

C) LA POLÍTICA I LA PLANIFICACIÓ LINGÜÍSTIQUES

LA INTERFERÈNCIA EN LA INNOVACIÓ LÈXICA DEL CATALÀ: ANÀLISI DELS NEOLOGISMES AMB CIBER- I HOMO-

Judit Freixa Aymerich*

Resumen al final del artículo.

The abstract can be found at the end of the article.

Aquest article pretén donar resposta a la pregunta següent: fins a quin punt els neologismes que incorpora actualment el català són fruit dels seus propis mecanismes d'innovació lèxica? L'anàlisi, centrada en totes les unitats neològiques formades amb *ciber-* i *homo-* recollides a l'Observatori de Neologia durant els últims anys, es planteja el grau de dependència que el català presenta respecte del castellà. Aquesta línia d'investigació vol ser d'utilitat per a la planificació lingüística perquè detectar i conèixer les vies d'interferència d'una llengua és imprescindible per activar les mesures correctores que es considerin necessàries.

Paraules clau: neologia; morfologia lèxica; prefixació; interferència (català, castellà/espanyol); planificació lingüística.

* Judit Freixa Aymerich, Observatori de Neologia, Institut Universitari de Lingüística Aplicada. Universitat Pompeu Fabra. judit.freixa@upf.edu.

Data de recepció de l'article: 1.07.2008; data d'avaluació: 30.12.2008; data d'acceptació de la versió final: 12.01.2009.

Sumari

1. Introducció
2. Antecedents
3. Objectius i hipòtesis
4. Anàlisi de dades
 - 4.1. Neologismes amb el formant *homo-*
 - 4.2. Neologismes amb el formant *ciber-*
 - 4.3. Primers indicadors d'(in)dependència aïllats
5. Conclusions
6. Bibliografia

1. Introducció

Durant els últims vint anys el català ha incorporat una mitjana anual de gairebé 3.000 neologismes, segons els resultats de l'Observatori de Neologia (Obneo),¹ i a partir d'aquesta dada es podria afirmar que el català té una gran vitalitat lèxica. Hi ha, però, diversos aspectes que s'han de tenir en compte en la interpretació d'aquests resultats: el primer, què s'entén per *neologisme* i per *incorporar* en la frase del començament.

En el marc de l'Observatori, *incorporar* vol dir que una unitat s'ha documentat en textos catalans (bàsicament en textos de premsa d'àmplia difusió), i cal aclarir que un percentatge molt elevat d'aquests neologismes catalans més que *incorporar-se* a la llengua, s'han utilitzat en una ocasió. Res no treu, però, que aquestes unitats existeixen i han estat documentades i, doncs, aquest fet parla positivament de la capacitat del català per crear paraules noves. D'altra banda, *neologisme* vol dir, per a l'Observatori, «unitat lèxica no documentada en el *Diccionari de la llengua catalana* de l'IEC i el *Gran Diccionari de la Llengua Catalana* d'Enciclopèdia Catalana», i sabem que alguns d'aquests neologismes només ho són des del punt de vista lexicogràfic, ja que no representen cap mena de novetat a la llengua: unitats que no apareixen als diccionaris perquè són interferències del castellà, o bé totalment predictibles des d'un punt de vista morfològic i semàntic, etc.

En aquest moment no sabem quin percentatge del total de neologismes lexicogràfics es poden considerar realment unitats noves; no podem dir a partir de

1. L'Observatori de Neologia és un projecte de recerca de l'Institut Universitari de Lingüística Aplicada de la Universitat Pompeu Fabra.

quan podem considerar que una unitat està incorporada a la llengua (és a dir, quan deixa de ser nova) i tampoc fer afirmacions categòriques sobre què està passant en el conjunt de la llengua perquè els corpus textuais orals i escrits d'on s'han extret els neologismes no reflecteixen la totalitat dels usos i, doncs, per exemple, no tenim dades sistemàtiques sobre la innovació lèxica dels joves en usos espontanis o dels especialistes en les seves comunicacions científiques. No obstant això, l'Obneo ofereix el corpus de neologismes més important de què disposem; un corpus que, tot i que no reflecteix la totalitat d'usos lingüístics, sí que permet fer estudis fiables i sistemàtics sobre la neologia lèxica del català.

En aquest article volem referir-nos a un aspecte que caldria estudiar a fons per poder mesurar el valor real de l'afirmació inicial sobre la vitalitat lèxica del català. La pregunta a la qual volem donar resposta és la següent: fins a quin punt els neologismes que incorpora actualment el català són fruit dels seus propis mecanismes d'innovació lèxica? I fins a quin punt són deutors de la interferència amb altres llengües? Pensem que aquesta línia d'investigació ha de tenir utilitat per a la planificació lingüística en el sentit que detectar i conèixer les vies d'interferència és imprescindible per activar les mesures correctores que es considerin necessàries.

2. Antecedents

Des de l'any 1988, l'Observatori de Neologia recull neologismes en català i en castellà amb un objectiu principal d'actualització dels diccionaris generals. El resultat d'aquests vint anys de treball² és un corpus de dades neològiques constituït per 50.175 neologismes per al català i 35.575 per al castellà consultable i explotable amb el cercador OBNEO (<http://obneo.iula.upf.edu/bobneo/index.php>).

El treball que avui presentem té l'antecedent més clar en l'article «Anàlisi contrastiva de la innovació lèxica en català i en castellà» publicat l'any 2001 a la revista *Caplletra* i elaborat per M. T. Cabré, J. Freixa i E. Solé en el marc de l'Observatori de Neologia. El supòsit d'aquell treball era determinar l'opció més defensible per a cadascun dels mecanismes d'innovació (creació del cata-

2. Els plantejaments inicials del projecte s'han mantingut fins a l'actualitat, si bé els canvis que ha viscut l'Obneo al llarg d'aquests anys han estat nombrosos. Per exemple, els textos dels quals s'extreuen els neologismes eren inicialment textos escrits de gènere periodístic editats a Barcelona i actualment es buiden també textos no periodístics (preferentment textos que no han passat per un procés de revisió lingüística), textos d'altres dialectes catalans, textos orals, etc.

là a partir dels seus propis recursos, còpia del castellà, còpia de l'anglès, etc.) i les dades analitzades van ser tots els neologismes de l'any 1998 per al català (3.932) i per al castellà (3.946). Els resultats obtinguts es poden resumir de la manera següent:

- El contrast percentual de processos i recursos (dins dels processos distingim la neologia morfològica, la sintàctica, la semàntica i els manlleus, i dins dels recursos es tenen en compte diferents tipus dins de cada procés; en el cas de la neologia morfològica, per exemple, es distingeix la derivació, la composició, la sintagmació, etc.) va permetre constatar una semblança màxima entre totes dues llengües. Aquest resultat, però, no té una única lectura ja que la màxima semblança pot ser un indicatiu que una llengua copia l'altra, però en dues llengües tan pròximes tipològicament no és prudent ni pertinent extreure directament aquesta conclusió.
- El contrast de la innovació en determinades regles que es van considerar més innovadores va permetre observar la coincidència dels fenòmens però no de les unitats i també aquest resultat es pot interpretar de maneres diferents.
- El contrast de la neologia semàntica va permetre constatar que coincideixen el 15% de les unitats, resultat que podria interpretar-se com la demostració d'un funcionament independent de les dues llengües.
- Finalment, el contrast de la neologia per manlleu va demostrar que català i castellà presenten una coincidència del 50% aproximadament pel que fa a manlleus de l'anglès i del francès. Els resultats quant als percentatges de manlleus coincidents d'altres llengües en català i en castellà van resultar poc similars i, per tant, apareix una vegada més el dubte sobre la interpretació que cal fer dels resultats.

A la vista d'aquestes dades, la conclusió final del treball fou que cal explicar els neologismes a partir del català perquè «no només és possible sinó que és plausible». Però la conclusió implícita que plana en tot el treball és que en realitat no tenim elements per interpretar quan les llengües creen independentment i quan són més o menys deutes de la interferència d'altres llengües.

3. Objectius i hipòtesis

El treball que presentem forma part d'un projecte que pretén avaluar el pes que el castellà, l'anglès i el conjunt de llengües romàniques pugui representar en la innovació lèxica del català, tenint en compte la globalització del nostre entorn i

la particular situació de contacte de llengües. Coneixem el percentatge exacte de manlleus (i la llengua de procedència) que l'Observatori de Neologia detecta actualment en els textos de buidatge catalans, i sabem que s'aproxima al 20%, però no sabem quin percentatge del 80% restant és creació pròpia del català o és també el resultat de la innovació lèxica d'altres llengües. Per obtenir respostes fiables, haurem d'analitzar a fons els diferents mecanismes que utilitza el català per innovar el lèxic i les unitats que en resulten; i comparar aquests mecanismes i aquestes unitats amb els que estan posant en joc les altres llengües.

La hipòtesi d'aquesta recerca és que es poden aïllar indicis sobre la interferència en la innovació lèxica que permetin avaluar el grau de dependència o d'independència d'una llengua respecte d'una altra o unes altres, i per dur a terme la recerca ens proposem d'explorar els indicadors potencials d'interferència amb l'anàlisi de la semblança o la coincidència que el català presenta en la seva innovació lèxica amb el castellà i altres llengües (tant pel que fa a les regles i als mecanismes de formació de paraules, com a les unitats concretes). Volem arribar a aïllar elements que es puguin considerar efectivament *indicadors d'(in)dependència*. La idea fonamental és que no tot allò que sigui semblant o idèntic ha de ser considerat índex de dependència ja que les llengües (especialment aquelles que són tipològicament pròximes) comparteixen mecanismes per innovar el lèxic i, per què no, també poden compartir moltes unitats.

En el marc d'aquest projecte de recerca presentem ara aquest article que té per objectiu l'anàlisi de les unitats formades amb *ciber-* i *homo-*, és a dir, unitats formades amb dues de les regles més innovadores de formació de mots. El contrast es realitza només entre el català i el castellà, i més endavant s'estendrà cap a les altres llengües romàniques i l'anglès. L'anàlisi que fem té un enfocament més sociolingüístic que no pas morfològic, i les raons per les quals hem seleccionat precisament les regles de formació de mots amb *ciber-* i *homo-* són diverses: l'aparició de nous formants (també anomenats *pseudoformants*, *prefixoides*, etc.) ha estat sovint considerada com un dels fenòmens més innovadors de la llengua des d'un punt de vista morfològic, i precisament ens interessa centrar-nos en aquells aspectes en què les llengües estan manifestant la seva vitalitat. Ens calia, però, que fossin regles per a les quals tinguéssim una quantitat acceptable d'exemples, i que haguessin aparegut a la llengua abans de l'any 2000 per poder tenir també una visió diacrònica del fenomen.³

3. Disposem de diversos estudis sobre l'aparició de nous formants en català: FELIU *et al.* (2000), RULL (2001), VALLÈS (2002), entre d'altres.

4. Anàlisi de dades

Aquesta recerca concreta es basa en els neologismes catalans i castellans detectats a l'Observatori de Neologia durant els últims 18 anys (període del 1990 al 2007 inclosos). Es tracta exactament d'un total de 48.190 neologismes diferents per al català (119.209 ocurrences) i 34.207 per al castellà (75.547 ocurrences). És important distingir els *neologismes* i les *ocurrences*: com que l'Observatori recull el mateix neologisme cada vegada que el documenta en una nova font de buidatge, el nombre de neologismes (diferents, és a dir, *type*) sempre serà menor que el nombre d'ocurrences (neologismes amb repeticions, és a dir, *token*). A la taula següent pot veure's com es distribueixen al llarg dels anys les ocurrences per al català i el castellà:

TAULA 1. Distribució anual de les ocurrences dels neologismes per al català i el castellà

| <i>Any</i> | <i>Total d'ocurrences dels neologismes catalans</i> | <i>Total d'ocurrences dels neologismes castellans</i> |
|------------|---|---|
| 2007 | 5.687 | 4.110 |
| 2006 | 6.973 | 5.112 |
| 2005 | 8.381 | 6.556 |
| 2004 | 6.739 | 5.230 |
| 2003 | 6.439 | 1.057 |
| 2002 | 7.027 | 2.170 |
| 2001 | 4.078 | 3.606 |
| 2000 | 5.839 | 5.019 |
| 1999 | 7.316 | 2.286 |
| 1998 | 5.814 | 5.655 |
| 1997 | 7.127 | 7.188 |
| 1996 | 7.555 | 5.099 |
| 1995 | 10.965 | 3.215 |
| 1994 | 7.221 | 6.159 |
| 1993 | 6.051 | 3.741 |
| 1992 | 5.164 | 1.464 |
| 1991 | 7.987 | 5.058 |
| 1990 | 2.846 | 2.822 |
| Total | 119.209 | 75.547 |

Com es pot veure a la taula, la quantitat d'unitats recollides per al català és sempre més gran que per al castellà. La raó d'aquesta diferència no té a veure amb el comportament neològic de les dues llengües: la quantitat d'unitats recollides està directament determinada pel nombre de textos (bàsicament de premsa) que s'han buidat per a cada llengua. Com que l'Observatori de Neologia ha rebut més suport econòmic per al treball en català que per al castellà, el nombre de textos i, doncs, d'unitats és superior per al català.

4.1. Neologismes amb el formant *homo-*

Com es pot veure a la taula 2, en català s'han documentat 27 neologismes diferents amb el formant *homo-*. Si tenim en compte les repeticions d'aquests neologismes, la quantitat puja a 95, xifra que representa el 0,079% sobre el total d'ocurrències (és a dir, neologismes amb repeticions). Per al castellà, el nombre de neologismes detectats és idèntica que la del català (27) però si hi sumem les repeticions el nombre és més baix que el del català (75) i, en canvi, és més alt el percentatge sobre el total d'ocurrències (0,099%).

Aquests resultats indiquen una coincidència molt important entre català i castellà: la coincidència és total quant al nombre de neologismes i la diferència del nombre d'ocurrències no sembla que indiqui una diferència rellevant en la formació de neologismes amb *homo-* en català i castellà.

TAULA 2. Contrast numèric de neologismes amb *homo-* (català i castellà)

| | <i>Total d'ocurrències de neologismes amb homo-</i> | <i>Percentatge sobre el total d'ocurrències</i> | <i>Total de neologismes amb homo-</i> |
|----------|---|---|---------------------------------------|
| Català | 95 | 0,079 | 27 |
| Castellà | 75 | 0,099 | 27 |

Si centrem l'anàlisi en les unitats formades amb el nou formant *homo-*, és a dir, aquelles unitats en què *homo-* ja no significa “igual” ni “home” sinó “homosexual” podrem afinar més les semblances i diferències entre les dues llengües. A la taula 3 es presenten totes les unitats amb *homo-* distribuïdes en els tres significats documentats (“home”, “igual” i “homosexual”).

En la primera columna d'exemples trobem un total de 14 neologismes (els neologismes de la primera fila s'han de multiplicar per dos perquè tenim la forma catalana i la castellana), alguns dels quals només són neologismes des

TAULA 3. Distribució de les unitats amb *homo-* segons el significat

| <i>Llengua on s'ha documentat</i> | = <i>home</i> (14 neologismes) | = <i>igual</i> (16 neologismes) | = <i>homosexual</i> (24 neologismes) |
|-----------------------------------|-----------------------------------|------------------------------------|---|
| Català | homo erectus (nom) | homogeneitzable (adj) | homoboda (nom) |
| | homo faber (nom) | homogenitzar (verb) | homofolkloric -a (adj) |
| | homo non-videns (nom) | homòleg (nom) | homo-sexual (adj) |
| | homo politicus (nom) | homologable (adj) | homosexualista (adj) |
| | homo velox (nom) | homologadora (nom) | |
| | homo viator (nom) | homologuista (nom) | |
| | homopatogen (adj) | | |
| Castellà | homo automovilis (nom) | homoetario -ria (adj) | homofamilia (nom) |
| | homo chilensis (nom) | homofágico -ca (adj) | homoparental (adj) |
| | homo consumens (nom) | homofónico -ca (adj) | homosentimental (adj) |
| | homo economicus (nom) | homogeneizador -ra (adj) | homosexualismo (nom) |
| | homo-sapiens-demens (nom) | homogeneizante (adj) | |
| | homo-turisticus (nom) | homotopía (nom) | |
| | homotransfusión (nom) | | |
| Català i castellà | homo sapiens (nom) | homologador -a (adj) | homo (adj) |
| | | | homoeròtic -a (adj) |
| | | | homoerotisme (nom) |
| | | | homòfil -a (adj) |
| | | | homòfob -a (adj) |
| | | | homòfob -a (nom) |
| | | | homofòbia (nom) |
| | | | homofòbic -a (adj) |

d'un punt de vista lexicogràfic, ja que són unitats que no representen cap novetat a la llengua i procedeixen de la fraseologia llatina (*homo erectus*, per exemple). Algunes d'aquestes unitats són exemples de neologismes més o menys efímers formats per analogia amb les unitats anteriors. En la segona columna

d'exemples trobem 16 neologismes formats amb el formant tradicional *homo-* que significa "igual"; observem que són unitats denominatives (cap dels exemples correspon a unitats lúdiques, efímeres, etc.) i que es reparteixen en nombre igual en català i en castellà. Observem també que només la unitat *homologador* és compartida per les dues llengües, però convé puntualitzar que algunes de les unitats recollides com a neològiques per al català no apareixen en el castellà perquè ja surten als diccionaris de referència i, doncs, no es poden considerar neològiques (*homòleg*, *homologable*) i s'esdevé el mateix a la inversa (*homofónico*, *homogeneizador*, *homosexualismo*).

La tercera columna d'exemples, que és la que ens interessa per a l'anàlisi de la interferència en aquest article, és la més nombrosa (24 neologismes entre català i castellà) i el repartiment entre les dues llengües resulta totalment equilibrat: 4 unitats només documentades per al català, 4 unitats documentades només per al castellà i 8 unitats documentades per a totes dues llengües (per tant, 16 unitats). Actualment, doncs, en els textos d'àmplia difusió és el nou formant *homo-* el que té una rendibilitat major.

Si observem més atentament les unitats amb aquest nou formant compartides per català i castellà, veiem que el pes cau sobre la família formada per la unitat *homofòbia* i els seus derivats *homòfob -a* i *homofòbic -a* (per al català aquestes unitats sumen 35 ocurrences i per al castellà en sumen 25).⁴ Pensem que pot ser un bon indicador per a l'anàlisi de la (in)dependència del català la informació sobre el moment en què aquestes unitats van aparèixer per primera vegada en cada llengua i l'evolució de la seva freqüència (entre parèntesis s'ofereix la freqüència de cada exemple quan és superior a 1).

Si considerem determinant la llengua en què ha aparegut per primera vegada una unitat formada amb aquest nou *homo-* direm que el català no ha "copiat" el castellà i prendrem aquesta dada com a indicador de la independència del català en la innovació lèxica. Però si considerem determinant el fet que aparegui amb una certa estabilitat, direm que el castellà ha assolit una certa regularitat en la formació de neologismes amb *homo-* abans que el català. Per al castellà trobem una primera aparició l'any 1994 però l'estabilització es produeix a partir del 1997. Per al català, en canvi, tot i tenir-ne més ocurrences, es van alternant els anys amb presència i absència d'aquestes unitats fins l'any 2002.

4. Aquesta freqüència més elevada en català té a veure amb el fet que per al català tenim més dades, però sobretot amb el fet que s'han buidat alguns exemplars de la revista *Lambda*, del col·lectiu de gais i lesbianes.

TAULA 4. Distribució de les unitats amb el nou formant *homo-* al llarg dels anys

| <i>Any</i> | <i>Català</i> | <i>Castellà</i> |
|------------|------------------------------------|-----------------------------|
| 1990 | | |
| 1991 | homofòbic -a | |
| 1992 | | |
| 1993 | | |
| 1994 | | homofobia homofóbico -ca |
| 1995 | homòfob -a homofòbia | |
| 1996 | | |
| 1997 | homòfob -a homofóbico -ca (2) | homofobia |
| 1998 | homòfob -a homófobo -ba | homofobia |
| 1999 | | homófobo -ba |
| 2000 | homòfob -a homofòbic -a | |
| 2001 | | homofobia (2) |
| 2002 | homòfob -a (2) homófobo -ba | homofóbico -ca |
| 2003 | homofòbic -a | homófobo -ba |
| 2004 | homofòbic -a (4) | homófobo -ba (3) |
| 2005 | homofòbic -a homòfob -a (10) | homófobo -ba (3) |
| 2006 | homofòbic -a (2) homòfob -a (5) | homófobo -ba |
| 2007 | homòfob -a (2) homofòbic -a | homófobo -ba (5) |

Què podem concloure després d'aquesta primera anàlisi? Que la semblança de la regla per formar paraules amb el nou *homo-* entre català i castellà és molt elevada. Potser les primeres unitats que es van detectar (*homofòbia*, *homòfob* i *homofòbic*) tant per al català com per al castellà no eren sinó traduccions d'una altra llengua (en el català podien ser-ho del castellà però també de l'anglès, i igualment les unitats castellanes podien ser traduccions de l'anglès), però la varietat actual de paraules formades amb *homo-* indica l'existència de la regla en totes dues llengües. Segons la interpretació que se'n vulgui fer, podem trobar algun indicador de la dependència del català respecte del castellà (per exemple, l'estabilització del recurs sembla que es va produir abans en castellà que en català) però també en trobem de la seva independència (la primera unitat documentada en les dades analitzades és en català). Potser la lectura més neutra (i menys atrevida també) és considerar que no hi ha cap evidència definitiva de la dependència del català respecte del castellà en aquesta regla.

Les altres conclusions que podem extreure de l'anàlisi són de caràcter metodològic i les reprendrem en les conclusions finals: en primer lloc, és evident que per continuar aquesta línia d'investigació és necessari establir criteris per interpretar i valorar els indicadors; i, en segon lloc, caldrà determinar uns mínims per assegurar la representativitat de les dades. En el cas de paraules formades amb *homo-*, per exemple, no sembla que les dades de l'Observatori siguin encara suficients per treure'n conclusions extrapolables.

4.2. Neologismes amb el formant *ciber-*

Des de l'any 1990 fins al 2007 s'han documentat 73 neologismes diferents amb el formant *ciber-* per al català. El nombre d'unitats comptant-hi les repeticions puja a 144, xifra que representa el 0,12% del total d'ocurrències de l'Obneo per al català. Com es pot veure a la taula 5, per al castellà s'han documentat 112 neologismes diferents amb el formant *ciber-*, i el nombre d'unitats comptant-hi les repeticions puja a 232, xifra que representa el 0,30% del total d'ocurrències de l'Obneo per al castellà.

TAULA 5. Contrast numèric de neologismes amb *ciber-* (català i castellà)

| | <i>Català</i> | <i>Castellà</i> |
|-------------------------------|---------------|-----------------|
| Ocurrències amb <i>ciber-</i> | 144 | 232 |
| % sobre el total | 0,12 | 0,30 |
| Neologismes diferents | 73 | 112 |

Amb aquests resultats, podem iniciar l'anàlisi afirmant que la regla per formar noves unitats amb *ciber-* existeix en totes dues llengües i en totes dues és rendible, si bé en castellà ho és més perquè té molts més neologismes diferents i el percentatge que representa sobre el total d'ocurrències gairebé triplica el del català (0,12% / 0,30%).

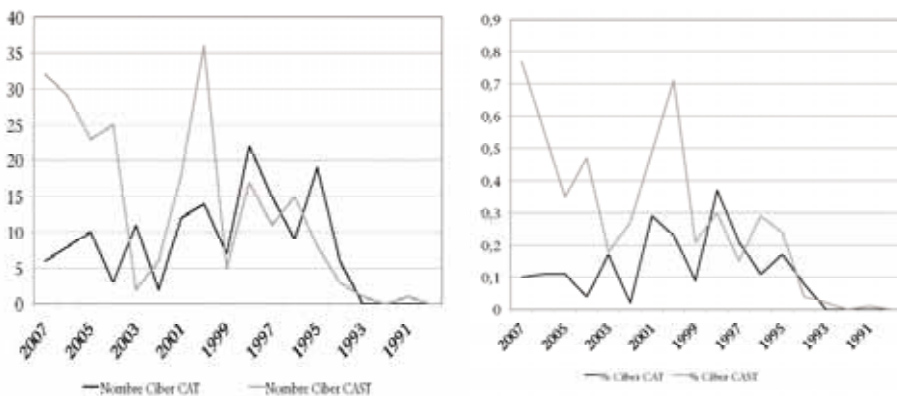
Per conèixer una mica més el funcionament d'aquesta regla observarem com es distribueixen al llarg dels anys les unitats que n'han resultat. Farem el contrast segons el nombre d'unitats però també segons el percentatge que representen cada any per a cada llengua (atès que, com ja s'ha dit en la presentació de les dades, el nombre de neologismes anuals és força variable).

TAULA 6. Contrast de neologismes amb *ciber-* sobre el total de dades anuals

| <i>Any</i> | <i>Neologismes amb ciber- català / castellà</i> | <i>Total de neologismes català / castellà</i> | <i>Percentatge català / castellà</i> |
|------------|---|---|--------------------------------------|
| 2007 | 6 / 32 | 5.687 / 4.110 | 0,10 / 0,77 |
| 2006 | 8 / 29 | 6.973 / 5.112 | 0,11 / 0,56 |
| 2005 | 10 / 23 | 8.381 / 6.556 | 0,11 / 0,35 |
| 2004 | 3 / 25 | 6.739 / 5.230 | 0,04 / 0,47 |
| 2003 | 11 / 2 | 6.439 / 1.057 | 0,17 / 0,18 |
| 2002 | 2 / 6 | 7.027 / 2.170 | 0,02 / 0,27 |
| 2001 | 12 / 18 | 4.078 / 3.606 | 0,29 / 0,49 |
| 2000 | 14 / 36 | 5.839 / 5.019 | 0,23 / 0,71 |
| 1999 | 7 / 5 | 7.316 / 2.286 | 0,09 / 0,21 |
| 1998 | 22 / 17 | 5.814 / 5.655 | 0,37 / 0,30 |
| 1997 | 15 / 11 | 7.127 / 7.188 | 0,21 / 0,15 |
| 1996 | 9 / 15 | 7.555 / 5.099 | 0,11 / 0,29 |
| 1995 | 19 / 8 | 10.965 / 3.215 | 0,17 / 0,24 |
| 1994 | 6 / 3 | 7.221 / 6.159 | 0,08 / 0,04 |
| 1993 | 0 / 1 | 6.051 / 3.741 | 0,00 / 0,02 |
| 1992 | 0 / 0 | 5.164 / 1.464 | 0,00 / 0,00 |
| 1991 | 0 / 1 | 7.987 / 5.058 | 0,00 / 0,01 |
| 1990 | 0 / 0 | 2.846 / 2.822 | 0,00 / 0,00 |
| Total | 144 / 232 | 119.209 / 75.547 | |

A partir de les dades de la taula 6 hem elaborat els gràfics següents que dibuixen en forma de línies l'evolució de la regla de formació amb *ciber-* per a les dues llengües. En el primer es pot veure com es distribueixen les unitats amb *ciber-* al llarg dels anys, i en el segon s'ofereix la informació del percentatge que representa anualment sobre el total anual de neologismes.

GRÀFICS 1 i 2. Contrast de l'evolució dels neologismes en *ciber-* (nombres i percentatges)



Tot i que els dos gràfics demostren una gran oscil·lació en els resultats obtinguts anualment, podem extreure dues informacions que podrien ser uns primers indicis de la dependència de la regla catalana respecte de la castellana: d'una banda, el fet que en català es comencen a documentar exemples amb *ciber-* dos anys més tard que en castellà i, de l'altra, que la línia més clara (que representa la del castellà) sempre (o gairebé sempre) passa per damunt de la línia més fosca, del català.

És cert, però, que només són uns primers indicis i que podem mirar de portar el contrast més enllà comprovant quines unitats formades amb *ciber-* coincideixen entre les dues llengües i quines no coincideixen. Els següents 20 neologismes catalans tenen l'equivalent en castellà: *ciber*, *ciber*, *cibercafé*, *ciberaddicte*, *ciberamant*, *ciberatac*, *cibercriminal*, *cibercultura*, *ciberdelinqüent*, *ciberespai*, *ciberespacial*, *ciberguerra*, *cibernauta*, *cibernàutic -a*, *ciberokupa*, *ciberpunk*, *ciberrealitat*, *cibersexe*, *ciberterrorisme*, *cibertext*. Així doncs, de les 73 unitats catalanes formades amb *ciber-*, 20 són compartides amb el castellà i 53 són "pròpies". Per al castellà, que en tenia 112, són 92 les unitats "pròpies", gairebé el doble que el català. Aquesta dada pot interpretar-se com una major independència de la regla del castellà, però la quantitat de neologismes

propis del català també és un símptoma d'independència en la formació d'unitats amb aquesta regla.

Les unitats no coincidents són majoritàriament hàpax per al català (95,9%) però no tant per al castellà (76,6%); en aquests percentatges no hi intervé el biaix d'unitats que ja puguin aparèixer en els diccionaris de referència d'una llengua i no en els de l'altra. En el DIEC de 1995 apareix *cibernètic -a* i en l'edició de 2007 apareix també *ciberespai*. En el diccionari de la RAE del 2001 apareixen les formes *ciberespacio*, *cibernauta* i *cibernético* i no hi ha noves incorporacions en versions posteriors. És precisament per això que l'any 2001 es produeix un descens important de les dades en *ciber-* del castellà (vegeu els gràfics 1 i 2).

A banda del que hem vist fins ara, la informació més rellevant la podem extreure de l'anàlisi de les unitats coincidents. A continuació ens centrem en l'anàlisi de la freqüència que tenen en català i en castellà les 20 unitats coincidents, i ens fixem també en la distribució que presenten per anys, de manera que podem fer una observació de caràcter diacrònic. La taula 7 resumeix els resultats que s'obtenen: hem tipologitzat els exemples encreuant els dos paràmetres que podrien indicar més o menys dependència: *a*) que en una llengua aparegui abans que en l'altra, i *b*) que en una llengua hi hagi una freqüència més elevada que en l'altra. Així, els codis alfabètics es refereixen al primer paràmetre (A significa que en català apareix primer; B significa que apareixen el mateix any en català i castellà, i C significa que apareix primer en castellà) i els codis numèrics a la freqüència (1 significa que en català té una freqüència major; 2 significa que català i castellà tenen la mateixa freqüència, i 3 significa que la freqüència del castellà és superior).

Si ens fixem en els casos A, quan la primera aparició és en català, sembla que hi ha un equilibri entre la freqüència que presenten les dues llengües, però en realitat hi ha un desequilibri a favor del castellà perquè la freqüència superior de *cibernauta* per al català és simplement perquè a partir de l'any 2001 el diccionari de la RAE incorpora *cibernauta* i, doncs, deixa de ser neologisme per a l'Observatori de Neologia (i, per tant, deixa de ser recollit i computat). Els casos B, quan en totes dues llengües es documenta un mateix neologisme el mateix any, són escassos, però la balança s'inclina a favor del castellà per nombre de casos i per freqüència de cada cas. Finalment, en els casos C el desequilibri a favor del castellà és manifest i, com abans, tant per nombre de casos com per freqüència de cada cas.

Taula 7. Anàlisi de les 20 unitats coincidents (català i castellà)

| <i>Típus</i> | <i>Subtípus</i> | <i>Casos</i> |
|-------------------------------------|---|--|
| A. Casos en què comença el català | A1: freqüència més elevada en català (4 casos) | cibercultura (5/1), cibernauta (10/7), cibernàutic (5/2), ciberpunk (3/1) |
| | A2: freqüència idèntica (2 casos) | ciberamant (1/1), cibertext (1/1) |
| | A3: freqüència més elevada en castellà (4 casos) | ciberatac (1/4), cibersexe (2/5), cibercriminal (1/2) i ciberdelinqüent (3/5) |
| B. Casos en què comencen igual | B1: freqüència més elevada en català (0 casos) | |
| | B2: freqüència idèntica (1 cas): | cíber (1/1) |
| | B3: freqüència més elevada en castellà (2 casos): | ciberrealitat (1/2), ciberterrorisme (3/7) |
| C. Casos en què comença el castellà | C1: freqüència més elevada en català (1 cas) | ciber (2/1) |
| | C2: freqüència idèntica (1 cas) | ciberaddicte (1/1) |
| | C3: freqüència més elevada en castellà (5 casos) | cibercafé (6/33), ciberespai (35/48), ciberespacial (1/7), ciberguerra (1/6), ciberokupa (1/4) |

Podríem considerar un indicador d'independència en la innovació lèxica del català el fet que en català es documentin les primeres ocurrències d'un neologisme, sobretot si a més presenta una freqüència superior o idèntica al castellà al llarg dels anys. Això succeeix en els exemples d'A1 i A2 (6 casos en total).

De la mateixa manera podríem considerar un indicador de dependència quan el neologisme català ha aparegut després del castellà i especialment si presenta una freqüència menor o igual, com succeeix en els exemples C2 i C3, i també el fet que el català presenti una freqüència superior quan es documenten per primera vegada el mateix any, com succeeix en els exemples B3.

En canvi, la interpretació no és òbvia quan els dos paràmetres s'oposen i la llengua que els ha documentat primer és la que té menys freqüència (4 exem-

ples d'A3 i 1 de C1) i és quan torna a fer-se palesa la necessitat d'establir criteris per valorar i interpretar els indicadors.

De tota l'anàlisi dels neologismes amb *ciber-* podem extreure'n com a mínim les conclusions següents:

- a) La regla existeix i és productiva en totes dues llengües, però sembla més productiva en castellà perquè ha format més neologismes diferents i el percentatge que representen les unitats formades amb aquesta regla sobre el total d'ocurrències gairebé triplica el del català (0,12% / 0,30%).
- b) Hi ha uns primers indicis de la dependència de la regla catalana respecte de la castellana: d'una banda, el fet que en català es comencen a documentar exemples amb *ciber-* dos anys més tard que en castellà i, de l'altra, que la freqüència d'aquests neologismes en castellà sol ser més elevada que la dels neologismes en català.
- c) En l'anàlisi de coincidència d'unitats es percep una major independència del castellà perquè de les 112 unitats són 92 les unitats "pròpies", gairebé el doble que el català (de les 73 unitats catalanes formades amb *ciber-*, 20 són compartides amb el castellà i 53 són "pròpies").
- d) Les unitats que no coincideixen són hàpax en el 95,9% dels casos per al català i en el 76,6% per al castellà. Les unitats que no són hàpax i, doncs, que estan més estabilitzades a la llengua només representen el 4,1% per al català i en canvi representen el 23,4% per al castellà. Això no obstant, el català presenta una quantitat prou significativa d'unitats no coincidents amb el castellà, dada que demostra que la regla catalana té una certa independència.
- e) Les unitats coincidents en general són més estables (més freqüents) en castellà que en català, i també és superior el nombre de casos en què les unitats castellanen van aparèixer abans que les catalanes. Tanmateix, en català es documenten per primera vegada algunes de les unitats coincidents amb el castellà i, fins i tot, n'hi ha que tenen una freqüència superior.

4.3. Primers indicadors d'(in)dependència aïllats

En el projecte en què s'insereix aquest treball serà on s'aniran establint els indicadors que s'han d'observar per a cada tipus de neologisme, atès que segons el mecanisme de formació de les unitats s'hauran de tenir en compte indicadors diferents. També serà allà on es determinaran els paràmetres i criteris que

han de permetre avaluar i interpretar correctament els indicadors, i on es definiran metodològicament els conceptes d'*interferència*, de *dependència* i d'*independència* aplicats a la innovació lèxica de les llengües.

No obstant això, l'anàlisi realitzada amb els neologismes formats amb *homo-* i amb *ciber-* ja ens han permès identificar el següent conjunt d'indicadors d'(in)dependència en la innovació lèxica:

1. *Nombre i percentatge de neologismes diferents formats amb la regla que s'analitza.* Amb aquest indicador es contrasta el pes que té la regla analitzada per al català i per al castellà. Un pes molt superior en una llengua que en l'altra s'interpretaria com un índex d'independència en la primera i de dependència en la segona. De totes maneres, cal definir quin seria el percentatge a partir del qual es pot considerar rellevant la diferència entre les dues llengües i cal també establir quin seria el pes d'aquest indicador sobre el resultat total.
2. *Nombre i percentatge d'ocurrències dels neologismes formats amb la regla que s'analitza.* Amb aquest indicador s'obté una informació que complementa l'anterior i que pot ajudar a fer-ne una interpretació més adequada perquè el nombre de neologismes varia molt d'un any a l'altre i també entre llengües.
3. *Data d'aparició de les primeres ocurrències de neologismes formats amb la regla que s'analitza.* Amb aquest indicador no s'avaluen encara les unitats concretes sinó la regla perquè el que es contrasta no és l'aparició d'un neologisme concret sinó dels primers *token* en cada llengua. Com hem dit en els casos anteriors, cal determinar el pes d'aquest indicador dins el conjunt i correlacionar-lo amb els indicadors que es basen en l'anàlisi de l'aparició de les unitats compartides entre les dues llengües.
4. *Evolució de la freqüència al llarg dels anys de les ocurrències de neologismes formats amb la regla que s'analitza.* Amb aquest indicador es pot observar per a cada llengua en quin moment es pot considerar que la regla s'ha estabilitzat; la noció d'*estabilització* es definiria i s'avaluaria a partir de la freqüència i la sistematicitat al llarg dels anys.
5. *Nombre d'unitats coincidents i no coincidents entre les dues llengües i percentatge que representa sobre el total de neologismes formats amb la regla que s'analitza.* Aquest indicador permet observar quin grau d'independència presenta la regla en les dues llengües segons si forma més o menys unitats que no són presents a l'altra llengua. Així, per exemple, el fet que les unitats coincidents representi un tant per cent baix sobre el total de neologismes formats amb la regla en una llengua concreta i que, en canvi, per a l'altra llengua representi un tant per

cent molt elevat indicaria que la primera és molt més independent que la segona. També en aquest cas caldrà establir el valor a partir del qual es considerarà que la coincidència és baixa o alta i, com sempre, també el pes d'aquest indicador dins del conjunt.

6. *No neologicitat en una de les dues llengües de les unitats no coincidents.* Amb aquest indicador es controlaria un possible error d'interpretació de l'indicador anterior: pot succeir que en una llengua no aparegui una unitat que sí que apareix en l'altra simplement perquè ja és una forma recollida en els diccionaris i, doncs, no es pot considerar neologisme lexicogràfic. Això no s'hauria d'interpretar de cap manera com un indicatiu d'independència en la llengua que presenta aquesta unitat no coincident (com es desprendria de l'indicador anterior) sinó que, més aviat, seria la llengua que ja recull la unitat en els diccionaris la que presentaria major independència en la innovació lèxica perquè la unitat ja ha assolit el nivell màxim d'estabilització.

7. *Característiques de les unitats no coincidents en les dues llengües.* Aquest indicador també té com a funció matisar la interpretació de l'indicador cinquè. La importància que cal donar al fet que una llengua hagi començat a crear paraules amb una regla nova és diferent segons si es tracta d'unitats que tenen més o menys possibilitats d'estabilitzar-se en l'ús i, doncs, amb aquest indicador s'ha de controlar el valor denominatiu i expressiu de les unitats.

8. *Data d'aparició de les primeres ocurrences dels neologismes coincidents en les dues llengües.* Aquest indicador és aparentment un dels més eloqüents: si la situació habitual és que les primeres aparicions dels neologismes formats amb una regla es produeix sistemàticament en una llengua, resulta evident la dependència de l'altra llengua respecte d'aquesta.

9. *Evolució de la freqüència al llarg dels anys dels neologismes coincidents en les dues llengües.* La informació que dona aquest indicador és complementària amb l'anterior: l'encreuament d'aquests dos paràmetres dona una informació força evident de la dependència d'una llengua respecte de l'altra (quan les unitats han aparegut abans en l'altra llengua i, a més, estan més estabilitzades), però no la dona quan aquests dos paràmetres entren en contradicció i la llengua que ha introduït primer una o diverses unitats no és la que més les utilitza.

Com es pot observar, els primers indicadors analitzen el funcionament de les regles (quant es generen unitats —*type* o *token*— han generat, quan han aparegut i com han evolucionat en freqüència al llarg dels anys), però a partir del sisè indicador ja s'observen les unitats concretes que han resultat de les regles. I tant en

els primers com en els últims és imprescindible, com s'ha anat dient, establir els criteris que han de permetre interpretar adequadament els resultats.

5. Conclusions

Si bé considerem que és molt important haver començat a identificar els indicadors d'(in)dependència en la innovació lèxica, creiem que és fonamental avançar en aquesta línia per poder avaluar i interpretar adequadament els resultats que ens ofereixen. Així, per exemple, en l'indicador 8 hem dit que el fet que en una llengua es documenti una unitat abans que l'altra sembla determinant per establir si una llengua depèn de l'altra en la creació d'aquesta mena d'unitats. Però si arribem a la conclusió que una llengua ha "copiat" la regla a una altra llengua (que probablement l'ha "copiada" d'una tercera llengua) i els altres indicadors demostren que la regla ha adquirit independència i forma unitats pròpies i estables, quina valoració en faríem? La resposta, és clar, s'obtidria a partir d'un estudi més ampli: no és anòmal que una llengua copii unitats i regles d'una altra llengua (o, més ben dit, copii unitats i n'extregui la seva pròpia regla, per bé que sigui idèntica a la de l'altra llengua o de les altres llengües). És a dir, no és anòmal que una llengua presenti símptomes de dependència en la seva innovació lèxica: anòmal és si una llengua **NOMÉS** presenta símptomes de dependència. Probablement, en el primer cas caldrà parlar d'interferència i distingir graus i tipus d'interferència. En canvi, seria en el segon cas on parlaríem de dependència.

Així, en el cas dels neologismes formats amb *homo-*, tindríem poc marge per determinar quina és la situació actual per al català perquè comptem amb un nombre reduït d'exemples amb el nou formant i les tendències són poc clares. Seria plausible considerar que el català va introduir unes primeres paraules (*homòfob*, *homofòbia*) directament traduïdes d'una altra llengua (de l'anglès, probablement a través del castellà) i que més tard va posar la regla en funcionament d'una manera relativament autònoma. Però la manca d'unitats pròpies creades amb aquesta regla (només *homoboda* i dues unitats efímeres), no ens permet afirmar que la regla per al català estigui definitivament incorporada a la llengua.

En canvi, per a *ciber-* tenim més exemples i més indicadors. El fet que uns indicadors puguin interpretar-se com indicis de dependència i uns altres com indicis d'independència ens aboca a la situació que comentàvem més amunt. No podem, doncs, fer cap afirmació sòlida fins que no s'hagin estudiat amb més profunditat els indicadors i no s'hagin establert valors concrets, però no

saltres ens inclinem a considerar que el català va prendre unes primeres unitats amb *ciber-* (com abans, pensem que les va prendre de l'anglès, amb el castellà com a intermediari en moltes ocasions) i en va extreure una regla pròpia que, en aquest moment, ja té un funcionament autònom. En aquest cas, parlaríem doncs d'interferència més que no pas de dependència.

Més enllà de les conclusions que es desprenen de l'anàlisi feta en aquest treball, voldríem estendre la nostra reflexió final cap a la justificació de la via d'anàlisi que hem encetat, una via d'anàlisi que proposa una lectura sociolingüística de les dades de l'Observatori de Neologia. En primer lloc, pensem que aquesta recerca podria representar una aportació a la lingüística i a la sociolingüística generals en el sentit d'aïllar indicadors que permetin establir quan, i en quins aspectes concrets, una llengua demostra (in)dependència lingüística d'una altra. I, més concretament, podria ser una aportació a la lingüística i a la sociolingüística catalanes pel fet d'arribar a establir indicadors d'(in)dependència del català respecte del castellà i altres llengües quant a la innovació lèxica. És prou sabut que existeix un debat (no només acadèmic sinó també social) sobre la genuïnitat del català en les seves innovacions lingüístiques i els resultats del projecte aportarien dades i mesures objectives en aquest sentit.

És evident que aquests indicadors d'(in)dependència serien aplicables a altres llengües; alguns ho serien directament però caldria identificar altres indicadors específics per a cada parella lingüística concreta: qualsevol llengua romànica respecte de l'anglès, el galleg respecte del castellà, etc.

El projecte que hem iniciat podria aportar als estudis de lingüística catalana una descripció detallada dels aspectes relatius a la innovació lèxica; si fins ara la producció lèxica s'explicava des del punt de vista dels recursos del català, la incorporació de biaixos d'ordre sociolingüístic enriquiria substancialment la descripció (així les unitats no estarien només tractades morfològicament, ni considerades al marge de fenòmens sociolingüístics que, de fet, n'expliquen la procedència i la formació). En el panorama català, i des d'un punt de vista aplicat, una de les aportacions més evidents seria en el terreny de la planificació lingüística, i concretament en la normalització lingüística ja que s'oferirien dades i informes d'utilitat per a les institucions que vetllen per la llengua per intervenir, si es considerés pertinent i necessari, en aquelles àrees on la innovació lèxica catalana es valorés com a més dependent d'altres llengües.

6. Bibliografia

- CABRÉ, M. T.; FREIXA, J.; SOLÉ, E. (ed.) (2002). *Lèxic i neologia*. Barcelona: Observatori de Neologia, Institut Universitari de Lingüística Aplicada, Universitat Pompeu Fabra. (Sèrie Monografies, 5).
- CABRÉ, M. T.; FREIXA, J.; SOLÉ, E. (1997). «À la limite des mots construits possibles». *Silexicales 1*. Lille: Conseil de laboratoire de l'URA 382 du CNRS. 65-78. [Mots possibles et mots existants. Forum de morphologie. Actes du colloque de Villeneuve d'Ascq (28-29 avril 1997).]
- CABRÉ, M. T.; BAYÀ, R.; BERNAL, E.; FREIXA, J.; SOLÉ, E.; VALLÈS, T. (2000). «Evaluación de la vitalidad de una lengua a través de la neología: A propósito de la neología espontánea y de la neología planificada». Dins CHEVALIER, J. C.; DELPORT, M. F. (dir.). *La fabrique des mots. La néologie ibérique*. París: Presses de l'Université de Paris-Sorbonne. (Sèrie Essais, 2). 91-130.
- CABRÉ, M. T.; FREIXA, J.; SOLÉ, E. (2001). «Anàlisi contrastiva de la innovació lèxica en català i en castellà». *Caplletra*, 30. Número monogràfic. 199-212.
- FELIU, J.; RAMÍREZ, S.; TALAMINO, Ò. (2000). «El formant *ciber-*: funcionament i distribució». Dins CABRÉ, M. T.; FREIXA, J.; SOLÉ, E. (ed.). *La neologia en el tombant de segle*. Barcelona: Observatori de Neologia, Institut Universitari de Lingüística Aplicada, Universitat Pompeu Fabra. (Sèrie Activitats, 5). 183-192.
- FREIXA, J. (2004). «Neologismes formats per prefixació». Dins OBSERVATORI DE NEOLOGIA (2004). *Llengua catalana i neologia*. Barcelona: Meteora.
- FREIXA, J. (2007). «Neologismos frecuentes del español». *Donde dice... Boletín de la Fundación del Español Urgente*. Fundéu, 9, 7-11
- FREIXA, J.; SOLÉ, E.; CABRÉ, M. T. (1998). «Observació de la variació i el contacte de llengües en els neologismes». Dins BOIX, E. et al. (ed.). *El contacte i la variació lingüístics: Descripció i metodologia*. Barcelona: Cooperativa Universitària Sant Jordi; Secció de Lingüística Catalana, Departament de Filologia Catalana, Universitat de Barcelona. 89-104.
- OBSERVATORI DE NEOLOGIA. (2004). *Llengua catalana i neologia*. Coordinació de J. Freixa i E. Solé. Barcelona: Meteora.
- RULL, X. (2001). «Innovació en el sistema afixal: nous prefixos i sufixos del català». *Catalan Review. International Journal of Catalan Culture XIV*: 1. 95-116.
- (2002). *La formació de mots. Qüestions de normativa*. Barcelona: Eumo Editorial.
- VALLÈS, T. (2002). «Anàlisi cognitiva del prefix *euro-*». Dins CABRÉ, M. T.; FREIXA, J.; SOLÉ, E. (ed.). *Lèxic i neologia*. Barcelona: Observatori de Neologia, Institut Universitari de Lingüística

Aplicada, Universitat Pompeu Fabra.
(Sèrie Monografies, 5). 115-120.

VALLÈS, T. (2004). *La creativitat lèxica en un model basat en l'ús: una aproximació cognitiva a la neologia i la productivitat*.

Barcelona: Publicacions de l'Abadia de Montserrat.

WEINREICH, U. (1953). *Languages in Contact. Findings and Problems*. The Hague-Paris: Mouton.

Resumen

La interferencia en la innovación léxica del catalán: análisis de los neologismos con *ciber-* i *homo-*

Judit Freixa Aymerich

Este artículo pretende dar respuesta a la pregunta siguiente: ¿hasta qué punto los neologismos que incorpora actualmente el catalán son fruto de sus propios mecanismos de innovación léxica? El análisis, centrado en todas las unidades neológicas formadas con *ciber-* y *homo-* recogidas en el Observatorio de Neología durante los últi-

mos años, se plantea el grado de dependencia que el catalán presenta respecto del castellano. Esta línea de investigación quiere ser de utilidad para la planificación lingüística porque detectar y conocer las vías de interferencia de una lengua es imprescindible para activar las medidas correctoras que se consideren necesarias.

Palabras clave: neología; morfología léxica; prefijación; interferencia (catalán, castellano/español); planificación lingüística.

Abstract

Interference in Lexical Innovation in Catalan: an Analysis of the Neologisms with *ciber-* and *homo-*

Judit Freixa Aymerich

This article aims at responding to the following question: to what point are the neologisms now being incorporated into Catalan the fruits of its own mechanisms of lexical innovation? The analysis, focussed on all the neological units composed either of *ciber-* or of *homo-* collected at the Neological Observatory in recent years, considers

the degree of dependence shown by Catalan with respect to Spanish. This line of investigation aspires to be of use to language planners, because detecting and being aware of the avenues of interference for a language is indispensable for launching the corrective measures considered necessary.

Key words: neology; lexical morphology; prefixation; interference (Catalan, Spanish); language planning.