüística</i>	1	-	1
<i>reglament municipal d'usos lingüístics</i>	1	-	1
<i>confiança lingüística</i>	1	-	1
<i>rellevància sociolingüística</i>	1	-	1
<i>moviment lingüístic</i>	1	-	1
<i>bipolaritat lingüística</i>	1	-	1
<i>conflicte lingüístic</i>	1	-	1
<i>bones notícies en matèria lingüística</i>	1	-	1
<i>autoritat lingüística</i>	1	-	1
<i>sensibilitat lingüística</i>	1	-	1

Sintagmes amb <i>lingüístic -a</i> (i adjectius derivats)	Premsa en CA	Premsa en ES	Total
<i>consens lingüístic</i>	1	-	1
<i>servei d'assessorament lingüístic</i>	1	-	1
<i>nou marc de convivència lingüística</i>	1	-	1
<i>divulgador lingüístic</i>	1	-	1
<i>objectiu lingüístic</i>	1	-	1
<i>docent lingüísticament capacitat</i>	1	-	1
<i>oferta lingüística</i>	1	-	1
<i>sociolingüística</i>	1	-	1
<i>consensos de 1983 i 1986 en matèria lingüística</i>	1	-	1
<i>substitució lingüística</i>	1	-	1
<i>oposició lingüística</i>	1	-	1
<i>document lingüístic</i>	1	-	1
<i>paisatge lingüístic</i>	-	1	1
<i>territori del domini lingüístic</i>	1	-	1
<i>conseqüència lingüística</i>	1	-	1
<i>búnker nacional-sociolingüista</i>	-	1	1
<i>context sociolingüístic</i>	1	-	1
<i>transmissió lingüística</i>	1	-	1
<i>perfil lingüístic</i>	-	1	1
<i>tria lingüística</i>	1	-	1
<i>persecució política, lingüística i cultural</i>	1	-	1
<i>unitat lingüística</i>	1	-	1

Sintagmes amb <i>lingüístic -a</i> (i adjectius derivats)	Premsa en CA	Premsa en ES	Total
<i>control lingüístic</i>	-	1	1
<i>varietat lingüística</i>	1	-	1
<i>polèmica lingüística</i>	-	1	1
<i>vincles culturals i lingüístics</i>	-	1	1
<i>política lingüicida i colonitzadora</i>	1	-	1
<b>*Noms - Declaració Universal dels Drets Lingüístics</b>	1	-	1
<i>convivència lingüística</i>	-	1	1
<i>esdeveniment pedagògic i lingüístic</i>	1	-	1
<i>debat lingüístic</i>	-	1	1
<i>estudi lingüístic</i>	1	-	1
<i>polítiques de l'Estat en matèria lingüística i educativa</i>	1	-	1
<b>Total</b>	<b>272<sup>13</sup></b>	<b>135</b>	<b>407</b>

Taula 23. Sintagmes que inclouen *lingüístic -a* (i adjectius derivats) (#LancsBox)

Constatem que entre les unitats més freqüents consten usos relatius, de nou, a la terminologia per a tractar qüestions lingüístiques amb relació a l'ensenyament i la sociolingüística (*immersió lingüística, model lingüístic, projecte lingüístic, ús lingüístic, domini lingüístic, comunitat lingüística, realitat lingüística/sociolingüística...*), a les campanyes, polítiques i actuacions institucionals de promoció de la llengua catalana (*política lingüística, drets lingüístics, requisit lingüístic, normalització lingüística, parella lingüística...*).

A més, és remarcable identificar a la franja alta de freqüències referències a conceptes connotats positivament, com *diversitat lingüística, pluralitat lingüística* o *riquesa lingüística*. Cal destacar també que la premsa en català inclou referències al sintagma

<sup>13</sup> La diferència amb la xifra de la Taula 22 (205) rau en el fet que aquest estudi de concordances inclou noms propis i variants que inclouen *lingüístic -a* (com *sociolingüístic -a* o *monolingüístic -a*).

*emergència lingüística*<sup>14</sup>, un concepte d'aparició relativament recent a informes de Plataforma per la Llengua que ha fet fortuna, tal com il·lustra l'exemple següent:

En aquest context, Òscar Escuder, president de Plataforma per la Llengua, ha alertat que sestà davant una emergència lingüística i ha demanat que el català sigui llengua oficial d'Estat.

[...]

L'InformeCAT recorda que, segons la Unesco, una llengua presenta una situació de vulnerabilitat i està en perill quan el nombre de parlants es converteix en una comunitat petita, ja que el risc que es perdi la llengua i la cultura és alt en benefici d'una altra que pugui absorbir els parlants. Per això, una de les preocupacions és el baix ús social de la llengua a les grans ciutats dels territoris de parla catalana.

(*El Periódico*, 17/06/2020)

Alguns dels sintagmes que només trobem a la premsa en català són, entre d'altres, *sensibilització lingüística*, *dret d'opció lingüística*, *catàleg de drets i deures lingüístics*, *inseguretat lingüística*, *autoestima lingüística*, *qualitat lingüística*, *Matrix lingüístic*, *confiança lingüística*, *sensibilitat lingüística* o *consens lingüístic*.

Pel que fa als sintagmes més emprats a la premsa en castellà, es constata que unitats com *immersió lingüística*, *política lingüística*, *projecte lingüístic* i *model lingüístic* encapçalen els llistats de freqüències en castellà. És significatiu, a més, que en castellà una de les unitats més freqüents és *casilla lingüística*:

File	Left	Node	Right
ES_2018_ABI	públicamente sobre la posibilidad de implantar la	casilla,	algo que a la vez puso en
ES_2018_PAI	ha pedido al ministerio que incluya esta	casilla	como arranque para garantizar que el castellano
ES_2018_PAI	al calor de la polémica lingüística. "La	casilla	debe existir pero no basta, no se
ES_2018_ABI	entendí como la posibilidad de implantar la	casilla.	El Gobierno matizó posteriormente esta idea, al
ES_2018_ABI	implantó. Con la negativa a implantar la	casilla	este curso, va a resultar imposible, de
ES_2018_PAI	realidad. Una posibilidad sería incluir la llamada	casilla	lingüística en los folletos de preinscripción escolar.
ES_2018_ABI	Parlament El Gobierno no recurrirá a la	casilla	lingüística para reforzar la presencia del castellano
ES_2018_ABI	del Supremo-que le obligan a restablecer la	casilla	lingüística. Los tribunales han dado varios ultimátum
ES_2018_ABI	Sin	casilla	lingüística.El Gobierno no aplica el 155 para
ES_2018_PAI	incluir en las hojas de inscripción una	casilla	que permita elegir a los padres una
ES_2018_PAI	en este sentido. La polémica sobre la	casilla	se encendió a mediados de febrero tras
ES_2018_ABI	hecho público, hay rastro de la citada	casilla.	Una opción que debería permitir a los

**Figura 10. Contextos d'ús de casilla (#LancsBox)**

La *casilla lingüística* és una opció que sorgeix com a estratègia per a intentar fer que les famílies poguessin triar la seva llengua de preferència en els formularis de preinscripció escoltar tot seleccionant-hi la casella de llengua preferent (castellà o català), tal com el context següent exposa:

<sup>14</sup> Una cerca a *Google* permet trobar 2.320 resultats d'aquest sintagma. Segons Carles Prats, però, és un concepte alarmista sobre el qual no hi ha consens entre els experts: "Les dades d'ús del català no deixen de caure. Una situació que ha portat Plataforma per la Llengua a comparar el català amb la crisi climàtica i que parlin d' "emergència lingüística". Una expressió que no agrada als experts. N'hi ha que la consideren massa alarmista i que pot desmotivar tant els catalanoparlants com les persones que potser s'animarien a aprendre la llengua si no fos "una llengua de perdedors"". (Font: <https://www.ccma.cat/tv3/sense-ficcio/emergencia-linguistica-o-alarmisme-que-pot-desmotivar/noticia/3168817/>).



*La polémica sobre la casilla se encendió a mediados de febrero tras una reunión entre el presidente Mariano Rajoy y Sociedad Civil Catalana (SCC), cuyos dirigentes, al acabar la misma, señalaron que el Gobierno se planteaba la posibilidad de modificar el folleto de preinscripción en Cataluña para acabar con los agravios al **castellano**. Una opción que asociaciones en defensa del **bilingüismo** como la propia SCC, Convivencia Cívica Catalana, o la Asociación por una Escuela Bilingüe llevan años reclamando. Desde el Gobierno se recurrió a circunloquios para responder públicamente sobre la posibilidad de implantar la casilla, algo que a la vez puso en jaque a la comunidad educativa **catalana**. Ante la amenaza del Gobierno de «atacar» la lengua, el «hardcore» nacionalista, docentes y padres se movilizaron.*

(ABC, 05/04/2018)

Alguns dels sintagmes que només s'han documentat a la premsa en castellà amb l'adjectiu *lingüístic-a* són *imposició lingüística, homogeneïtat lingüística, dictadura lingüística, imperialisme lingüístic, colonització lingüística, barbaritat lingüística, segregació lingüística, integrisme lingüístic, supremacisme lingüístic, caixa de pandora lingüística o litigi polític lingüístic*. Un altre sintagma relacionat (s'hi empra l'adjectiu sociolingüista) és *búnker nacional-sociolingüista*.

Es constata que algunes d'aquestes unitats tenen una forta connotació negativa i fan servir un llenguatge molt sensacionalista. Constatem que el diari que més utilitza aquesta mena de vocabulari és l'ABC.

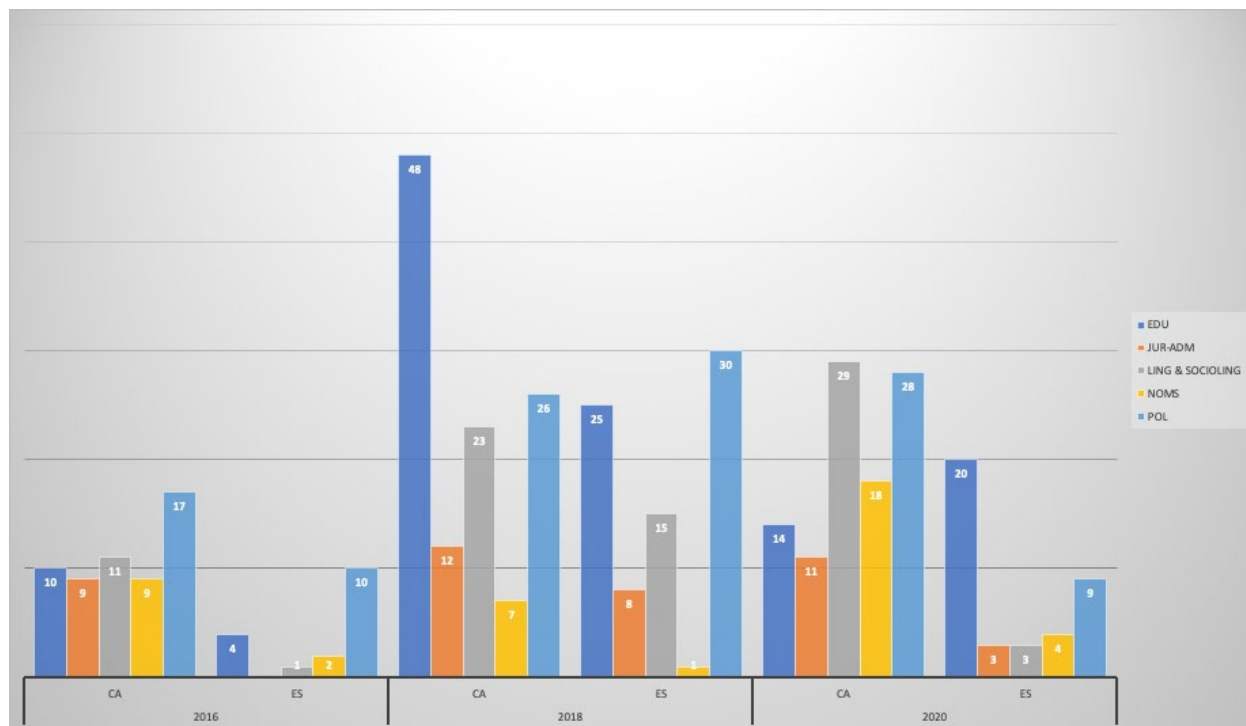
Els textos acostumen a descriure un règim lingüístic de sotmetiment, imposició i adoctrinament a través del sistema educatiu, sobretot a Catalunya (però també extensible a altres territoris amb llengües cooficials), tal com il·lustren els contextos següents:

*Mucho se ha dicho y escrito de las barbaridades lingüísticas perpetradas en Cataluña por los sucesivos gobiernos independentistas, en flagrante violación del ordenamiento jurídico vigente e impune desacato reiterado a las sentencias del Tribunal Constitucional. Tantas han sido las denuncias como nulos los intentos del Ejecutivo central por amparar los derechos conculcados. Pero, si grave es lo que sucede en el feudo del secesionismo catalán, mucho peor es el drama silenciado que viven millares de niños y padres en el País Vasco. Niños y padres sometidos a una euskaldunización tan forzosa como forzada en lugares de los que el vascuence desapareció hace siglos, lo que convierte esa lengua en un vehículo de expresión completamente extraño y por ende hostil para los alumnos obligados a estudiar materias ya de por sí complejas en un idioma endiablado que jamás han oído hablar en su entorno. (ABC, 15/12/2016)*

*La directora creativa de Mascaró denuncia en ABC la «dictadura lingüística impuesta por el separatismo» y señala a los políticos que han sido sus cómplices durante décadas. Sin miedo a un boicot empresarial, Úrsula Mascaró aboga por la españolidad irrenunciable del archipiélago balear. (ABC, 18/03/2018)*

*El PSOE, sí. El antiguo partido constitucionalista que Sánchez ha convertido en la carcasa de su caudillismo, el que apoyó de boquilla la aplicación del Artículo 155, entrega ahora a sus aliados sediciosos la llave del supremacismo lingüístico. (ABC, 05/11/2020)*

Hem revisat els 407 contextos d'ús i els hem classificat en funció si, quan parlen de llengua, refereixen aspectes educatius (EDU), juridicoadministratius (JUR-ADM), lingüístics i sociolingüístics (LING & SOCIOLING), noms propis (NOMS) o polítics (POL). La Gràfica 6 en presenta les dades resumides.



**Gràfica 6. Classificació temàtica de les concordances de *lingüístic -a***

Les dades de la gràfica anterior es desglossen a la Taula 24.

	EDU	JUR-ADM	LING & SOCIOLING	NOMS	POL	Total
<b>2016</b>	<b>14</b>	<b>9</b>	<b>12</b>	<b>11</b>	<b>27</b>	<b>73</b>
CA	10	9	11	9	17	56
ES	4	-	1	2	10	17
<b>2018</b>	<b>73</b>	<b>20</b>	<b>38</b>	<b>8</b>	<b>56</b>	<b>195</b>
CA	48	12	23	7	26	116
ES	25	8	15	1	30	79

<b>2020</b>	<b>34</b>	<b>14</b>	<b>32</b>	<b>22</b>	<b>37</b>	<b>139</b>
CA	14	11	29	18	28	100
ES	20	3	3	4	9	39
<b>Total</b>	<b>121</b>	<b>43</b>	<b>82</b>	<b>41</b>	<b>120</b>	<b>407</b>

Taula 24. Classificació desglossada de les concordançes de *lingüístic -a*

#### 5.4.3. Els 25 sintagmes nominals més freqüents

Una altra de les funcionalitats del programari #LancsBox és la detecció d'*n-grams* al corpus<sup>15</sup>. En aquest cas, hem hagut de filtrar-ne manualment els resultats i seleccionar els sintagmes nominals més utilitzats (vegeu la Taula 25).

	n-gram	Freq_abs	Freq_rel		n-gram	Freq_abs	Freq_rel
1	<i>llengua catalana</i>	58	6,0		<i>lengua vehicular</i>	40	6,3
2	<i>llengua pròpia</i>	29	3,0		<i>inmersión lingüística</i>	29	4,6
3	<i>immersió lingüística</i>	29	3,0		<i>lengua oficial</i>	27	4,3
4	<i>llengua vehicular</i>	28	2,9		<i>ley celaá</i>	25	4,0
5	<i>govern espanyol</i>	27	2,8		<i>sistema educativo</i>	23	3,6
6	<i>política lingüística</i>	27	2,8		<i>comunidad autónoma</i>	17	2,7
7	<i>estat espanyol</i>	20	2,1		<i>lengua catalana</i>	17	2,7
8	<i>escola catalana</i>	19	2,0		<i>comunidades autónomas</i>	16	2,5
9	<i>normalització lingüística</i>	19	2,0		<i>lengua extranjera</i>	14	2,2
10	<i>model lingüístic</i>	15	1,5		<i>ley educativa</i>	13	2,1
11	<i>nostra llengua</i>	12	1,2		<i>política lingüística</i>	13	2,1
12	<i>tribunal constitucional</i>	12	1,2		<i>lengua castellana</i>	13	2,1

<sup>15</sup> En aquest cas, n'hem hagut de filtrar els resultats manualment, ja que les deu seqüències textuais més freqüents són: *de la, a la, la llengua, de les, que el, que no, la seva, en català, en el i i la*.

	n-gram	Freq_abs	Freq_rel		n-gram	Freq_abs	Freq_rel
13	<i>llengües oficials</i>	11	1,1		<i>centros educativos</i>	12	1,9
14	<i>unió europea</i>	11	1,1		<i>tribunal constitucional</i>	12	1,9
15	<i>llengua materna</i>	11	1,1		<i>lengua española</i>	11	1,7
16	<i>sistema educatiu</i>	11	1,1		<i>lenguas oficiales</i>	11	1,7
17	<i>llengua oficial</i>	11	1,1		<i>lengua cooficial</i>	11	1,7
18	<i>televisió pública</i>	10	1,0		<i>lenguas cooficiales</i>	10	1,6
19	<i>llengua habitual</i>	10	1,0		<i>comunidad valenciana</i>	10	1,6
20	<i>llengua castellana</i>	10	1,0		<i>lengua común</i>	10	1,6
21	<i>comunitat educativa</i>	10	1,0		<i>proyectos lingüísticos</i>	10	1,6
22	<i>xarxes socials</i>	10	1,0		<i>alto tribunal</i>	9	1,4
23	<i>drets lingüístics</i>	9	0,9		<i>nuestra lengua</i>	8	1,3
24	<i>direcció general</i>	8	0,8		<i>escuela bilingüe</i>	8	1,3
25	<i>cultura catalana</i>	8	0,8		<i>escuela catalana</i>	8	1,3

Taula 25. Els 25 sintagmes nominals més freqüents (#LancsBox)

De nou, constatem que a ambdós blocs predominen els sintagmes relacionats amb el món de l'ensenyament. En català, de nou, entre els sintagmes més freqüents hi ha *llengua pròpia*, *immersió lingüística*, *llengua vehicular*, *escola catalana*, *model lingüístic*, *sistema educatiu* i *comunitat educativa* (així com el sintagma *[la] nostra llengua*, però no és un terme com la resta de casos).

Constatem també tot de sintagmes relacionats amb la política lingüística i l'estatus de les llengües, com *política lingüística*, *normalització lingüística*, *llengües oficials*, *llengua materna*, *llengua oficial* o *llengua habitual*. És rellevant, a més, observar com són de freqüents les referències als sintagmes *govern espanyol*, *estat espanyol* i *Tribunal Constitucional*. Altres sintagmes com *Unió Europea*, *xarxes socials*, *drets lingüístics*, *direcció general*, *cultura catalana* i *televisió pública* apareixen al llistat d'*n-grams* més freqüents.

Pel que fa al llistat d'unitats en castellà, de nou, consten nombrosos sintagmes relacionats amb l'ensenyament de les llengües, com *lengua vehicular, inmersión lingüística, sistema educativo, ley educativa, centros educativos, proyectos lingüísticos, escuela bilingüe* i *escuela catalana*, i tot de formes sinònimes per a referir-se al castellà (*lengua castellana, lengua española, lengua común, nuestra lengua*). També són freqüents les referències a *lengua oficial, lenguas oficiales, lengua cooficial, lenguas cooficiales* i *política lingüística*.

A la part alta de les freqüències també trobem *Tribunal Constitucional* i *Alto Tribunal*, així com *ley celtaá, comunidad autónoma, comunidades autónomas, Comunidad Valenciana* i *lengua extranjera*.

És significatiu que en ambdues llengües el sintagma *Tribunal Constitucional* (i la variant *Alto Tribunal*) tinguin una posició tan predominant. En castellà les al·lusions al TC són cites d'autoritat de la jurisprudència constitucional (naturalment, en matèria lingüística i educativa). De fet, la constitucionalitat és un argument que emergeix sovint a la premsa en castellà (vegeu l'exemple següent) i que, segons els discursos d'alguns diaris (*ABC* i *El Mundo*, sobretot), es troba en perill per l'amenaça del bilingüisme i la marginació (inexistent) del castellà a l'escola:

*La proposición del PPC, adelantada ayer por ABC, pretende situar de una vez el sistema educativo en esta comunidad dentro del marco constitucional. Según la propuesta, con el actual reparto de mayorías parlamentarias es imposible que prospere, se pondría fin al actual sistema monolingüe en catalán para avanzar hacia uno nuevo de conjunción lingüística o «bilingüismo integral», con castellano y catalán en pie de igualdad. (ABC, 05/04/2018)*

*La Ley Celaá pisotea los pilares de la Constitución. Se caracteriza por su sectarismo y no se ha contado con la comunidad educativa", añadió el portavoz del PP en el Ayuntamiento de Barcelona, Óscar Ramírez. Mientras, el presidenciable de Vox, Ignacio Garriga, ofrecía a su formación para "recuperar la Cataluña arrebatada y defender hasta el final que los padres puedan escoger la educación de sus hijos en castellano. (El Mundo, 21/12/2020)*

#### 5.4.4. Macroestructures

Aquest apartat analitza els àmbits (geogràfics i temàtics) i les macroestructures dels 286 articles del corpus d'anàlisi d'aquest treball. La lectura de cada article i l'atribució manual de cada categoria segons el contingut del text, i la validació dels resultats per part dels dos investigadors del projecte, ha permès obtenir les dades quantitatives que hi presentem.

Hem pres el concepte de macroestructura textual de van Dijk (1977,1978,1980) per referència al contingut semàntic global que representa un text. És, per tant, el tema principal del text o, dit en termes de l'anàlisi del discurs, l'estructura semàntica del conjunt

del text. Aquesta tasca ha estat feta de manera manual, per als 286 textos de la mostra estratificada i representativa analitzada qualitativament. Hem identificat i categoritzat, doncs, els temes principals de cada text. En aquest apartat abordem l'anàlisi de les macroestructures tant des del punt de vista quantitatiu com qualitatiu.

Com s'ha indicat a l'inici, hem analitzat premsa en llengua catalana i en llengua castellana. Tot i que l'Annex 2 inclou el llistat complet de l'anàlisi de les macroestructures, en aquest apartat analitzem qualitativament les categories amb una freqüència igual o superior a cinc (5) en cadascuna de les llengües.

En total, hem identificat 381 macroestructures diferents: 270 macroestructures a la premsa en llengua catalana i 187 a la premsa en castellà. Naturalment, hi ha coincidència entre els descriptors de les macroestructures d'ambdues llengües.

La Taula 26 presenta els descriptors de macroestructures més habituals (en ordre descendent) en ambdues llengües.

	<b>Prensa en CA</b>	<b>Prensa en ES</b>	<b>Total</b>
<b>immersió (lingüística)</b>	28	32	60
<b>ús</b>	38	2	40
<b>llengua</b>	30	8	38
<b>escola</b>	16	15	31
<b>controvèrsia</b>	-	28	28
<b>educació/ensenyament</b>	5	21	26
<b>normalització (lingüística)</b>	23	2	25
<b>imposició</b>	-	24	24
<b>Estat</b>	24	-	24
<b>llengua vehicular</b>	7	15	22
<b>atac/atacar</b>	16	4	20
<b>castellà/espanyol</b>	8	12	20
<b>discriminació</b>	3	16	19

	<b>Prensa en CA</b>	<b>Prensa en ES</b>	<b>Total</b>
<b>polèmica</b>	18	-	18
<b>utilitat</b>	5	12	17
<b>literatura</b>	12	4	16
<b>dret(s) lingüístic(s)</b>	12	2	14
<b>(llei) Celaá</b>	3	11	14
<b>llei</b>	10	3	13
<b>audiovisual</b>	12	-	12
<b>bilingüisme</b>	8	4	12
<b>nacionalisme</b>	2	9	11
<b>oportunitat</b>	2	9	11
<b>internacionalització</b>	6	5	11
<b>TV3</b>	9	2	11
<b>cultura</b>	4	7	11
<b>minorització</b>	11	-	11
<b>llengua catalana</b>	10	-	10
<b>anglès</b>	1	8	9
<b>sentència</b>	-	9	9
<b>castellanisme</b>	8	1	9
<b>coneixement</b>	5	3	8
<b>oficialitat (lingüística)</b>	5	3	8
<b>protecció</b>	7	1	8
<b>aprenentatge</b>	4	4	8
<b>model (educatiu)</b>	7	1	8

	Prensa en CA	Prensa en ES	Total
integració	5	2	7
immigració	5	2	7
tribunal	2	5	7
Espanya	7	-	7
promoció	6	1	7
independentisme	1	6	7
català	6	1	7
ortografia	4	2	6
Koiné	2	4	6
radicalisme	-	6	6
separatisme	-	6	6
vitalitat (lingüística)	6	-	6
Ciudadans/Ciudadanos	5	1	6
justícia/administració de justícia	6	-	6
desaparició	5	1	6
premi	4	1	5
divulgació	5	-	5
Plataforma per la Llengua	5	-	5
obligació/obligatorietat	1	4	5
Generalitat	2	3	5
model lingüístic	1	4	5
multilingüisme	4	1	5
expansió	-	5	5



	Premsa en CA	Premsa en ES	Total
<b>Partit Popular</b>	5	-	5
<b>irrellevant</b>	1	4	5

**Taula 26. Freqüències de les macroestructures identificades**

Constatem, en primer lloc, que en ambdós blocs la *immersió (lingüística)* és un dels aspectes més tractats a premsa, juntament amb *escola* i, en menor grau, *coneixement* o *oficialitat (lingüística)*. Aquests resultats posen en relleu, de nou, com l'aspecte i la gestió lingüística en el món educatiu és un dels temes més recurrents als diaris. Malgrat que el grau elevat de coincidència en macroestructures, la taula denota que els diaris en cada llengua emmarquen la qüestió lingüística amb *frames* diferents.

Els articles en català, per exemple, al·ludeixen a macroestructures com les següents apareixen amb una freqüència molt més alta que a la premsa en castellà: *ús* (38), *llengua* (30), *normalització (lingüística)* (23), *atac/atacar* (16), *literatura* (12), *dret(s) lingüístic(s)* (12), *lleï* (10), *TV3* (9), *bilingüisme* (8), *castellanisme* (8), *protecció* (7), *model (educatiu)* (7), *internacionalització* (6), *promoció* (6), *català* (6), *coneixement* (5), *oficialitat (lingüística)* (5), *integració* (5), *immigració* (5), *Ciutadans/Ciudadanos* (5) i *desaparició* (5). N'hi ha d'altres que apareixen exclusivament als textos en català, com *Estat* (24), *polèmica* (18), *audiovisual* (12), *minorització* (11), *llengua catalana* (10), *Espanya* (7), *vitalitat (lingüística)* (6), *justícia/administració de justícia* (6), *divulgació* (5), *Plataforma per la Llengua* (5) o *Partit Popular* (5).

Hi destaquem, per exemple, com emergeixen i s'expliciten conceptes clau de la defensa i la supervivència (contra l'*atac* i la *desaparició*, paraules clau que apareixen a l'enumeració anterior) dels drets lingüístics dels catalans, amb tot d'unitats que fan referència a la situació sociolingüística: *ús*, *llengua*, *normalització (lingüística)*, *dret(s) lingüístic(s)*, *lleï*, *bilingüisme*, *protecció*, *model (educatiu)*, *promoció*, *català*, *coneixement*, *oficialitat (lingüística)*, *vitalitat (lingüística)*, *audiovisual* i *minorització*.

Hem constatat que nombrosos articles refereixen les dificultats i reptes que té en l'actualitat la llengua catalana, alhora que també emergeixen les referències, les efemèrides i les fites de les polítiques lingüístiques endegades a principis dels anys vuitanta en els diversos territoris catalanoparlants. Alguns fragments il·lustratius són:

*Què és el que cal no oblidar en aquest 30è aniversari de la Llei de Normalització Lingüística? Cal recordar, per exemple, que l'Estatut d'autonomia de 1983 va representar la primera porta des de l'any 1908 a Menorca i 1717 a la resta d'illes, per començar a recobrar l'ús de la llengua catalana i capgirar així el procés de substitució lingüística. És l'Estatut, tot i que després la LNL ho reiterarà*

*i ho justificarà, que estableix per a la llengua catalana, i no per a la castellana, l'estatus de "pròpia de les Illes Balears", i en reconeix així l'oficialitat a partir del principi de territorialitat. (Ara Balears, 10/01/2016)*

*La Plataforma va tornar a insistir ahir que en l'àmbit de la justícia és on més problemes està trobant la normalització del català. "En aquest àmbit no hem estat capaços de trobar-hi cap dada positiva", va dir gràficament Escuder. (El Punt Avui, 04/05/2016)*

Així mateix, s'hi identifiquen agents clau amb relació a la promoció de la llengua catalana (TV3 i *Plataforma per la Llengua*) o contraris a aquesta (*Estat, Espanya, Partit Popular* i *Ciudadans/Ciudadanos*). Són interessants també que s'hi tracten nocions com *castellanisme, internacionalització, integració* o *immigració*, clau per a gestionar i garantir la salut de la llengua catalana.

Amb relació a les macroestructures dels articles en castellà, hem constatat que algunes, tot i que també són presents als textos en català, en castellà ocupen una posició més alta al rànquing de freqüències: *discriminació* (16), *llengua vehicular* (15), *utilitat* (12), *castellà/espanyol* (12), *(Ilei) Celaá* (11), *oportunitat* (9), *anglès* (8), *cultura* (7), *independentisme* (6), *tribunal* (5), *obligació/obligatorietat* (4) i *aprenentatge* (4). Altres temes són exclusius a la premsa en castellà, com *controvèrsia* (28), *sentència* (9), *radicalisme* (6), *separatisme* (6) o *expansió* (5).

El llistat de macroestructures clau de la premsa en castellà fa que emergeixin unitats com *obligació/obligatorietat, controvèrsia, radicalisme* o *separatisme*, nocions que s'associen habitualment amb la llengua catalana, tal com s'exemplifica amb els contextos següents:

*El integrismo lingüístico, ese subproducto gramático de los nacionalismos indigenistas que encuentra su raíz ideológica última en tratar de desposeer de derechos a los hablantes de los idiomas para traspasar tales fueros a los idiomas mismos, fijación fundamentalista una de cuyas manifestaciones más radicales es la inmersión obligatoria que hoy se practica en las escuelas de Cataluña y sólo de Cataluña (en ningún otro país del mundo occidental, incluido el Quebec canadiense, rige semejante sistema), nunca fue, contra lo que tantos quieren creer, una bandera histórica de la izquierda catalana de obediencia socialista. (El Mundo, 23/02/2018)*

*Promotora de Mos Movem, organización cívica que se ha levantado contra el pancatalanismo en Baleares, la directora creativa de Mascaró denuncia en ABC la «dictadura lingüística impuesta por el separatismo» y señala a los políticos que han sido sus cómplices durante décadas. Sin miedo a un boicot empresarial, Úrsula Mascaró aboga por la españolidad irrenunciable del archipiélago balear. (ABC, 18/03/2018)*

Altres temes com *aprenentatge, Koiné, ortografia, model lingüístic, Generalitat, premi* i *multilingüisme* apareixen en proporcions similars als dos blocs analitzats. Volíem identificar també, però, de quina (o quines) llengua tracta cada article. Tal com s'observa a la Taula 27, el 84,3% d'articles de la premsa en català tracta sobre la llengua catalana;

en paral·lel, la premsa en castellà hi dedica un 42,1% dels articles analitzats. La llengua de Mercè Rodoreda és, per tant, la llengua que en termes relatius i absoluts de la qual parlen més els mitjans: els diaris en català, però, semblen centrar-s'hi de manera quasi exclusiva. La llengua castellana ocupa el segon lloc al rànquing de llengües sobre les quals tracta la premsa en català, però només el 4,7% dels articles apareguts als quatre diaris catalans analitzats en parlen. En canvi, la premsa en llengua castellana dedica a l'espanyol el 28,1% dels articles analitzats.

	Premsa en CA		Premsa en ES		Total	
	Articles	%	Articles	%	Articles	%
<b>CA - Català</b>	145	<b>84,3%</b>	48	<b>42,1%</b>	193	<b>67,5%</b>
<b>ES - Castellà</b>	8	<b>4,7%</b>	32	<b>28,1%</b>	40	<b>14,0%</b>
<b>Bilingüisme/Trilingüisme/ Multilingüisme</b>	5	<b>2,9%</b>	9	<b>7,9%</b>	14	<b>4,9%</b>
<b>Llengües espanyoles/ cooficials<sup>16</sup></b>	3	<b>1,7%</b>	9	<b>7,9%</b>	12	<b>4,2%</b>
<b>ES &amp; CA/EU</b>	4	<b>2,3%</b>	3	<b>2,6%</b>	7	<b>2,4%</b>
<b>EU - Èuscar</b>		<b>0,0%</b>	6	<b>5,3%</b>	6	<b>2,1%</b>
<b>Altres llengües</b>	3	<b>1,7%</b>	2	<b>1,8%</b>	5	<b>1,7%</b>
<b>N/A<sup>17</sup></b>	1	<b>0,6%</b>	4	<b>3,5%</b>	5	<b>1,7%</b>
<b>EN - Anglès</b>	3	<b>1,7%</b>	1	<b>0,9%</b>	4	<b>1,4%</b>
<b>Total</b>	<b>172</b>	<b>100,0%</b>	<b>114</b>	<b>100,0%</b>	<b>286</b>	<b>100,0%</b>

Taula 27. Llengua o llengües a què fan referència els contextos

És remarcable posar en relleu el fet que el 13,2% dels articles en castellà tracten sobre alguna de les llengües cooficials de l'Estat i que un 7,9% tracten genèricament el multilingüisme, el bilingüisme o el trilingüisme.

<sup>16</sup> Són, en general, textos que tracten de l'estatus oficial lingüístic del català, el galleg i el basc.

<sup>17</sup> El text dels articles no permet assignar una llengua concreta a aquests textos perquè aborden qüestions lingüístiques de manera abstracta i general.

En ambdós casos, el percentatge d'articles sobre la llengua anglesa és molt minso (1,7% i 0,9%, respectivament, a la premsa en català i en castellà), equiparable als que tracten sobre altres llengües, com el cors, el rus, l'ucraïnès o l'àrab. Per exemple, l'article *El àrab aterrizo como optativa en el instituto* (*El País*, 27/10/2018) presenta, en termes positius, la prova pilot que l'IES Pompeu Fabra de Martorell, juntament amb tres centres públics de Terrassa, Badalona i Sant Feliu de Llobregat, van oferir-la com a assignatura optativa.

*Wissal es una de las estudiantes que puede hablar árabe con su familia, y decidió matricularse en la asignatura para perfeccionar su nivel. "El instituto hace bien ofertándola porque es un idioma que nos puede abrir muchas puertas", opina Wissal. Mientras, Pol vio esta asignatura como un reto. "Me apunté porque es una lengua muy diferente. Con otras puedes llegar a imaginar la pronunciación, pero con el alfabeto árabe es imposible", señala.*

[...]

*En el centro hay un 7,3% de población extranjera, la mayoría marroquí, pero el director Xavier Joseph señala que la cifra ascendería a alrededor de un 12% contando a los alumnos nacidos en España de padres marroquíes. "La asignatura es una muy opción para asegurar que aprendan árabe con un profesor de perfil académico. Esperamos tener una demanda superior para el próximo año", afirma Joseph.*

(*El País*, 27/10/2018)

La premsa catalana se centra, per tant, a parlar sobretot de la llengua catalana, mentre que la premsa en castellà analitzada tracta, a més d'aspectes del català i el castellà, qüestions en genèric de bilingüisme/trilingüisme/multilingüisme, les llengües espanyoles/cooficials o l'eusquera.

A més, hem identificat a quin àmbit geogràfic es refereixen aquests 286 articles quan parlen de temes lingüístics (Taula 28).

	Premsa en CA		Premsa en ES		TOTAL	
	Articles	%	Articles	%	Articles	%
<b>CA<sup>18</sup> espanyola (règim multilingüe)</b>	<b>129</b>	<b>75,0%</b>	<b>58</b>	<b>50,9%</b>	<b>187</b>	<b>65,4%</b>
Catalunya	109		42		151	
Illes Balears	14		8		22	
País Valencià	4		1		5	

<sup>18</sup> En aquesta columna, l'abreviatura CA equival a "comunitat autònoma".

	Premsa en CA		Premsa en ES		TOTAL	
	Articles	%	Articles	%	Articles	%
Navarra	1		3		4	
Euskadi			2		2	
Galícia			2		2	
Aragó	1				1	
<b>Espanya</b>	<b>17</b>	<b>9,9%</b>	<b>30</b>	<b>26,3%</b>	<b>47</b>	<b>16,4%</b>
Espanya	17		30		47	
<b>Internacional<sup>19</sup></b>	<b>7</b>	<b>4,1%</b>	<b>7</b>	<b>6,1%</b>	<b>14</b>	<b>4,9%</b>
Internacional	7		7		14	
<b>Països tercers</b>	<b>6</b>	<b>3,5%</b>	<b>3</b>	<b>2,6%</b>	<b>9</b>	<b>3,1%</b>
França	4				4	
Ucraïna	1				1	
Mèxic			1		1	
Còrsega	1				1	
Beirut			1		1	
Damasc			1		1	
<b>Múltiples CA/països</b>	<b>4</b>	<b>2,3%</b>	<b>4</b>	<b>3,5%</b>	<b>8</b>	<b>2,8%</b>
Catalunya/Espanya	1		2		3	
La Rioja/Espanya	1				1	
Balears/Espanya			1		1	
Amèrica Llatina/Espanya			1		1	

<sup>19</sup> Categoria que agrupa aquells articles que tracten el fet lingüístic i el seu estatus oficial des d'una perspectiva internacional i abstracta.

	Premsa en CA		Premsa en ES		TOTAL	
	Articles	%	Articles	%	Articles	%
Alemanya/Catalunya	1				1	
Balears/Catalunya	1				1	
<b>CA espanyola (règim monolingüe)</b>	<b>1</b>	<b>0,6%</b>	<b>6</b>	<b>5,3%</b>	<b>7</b>	<b>2,4%</b>
Andalusia			2		2	
Madrid			2		2	
La Rioja			2		2	
Múrcia	1				1	
<b>Altres territoris catalanoparlants</b>	<b>5</b>	<b>2,9%</b>	<b>1</b>	<b>0,9%</b>	<b>6</b>	<b>2,1%</b>
Països Catalans	4		1		5	
Catalunya del Nord	1				1	
<b>N/A</b>	<b>1</b>	<b>0,6%</b>	<b>4</b>	<b>3,5%</b>	<b>5</b>	<b>1,7%</b>
N/A	1		4		5	
<b>Europa</b>	<b>1</b>	<b>0,6%</b>	<b>1</b>	<b>0,9%</b>	<b>2</b>	<b>0,7%</b>
Europa	1		1		2	
<b>UE</b>	<b>1</b>	<b>0,6%</b>		<b>0,0%</b>	<b>1</b>	<b>0,3%</b>
UE	1				1	
<b>Total</b>	<b>172</b>	<b>100,0%</b>	<b>114</b>	<b>100,0%</b>	<b>286</b>	<b>100,0%</b>

Taula 28. Àmbit geogràfic a què fan referència els textos

Segons la Taula 28, el 77,9% dels articles publicats en llengua catalana es refereixen a notícies que tracten exclusivament sobre Catalunya, les Illes Balears, el País Valencià, els Països Catalans o la Catalunya del Nord. El segon àmbit geogràfic més tractat als articles en català, en un 9,9% dels textos analitzats, és Espanya. Després, hi ha les referències a aspectes lingüístics internacionals en abstracte (4,1%) i a països tercers (3,5%). En aquest darrer bloc predomina França (a quatre de sis articles).

Pel que fa a les dades de la premsa en llengua castellana, s'observa a la Taula 28 que, igual que passava en l'apartat anterior, hi ha una distribució més equitativa entre les diferents subcategories. Així, la premsa en castellà analitzada es refereix, en primer lloc, a regions multilingües de l'Estat, sobretot a Catalunya (el 36,8% dels articles se centren de manera exclusiva en aquest àmbit geogràfic). El segon àmbit geogràfic més tractat és el conjunt de l'Estat espanyol, sobre el qual tracta el 26,3% dels textos analitzats. De nou, el bloc internacional ocupa la tercera posició (amb el 6,1% dels documents estudiats), seguit dels articles que presenten i analitzen situacions lingüístiques aplicades a comunitats autònomes amb un règim oficial monolingüe, com Andalusia, Madrid, La Rioja o Múrcia.

Per tant, com a l'apartat anterior, constatem que la premsa en llengua catalana és molt autoreferencial i posa èmfasi, sobretot, en aspectes relacionats amb la llengua catalana als territoris catalanoparlants nacionals o de països tercers. La premsa en castellà també tracta, sobretot, aspectes relacionats amb l'àmbit català, però en una proporció inferior, i compensa aquella diferència percentual amb articles que se centren en Espanya i regions monolingües. És interessant esmentar que aquests textos presenten proves pilot i projectes educatius que fomenten el multilingüisme, habitualment amb programes bilingües que treballen la llengua nacional (el castellà) i una llengua estrangera (l'anglès). Aquests articles en són il·lustratius, atès que reflecteixen les bondats que els poders públics atribueixen als programes (o seccions) bilingües que incorporen l'anglès:

*La Junta lleva el segundo idioma al cole sin cumplir su plan de centros bilingües (ABC, 01/04/2016)*

*Los alumnos del programa bilingüe sacan mejores notas (El País, 21/12/2016)*

#### 5.4.4.1. Llengua i educació

El binomi llengua-educació ha estat, amb escreix, l'àmbit discursiu més recurrent al llarg de tota la mostra analitzada, emmarcat en una conflictivització constant, exacerbada sobretot quan es produeixen els dos fets concrets següents:

- a. L'aprovació de la *Ley Orgánica de Modificación de la Ley Orgánica de Educación* (LOMLOE), també denominada Llei Celaá (per la ministra que li va donar forma). El sector conservador titllava d'atac contra el castellà com a llengua vehicular a l'àmbit educatiu, a favor de la llengua catalana. Alguns exemples il·lustratius en són:

*La «famosa» Ley Celaá está sonando mucho últimamente en las redes sociales y en los medios de comunicación. Esta ley, aprobada desde el pasado jueves, propone como uno de sus puntos principales la eliminación del **castellano** como lengua vehicular en la*

*enseñanza. Bajo mi punto de vista, esta medida es anticonstitucional y una cesión al independentismo catalán. (ABC, 21/11/2020)*

*La llamada Ley Celaá no puede ser entendida sino como un eslabón más de las cesiones al independentismo en su pretensión de eliminar todo aquello que, como nuestra lengua, contribuya a la articulación de España. (El Mundo, 19/11/2020)*

*La Justicia ha hablado claramente y ha dicho que la Generalitat está excluyendo deliberadamente el español de la escuela. A ese sistema injusto, ilegal y anticonstitucional está intentando ampararlo la nueva Ley Celaá. Que dejen de mentirnos. Esta nueva ley no va a pasar por encima de la última sentencia del TSJC, ni de la Constitución", advirtió la presidenta de la Asamblea por la Escuela Bilingüe, Ana Losada, mientras los cláxones de los 300 vehículos conducidos por los manifestantes ensordecían sus palabras. (El Mundo, 21/12/2020)*

- b. L'aparició de sentències en l'àmbit educatiu, sobretot la sentència relacionada amb la implantació del 25% de castellà a l'escola catalana. En aquest sentit, la premsa es feia ressò de les reaccions polítiques i socials altament polaritzades, a favor i en contra, amb uns marcs discursius completament oposats.

*El texto de Ayuso blinda la educación concertada , defiende la libertad de los padres a elegir la educación de sus hijos y mantiene el español como lengua vehicular . El portavoz del Ejecutivo autonómico, Enrique Ossorio, defendió en su día que se trata de defender la libertad de los padres para elegir el tipo de enseñanza que quieren para sus hijos. (ABC, 29/09/2021)*

*Los centros educativos en Cataluña intentan que los padres admitan que la única lengua de relación posible con la administración educativa es el catalán. Esto es igual de ilegal que la inmersión. Todas las comunicaciones deberían ser bilingües catalán y castellano. (ABC, 29/09/2021)*

*Se rebeló la resistencia cívica contra el blindaje de la inmersión. Marchó por Barcelona contra el proyecto de Cataluña monolingüe. Protestó con vehemencia contra la Ley Celaá y su pretensión de suprimir el español como lengua vehicular y oficial del Estado. Y celebró con júbilo contenido la sentencia que obliga a impartir un 25% de materias en castellano en las escuelas catalanas.El histórico fallo judicial que derrumba los cimientos del proyecto de "normalización" lingüística de la Generalitat y acredita el uso "residual" del castellano en las aulas otorgó un vigor añadido a la manifestación. (El Mundo, 21/12/2020)*

*Fins ara, el TSJC ja havia publicat sentències obligant a garantir un 25% de l'ensenyament en castellà però amb una diferència cabdal: les denúncies d'aquelles causes eren de pares d'escoles concretes i, per tant, només calia aplicar la sentència a aquells centres. Ara la denúncia provenia del ministeri d'Educació espanyol i afecta tot el sistema públic d'ensenyament. (ARA, 17/12/2020)*

*El Constitucional també va anul·lar l'article pel qual les comunitats pagarien el centre concertat o privat a aquelles famílies que no poguessin escolaritzar els seus fills en una*



*escola pública amb el castellà com a llengua vehicular. L'anomenat xec Wert va ser sol·licitat per 154 famílies, que representen al 0,04% de l'alumnat. Una sentència judicial va obligar els centres a donar un 25% d'hores en castellà a les famílies que així ho requerissin. En l'últim curs escolar, 12 famílies ho van sol·licitar (cosa que representa un 0,35% dels centres). (El Periódico, 03/11/2020)*

En tots dos casos, detectem la figura retòrica de l'oxímoron com a manera de polaritzar discursivament el debat. Dit d'una altra manera: la positivitat aplicada a un fet (la llengua catalana o la castellana) es defensa a través de l'atac de l'altra. S'estableix constantment un *nosaltres vs. ells*, especialment en els diaris *El Mundo* i *ABC*, tot i que no exclusivament.

Pel que fa a la llengua castellana, observem les referències constants al vessant internacional i a la política exterior de la llengua (les bondats de l'espanyol com a llengua comuna, llengua parlada per 500 milions de parlants, llengua d'oportunitats, llengua que t'obre portes) en contraposició a una "política interior". En aquest context nacional, la llengua espanyola es presenta discursivament com a llengua perseguida [per la llengua catalana o més aviat per les autoritats catalanes].

*Pero, si grave es lo que sucede en el feudo del secesionismo catalán, mucho peor es el drama silenciado que viven millares de niños y padres en el País Vasco. Niños y padres sometidos a una euskaldunización tan forzosa como forzada en lugares de los que el vascuence desapareció hace siglos, lo que convierte esa lengua en un vehículo de expresión completamente extraño y por ende hostil para los alumnos obligados a estudiar materias ya de por sí complejas en un idioma endiablado que jamás han oído hablar en su entorno. Bilbao, ambas márgenes del Nervión, la mayor parte de los núcleos urbanos de Álava y Guipúzcoa... Pueblos y ciudades en los que el castellano, lengua común de todos los españoles, es la que se emplea desde tiempo inmemorial en todas las facetas de la vida, excepto la escuela, donde el euskera se ha impuesto a golpe de sectarismo con fines de adoctrinamiento ideológico y consecuencias devastadoras. (ABC, 15/12/2016)*

*[La nueva ley] busca cortar a hachazos el cordón umbilical y los lazos de unión que ensamblan y cimientan el sentimiento de pertenencia a una misma nación, expropiando, además, a las generaciones futuras el patrimonio y el tesoro común que es el castellano, español en el mundo, una lengua universal. Ello no solo privará, a esas generaciones, de conocer en profundidad la riqueza idiomática del español -continúa el manifiesto- sino que las desconectará paulatinamente del resto de compatriotas desde el punto de vista histórico y emocional. (El Mundo, 11/12/2020)*

*Iniciativas que pretendían que lo importante, la lucha contra el virus nacionalista, no se convirtiera en lo urgente, el aplacamiento al independentismo más feroz que pusiera en jaque nuestro Estado de derecho. Se ha convertido en imperativo ya, cuando hasta nuestras principales autoridades se vanaglorian de no saber el idioma común que tenemos con el resto de españoles y quieren "plantear debates" para acabar con nuestras instituciones, que los partidos políticos constitucionalistas hagan de Baleares su prioridad y sean conscientes de que aún estamos a tiempo de hacer que las islas no sean el siguiente capítulo de la lucha de España contra su autodestrucción. (El Mundo, 10/08/2018)*

*La izquierda se unió a ellos contra la idea nacional de España que identificaron, en terrible y trágico error histórico, con el régimen de Franco. La derecha, con su mala conciencia y cobardía proverbial, no se atrevió a hacerles frente para no ser acusada de franquista. Así llegamos a la indignidad y postración de la lengua común, perseguida impunemente en muchas regiones. Ante la estupefacción de cualquier visitante extranjero, no solo de los 500 millones de hispanohablantes. Leyes de vergüenza aprobadas en las Cortes españolas obligan a usar en documentos españoles los nombres no españoles de ciudades españolas. Ahí están esos carteles ridículos de Girona o Lleida o A Coruña. El BOE publica grotescos nombres inventados solo para hacer olvidar los auténticos nombres españoles. (ABC, 07/01/2018)*

Constatem igualment polaritats dicotòmiques a través d'estructures valoratives com per exemple llengua àmpliament difosa-llengua perseguida; expansió universal-persecució local; ús i promoció internacional-espionatge lingüístic local; llibertat/democràcia-imposició; llei-incompliment. A través d'aquestes polaritats, es forgen uns marcs mentals contraposats mitjançant aquestes pràctiques discursives persuasives que fan que el lector hagi de resoldre dues preguntes implícites: Com es pot degradar el valor d'una llengua universal en favor d'una llengua local? Com es pot anar en contra de la llibertat i de la democràcia?

La construcció d'aquest marc mental és el que Swiggers (2018, p. 85) ha denominat *mecanisme d'acomodació*, definit com la presentació de contingut ideològic com a evident, a través de diverses figures discursives com ara l'oxímoron.

*El Gobierno que preside Ximo Puig establecerá bonificaciones fiscales para las empresas que usen el valenciano en la rotulación de sus negocios y en la página web. La reducción de impuestos para discriminar de forma positiva una de las dos lenguas oficiales frente al castellano se regulará a través de un «Sello de Calidad Lingüística» que comenzará a expedirse el próximo año. La medida «tiene como objetivo concienciar a pymes, comercios y firmas privadas del valor de utilizar el valenciano». (ABC, 10/02/2016)*

*El otro gran punto de discusión fueron las multas a los comerciantes que no rotulan en catalán. Carlos Silva, de Impulso Ciudadano, criticó lo que estima una estrategia por parte de los sucesivos gobiernos nacionalistas para excluir el español del espacio público. "Miles de empresarios, grandes y pequeños, han sido sometidos a controles lingüísticos y han sido sancionados. Los inspectores de consumo se han convertido en guardianes de la lengua catalana", lamentó. Además, alertó de la persecución a la que se ven sometidos los comerciantes que optan por rotular la mitad en español y la mitad en catalán. (El País, 25/04/2018)*

*–Cuando aprendimos un idioma imperial, caímos en la servidumbre. Cuando lo desaprendamos, seremos otra vez libres.*

*La Nueva Censura es la Inquisición sin lactosa, y así la establece el BOE, que también ha prohibido el español en España, con lo cual quien quiera pecar o delinquir habrá de hacerlo por señas del mus.*

[...]

*Contra el español, «lengua del imperio», se dirige el Glavlit de Sánchez, que tiene menos oposición que Stalin. Dombrovski recuerda que en el encuentro con los escritores en casa de Gorki un viejo idiota abordó al dictador: «El Glavlit y sus redactores son demasiado duros, camarada Stalin. Por ejemplo, usted, Iósif Vissariónovich, tiene la cara picada de viruelas, pero nos prohíben mencionarlo».*

[...]

*Tanto progreso, y hemos acabado en la Constantinopla de Julio Camba, donde los Jóvenes Turcos concedieron la libertad de prensa, pero a condición de no usar de ella. El Ministerio de la Verdad es una petición centrista de hace tres años. (ABC, 07/11/2020)*

*Quienes estrangulan a la Cultura con menos del 2 por ciento de un presupuesto para priorizar propaganda y embajadas, ignoran que la lengua -sostenida por la inmersión escolar, multas, subvenciones mediáticas y clientelismo político- no es Cataluña: el castellano es la lengua habitual del 60,9 por ciento de los jóvenes entre 18 y 24 años y el catalán mantiene un equilibrio con el castellano en la población mayor de 50 años. (ABC, 20/09/2020)*

*En un escenario de absoluto drama para la lengua oficial del Estado, además, las Administraciones Públicas baleares, con el incomprensible apoyo de partidos constitucionalistas, apoyan subvenciones para que el ya de por sí predominante catalán sea el idioma principal en los rótulos de los comercios de las islas. Como en los peores tiempos de Pujol y del tripartito. Ya estamos ahí. Todo lo anterior tiene una conclusión evidente, y es que el futuro que nos espera en Baleares, si no hacemos nada por remediarlo, es acabar siendo en unos años la nueva Cataluña. (El Mundo, 10/08/2018)*

#### 5.4.4.2. Llengua i justícia

Les dades analitzades ens indiquen que el binomi llengua-justícia és transversal i afecta, principalment, tres camps en concret: la sanitat, l'Administració pública i, molt especialment, l'educació.

Pel que fa a la sanitat i a l'Administració, trobem una categorització constant d'*imposició*, exclusivament en la premsa espanyola conservadora, a totes aquelles iniciatives polítiques adreçades a garantir els drets lingüístics en català, sobretot al País Valencià, i en èuscar a Navarra i Euskadi. En altres paraules, totes aquelles propostes de llei adreçades a fer de la llengua catalana i basca un requisit per poder optar a una plaça en l'administració pública, inclosa la sanitat, són titllades discursivament d'impositives.

*Educación oferta 320 plazas de profesor, el 70% de ellas en euskera*

*La oposición denuncia la discriminación y el sectarismo nacionalista de la medida*

*El Gobierno de Navarra ha aumentado a 320 el número de plazas de profesor que saldrán a concurso el año que viene. Eso sí, quienes sepan euskera tendrán ventaja ya que el 71,25% de las plazas son para profesores en ese idioma. La oferta de plazas anunciada ayer por el consejero de Educación, José Luis Mendoza, llamó la atención teniendo en cuenta que más del 80% de la población navarra no sabe euskera.*

*Y de las 92 plazas ofertadas en castellano, más de la mitad, 50, están destinados a profesores de Pedagogía Terapéutica, es decir, a profesores de apoyo. El resto de plazas en castellano van*

*destinadas a las especialidades de Audición y Lenguaje (14), Educación Física (16) y Música (12). (ABC, 07/11/2015)*

*Promovimos una ley de símbolos que acababa con toda simbología pancatalanista habida y por haber. Propugnamos que el catalán fuera un mérito y no un requisito en el acceso a la función pública y yo, personalmente y como presidente, pedí más inspección educativa de ámbito nacional para acabar con el adoctrinamiento. (El Mundo, 10/08/2018)*

*Pedagogos y lingüistas coinciden en que el modelo de lenguas catalán, único en Europa, «es discriminatorio y perjudica a los alumnos castellanohablantes» y en que «es incorrecto hablar de inmersión porque, al aplicarla, no hay respeto manifiesto hacia el resto de lenguas del sistema». Por eso, el modelo es de sumersión. «Su único fin es anular al castellano», explican. (ABC, 22/03/2021)*

L'àmbit educatiu ha estat, tal com hem vist en l'apartat anterior, l'autèntic camp de batalla judicial, polític i mediàtic que ha copsat de manera més notòria la premsa analitzada. Observem, molt especialment en l'educació però també en l'àmbit de la Sanitat, la superposició d'un discurs judicial hostil cap a la llengua catalana i èuscar, contextualitzades en un marc de permanent devaluació discursiva en referència a la llengua castellana.

*El rodillo lingüístico de la Generalitat llega al Inem catalán. Asistentes a cursos de formación para desempleados financiados por el Servicio Catalán de Empleo (SOC) y el Ministerio de Trabajo, denuncian haber sido discriminados por no dominar el catalán. Se sienten especialmente indignados por el hecho de que tras estos cursos que les «marginan por hablar la lengua oficial del Estado esté también el Gobierno, que no hace nada para proteger nuestros derechos». (ABC, 25/02/2021)*

*Escrache lingüístico en un centro de salud por no hablar en catalán  
Un activista exige el despido de una doctora que atendió en castellano a una paciente  
Nuevo capítulo de persecución lingüística en Cataluña. En esta ocasión en un centro de atención médica primaria (CAP) del barrio de Les Corts de Barcelona. Hasta allí se desplazó un grupo de activistas ultranacionalistas encabezados por el empresario Santiago Espot, conocido por, entre otras cosas, haber promovido la pitada al Rey en la final de la Copa de 2015, para increpar y pedir explicaciones al director del centro por un supuesto caso de discriminación lingüística. (ABC, 10/10/2019)*

#### 5.4.4.3. Llengua i política

El binomi llengua-política és, també, transversal i impregna totes les dades analitzades, en la mateixa línia de l'anàlisi feta als apartats anteriors. Políticament, es fa servir el fantasma de la desaparició del castellà a Catalunya per part dels partits de dreta i ultradreta espanyols, uns arguments recollits mediàticament per la premsa conservadora espanyola.

S'alimenta, així, una visió conflictiva de la cohabitació de llengües, sobretot a Catalunya, País Valencià, Balears, Euskadi i Navarra (territoris on governen les esquerres autòctones, o els socialistes). Casualment, aquest marc de conflicte no és aplicat a Galícia, on no hem trobat cap notícia en els mateixos termes de conflicte i problematització que per a la resta de territoris. Cal fer una anàlisi política succinta per poder trobar una explicació a aquesta realitat discursiva: a Galícia, hi governa el Partit Popular, partit conservador en plena sintonia amb els diaris que més conflictivitzen les llengües (*El Mundo* i *ABC*).

L'anàlisi fa ressaltar una preeminència del castellà, vehiculat discursivament com a element "natural" i consubstancial a l'espanyolitat. Observem, per tant, la creació d'un marc de subordinació lingüística, en què el castellà ha de ser la llengua comuna de tothom. Qualsevol modificació d'aquest marc és titllat de "supremacisme nacionalista" (principalment català i basc) l'objectiu del qual és trencar un ordre polític i social, amb una "harmònica distribució" de llengües derivada de l'ordre constitucional espanyol. Un marc discursiu que, en definitiva, vehicula un supremacisme de la llengua castellana, emmarcat en termes de naturalitat.

#### 5.4.4.4. Interpretació dels resultats

L'anàlisi, tant quantitativa com qualitativa, ens ha mostrat les tensions discursives (altament polaritzades) que no són més que les representacions mediàtiques de la conflictivització política i jurídica del fet lingüístic, molt especialment durant el període analitzat, coincident amb el procés d'independència de Catalunya.

D'una banda, observem uns discursos (vehiculars molt especialment per part dels diaris conservadors) que conceptualitzen les llengües de l'Estat des del punt de vista de la subordinació lingüística. Dit en altres paraules, estableixen una jerarquia lingüística, amb unes llengües conceptualitzades com subordinades al castellà. Són representacions conceptuals diglòssiques que treuen a la llum unes "ideologies diglòssiques", citant el terme de R: Lluís Ninyoles (1971), que emmarquen la realitat lingüística de l'Estat a través de dicotomies, contraposicions i dualitats.

El que hem vist a través de la nostra anàlisi és l'exemple empíric discursiu del les oposicions que Gal (2018, p. 233) estableix entre propietats atribuïdes a la llengua estàndard, d'una banda (el castellà en la nostra anàlisi), i a les formes lingüístiques subordinades, de l'altra (el català i l'èuscar). Així, les llengües són emmarcades en dos valors dicotòmics, que Gal resumeix en el quadre següent:

Eix modernista de valors/propietats oposats

<i>Llengua estàndard</i>	<i>Dialectes, patois, llengües minoritàries i indígenes</i>
anonimat	autenticitat
universal	particular/localitzada
raó/economia	emoció
progrés	tradició/endarreriment
alfabetitzat/instruït	oral
centre	perifèria
homogènia/unificada	diversa

**Figura 11. Eix de valors dicotòmics (Gal 2018)**

L'anàlisi també ens ha mostrat que qualsevol intent de trencar aquesta dicotomia ha estat contestat políticament i jurídicament i vehiculat mediàticament a través d'una deslegitimació discursiva i conceptual, que titllen les llengües catalana i èuscar - o més aviat els que representen políticament - en termes de radicalitat, imposició i manca de llibertat.

Com hem esmentat, és, amb escreix, l'àmbit educatiu, seguit de l'àmbit de l'Administració pública i l'àmbit sanitari, els àmbits amb un índex de conflictivització mediàtica, política i judicial. Qualsevol intent polític d'adoptar un marc normatiu que garanteixi els drets lingüístics dels parlants de català i èuscar en aquests tres àmbits ha estat combatut des d'aquestes tres esferes. Podem afirmar, doncs, que l'anàlisi discursiva sobre les llengües no pot anar dissociada d'una anàlisi política i jurídica.

## 6. Conclusions preliminars

La lingüística de corpus, la lexicometria i l'anàlisi del discurs ens han permès identificar les tendències i representacions discursives que emergeixen a set diaris generalistes d'àmbit català i espanyol quan tracten notícies sobre llengua i afers lingüístics.

Com ja s'ha assenyalat a l'apartat 3, el projecte ha estat molt més ambiciós del que s'havia previst inicialment, motiu pel qual només s'han pogut analitzar parcialment el conjunt de dades recollides, tractades i sistematitzades. Per aquest motiu, presentem les conclusions principals de la feina duta a terme en aquest treball i proposem algunes línies de treball futures que volem endegar.

Tot seguit, per tant, presentem les conclusions preliminars que podem extreure d'aquesta primera explotació quantitativa i qualitativa del corpus amb relació al corpus, l'anàlisi i els resultats obtinguts:

### 6.1. Sobre el corpus d'anàlisi

Aquest estudi ha partit d'una anàlisi que podem inserir en la tradició de la Lingüística de Corpus que combina criteris quantitius i qualitius.

Com ja hem assenyalat a l'apartat 4, atès el gran soroll obtingut amb la cerca a la base de dades *Factiva* amb cerques booleanes (vegeu l'apartat 4.1.2), hem revisat manualment els 46.960 titulars (vegeu la Taula 7). Quan no n'hi havia prou per a escatir si eren pertinents o no, n'hem revisat el context addicional dels resultats de *Factiva* (o el text complet de l'article) per tal d'identificar els 11.481 articles rellevants.

	CA				Total CA	ES			Total ES	Total
Diari	ARA	PER	PUN	VAN		ABC	MUN	PAI		
<b>Brut</b>	<b>10.347</b>	<b>2.589</b>	<b>7.606</b>	<b>5.804</b>	<b>26.346</b>	<b>7.131</b>	<b>5.688</b>	<b>7.795</b>	<b>20.614</b>	<b>46.960</b>
<b>Net</b>	<b>3.969</b>	<b>675</b>	<b>1.221</b>	<b>897</b>	<b>6.762</b>	<b>2.544</b>	<b>780</b>	<b>1.395</b>	<b>4.719</b>	<b>11.481</b>
<b>%</b>	<b>38,4%</b>	<b>26,1%</b>	<b>16,1%</b>	<b>15,5%</b>	<b>25,7%</b>	<b>35,7%</b>	<b>13,7%</b>	<b>17,9%</b>	<b>22,9%</b>	<b>24,4%</b>

Tot i que això suposa només el 24,4% del corpus brut inicial, era necessari dur a terme aquest filtratge manual per garantir la sistematicitat, l'exhaustivitat i la compleció de les dades de partida per a una recollida, un tractament i una sistematització dels articles pertinents per al nostre estudi.

La constitució del corpus és, per tant, una de les fites assolides amb aquest projecte. Com ja s'ha indicat a l'apartat 4.1.2, només la descàrrega dels 11.481 articles hauria comportat una quantitat d'hores de feina inassolibles en el marc d'aquesta recerca. Els esforços esmerçats, però, a filtrar els resultats i a identificar-hi els articles pertinents ha pagat la pena.

Entre d'altres, això ens ha permès identificar els pics (Gràfica 4) més importants de notícies sobre llengua i multilingüisme, que ha suposat la base empírica de constitució del subcorpus estratificat i representatiu analitzat a l'apartat 5.4.

L'elevat nombre d'articles rellevants identificats –més d'11.000 durant període 2015-2021– palesa l'alta mediatització del fet lingüístic a la premsa catalana i espanyola. La xifra ha depassat, de llarg, les nostres expectatives inicials. Per qüestions de temps i d'espai, hem concentrat la nostra anàlisi en els quatre pics principals del període cronològic analitzat en què les notícies eren més recurrents.

La constitució del corpus ha implicat tota una sèrie de reptes tècnics (vegeu l'apartat 4.1.3) que han fet que l'anàlisi s'hagi hagut d'articular en dues dimensions. D'una banda, s'han estudiat lexicomètricament –amb el programa *#LancsBox*– i qualitativament els titulars dels 11.481 resultats. Com s'ha observat a l'apartat 5.3, la constitució d'aquest subcorpus permet identificar, de manera global, les tendències amb relació a l'ús del vocabulari que es confirmen amb l'anàlisi del corpus estratificat i representatiu de 286 textos.

Els titulars suposen una font d'informació discursiva inicial molt valuosa, amb un esforç relatiu baix pel que fa a la seva constitució (en comparació amb l'altre subcorpus). Tanmateix, aquest subcorpus no és suficient quan s'han de copsar els matisos i les idiosincràsies de cada llengua i diari amb relació a les llengües i el multilingüisme.

De l'altra, hem hagut de limitar la nostra anàlisi quantitativa i qualitativa aprofundida a una mostra estratificada i representativa del corpus. L'anàlisi dels textos de l'apartat 5.4 (descarregats manualment i analitzats amb *#LancsBox* ha confirmat les tendències generals observades amb l'anàlisi dels titulars de l'apartat 5.3, fet que valida la metodologia emprada per a constituir la mostra representativa i estratificada del corpus d'anàlisi. Aquesta mostra del corpus, integrada pels 286 articles analitzats amb un enfocament més qualitatiu, fa que emergeixin les representacions discursives i les singularitats en cada llengua i diari.



Com diu l'aforisme atribuït a Mies van der Rohe, "el dimoni és als detalls". L'anàlisi qualitativa dels 11.481 articles identificats, naturalment, ens permetrà identificar, a partir de l'aplicació generalitzada de les variables diari i any, els posicionaments i les representacions discursives que difonen cadascuna de les set capçaleres analitzades. En conseqüència, constatem que el corpus de textos pertinents identificats, malgrat les diferències quantitatives, tant en termes absoluts com relatius (vegeu la Taula 9), és una mina inesgotable que hem d'explotar amb recerques futures.

L'anàlisi qualitativa del corpus s'ha centrat, sobretot, en la representació de les llengües (especialment del català) i del multilingüisme a la premsa en castellà. Aquest bloc presenta un discurs molt més connotat i polèmic que el que s'identifica a la premsa en català, molt més uniforme. Això es fa evident, per exemple, amb el llistat de freqüències i el fet que paraules genèriques (des del punt de vista temàtic) com *any*, *manera*, *cas*, *cosa*, *fet* o *part* apareguin al llistat dels 25 substantius més freqüents (vegeu l'apartat 5.4.1), mentre que tot el vocabulari en castellà (a excepció, potser, d'*año*) és molt més concret.

Com es pot constatar, a més, gran part dels exemples que il·lustren les afirmacions de l'informe provenen, sobretot, de dues fonts: l'*ABC* i *El Mundo*. Això es deu al fet que la lectura atenta i l'anàlisi qualitativa dels corpus ha fet palès que aquests dos diaris són els que empen un vocabulari més sensacionalista i conté unes afirmacions més polèmiques i enceses. Així, per exemple, a *El Mundo* es documenten, amb relació al català (o, en alguns casos, a alguna altra llengua cooficial de l'Estat) sintagmes com *integrismo lingüístico* (sobre el model d'immersió lingüística català) o *búnker nacional-sociolingüista* (com a qualificador d'Òmnium Cultural).

Malgrat que no hem pogut quantificar-ho de manera global, les freqüències de l'apartat 5.4.2.1 sí demostren que el diari *ABC* hi recorre de forma habitual. Així, per exemple, hi consten expressions com *segregación lingüística*, *barbaridades lingüísticas*, *supremacismo lingüístico*, *exclusión lingüística*, *política monolingüística* o *colonización lingüística* amb relació, de nou, al model lingüístic de l'escola catalana.

Malgrat l'interès discursiu obvi d'aquest tipus d'expressions, però, considerem que, per a tenir una panoràmica més fidel del paisatge mediàtic en català i en castellà sobre les llengües i el multilingüisme, és important mantenir tots els blocs del corpus. Per tant, el futur de la nostra recerca sobre llengua i multilingüisme a la premsa en català i en castellà es proposa:

- Descarregar els 11.481 textos pertinents identificats i explotar-los quantitativament i qualitativament amb aproximacions teòriques i metodològiques similars i complementàries a les d'aquest primer estudi.

- Constituir una base de dades de consulta i d'explotació d'aquests 11.481 textos, per tal que l'equip actual (amb la col·laboració d'altres investigadors futurs) pugui ampliar la recerca sobre les representacions discursives de la llengua i del multilingüisme a la premsa en català i en castellà. En especial, volem continuar investigant de quina manera s'hi representa el català, tot i que també volem incloure les de la resta de llengües cooficials a l'estat i de l'anglès).
- Enriquir, amb una periodicitat determinada, el corpus textual de notícies pertinents per tal de continuar alimentant aquesta recerca i, així, detectar-hi variacions (segons variables com la diacronia, l'autoria o la llengua).

## 6.2. Sobre l'anàlisi de les dades

El doble enfocament del treball ha estat clau per poder fer una anàlisi acurada de les dades. D'una banda, un enfocament quantitatiu a partir de l'exploració lexicomètrica, les taules de freqüències de substantius, adjectius i sintagmes nominals. La metodologia quantitativa aplicada d'anàlisi global dels titulars permet obtenir tendències i resultats prou sòlids.

L'anàlisi lexicomètrica del corpus de 286 articles ha confirmat les tendències observades amb l'anàlisi dels 11.481 titulars de premsa pel que fa a temàtiques i vocabulari. Evidentment, però, quan cal extreure exemples il·lustratius de les representacions discursives que se'n deriven, el corpus de titulars presenta limitacions. Tanmateix, és un bon indicador per a caracteritzar des de manera general els usos discursius en la premsa en les dues llengües analitzades.

Aquest treball exploratori, ateses les limitacions temporals i de recursos d'aquesta primera anàlisi, només ha permès explotar els resultats sistemàticament a partir de la variable llengua (català/castellà) i any. Així doncs, la creació del corpus presenta encara un potencial ingent, l'anàlisi del qual creiem que proporcionarà més contrastos encara, si emprem com a variables de l'anàlisi la font (diferències entre diaris) o el gènere discursiu (diferències entre gèneres informatius i d'opinió). La sistematització de les dades ens permet continuar la recerca amb una explotació més acurada.

L'anàlisi qualitativa ha estat cabdal a l'hora d'entendre els contextos d'usos. A través de l'aproximació de Thompson, s'ha posat en relació l'anàlisi textual amb l'anàlisi contextual. Qualsevol anàlisi del discurs ha d'incloure l'anàlisi del context en què es produeixen. Seguint Thompson, les dades qualitatives han permès esbrinar la dimensió ideològica dels discursos sobre el multilingüisme, les temàtiques –a través de les macroestructures– les cadenes de raonament i els argumentaris. La combinació de totes dues tècniques – l'anàlisi lexicomètrica i l'anàlisi contextual– ha proporcionat una radiografia acurada i

sistemàtica de com la premsa conceptualitza les notícies sobre llengües als diaris en catalans i espanyols analitzats.

### 6.3 Sobre els resultats

L'anàlisi dels 11.481 titulars i dels 286 articles ha permès identificar els temes principals (analitzats a través de les macroestructures i l'anàlisi contextual) que la premsa associa als debats lingüístics.

L'estudi ens permet constatar que, quan la premsa catalana tracta qüestions lingüístiques, se centra quasi exclusivament en els àmbits catalanoparlants, sobretot Catalunya, València i les Illes Balears. Quan la premsa en castellà tracta aspectes lingüístics, parla, sobretot, de Catalunya i de la llengua en l'àmbit educatiu, i, en menor mesura, en l'Administració i la sanitat. És interessant remarcar les referències de la premsa en castellà – sobretot els diaris *ABC* i *El Mundo* – a Balears i País Valencià, que compara constantment amb Catalunya i crea, així, una extensió semàntica dels termes del debat. De manera molt més residual, la premsa castellana parla també, tot i que de manera molt menys recurrent, de les iniciatives de programes bilingües anglès-castellà en l'àmbit educatiu en la resta de l'Estat, especialment a Madrid.

L'anàlisi més qualitativa ha constatat l'alta polarització discursiva basada en una alta polarització discursiva, que categoritza les llengües de manera oposada i dicotòmica a través d'estructures valoratives jerarquitzants. El castellà, com a llengua comuna i “naturalment” cohesionadora, i les llengües pròpies dels diferents territoris –català i èuscar– com a llengües hostils i imposades per les autoritats públiques catalanes en l'àmbit de l'educació sobretot, però també de l'administració pública i la sanitat. Aquesta línia de raonament emergeix més evidentment en els diaris més conservadors *El Mundo* i *ABC*. De manera molt més subtil i modulada, també la trobem al diari *El País*. En canvi, els diaris d'àmbit català recorren a un marc discursiu defensiu i de regressió de drets.

Les dades confirmen que no podem dissociar l'anàlisi del discurs sobre llengües de l'anàlisi contextual política i social. L'alta judicialització de la llengua ha estat una constant observada a tota la premsa, en català i en castellà, progressista i conservadora, en un àmbit molt concret: l'Educació. La franja temporal que hem analitzat ha coincidit amb el procés independentista català, i la llengua ha estat a la diana dels mitjans de comunicació.

Tancàvem aquest informe, a finals de 2022, amb la detecció d'un altre pic remarcable de notícies sobre el català l'àmbit educatiu, altament polaritzat i amb un establiment cada cop més exacerbant entre el *nosaltres vs. ells*, aquesta vegada amb una clara estratègia

d'europèitziació dels debats sobre llengües al sistema educatiu català. *Los 10 agravios al español que constatará la comisión de eurodiputados en su visita a Cataluña*, llegíem el 22 d'octubre a *El Mundo* que, entre molts altres diaris, es feien ressò de la decisió del Comitè de Peticions del Parlament Europeu, liderat per l'eurodiputada del Partit Popular, Dolors Montserrat, d'enviar un grup d'eurodiputats a Catalunya per analitzar la gestió de les llengües a les escoles.

Fer un seguiment de la conceptualització discursiva de les llengües als mitjans de comunicació és crucial per diverses raons:

- Les nombroses dimensions de la societat es construeixen, sobretot, a través del discurs. Per tant, la manera en què ens arriba informació sobre les llengües a través dels mitjans de comunicació determinarà com les conceptualitzem.
- El discurs és clau en la (re)producció de les cognicions socials com ara els coneixements, les ideologies, les normes i els valors que compartim com a membres d'un grup i que, al seu torn, regulen i controlen els actes i les interaccions. Per tant, el tipus de missatge que els mitjans de comunicació vehiculen sobre les llengües passarà, en gran part, a formar part de les ideologies lingüístiques de la nostra societat.
- El discurs construeix, constitueix, canvia, defineix i contribueix a formar determinades maneres d'entendre un fet social. Per tant, el tipus de discurs que ens arriba a través dels mitjans de comunicació condicionarà i definirà la nostra manera d'entendre la realitat lingüística. L'anàlisi del discurs, en conclusió, és cabdal per a identificar les representacions i els marcs discursius que s'hi activen.

## 7. Referències bibliogràfiques

- Ambye, Philippei and Richards, Mary. (2012). The paradoxical visions of multilingualism in education: The ideological dimension of discourses on multilingualism in Belgium and Canada. *International Journal of Multilingualism*, 9(2), 163–185.
- Arboix, E. A. (1978). Lenguaje y elecciones: Estudio lexicométrico de programas y manifestos electorales. In *Facultat de Filologia*. Universitat de Barcelona.
- Bambi B. Schieffelin (coord.), Kathryn A. Woolard (coord.), Paul V. Kroskrity (coord.). (1998). *Language Ideologies: Practice and Theory*. Oxford University Press.
- Bell, A. (2014). *The Guidebook to Sociolinguistics* (Wiley Blackwell). Wiley Blackwell.
- Benoit, R. (1982). Le lexique communiste (1932-1946): Description d'une évolution. In D. [et al.] Bonnaud-Lamotte (Ed.), *Actes du 2<sup>o</sup> Colloque de Lexicologie Politique (Saint-Cloud, 15-20 septembre 1980)*. Librairie Klincksieck.
- Benzécri, J. P. (1980). *Pratique de l'Analyse des données*. Dunod.
- Bigey, M. (2007). *Rencontres en séries: Matériel romanesque, lecteurs et lexique ; une analyse du roman sentimental de type sériel (1942-2004)*. Université de Franche-Comté Besançon.
- Bordieu, Pierre. (1981). *El sentido práctico*. Taurus.
- Bordieu, Pierre. (1991). *Sociología y cultura*. Grijalbo.
- Bouzas, Paula. (2016). Análisis del discurso sobre las lenguas del estado español en la prensa hegemónica durante la era Aznar. In *Redes de Comunicación. Estudios sobre la prensa en el mundo hispanohablante* (pp. 157–174). Frank i Timme.
- Brezina, V., Weill-Tessier, P., & McEnery, A. (2020). #LancsBox v. 6.0.0. [Programari].
- Cabré Castellví, M. T. (1977). Lenguajes especiales: Estudio léxico-semántico de los debates parlamentarios [Tesi de doctorat no publicada]. In *Facultat de Filologia*. Universitat de Barcelona.
- Calafat, R. (2021). La ideologia en el discurs de la història de la llengua: El tractament del procés sobiranista. *Caplletra: Revista Internacional de Filologia.*, 71, 237–258.
- Camargo Fernández, L. (2018). El nuevo orden discursivo de la extrema derecha española: De la deshumanización a los bulos en un corpus de tuits de Vox sobre la inmigración. *Cultura, Lenguaje y Representación: Revista de Estudios Culturales de La Universitat Jaume I*, 26, 63–82.
- Castellà, J. M. (1992). *De la frase al text*. Empúries.
- Centre de Lexicometria. Divisió de Ciències Humanes i Socials. (1993). *La Lexicometria*. PPU.
- de Bazin de Bezons, J. (1967). *Vocabulaire de «La Princesse de Clèves»*. Nizet.
- Demonet, M., Geffroy, A., Gouazé, J., Lafon, P., Mouillaud, M., & Tournier, M. (1975). *Des tracts en mai 68. Mesures de vocabulaire et de contenu*. Armand Colin et Presses de la Fondation Nationale des Sciences Politiques.

- [GEC] Enciclopèdia catalana. (2022). Lingüística de corpus. In [GEC] *Gran enciclopèdia catalana*. Enciclopèdia catalana.
- Factiva. (n.d.). <https://search.library.wisc.edu/catalog/999795312502121>
- Fairclough, Norman. (1989). *Critical Discourse Analysis: The critical study of language*. Longman.
- Fairclough, Norman. (2003). *Analysing Discourse*. Routledge.
- Fernández Moreno, L. (1998). *L'Espagne à la Une du Monde: 1969-1985: Le discours médiatique et ses filtres interculturels*. L'Harmattan.
- Flors-Mas, A., Nussbaum, L., Sabaté-Dalmau, M., Solé i Mireia, M., & Trenchs Parera, M. (2022). *Síntesi de les reflexions del cicle sobre el model lingüístic de l'educació obligatòria a Catalunya*. 32, 311–329.
- Gal, S. (2018). Visions and Revisions of Minority Languages. In P. Lane, J. Costa, & H. De Korne (Eds.), *Standardizing minority languages: Competing ideologies of authority and authenticity in the global periphery* (First edition, pp. 222–242). Routledge.
- Gal, Susan. (2019). *Signs of Difference: Language and Ideology in Social Life*. Cambridge University Press.
- Gifreu, J. (2018). *El quart poder i la independència. La batalla mediàtica per l'1 d'octubre*. Angle Editorial.
- Grasset, E. (2011). *Modernitat i canvi de llengua el pas del castellà al català a l'obra de Pere Gimferrer (aspectes crítics, teòrics i lexicomètrics)* [Tesi de doctorat no publicada]. Universitat de Girona.
- Guiraud, P. (1953). *Langage et versification d'après l'œuvre de Paul Valéry*. Klincksieck.
- Guiraud, P. (1960). *Problèmes et méthodes de la statistique linguistique*. Presses Universitaires de France.
- Heller, Monica. (2011). *Paths to postnationalism. A critical ethnography of language and identity*. Oxford University Press.
- Jaworska, S., T., C. (2018). Public discourses on multilingualism in the UK: Triangulating a corpus study with a sociolinguistic attitude survey. *Language in Society*, 47(1), 57–88. <https://doi.org/10.1017/S0047404517000744>
- Lakoff, G. (2004). *No pienses en un elefante. Lenguaje y debate político*. Editorial Complutense.
- Lebart, L., & Salem, A. (1994). *Statistique textuelle*. Dunod.
- Lleal Galceran, C. (1980). Análisis del vocabulario político a través de una década de panfletos. In *Facultat de Filologia*. Universitat de Barcelona.
- Marimón Llorca, C. (2021). “La maldición de Babel”. Crónicas periodísticas del nacionalismo lingüístico español. *Revista de Llengua i Dret*, 76, 79–96. <https://doi.org/10.2436/rld.i76.2021.3668>
- Martin Rojo, Luisa; Ruiz, Elvira. (2019). EDiSo ante los desafíos discursivos de nuestro tiempo: Herramientas para neutralizar el fascismo universal”. *Sesión Plenaria Colectiva*.

*IV Simposio Internacional EDiSo. Veus, Silencis i Silenciaments En Els Estudis Del Discurs.*

Martínez de Luna, I. (2013). *Euskarak duen Framing Berri baten premia asetzeko proposamena*. Bat Soziolinguistika aldizkaria.

Martinez, W. (2003). Contribution à une méthodologie de l'analyse des cooccurrences lexicales multiples dans les corpus textuels [Tesi de doctorat no publicada]. In *École de langage et langues*. Université Paris III–Sorbonne Nouvelle.

Meyer, C. F. (2002). *English corpus linguistics: An introduction*. Cambridge University Press.

Morales Moreno, A. (2015). Estudi lexicomètric del procés de redacció de l'Estatut d'Autonomia de Catalunya (2006) [Tesi doctoral no publicada]. In *Departament de Traducció i Ciències del Llenguatge*. Universitat Pompeu Fabra.

Morales Moreno, A. (2018). *Estudi lexicomètric contrastiu dels estatuts d'autonomia de Catalunya de 1932, 1979 i 2006*. Generalitat de Catalunya. Institut d'Estudis de l'Autogovern.

Morales Moreno, A. (2021). Un any de lluita discursiva i legislativa contra la COVID-19 (2020-2021). *Allò que hem après de la COVID-19: informe de l'Institut d'Estudis Catalans*, 175–184. <https://doi.org/10.2436/15.2000.69.30>

Muller, C. (1967). *Étude de statistique lexicale, le vocabulaire du théâtre de Pierre Corneille*. Larousse.

Muller, C. (1974). La lemmatisation. Essai d'analyse mathématique. *Travaux de Linguistique et de Littérature*, XII, 189–208.

Muller, P. (1984). Preface. De la lemmatisation. In P. Lafon (Ed.), *Dépouillements et statistique en lexicométrie*. Slatkine.

Niklas-Salminen, A. (1997). *La lexicologie*. Armand Colin/Masson.

Ninyoles, R.LL. (1971). *Idioma i prejudici*. Editorial Moll.

Piller, I; Cho, J. (2013). Neoliberalism as language policy. *Language in Society*, 42(1), 23–44.

Reig, R. (2011). *Los dueños del periodismo: Claves de la estructura mediática mundial en España* (1. ed). Editorial Gedisa.

Sabaté-Dalmau, Maria. (2020). Neoliberal language policies and linguistic entrepreneurship in Higher Education: Lecturers' perspectives. *Language, Culture and Society*, 2(2), 175–197.

Salem, A. (1993). De travailleurs à salariés. Repères pour une étude de l'évolution du vocabulaire syndical (1971-1990). *Mots. Les Langages Du Politique*, 36, 74–83.

Salem, A., Cabré, M. T., & Romeu, L. (1990). *Vocabulari de la lexicometria: Català, castellà, francès*. Centre de Lexicometria, Divisió de Ciències Humanes i Socials.

Savoy, J. (2014). La voix du Président américain (1934-2014). In E. Née, J.-M. Daube, M. Valette, & S. Fleury (Eds.), *Proceedings-Actes. 12th International Conference on*

- Textual Data Statistical Analysis. 12es Journées Internationales d'Analyse Statistique des Données Textuelles. Paris. June 3-6, 2014. 3-6 Juin 2014* (pp. 593–604). JADT.org.
- Schieffelin, Bambi B., Kathryn A. Woolard i Paul V. Kroskrity (Eds.). (2012). *Ideologías lingüísticas: Práctica y teoría*. Los libros de la Catarata.
- Stefanowitsch, Anatol. (2020). *Corpus linguistics: A guide to the methodology*. Zenodo. <https://doi.org/10.5281/ZENODO.3735822>
- Van Dijk, T. (1998). *Ideology. A Multidisciplinary approach*. Sage Publications.
- Wagner, R. L. (1949). *Clitandre. Tragi-Comedie. Edition du texte de 1632 avec les variantes et un lexique*. Librairie Droz.
- Woolard, Katherine. (2016). Shifting Discourses of Language in Catalan Politics and Media. In *Singular and Plural: Ideologies of Linguistic Authority in 21st Century Catalonia* (pp. 97–117). Oxford University Press.



## 8. Annexos

### Annex 1. Llistats de freqüències

Els arxius del corpus i les matrius de dades són a [aquesta carpeta de Google Drive](#).

## Annex 2. Resultats de l'anàlisi de les macroestructures

Descriptor (macroestructura)	CA	ES	Total
immersió (lingüística)	28	32	60
ús	38	2	40
llengua	30	8	38
escola	16	15	31
controvèrsia		28	28
educació/ensenyament	5	21	26
normalització (lingüística)	23	2	25
imposició		24	24
Estat	24		24
llengua vehicular	7	15	22
atac/atacar	16	4	20
castellà/espanyol	8	12	20
discriminació	3	16	19
polèmica	18		18
utilitat	5	12	17
literatura	12	4	16
dret(s) lingüístic(s)	12	2	14
(Llei) Celaá	3	11	14
lleis	10	3	13
audiovisual	12		12
bilingüisme	8	4	12
nacionalisme	2	9	11

oportunitat	2	9	11
internacionalització	6	5	11
TV3	9	2	11
cultura	4	7	11
minorització	11		11
llengua catalana	10		10
anglès	1	8	9
sentència		9	9
castellanisme	8	1	9
coneixement	5	3	8
oficialitat (lingüística)	5	3	8
protecció	7	1	8
aprenentatge	4	4	8
model (educatiu)	7	1	8
integració	5	2	7
immigració	5	2	7
tribunal	2	5	7
Espanya	7		7
promoció	6	1	7
independentisme	1	6	7
català	6	1	7
ortografia	4	2	6
Koiné	2	4	6
radicalisme		6	6

separatisme		6	6
vitalitat (lingüística)	6		6
Ciutadans/Ciudadanos	5	1	6
justícia/administració de justícia	6		6
desaparició	5	1	6
premi	4	1	5
divulgació	5		5
Plataforma per la Llengua	5		5
obligació/obligatorietat	1	4	5
Generalitat	2	3	5
model lingüístic	1	4	5
multilingüisme	4	1	5
expansió		5	5
Partit Popular	5		5
irrellevant	1	4	5
greuge	4		4
desús	4		4
mitjans audiovisuals/de comunicació	4		4
Lleida	4		4
llengua estrangera	3	1	4
convivència		4	4
administració	3	1	4
normativa	4		4
desintegració		4	4

regionalisme		4	4
absència	4		4
cohesió (social)	4		4
155	2	2	4
inutilitat		4	4
progrés		3	3
tecnologia	1	2	3
sectarisme		3	3
correcció	3		3
Països Catalans	1	2	3
democràcia	2	1	3
reconeixement	2	1	3
descens	3		3
societat	3		3
funció pública		3	3
valor		3	3
idioma		3	3
plurilingüe/plurilingüisme	3		3
campanya	2	1	3
Ramon Llull (Institut)	3		3
adoctrinament	1	2	3
resultat	1	2	3
llengua pròpia	3		3
confrontació	2	1	3

manifestació	2	1	3
suport	3		3
Catalunya		3	3
universal/universalitat		3	3
Cervantes		3	3
cinema	2	1	3
menyspreu	3		3
unitat	2		2
presència	2		2
novel·la	2		2
àrab	1	1	2
religió	1	1	2
dialectologia	2		2
monolingüisme		2	2
associacionisme	2		2
competència		2	2
doblatge	2		2
condició		2	2
dret	2		2
tancament	2		2
escriptor	1	1	2
marginació	1	1	2
esport	2		2
norma	2		2

esquerra		2	2
pedagogia	2		2
Estatut (d'autonomia)	1	1	2
política		2	2
etnicisme		2	2
defensa	2		2
Euskadi		2	2
regla	1	1	2
exclusió		2	2
salut	1	1	2
foment	2		2
transmissió intergeneracional	2		2
francès	1	1	2
ús social	2		2
fraseologia	2		2
catalanització	1	1	2
gramàtica	1	1	2
nombre parlants	2		2
batalla		2	2
normalitat	2		2
història	1	1	2
oposició		2	2
IEC	2		2
perill	2		2

independència	2		2
compliment		2	2
bilingüïtzació	1	1	2
politització	1	1	2
blindatge	1	1	2
préstec	2		2
intolerància	1	1	2
qualitat	2		2
Bressola	2		2
recuperació	2		2
lectura	2		2
regressió	2		2
llengua comuna		2	2
respecte	2		2
llengua en perill	2		2
conflicte	1	1	2
llengua oficial	2		2
Corona	1	1	2
crisi	1	1	2
trilingüisme	2		2
llibertat		2	2
unitat lingüística	2		2
manifest		2	2
croada	1	1	2



música		2	2
forma	1		1
descomposició		1	1
gallec	1		1
homogeneïtat lingüística		1	1
plurinacional	1		1
identitat	1		1
acreditació	1		1
ideologia		1	1
aplicació		1	1
correcció política	1		1
xarnego		1	1
ERC	1		1
compromís	1		1
beca	1		1
raça		1	1
Bíblia		1	1
reivindicació	1		1
immigrants	1		1
socialisme		1	1
importància		1	1
Consell per la República Catalana	1		1
Bildu		1	1
desconeixement		1	1

incomprensió		1	1
persecució política	1		1
errada		1	1
política lingüística	1		1
administració estatal	1		1
presos		1	1
inexistència	1		1
pronúncia	1		1
informació	1		1
reciprocitat	1		1
Inquisició		1	1
convivència lingüística	1		1
Institut Cervantes		1	1
requisit lingüístic	1		1
Instituto Cervantes	1		1
barrera		1	1
abandonament	1		1
subordinació lingüística	1		1
integrisme		1	1
televisió	1		1
escola catalana	1		1
UE	1		1
escolarització	1		1
hispanofòbia		1	1

Bombers		1	1
Constitució		1	1
jerarquia lingüística	1		1
pelegrí		1	1
dictadura lingüística		1	1
èxit	1		1
Camí de sant Jaume		1	1
barreja	1		1
LAPAO	1		1
feblesa	1		1
Laura Borràs	1		1
preocupació	1		1
Església	1		1
problema		1	1
lingüística computacional		1	1
projecció	1		1
cosmovisió		1	1
concurs		1	1
literatura infantil	1		1
dèficit	1		1
canvi lingüístic		1	1
demanda	1		1
lleis d'educació estatal	1		1
anglicisme		1	1

casella	1		1
Reial Madrid	1		1
CPNL	1		1
repressió	1		1
llengua	1		1
demanda política	1		1
afebliment	1		1
cooperació	1		1
espai comunicatiu compartit	1		1
situació sociolingüística	1		1
automoció	1		1
Societat Civil Catalana	1		1
llengua espanyola		1	1
subtitulació	1		1
castellanització		1	1
gestió	1		1
llengua minoritzada espanyola	1		1
transició	1		1
difusió	1		1
Consell Social	1		1
llengua optativa		1	1
enverinament		1	1
assaig	1		1
consellera	1		1

llengua universal		1	1
utilització	1		1
Alta Cerdanya	1		1
Vic	1		1
llengües cooficials		1	1
Vox		1	1
essencialisme		1	1
xifra	1		1
llibre	1		1
diccionari	1		1
magistratura	1		1
plataforma (audiovisual)	1		1
estadística	1		1
decret plurilingüisme	1		1
Babel		1	1
poesia	1		1
manlleu	1		1
família		1	1
estat espanyol	1		1
política: història	1		1
Menorca	1		1
predomini		1	1
constitució catalana		1	1
premi Planeta	1		1

Congrés de la Llengua Espanyola	1		1
festival	1		1
Tabàrnia		1	1
dictadura		1	1
minoria lingüística	1		1
professorat	1		1
catalanisme	1		1
prohibició		1	1
estat plurilingüe	1		1
comunitat		1	1
mobilitat		1	1
propaganda		1	1
Any Pompeu Fabra	1		1
fracàs		1	1
censura		1	1
França	1		1
diversitat		1	1
rebuig	1		1
Mos Movem		1	1
recompensa		1	1
multiculturalitat	1		1
recull	1		1
CUP		1	1
referèndum d'autonomia	1		1

accès	1		1
audiollibre	1		1
nació	1		1
rei		1	1
davallada	1		1
reinterpretació		1	1
nacisme		1	1
futur	1		1
negociació	1		1
requeriment	1		1
nivell	1		1
Galícia		1	1
estranger		1	1
rupturisme		1	1
estudi		1	1
sanitat	1		1
ETA	1		1
seguretat alimentària		1	1
normalitat	1		1
denúncia	1		1
Amèrica Llatina		1	1
sobre nacionalismes	1		1
code switching	1		1
desafiament	1		1

notaria	1		1
software	1		1
Nova Planta	1		1
substitució lingüística	1		1
diversitat lingüística	1		1
Súper 3	1		1
noves tecnologies	1		1
supremacisme	1		1
codi de consum		1	1
govern	1		1
observatori lingüístic	1		1
traductor	1		1
ocupabilitat	1		1
emergència lingüística	1		1
accent	1		1
gravetat		1	1
Òmnium	1		1
Ucraïna	1		1
opció lingüística		1	1
unió		1	1
Colom (Cristòfor)		1	1
hispanitat		1	1
dialecte	1		1
universitat	1		1



opressió	1		1
ús decreixent	1		1
colpista		1	1
consens	1		1
país	1		1
valencià		1	1
país valencià	1		1
vertebració		1	1
decret		1	1
vigor		1	1
pancatalanisme		1	1
Voluntaris per la Llengua	1		1
parlant	1		1
vulnerabilitat	1		1
comerç	1		1
xenofòbia		1	1
patrimoni		1	1
equiparació	1		1
euskera	1		1
aniquilació	1		1
mestissatge		1	1
anomalia	1		1
millora	1		1
Total	742	521	1263



### Annex 3. Els 25 noms propis més freqüents

	Lemma	Freq_abs	Freq_rel		Lemma	Freq_abs	Freq_rel
1	<i>catalunya</i>	256	26,4		<i>cataluña</i>	272	45,4
2	<i>barcelona</i>	76	7,8		<i>gobierno</i>	155	25,9
3	<i>govern</i>	74	7,6		<i>españa</i>	144	24,0
4	<i>tv3</i>	62	6,4		<i>pp</i>	143	23,9
5	<i>generalitat</i>	62	6,4		<i>generalitat</i>	128	21,4
6	<i>pp</i>	55	5,7		<i>celaá</i>	111	18,5
7	<i>jordi</i>	48	4,9		<i>psoe</i>	107	17,9
8	<i>parlament</i>	43	4,4		<i>baleares</i>	89	14,9
9	<i>espanya</i>	39	4,0		<i>cervantes</i>	87	14,5
10	<i>plataforma</i>	38	3,9		<i>barcelona</i>	80	13,4
11	<i>consell</i>	34	3,5		<i>rae</i>	80	13,4
12	<i>ciutadans</i>	33	3,4		<i>govern</i>	71	11,9
13	<i>josep</i>	33	3,4		<i>cs</i>	62	10,4
14	<i>sant</i>	32	3,3		<i>sánchez</i>	60	10,0
15	<i>torra</i>	32	3,3		<i>madrid</i>	56	9,4
16	<i>illes</i>	31	3,2		<i>rey</i>	50	8,4
17	<i>balears</i>	30	3,1		<i>psc</i>	39	6,5
18	<i>fabra</i>	29	3,0		<i>tv3</i>	38	6,3
19	<i>catalana</i>	29	3,0		<i>valencia</i>	37	6,2
20	<i>llull</i>	27	2,8		<i>sevilla</i>	33	5,5
21	<i>psoe</i>	27	2,8		<i>justicia</i>	32	5,3

<b>22</b>	<b>erc</b>	26	<b>2,7</b>	<b>puigdemont</b>	27	<b>4,5</b>
<b>23</b>	<b>madrid</b>	25	<b>2,6</b>	<b>tsjc</b>	25	<b>4,2</b>
<b>24</b>	<b>ramon</b>	25	<b>2,6</b>	<b>américa</b>	25	<b>4,2</b>
<b>25</b>	<b>cs</b>	25	<b>2,6</b>	<b>europa</b>	25	<b>4,2</b>

## Annex 4. Els 25 verbs més freqüents

	verb	Freq_abs	Freq_rel		verb	Freq_abs	Freq_rel
1	<i>haver</i>	929	95,6		<i>ser</i>	514	85,8
2	<i>fer</i>	764	78,6		<i>haber</i>	197	32,9
3	<i>tenir</i>	722	74,3		<i>pedir</i>	162	27,1
4	<i>dir</i>	651	67,0		<i>hablar</i>	145	24,2
5	<i>parlar</i>	614	63,2		<i>poder</i>	120	20,0
6	<i>veure</i>	530	54,6		<i>tener</i>	111	18,5
7	<i>ser</i>	487	50,1		<i>hacer</i>	110	18,4
8	<i>donar</i>	464	47,8		<i>querer</i>	89	14,9
9	<i>passar</i>	455	46,8		<i>dar</i>	82	13,7
10	<i>voler</i>	382	39,3		<i>denunciar</i>	76	12,7
11	<i>caldre</i>	349	35,9		<i>ir</i>	71	11,9
12	<i>estar</i>	347	35,7		<i>ver</i>	67	11,2
13	<i>viure</i>	306	31,5		<i>exigir</i>	65	10,9
14	<i>recordar</i>	289	29,7		<i>imponer</i>	65	10,9
15	<i>considerar</i>	258	26,6		<i>estar</i>	64	10,7
16	<i>permetre</i>	257	26,5		<i>decir</i>	62	10,4
17	<i>treballar</i>	224	23,1		<i>defender</i>	59	9,9
18	<i>servir</i>	224	23,1		<i>llegar</i>	50	8,4
19	<i>arribar</i>	222	22,9		<i>aprobar</i>	48	8,0
20	<i>posar</i>	221	22,7		<i>poner</i>	47	7,8

<b>21</b>	<b>conèixer</b>	217	22,3	<b>usar</b>	47	7,8
<b>22</b>	<b>rebre</b>	210	21,6	<b>abrir</b>	43	7,2
<b>23</b>	<b>quedar</b>	200	20,6	<b>presentar</b>	43	7,2
<b>24</b>	<b>explicar</b>	184	18,9	<b>deber</b>	41	6,8
<b>25</b>	<b>venir</b>	166	17,1	<b>estudiar</b>	41	6,8

## Annex 5. Llistat de taules

Taula 1. Dades quantitatives del "Proyecto Lengua y Prensa" (Universidad de Málaga).....	14
Taula 2. Diaris analitzats .....	18
Taula 3. Fonts disponibles en català i castellà d'àmbit espanyol ( <i>Factiva</i> ) .....	19
Taula 4. Els 10 diaris més llegits en català (Font: BOP – Dades 2020 (3a onada)).....	20
Taula 5. Els 10 diaris més llegits en castellà (Font: EGM – Dades 2020 (3a onada)).....	21
Taula 6. Operadors booleans emprats per a obtenir les dades ( <i>Factiva</i> ) .....	21
Taula 7. Distribució dels articles obtinguts amb els operadors booleans ( <i>Factiva</i> ) .....	23
Taula 8. Distribució dels articles pertinents per a l'anàlisi.....	26
Taula 9. Documents i còmput de paraules dels articles pertinents per a l'anàlisi.....	27
Taula 10. Articles filtrats (còmput de paraules decreixent) .....	28
Taula 11. Dades quantitatives del corpus de l'anàlisi qualitativa .....	29
Taula 12. Dades lexicomètriques del corpus de titulars ( <i>#LancsBox</i> ).....	38
Taula 13. Els 25 substantius més freqüents (corpus de titulars).....	39
Taula 14. Comparació de freqüències relatives ( <i>#LancsBox</i> ).....	39
Taula 15. Coocurrències lèxiques de <i>lleí</i> i <i>ley</i> ( <i>#LancsBox</i> ).....	43
Taula 16. Lemes més freqüents en català i castellà (corpus textual).....	45
Taula 17. Sintagmes en què es documenta <i>llengua</i> ( <i>#LancsBox</i> ) .....	52
Taula 18. Sintagmes de <i>llengua</i> documentats només en CA o ES.....	52
Taula 19. Llengües a què fan referència les concordances.....	54
Taula 20. Contextos d'ús de <i>monolingüisme</i> , <i>bilingüisme</i> , <i>trilingüisme</i> , <i>multilingüisme</i> i <i>plurilingüisme</i> .....	57
Taula 21. Llengües a què fan referència <i>monolingüisme</i> , <i>bilingüisme</i> , <i>trilingüisme</i> , <i>multilingüisme</i> i <i>plurilingüisme</i> .....	59
Taula 22. Els 25 adjectius més freqüents en català i en castellà ( <i>#LancsBox</i> ) .....	61
Taula 23. Sintagmes que inclouen <i>lingüístic -a</i> (i adjectius derivats) ( <i>#LancsBox</i> ) .....	70
Taula 24. Classificació desglossada de les concordances de <i>lingüístic -a</i> .....	74
Taula 25. Els 25 sintagmes nominals més freqüents ( <i>#LancsBox</i> ).....	75
Taula 26. Freqüències de les macroestructures identificades .....	80
Taula 27. Llengua o llengües a què fan referència els contextos .....	82
Taula 28. Àmbit geogràfic a què fan referència els textos .....	85

## Annex 6. Llistat de figures

Figura 1. Interfície ( <i>Factiva</i> ) .....	22
Figura 2. Interfície de resultats i opcions d'exportació ( <i>Factiva</i> ) .....	29
Figura 3. Exemple de neteja d'un text ( <i>Ara</i> ) .....	31
Figura 4. Exemple d'un text amb errors de codificació ( <i>El Periódico</i> ) .....	32
Figura 5. Concordances del lema <i>any</i> ( <i>#LancsBox</i> ) .....	40
Figura 6. Coocorrents de <i>lleí</i> ( <i>#LancsBox</i> ) .....	43
Figura 7. Coocorrents de <i>ley</i> ( <i>#LancsBox</i> ) .....	43
Figura 8. Contextos d'ús de <i>primer -a</i> ( <i>#LancsBox</i> ).....	62
Figura 9. Contextos d'ús de <i>constitucional</i> ( <i>#LancsBox</i> ) .....	63
Figura 10. Contextos d'ús de <i>casilla</i> ( <i>#LancsBox</i> ).....	71
Figura 11. Eix de valors dicotòmics (Gal 2018, 233) .....	93



## Annex 7. Llistat de gràfiques

Gràfica 1. Distribució d'articles per diari (dades sense filtrar).....	24
Gràfica 2. Distribució cronològica d'articles per diari (dades sense filtrar).....	24
Gràfica 3. Distribució d'articles per diari (dades filtrades).....	26
Gràfica 4. Distribució cronològica d'articles per diari (dades filtrades).....	28
Gràfica 5. Distribució cronològica de les concordances de <i>lleí</i> i <i>ley</i> .....	41
Gràfica 6. Classificació temàtica de les concordances de <i>lingüístic -a</i> .....	73