

**Belén López Meirama
María Álvarez de la Granja**

***Léxico dispoñible
do galego***

Verba

Anexo 71

2014

BELÉN LÓPEZ MEIRAMA
MARÍA ÁLVAREZ DE LA GRANJA

LÉXICO DISPOÑIBLE DO GALEGO

Verba, ANUARIO GALEGO DE FILOLOXÍA
ANEXO 71

2014
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA



Esta obra atópase baixo unha licenza internacional Creative Commons BY-NC-ND 4.0. Calquera forma de reprodución, distribución, comunicación pública ou transformación desta obra non incluída na licenza Creative Commons BY-NC-ND 4.0 só pode ser realizada coa autorización expresa dos titulares, salvo excepción prevista pola lei. Pode acceder Vde. ao texto completo da licenza nesta ligazón: <https://creativecommons.org/licenses/by-nc-nd/4.0/deed.gl>

© Universidade de Santiago de Compostela, 2014

Edita

Servizo de Publicacións
e Intercambio Científico
Campus Vida
15782 Santiago de Compostela
www.usc.es/publicacions

DOI <https://dx.doi.org/10.15304/9788415876687>

ISSN 2341-1198

ISBN 978-84-15876-68-7

AGRADECEMENTOS

O estudo que presentamos neste libro é froito de varios anos de dedicación a un proxecto de investigación que se puido levar a cabo grazas ao financiamento da Consellería de Educación e Ordenación Universitaria da Xunta de Galicia entre os anos 2008 e 2011¹, no que contamos coa inestimable contribución de Mercedes Luaces Carreño na transcripción dos materiais e coa de Marta Blanco e Francisco García Gondar no traballo de lematización. Así mesmo, debemos a Guillermo Rojo a creación e manipulación dun programa informático que facilitou enormemente as sucesivas revisións a que sometemos os datos, ademais de incrementar as posibilidades de explotación dos mesmos.

Este traballo, con todo, non sería posible sen a colaboración desinteresada dos departamentos de Lingua Galega e Literatura que participaron activamente na recollida dos datos dirixindo a realización das enquisas nos cincuenta centros educativos de toda Galicia seleccionados para elaborar a mostra, e sobre todo sen as palabras que nos regalaron os centos de estudantes deses centros que asistiron ás aulas de 2º de bacharelato no curso 2008-2009. Vaia para todos eles o noso máis profundo agradecemento.

¹ Disponibilidade léxica en Galicia (Dispogal). Código INCITE08PXIB204095PR. Este traballo tamén se beneficiou de dúas axudas da Consellería de Cultura, Educación e Ordenación Universitaria da Xunta de Galicia, a primeira no ámbito do programa Consolidación e estruturación de unidades de investigación competitivas (GRC2013-040) e a segunda no ámbito da Rede de Lexicografía-RELEX (CN2012/290), con financiamento parcial en ambos os casos do Fondo Europeo de Desenvolvemento Rexional (FEDER).

ÍNDICE

1	INTRODUCCIÓN	7
2	ASPECTOS METODOLÓXICOS	11
2.1	A enquisa e os centros de interese.....	12
2.2	A mostra e as variables de estratificación social.....	14
2.3	Cráterios de edición dos materiais.....	31
2.3.1	Formas eliminadas e formas conservadas	32
2.3.2	Cráterios xerais de lematización	37
2.3.2.1	O xénero	38
2.3.2.2	O número.....	40
2.3.2.3	As palabras con morfemas avaliativos.....	42
2.3.2.4	Os verbos	42
2.3.2.5	Os acurtamentos, as abreviaturas gráficas, as siglas e acrónimos e os símbolos.....	43
2.3.2.6	Os nomes de marcas, empresas e produtos comerciais	44
2.3.2.7	Os sintagmas.....	45
2.3.2.7.1	Determinantes e cuantificadores.....	47
2.3.2.7.2	Preposicións	49
2.3.2.7.3	Orde	50
2.3.2.7.4	Supresión e mantemento de constituíntes	51
2.3.2.7.5	Recuperación de constituíntes	53
2.3.2.8	A separación / fusión gráfica	54
2.3.2.9	O emprego de maiúsculas	54
2.3.2.10	As variantes ortográficas, os lapsus cáلامي e as creacións dos informantes	55
2.3.2.11	Os estranxeirismos.....	56
2.3.2.12	As variantes galegas e as voces castelás	57
2.3.3	A lematización das variantes galegas fónicas e morfolóxicas e das voces procedentes do castelán	58
2.3.3.1	Variante fónicas e morfolóxicas de orixe galega.....	60
2.3.3.2	Voces do castelán cognadas de palabras galegas.....	64
2.3.3.3	Voces do castelán sen cognado galego	69
2.3.3.4	Supresión de variantes galegas e de cognados	71
2.3.4	Cráterios específicos de edición. Os centros de interese.....	73
2.3.4.1	CI 01. O corpo humano.....	73
2.3.4.2	CI 02. A roupa	74
2.3.4.3	CI 03. As partes da casa (sen mobles)	78
2.3.4.4	CI 04. Os mobles da casa	81
2.3.4.5	CI 05. Os alimentos e as bebidas	84
2.3.4.6	CI 06. Os obxectos colocados na mesa para a comida	87
2.3.4.7	CI 07. A cociña e os seus utensilios.....	89
2.3.4.8	CI 08. A escola: mobles e materiais.....	91
2.3.4.9	CI 09. A calefacción e a iluminación.....	94

2.3.4.10	CI 10. A cidade.....	96
2.3.4.11	CI 11. O campo.....	99
2.3.4.12	CI 12. Os medios de transporte	102
2.3.4.13	CI 13. Os traballos do campo e do xardín.....	104
2.3.4.14	CI 14. Os animais	109
2.3.4.15	CI 15. Os xogos e as distraccións	112
2.3.4.16	CI 16. As profesións e os oficios	117
2.3.4.17	CI 17. As cores	120
2.3.4.18	CI 18. O mar.....	123
2.4	Tratamento informático dos datos	127
2.4.1	O proxecto panhispánico.....	127
2.4.2	O proxecto galego.....	129
3	ANÁLISE DE RESULTADOS.....	133
3.1	Os centros de interese. Índices xerais de cuantificación.....	134
3.1.1	Número de palabras	134
3.1.2	Índice de cohesión	140
3.1.3	Índice de dispoñibilidade	143
3.1.4	Léxico galego e léxico castelán	153
3.1.5	Recapitulación	160
3.2	As variables de estratificación social.....	160
3.2.1	As variables sociais e a produción de palabras	161
3.2.1.1	O sexo	161
3.2.1.2	A lingua	163
3.2.1.3	O tipo de poboación.....	168
3.2.1.4	O tipo de centro	171
3.2.1.5	O nivel sociocultural	172
3.2.2	As variables sociais e a compatibilidade léxica.....	175
4	O DICCIONARIO DE LÉXICO DISPOÑIBLE: <i>DISPOGAL</i>.....	187
5	PRESENTACIÓN DE APÉNDICES	193
5.1	Listaxes de lemas e variantes por CI	193
5.2	Listaxe xeral de lemas ordenados por frecuencia.....	195
5.3	Listaxe xeral de lemas ordenados alfabeticamente	195
5.4	Mostra da enquisa	196
6	REFERENCIAS BIBLIOGRÁFICAS	197
7	ACCESO A APÉNDICES.....	201
	Apéndice I Listaxes de lemas e variantes por CI	
	Apéndice II Listaxe xeral de lemas ordenados por frecuencia	
	Apéndice III Listaxe xeral de lemas ordenados alfabeticamente	
	Apéndice IV Mostra da enquisa	

1 INTRODUCCIÓN

Os primeiros estudos sobre dispoñibilidade léxica xurdiron en Francia nos anos cincuenta do século pasado como unha contribución necesaria na descrición do vocabulario común, do vocabulario coñecido e empregado polos falantes do francés. O obxectivo era facilitarlles a adquisición desta lingua aos habitantes dos países que formaran parte da Union Française e aos moitos inmigrantes que daquela chegaban a Francia. Tendo en conta os coñecementos da época, o único factor que se consideraba ata entón nas análises estatísticas era o da frecuencia, de xeito que as unidades léxicas se clasificaban pola súa maior ou menor ocorrencia nos textos, o cal deixaba fóra un bo número de palabras de uso común, pero que non aparecían nas listaxes de vocabulario frecuente. As listaxes nutríanse de palabras gramaticais, de verbos, de adxectivos e de substantivos de carácter xeral, pero apenas contiñan palabras –basicamente substantivos– de significado máis concreto, aínda que fosen ben coñecidas e mesmo moi empregadas polos falantes. Estas palabras eran infrecuentes nos corpus porque a súa actualización estaba condicionada polo tema do discurso; eran, por tanto, palabras temáticas que non chegaban á estabilidade estatística das palabras atemáticas.

Ao quedar de manifesto a parcialidade dun léxico seleccionado exclusivamente empregando a frecuencia, xurdiu a necesidade de complementar o traballo cunha nova variable, a da dispersión, que supuña a división do universo léxico nunha serie de ámbitos temáticos (*vid.* Samper Hernández 2009). López Morales (1999) indica que foi Michéa (1953) o primeiro en opoñer as ‘palabras frecuentes’ ás ‘palabras dispoñibles’, de xeito que a dispoñibilidade léxica pasou a entenderse como o caudal léxico utilizable nunha situación comunicativa dada. Di o autor que

Lo que subyace a esta afirmación es que existe en el lexicón mental una serie de términos que no se actualizan a menos que sea necesario para comunicar

una información muy específica. Se trata de un léxico ‘disponible’, cuyo estudio no puede emprenderse manejando frecuencias, porque este factor es pertinente solo en el caso de las actualizaciones léxicas efectivas, no de las potenciales. (López Morales 1999: 11)

Dado que, de acordo con Michéa (1953), o léxico dispoñible se caracteriza porque se presenta na mente do falante de xeito inmediato e natural cando se trata de determinado tema, empeza a recollerse empregando uns estímulos verbais –os temas–, chamados ‘centros de interese’, a partir dos cales os suxeitos producen listaxes de unidades léxicas, cun método de asociación controlada.

O traballo realizado en Gougenheim *et al.* (1964) marca a pauta deste tipo de estudos, continuados en Canadá nun gran proxecto dirixido por Mackey (1971) e, sobre todo, no ámbito hispánico, a partir do labor pioneiro de López Morales, que desenvolve o modelo da investigación francesa nos escolares portorriqueños e edita en 1973 o primeiro traballo sobre dispoñibilidade léxica na zona metropolitana de San Juan. Este autor é o coordinador xeral do Proxecto panhispánico, que se vén realizando sobre todo a partir dos anos noventa do século XX co obxectivo último de elaborar un dicionario de léxico dispoñible do español que recolla a norma léxica a partir do vocabulario dos estudantes do derradeiro curso do ensino secundario². Contabilízanse numerosos traballos realizados no outro lado do Atlántico (ademais de Porto Rico, a República Dominicana, Uruguai, México, Chile, Costa Rica, Panamá, Colombia, Cuba...), aínda que non todos se axustan aos criterios e protocolos de edición xerais que si se están empregando nas investigacións levadas a cabo en España, establecidos en sucesivas reunións, iniciadas en Bilbao en 1999.

Neste gran proxecto enmárcase o *Léxico disponible en el español de Galicia* (en adiante, LDEG), que viu a luz hai seis anos (López Meirama 2008) e que desde os seus inicios, alá por 2005, foi concibido, dun lado, coma unha peza do macroproxecto hispánico e, doutro, coma unha das dúas metades da investigación que se iniciaba sobre a dispoñibilidade léxica en Galicia: tendo en conta que este campo da léxico-estadística abre multitude de camiños na investigación lingüística (ademais de no estritamente lexicolóxico, na dialectoloxía, na sociolingüística, na psicolingüística, na lingüística aplicada ao

² Para obter información sobre a historia da dispoñibilidade léxica no mundo hispánico pódense consultar, entre outros: López Morales (1995, 1999), Carcedo (1998), Samper Padilla / Bellón Fernández / Samper Hernández (2003) ou Samper Padilla / Samper Hernández (2006). A páxina web <http://www.dispollex.com>, organizada por José Antonio Bartol e Natividad Hernández Muñoz, é o banco de datos máis importante con que conta o Proxecto panhispánico, así como un foro de encontro de todos os investigadores e equipos involucrados no mesmo.

ensino e aprendizaxe das linguas etc.), resultaba imprescindible desenvolver o proxecto nas dúas linguas da nosa Comunidade.

Así pois, a segunda metade da investigación constitúea o *Léxico dispoñible do galego* (en diante LDG) que ofrecemos nesta publicación. O estudo que precede ás listaxes de palabras ábrese coa presentación dos aspectos metodolóxicos do traballo, no que se refire ás enquisas realizadas (§ 2.1), á mostra e á súa estratificación (§ 2.2), aos criterios de edición dos materiais (§ 2.3) e ao seu tratamento informático (§ 2.4). En § 3 levamos a cabo unha análise global, necesariamente breve, dos resultados obtidos, centrada tanto no seu estudo cuantitativo para cada centro de interese (§ 3.1), cunha breve comparación co LDEG, como na pescuda sobre a incidencia en tales resultados das diferentes variables de estratificación social consideradas na mostra (§ 3.2).

Este estudo amósanos que, máis alá do interese e aplicacións arriba mencionadas, inherentes a calquera traballo destas características, o léxico dispoñible do galego é de especial importancia porque nos permite facer unha avaliación do dominio do léxico galego por parte dos nosos estudantes, máxime tendo en conta que carecemos de traballos específicos centrados nesta liña de investigación. O feito de poder establecer cotexos con outros léxicos dispoñibles, nomeadamente co do castelán de Galicia, permite facer unha valoración dos resultados cuantitativos obtidos, valoración que se poderá enriquecer cunha análise cualitativa máis detallada³. Confiamos, así pois, en que esta publicación contribúa a un mellor coñecemento da competencia léxica dos alumnos de segundo de bacharelato de Galicia e a partir de aí, que esperte intereses e abra camiños.

³ En Álvarez de la Granja (no prelo) faise unha análise do dominio do léxico galego da variedade estándar a partir dos datos do LDG.

2 ASPECTOS METODOLÓXICOS

Dado que este estudo, dedicado á lingua galega, se enmarca nun proxecto máis amplo de análise da dispoñibilidade léxica en Galicia, nel empregamos, como é lóxico, unha metodoloxía común a ambos os subproxectos (o do léxico galego e o do léxico castelán), a cal, á súa vez, está determinada polo traballo previo realizado no ámbito do citado Proxecto panhispánico.

En xeral, tanto os dominios léxicos considerados (os *centros de interese*, dos que falaremos máis adiante) como as variables de estratificación social a través das que se estrutura a mostra coinciden cos empregados nos diversos estudos realizados no ámbito hispánico, e mesmo cos que no seu día se utilizaron nos pioneiros traballos sobre o francés. As variables son o sexo, o nivel sociocultural dos pais, o tipo de ensino (público ou privado), o tipo de poboación do centro de ensino e da residencia familiar e, nas comunidades bilingües, a lingua habitual.

Con todo, hai algunhas particularidades que convén destacar, que serán explicadas polo miúdo nas páxinas seguintes: canto aos *centros de interese*, o proxecto galego engade dous aos dezaseis comúns, un dedicado ás cores e outro ao mar; canto á mostra, a distribución das poboacións dos centros de ensino e dos domicilios familiares faise en catro grupos (poboación urbana, poboación periurbana, vila e poboación rural), en lugar de en dous (poboación urbana e poboación rural). No que respecta, en particular, ao subproxecto dedicado á lingua galega, incorporamos a variable lingua inicial, que non está recollida no LDEG; con todo, as diferenzas máis salientables localízanse na edición de materiais, xa que a lematización do léxico galego produce problemas diferentes aos orixinados na do léxico castelán, o cal determina que se deban tomar decisións tamén diferentes. En § 2.3.2 explicaremos pormenorizadamente todas as peculiaridades do proceso de lematización a que sometemos as enquisas elaboradas polos oitocentos informantes que constitúen a mostra da lingua galega.

2.1 A enquisa e os centros de interese

Como xa sinalamos, o léxico dispoñible é o que está preparado na memoria para ser empregado no momento no que se fala sobre determinado asunto, o cal funciona como un estímulo para a produción de palabras, de maneira que o xeito máis adecuado de recollelo é a través dunha enquisa, en forma de cuestionario, dividida en tantas partes como temas ou estímulos se consideren. Os temas, denominados, como vimos dicindo, *centros de interese* (CI), foron establecidos nos estudos franceses que se levaron a cabo nos anos sesenta (cfr. Gougenheim *et al.* 1964) e son tamén os empregados no Proxecto panhispánico dirixido por López Morales.

Por razóns obvias, o LDEG procurou acomodarse a este proxecto, aínda que engadiu dous aos dezaseis CI tradicionais, un referido ás cores, tamén presente noutros traballos dedicados a sintopías tanto españolas como americanas, e outro referido ao mar, que é un elemento fundamental da vida cotiá da maior parte dos galegos⁴. O LDG emprega esta mesma distribución, de xeito que os dezaioito CI que configuran o test asociativo son os seguintes:

- CI 01. O corpo humano
- CI 02. A roupa
- CI 03. As partes da casa (sen mobles)
- CI 04. Os mobles da casa
- CI 05. Os alimentos e as bebidas
- CI 06. Os obxectos colocados na mesa para a comida
- CI 07. A cociña e os seus utensilios
- CI 08. A escola: mobles e materiais
- CI 09. A calefacción e a iluminación
- CI 10. A cidade
- CI 11. O campo
- CI 12. Os medios de transporte
- CI 13. Os traballos do campo e do xardín
- CI 14. Os animais
- CI 15. Os xogos e as distraccións
- CI 16. As profesións e os oficios
- CI 17. As cores
- CI 18. O mar

⁴ Esta é unha práctica relativamente habitual nos distintos subproxectos que compoñen o panhispánico; por exemplo, no de Navarra engádense tres CI aos dezaseis comúns (17 'Cores', 18 'Virtudes e defectos' e 19 'Intelixencia'), coa aspiración de obter datos interesantes para estudos psicolingüísticos, sociolingüísticos e dialectais (*vid.* Saralegui / Taberner 2008: 748).

As enquisas, das cales se inclúe una mostra no apéndice 4, realizáronse no contexto dos centros educativos, na maior parte dos casos dirixidas polos profesores que desinteresadamente colaboraron no proxecto⁵, seguindo unhas pautas moi precisas que lles foron marcadas para asegurar a homoxeneidade na recollida dos datos. Constan de dúas partes, das que se obteñen dous tipos diferenciados de información: dun lado, reúnen os datos de tipo sociocultural dos alumnos (lingua habitual, lingua inicial, escolaridade e profesión dos pais, lugar de residencia familiar, tipo de centro educativo, sexo), que serven para establecer unhas variables independentes a través das que se estratifica a mostra; doutro lado, recollen a información nuclear, a estritamente lingüística, a través de listas de palabras correspondentes aos dezaioito CI.

Dado que os índices de dispoñibilidade serven para ordenar as unidades léxicas segundo a súa posibilidade de aparición en contextos específicos (os anteditos CI), o léxico dispoñible xerarquíxase atendendo non só á frecuencia de aparición dos ítems léxicos, senón tamén á orde de mención, polo que resulta imprescindible recoller as palabras ordenadamente, tal e como se van activando no lexicón mental; de aí a división de cada páxina en 48 cadros numerados, que o informante debe ir enchendo durante dous minutos despois de anotar o nome do CI, proporcionado inmediatamente antes polo profesor. Deste modo, as respostas son producións instantáneas, e non o resultado da reflexión, o que asegura a espontaneidade, característica esencial do concepto de dispoñibilidade.

Ademais dos dezaioito CI, incluímos un de mostra, dedicado aos deportes, do que nos servimos para que os profesores advertisen aos informantes dalgúns aspectos que consideramos importantes na recollida do léxico dispoñible: por exemplo, que debían procurar non demorarse tentando lembrar unha palabra que tardase en vir á memoria; que podían incluír compostos sintagmáticos, tipo *fóra de xogo*, *tarxeta amarela* ou *patinaxe artística*, préstamos (*córner*, *basket*, *sláalom*), acurtamentos (*bici*) etc., pero non marcas comerciais (*Nike*, *Adidas*) nin nomes propios (*Pau Gasol*, *Fernando Alonso*); que non debían interpretar o nome de cada CI exclusivamente como un hiperónimo do que ofrecer hipónimos (*baloncesto*, *tenis*, *fútbol*, *rugby* etc.), senón, máis xeralmente, como o xeito de facer referencia a un ámbito de uso de determinadas palabras, é dicir, a un tema sobre o que conversar, reflexionar etc., polo que no CI empregado como mostra cabían tamén palabras como *meta*, *campión*, *adestramento*, *dopaxe* etc. Neste sentido, consideramos importante indicarlles aos informantes que, aínda que na maior parte dos casos as palabras seleccionadas serían substantivos, tamén podían incluír verbos (*xutar*, *canear*, *montar a cabalo*) e adxectivos (*descualificado*, *lesionado*).

⁵ Só esporadicamente se solicitou a nosa presenza.

Finalmente, sinalámoslles aos profesores colaboradores que, debido ás características peculiares do léxico galego e ao feito de que a mostra se realizaría no ámbito do ensino, podería suceder que algún alumno se preguntase se debía poñer o que di habitualmente ou ben a forma “correcta” (o que “debería” dicir, ou o que escribe)⁶. A este respecto, consideramos oportuno que se lles recordase aos rapaces que non estaban ante unha proba, exercicio ou exame de Lingua Galega, e que o propósito da enquisa era recoller as palabras que se lles ocorresen poñéndose na situación de falar ou escribir en galego sobre o tema que representa cada CI.

2.2 A mostra e as variables de estratificación social

Como indicamos anteriormente, para facilitar os estudos comparativos decidimos que a mostra do léxico galego dispoñible fose idéntica á do castelán de Galicia, a cal, á súa vez, se adaptou ás bases metodolóxicas do Proxecto panhispánico, aínda que contemplando as particularidades da nosa comunidade. Así, e tendo en conta o número de habitantes de Galicia, de mozos de 17 anos e de alumnos matriculados en 2º curso de Bacharelato en xaneiro do ano 2006⁷, ciframos en oitocentos os informantes necesarios, xa que esta cantidade aseguraba unha representación mínima das provincias menos poboadas, Lugo e Ourense, e ademais supuña o 4,29% dos alumnos, o 3,02 dos mozos e mesmo o 0,029 da poboación total, porcentaxes todas elas superiores á establecida por Labov (1966: 170-174) para considerar a representatividade dunha mostra (0,025; véxase López Meirama 2008: 18). A levada a cabo para a lingua galega, xa que logo, está tamén composta de oitocentas enquisas realizadas a estudantes preuniversitarios galegos, neste caso durante o curso 2008-2009, coa colaboración dos departamentos de Lingua Galega e Literatura dos centros seleccionados e, especificamente, dos profesores encargados da docencia do segundo curso de bacharelato.

Na táboa 1 amosamos os cincuenta centros educativos que colaboraron na recollida das enquisas, indicando en cada caso a localidade, o tipo de poboación a que esta corresponde, a provincia⁸, a situación —costeira ou de in-

⁶ De maneira moi simplificada, podemos dicir que hoxe conviven en Galicia un galego popular, moi influído polo castelán no ámbito léxico, e con solucións propias de cada variedade xeográfica, e un galego estándar libre en boa medida de castelanismos e unificador da variación dialectal. O galego estándar, claro está, é a variedade empregada no ámbito do ensino.

⁷ As enquisas do LDEG realizáronse ao longo do ano 2006, a maioría nos dous últimos trimestres do curso 2005-2006 e unha parte sensiblemente menor no primeiro trimestre do curso 2006-2007.

⁸ Os datos do Instituto Galego de Estatística sobre a matrícula no bacharelato ordinario indican que a distribución dos estudantes nas catro provincias galegas apenas variou do curso 2005-2006, cando tomamos a mostra da lingua castelá, ao curso 2008-2009, cando recolleemos os datos da mostra da lingua galega, como se pode ver na seguinte táboa, que recolle as porcentaxes de cada provincia sobre o total de cada curso en Galicia:

terior— e o número de enquisas proporcionadas. As siglas que preceden aos nomes indican se se trata de centros públicos (IES, Instituto de Educación Secundaria) ou privados (CPR, Centro Privado).

Táboa 1. Centros de ensino que compoñen a mostra

	Centro escolar	Localidade	Poboac.	Prov.	Ubic.	Nº enq.
1	IES Agra do Orzán	A Coruña	urbana	CO	costa	16
2	IES Salvador de Madariaga	A Coruña	urbana	CO	costa	15
3	IES Eusebio da Guarda	A Coruña	urbana	CO	costa	17
4	CPR Cristo Rey	A Coruña	urbana	CO	costa	16
5	CPR Grande Obra de Atocha	A Coruña	urbana	CO	costa	16
6	CPR Santo Domingo	A Coruña	urbana	CO	costa	16
7	IES Saturnino Montojo	Ferrol	urbana	CO	costa	17
8	IES Rodolfo Ucha Piñeiro	Ferrol	urbana	CO	costa	15
9	IES Pontepedriña	Santiago	urbana	CO	interior	14
10	IES Antonio Fraguas Fraguas	Santiago	urbana	CO	interior	18
11	IES As Fontiñas	Santiago	urbana	CO	interior	16
12	CPR La Salle	Santiago	urbana	CO	interior	16
13	IES Fragas do Eume	Pontevedra	periurb.	CO	costa	19
14	IES Afonso X O Sabio	Cambre	periurb.	CO	costa	16
15	IES de Ames	Bertamiráns	periurb.	CO	interior	13
16	IES Isidro Parga Pondal	Carballo	vila	CO	costa	16
17	IES Campo de San Alberto	Noia	vila	CO	costa	16
18	IES Francisco Aguiar	Betanzos	vila	CO	interior	16
19	IES Punta Candieira	Cedeira	rural	CO	costa	16
20	IES Macías o Namorado	Padrón	rural	CO	interior	16
21	IES Terra do Xallas	Sta. Comba	rural	CO	interior	16
22	IES Politécnico de Lugo	Lugo	urbana	LU	interior	16
23	IES Xoan Montes	Lugo	urbana	LU	interior	16
24	IES Leiras Pulpeiro	Lugo	urbana	LU	interior	16
25	IES Francisco Daviña Rei	Monforte	vila	LU	interior	16
26	IES Pedregal de Irimia	Meira	rural	LU	interior	16
27	IES Porta da Auga	Ribadeo	rural	LU	costa	16
28	IES Otero Pedrayo	Ourense	urbana	OU	interior	18
29	IES A Carballeira	Ourense	urbana	OU	interior	14
30	CPR Santa María	Ourense	urbana	OU	interior	16
31	IES N.º 1 do Carballiño	O Carballiño	vila	OU	interior	16

curso ⇨	2005-2006	2008-2009
provincia ⇩	(subprox. do castelán)	(subprox. do galego)
A Coruña	40,94%	41,02%
Lugo	12,30%	12,29%
Ourense	10,96%	11,61%
Pontevedra	35,79%	36,08%

	Centro escolar	Localidade	Poboac.	Prov.	Ubic.	Nº enq.
32	IES Aquis Querquernis	Bande	rural	OU	interior	16
33	IES do Castro	Vigo	urbana	PO	costa	16
34	IES Alexandre Bóveda	Vigo	urbana	PO	costa	16
35	IES de Beade	Vigo	urbana	PO	costa	16
36	CPR María Inmaculada	Vigo	urbana	PO	costa	16
37	CPR Las Acacias	Vigo	urbana	PO	costa	16
38	CPR Apóstol Santiago	Vigo	urbana	PO	costa	16
39	IES Sánchez Cantón	Pontevedra	urbana	PO	costa	16
40	IES G. Torrente Ballester	Pontevedra	urbana	PO	costa	16
41	IES A Xunqueira	Pontevedra	urbana	PO	costa	16
42	IES Val Miñor	Nigrán	periurb.	PO	costa	16
43	IES Primeiro de Marzo	Baiona	periurb.	PO	costa	16
44	IES Ribeira do Louro	O Porriño	periurb.	PO	interior	16
45	IES de Rodeira	Cangas	vila	PO	costa	16
46	IES Castro Aobre	Vilagarcía	vila	PO	costa	16
47	IES Manuel García Barros	A Estrada	vila	PO	interior	16
48	IES Pedra da Auga	Ponteareas	vila	PO	interior	16
49	IES A Sangriña	A Guarda	rural	PO	costa	20
50	IES de Valga	Valga	rural	PO	interior	12

Como pode verse na táboa e como xa indicamos, diferenciamos catro tipos de poboación, que non só se contemplan na variable ‘Localización do centro de ensino’, senón tamén, como veremos máis adiante, na variable ‘Lugar da residencia familiar’. Este é un dos aspectos no que o LDEG se diferencia do Proxecto panhispánico, que só considera unha división en dous tipos, poboación urbana e poboación rural. En López Meirama (2008: 18-19) explícase como as particularidades na distribución poboacional de Galicia levaron a esta división, que permite diferenciar:

- i. Poboación urbana: abarca as cidades galegas con máis de 50.000 habitantes: A Coruña, Ferrol, Santiago de Compostela, Lugo, Ourense, Vigo e Pontevedra.
- ii. Poboación periurbana: abarca as conurbacións constituídas, dun xeito máis ou menos oficial, ao redor das cidades da Coruña, Ferrol, Santiago de Compostela, Vigo e Pontevedra. Estas áreas metropolitanas, aínda que resultan bastante extensas en relación co tamaño das cidades, están vinculadas a estas económica e socialmente, o que implica a coordinación de non poucos servizos públi-

cos, como os de transporte, protección civil, ordenación urbanística, acción cultural etc.

- iii. Vila: abarca os concellos de máis de 15.000 habitantes e mesmo algúns de ata 10.000 que, pola súa proximidade a núcleos urbanos importantes ou polo sector económico que caracteriza o seu territorio, non corresponden ao prototipo de poboación rural.
- iv. Poboación rural: abarca os concellos de menos de 15.000 habitantes, que se caracterizan por unha situación xeográfica de relativo illamento respecto dos núcleos urbanos máis poboados e por unha estrutura económica centrada no sector primario.

Os centros de ensino están distribuídos por provincias e por tipo de poboación exactamente do mesmo xeito ca no LDEG, onde se fixo unha proxeción dos datos dos alumnos matriculados en 2.º de bacharelato no curso 2005-2006 proporcionados pola Inspección Educativa da Consellería de Educación e Ordenación Universitaria (CEOU):

Táboa 2. Distribución dos centros por provincias e tipo de poboación

	<i>A Coruña</i>	<i>Lugo</i>	<i>Ourense</i>	<i>Pontevedra</i>	Total
pob. urbana	12	3	3	9	27
pob. periurb.	3	-	-	3	6
vila	3	1	1	4	9
pob. rural	3	2	1	2	8
total	21	6	5	18	50

Así mesmo, os informantes están distribuídos por provincias e por tipo de centro de ensino (público ou privado), segundo se indica na táboa 3, que recolle as cantidades absolutas e as porcentaxes de alumnos por tipo de centro⁹ en cada provincia e en total:

Táboa 3. Distribución dos informantes por provincias e tipo de centro

	<i>A Coruña</i>		<i>Lugo</i>		<i>Ourense</i>		<i>Pontevedra</i>		Total	
centro público	272	80,95%	96	100%	64	80%	240	83,33%	672	84%
centro privado	64	19,05%	-	-	16	20%	48	16,67%	128	16%
total	336		96		80		288		800	

Se contrastamos estas cifras coas proporcionadas pola Secretaría Xeral da CEOU relativas aos alumnos matriculados no bacharelato ordinario no

⁹ Neste caso, a proporción de estudantes formados en centros de ensino privados medrou con respecto ao curso 2005-2006; de aí que no subproxecto do galego contemos con 128 informantes (o 16% do total) repartidos en oito colexios, fronte aos 96 do subproxecto do castelán (o 12%), repartidos en seis colexios.

curso 2008-2009, observamos que as diferenzas, aínda que ineludibles pola necesidade de axustar o número de informantes aos cincuenta centros educativos, son mínimas. En todo caso, podemos comprobar que a columna da dereita da táboa 4 recolle unhas porcentaxes moi similares ás da mesma columna da táboa 3, e os gráficos que ofrecemos a continuación permiten constatar que a mostra se axusta á distribución da poboación estudiantil galega do curso en que foi recollida, tanto no que atinxe á distribución por provincias como no referido ao tipo de centro educativo.

Táboa 4. Distribución dos estudantes galegos por provincias e tipo de centro no curso 2008-2009

	<i>A Coruña</i>		<i>Lugo</i>		<i>Ourense</i>		<i>Pontevedra</i>		Galicia	
c. púb.	4971	77,94%	1800	94,14%	1234	77,46%	4478	81,95%	12483	83,68%
c. priv.	1407	22,06%	112	5,86%	359	22,54%	986	18,05%	2864	16,32%
total	6378		1912		1593		5464		15347	

Os gráficos 1 e 2 amosan as porcentaxes de estudantes en cada provincia de Galicia e na mostra no curso 2008-2009:

Gráfico 1. Distribución da poboación estudiantil por provincias

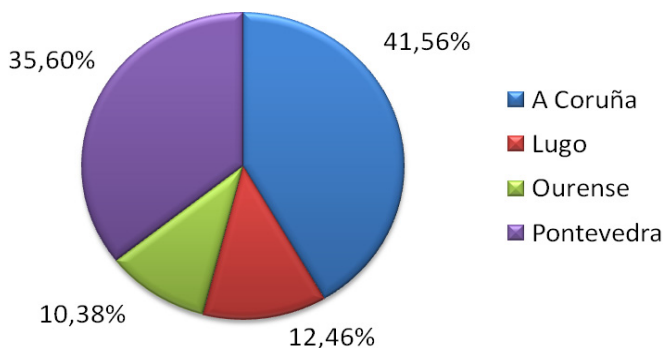
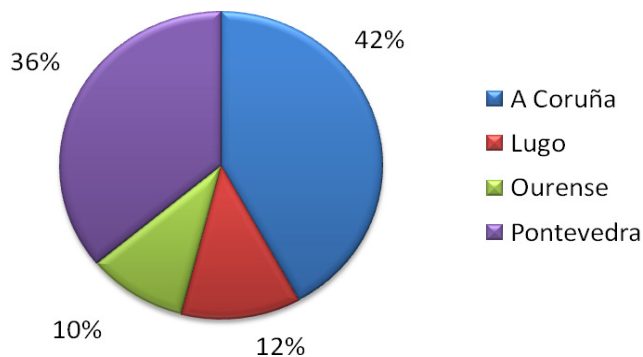
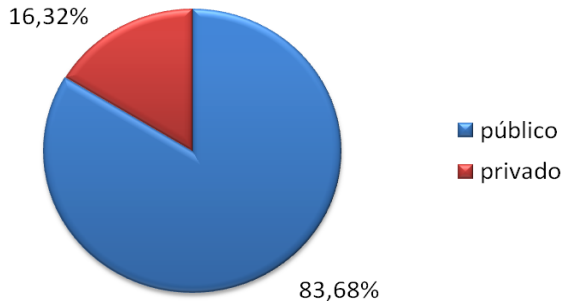


Gráfico 2. Distribución da poboación da mostra por provincias

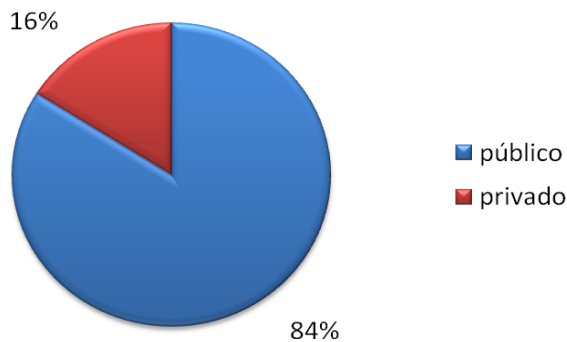


Os graficos 3 e 4 amosan as porcentaxes de estudantes en cada tipo de centro de ensino de Galicia e na mostra:

Grafico 3. Distribuci3n da poboaci3n estudantil por tipo de ensino



Grafico 4. Distribuci3n da poboaci3n da mostra por tipo de ensino



A elecci3n destes cincuenta centros escolares, no entanto, s3 asegura en principio a representatividade da mostra no que se refire as variables ‘Tipo de poboaci3n do centro’ e ‘Tipo de ensino’, xa que no resto a distribuci3n dos informantes esta feita dun xeito practicamente aleatorio, como explicaremos de seguido. Non obstante, o resultado constitue, igualmente, un retrato bastante fiel do segmento de poboaci3n estudado, o que poderemos comprobar contrastando os nosos datos cos obtidos do Instituto Galego de Estatistica (IGE) relativos ao ano 2009, salvo os referidos a lingua, que son de 2008.

En primeiro lugar, debemos comentar que a mostra e produto dunha escolma realizada sobre un total aproximado de 1.500 enquisas: de todas as elaboradas polos mozos preuniversitarios que colaboraron neste proxecto precisabamos unicamente dezaseis por cada centro de ensino, polo que, exceptuando

algún centro escolar no que non se acadou tal número (os axustes necesarios pódense consultar na columna da dereita da táboa 1), en xeral descartamos practicamente tantas como conservamos, e fixémoslo con criterios de lexibilidade e compleción¹⁰, sen manipular a selección baseándonos na información ofrecida en cada enquisa en relación co lugar de residencia familiar, a lingua inicial e habitual ou o nivel sociocultural dos pais. Si procuramos equilibrar, na medida do posible, a selección de enquisas no que se refire ao sexo dos informantes, e tamén descartamos a todos aqueles que declararon levar menos de dez anos vivindo en Galicia, porque consideramos que nese caso podería haber especiais deficiencias no coñecemento do léxico galego.

Canto ao sexo, vemos nos gráficos 5 e 6 que a mostra se desvía un par de puntos porcentuais respecto aos datos de poboación estudiantil proporcionados polo IGE, segundo o cal no ano 2009 houbo 17.943 rapazas e 14.306 rapaces matriculados no bacharelato ordinario:

Gráfico 5. Distribución da poboación estudiantil por sexo

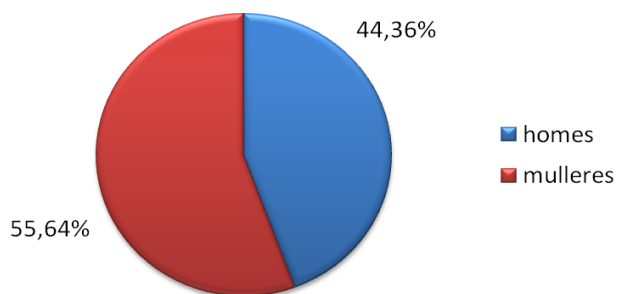
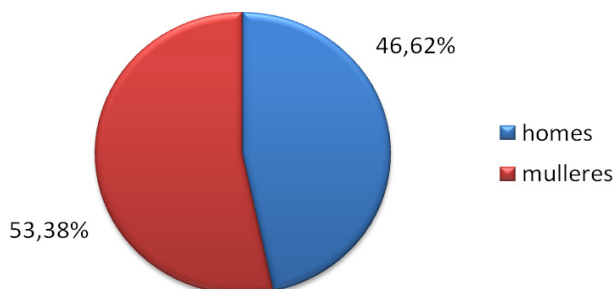


Gráfico 6. Distribución da poboación da mostra por sexo



¹⁰ Algúns informantes deixaron sen cubrir datos de estratificación social, ou mesmo algún centro de interese.

No que se refire á lingua inicial, as cifras que contrastamos están extraídas da *Enquisa de condicións de vida das familias* do IGE¹¹, e corresponden ao ano 2008. Nela, a distribución das persoas de quince a vinte e nove anos segundo a lingua que aprenderon a falar difire da dos nosos informantes, de xeito marcado na porcentaxe relativa aos que declararon que aprenderan nas dúas linguas, sensiblemente superior na mostra, como se pode ver nos gráficos 7 e 8:

Gráfico 7. Datos do IGE sobre a lingua inicial

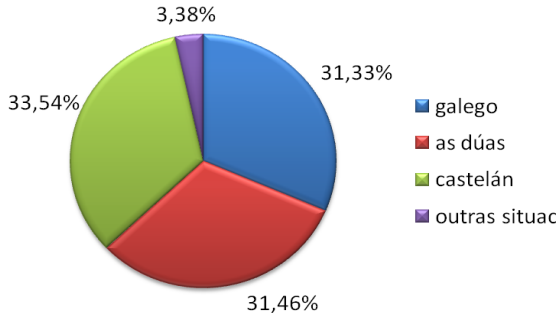
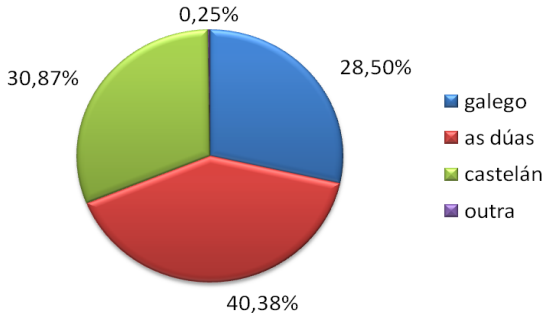


Gráfico 8. Datos da mostra sobre a lingua inicial



As diferenzas son aínda maiores nos datos relativos á lingua habitual, en primeiro lugar porque na mostra, seguindo a distribución feita no LDEG para propiciar as comparacións, establecemos un grupo específico cos informantes que declararon falar ambas as dúas linguas indistintamente, grupo inexistente na enquisa do IGE; en segundo lugar, porque o grupo dos que din falar máis castelán ca galego na nosa mostra é moi superior a todos os demais, que están repartidos de xeito proporcionado. Con todo, se lles sumamos a estes os que se declaran monolingües en castelán, o resultado acada o 57,75% dos informantes, porcentaxe case que idéntica á dos mozos de entre 15 e 29 anos que no ano 2008 manifestaron falar máis castelán ou sempre castelán (o 57,05%), e é probable que no contexto da elaboración dunha enquisa realizada no ámbito escolar e que recolle léxico galego os alumnos se amosasen máis reticentes a manifestar o seu monolingüismo en castelán. De feito, nestes dous grupos de

¹¹ Dispoñible en http://www.ige.eu/web/mostrar_actividade_estadistica.jsp?codigo=0205002

castelanfalantes a diferenza entre as porcentaxes da mostra e as da *Enquisa de condicións de vida das familias* e practicamente a mesma, en volta de 8 puntos, pero dáse, claro, en sentidos opostos.

Gráfico 9. Datos do IGE sobre a lingua habitual

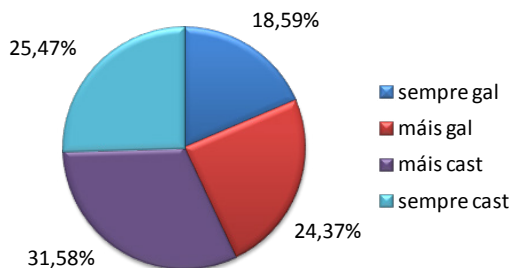
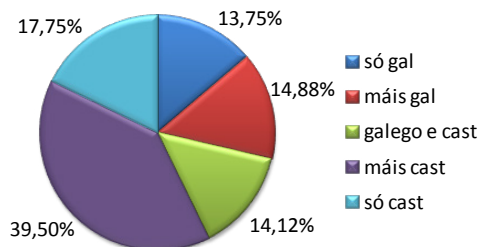


Gráfico 10. Datos da mostra sobre a lingua habitual



De todos os xeitos convén ter en conta que o segmento de poboación ao que se refiren os datos extraídos do IGE é moito maior ca o dos nosos informantes, aspecto este que talvez inflúa nas disimilitudes detectadas.

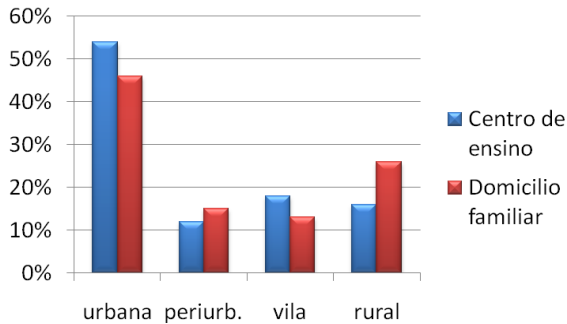
O cruzamento dos datos das dúas variables amósase na táboa 5, onde podemos apreciar que os rapaces adoitan conservar como habitual a súa lingua inicial, exceptuando os que declararon aprender a falar nas dúas linguas, que maioritariamente se decantan polo emprego do castelán como lingua de uso máis frecuente:

Táboa 5. Distribución dos informantes segundo a lingua inicial e habitual

<i>l. habitual</i> ⇨ <i>l. inicial</i> ⇩	<i>só galego</i>	<i>máis galego</i>	<i>ambas</i>	<i>máis castelán</i>	<i>só castelán</i>	total
galego	98	90	30	10	-	228
ambas	8	22	58	138	21	247
castelán	3	7	25	167	121	323
outra	1	-	-	1	-	2
total	110	119	113	316	142	800

Polo que respecta ao tipo de poboación onde se sitúa o domicilio familiar, os datos, como sería de esperar, non son moi disímiles dos relativos á variable ‘Tipo de poboación do centro educativo’, mais tampouco son absolutamente coincidentes. O gráfico 11 recolle, en porcentaxe, a cantidade de informantes que corresponde a cada un dos grupos establecidos:

Gráfico 11. Tipo de poboación



As diferenzas entre os datos de ambas as dúas variables explícanse, en xeral, polos desprazamentos de estudantes a que ás veces obriga a dispersión da poboación que caracteriza Galicia, desprazamentos que en xeral se levan a cabo desde as poboacións de menor tamaño ás máis populosas, como pode observarse na táboa 6, no que cruzamos os datos que corresponden ás dúas variables relativas ao tipo de poboación. Destaca o caso dos mozos que viven en núcleos rurais, dos que case que o 40% asisten a clase lonxe do seu domicilio:

Táboa 6. Distribución dos informantes segundo o tipo de poboación do centro educativo e do domicilio familiar

<i>pob. centro</i> ⇨ <i>pob. domicilio</i> ↓	<i>urbana</i>	<i>periurbana</i>	<i>vila</i>	<i>rural</i>	total
urbana	368	-	-	-	368
periurbana	32	83	5	-	120
vila	2	-	103	-	105
rural	30	13	36	128	207
total	432	96	144	128	800

No Proxecto panhispánico de léxico dispoñible a variable ‘Nivel socio-cultural’ refírese, como é obvio, ás familias dos informantes, e distribúese sempre en tres grupos (NSC alto, medio e baixo), pero tense recollido con sistemas parcialmente diferentes nuns subproxectos e noutros. No LDEG establecemos estes tres grupos empregando a información achegada polos rapaces acerca do nivel educativo dos seus pais, aínda que secundariamente

tamén tivemos en conta o ámbito profesional. Para que os niveis estivesen ben contrastados, establecemos uns límites moi marcados, de xeito que o NSC alto se limita aos pais con estudos universitarios e o NSC baixo corresponde aos pais sen estudos ou con estudos primarios (véxase López Meirama 2008: 28). Os datos do proxecto dedicado á lingua galega, que emprega os mesmos criterios, son os seguintes: NSC alto, 95 informantes (11,87%); NSC medio, 479 (59,88%); NSC baixo: 226 (28,25%).

Se facemos unha distribución das familias galegas en tres grupos baseada exclusivamente no criterio económico, empregando para iso os datos proporcionados polo IGE relativos ao ano 2009, comprobamos que as porcentaxes non están excesivamente afastadas das dos tres niveis da nosa mostra, e sobre todo que o NSC medio acada practicamente a mesma porcentaxe ca o grupo das familias con ingresos medios, como se ve nos gráficos 12 e 13:

Ingresos das familias:

Ata 1000 €: 22,27%

De 1000 a 3000 €: 60,28%

Máis de 3000 €: 17,45%

Gráfico 12. Distribución das familias galegas por ingresos medios (IGE)

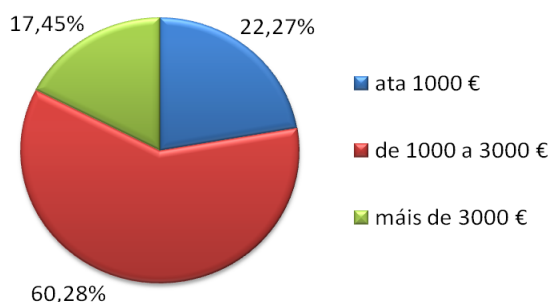
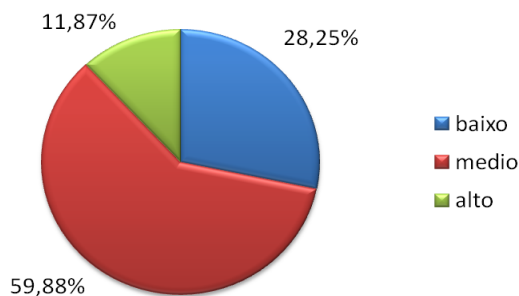


Gráfico 13. Distribución da mostra por NSC



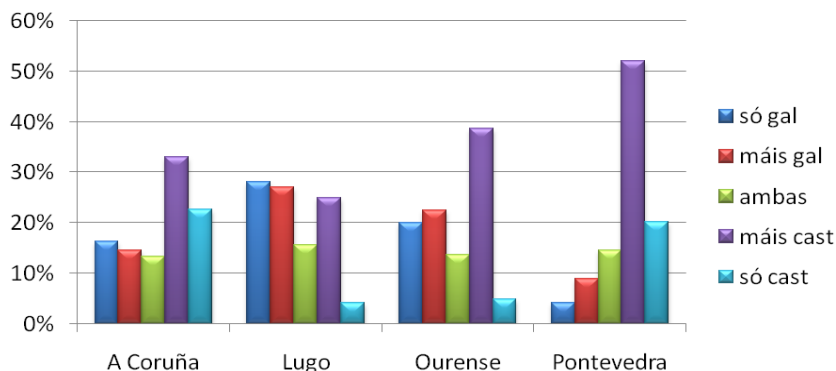
Para amosar un retrato o máis detallado posible da mostra, ofrecemos finalmente, en táboas e gráficos, os diferentes datos de estratificación social combinados co da lingua de uso habitual. Nas táboas recolleemos as cifras absolutas, e os gráficos elaborámosos sobre as porcentaxes (establecidas sobre cada grupo; por exemplo, porcentaxe de mulleres que falan exclusivamente galego sobre o total de mulleres, ou de informantes urbanos que falan só castelán sobre o total de informantes urbanos).

En primeiro lugar, a distribución dos falantes por provincia e situación —costeira ou interior— dos centros educativos debuxa un mapa xa elaborado pola sociolingüística galega dende hai anos: en xeral, o galego fálase máis no interior e nas provincias de Ourense e Lugo; na nosa mostra, sobre todo en Lugo, provincia na que os informantes que afirman falar exclusiva ou maioritariamente castelán só suman vinte e oito dun total de noventa e seis (é dicir, o 29,17%), fronte aos trinta e cinco de Ourense (o 43,75%), os cento oitenta e sete da Coruña (o 55,66%) e os douscentos oito de Pontevedra (o 72,22%). As diferenzas que atopamos entre estas dúas últimas provincias, moi marcadas no que atinxe ao emprego exclusivo do galego, que apenas chega ao 4% en Pontevedra, talvez se expliquen pola situación dos concellos que conforman a mostra (véxase táboa 1): tendo en conta que onde contabilizamos a maior parte destes falantes é nos concellos de interior, pode ser relevante o feito de que só se sumen catro municipios deste tipo en Pontevedra (con sesenta informantes, o 20,83% dos informantes da provincia) fronte aos oito da Coruña (con cento vinte e cinco informantes, o 37,20% dos informantes da provincia). Por exemplo, dos cincuenta e cinco falantes exclusivos de galego desta última, trece son informantes urbanos, todos estudantes de centros de Santiago de Compostela —cidade de interior—, e once son periurbanos, oito dos cales estudan en Ames, concello limítrofe co de Compostela (os tres restantes en Pontevedra). Fronte a esta situación, só hai un falante exclusivo de galego urbano e outro periurbano da provincia de Pontevedra, o primeiro de Vigo e o segundo do Porriño. Ademais, e seguindo nesta provincia, constatamos que a maioría dos informantes das vilas e dos concellos rurais que falan só galego (dez en total) residen no interior, dous deles na Estrada e seis en Valga. Na táboa 7 vemos os datos de distribución dos informantes por provincia e lingua habitual.

Táboa 7. Distribución dos informantes por provincia e lingua habitual

<i>l. habitual</i> ⇔ <i>provincia</i> ↓	<i>só gal.</i>	<i>máis gal.</i>	<i>ambas</i>	<i>máis cast.</i>	<i>só cast.</i>	total
A Coruña	55	49	45	111	76	336
Lugo	27	26	15	24	4	96
Ourense	16	18	11	31	4	80
Pontevedra	12	26	42	150	58	288

Gráfico 14. Distribución dos informantes por provincia e lingua habitual

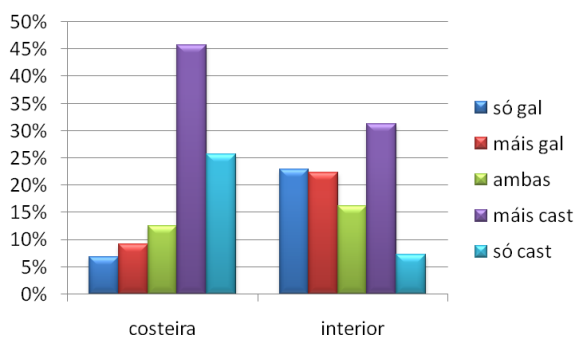


Os datos recollidos na táboa 8, referidos á distribución dos informantes segundo a lingua habitual e a situación xeográfica do centro educativo, demostran que a presenza do galego é significativamente menor na zona costeira ca no interior: 325 informantes dos 455 que estudan a carón do mar (o 71,43%) din falar só ou máis castelán, mentres que só 73 (o 16,04%) se declaran exclusiva ou maioritariamente galegofalantes; en contraste con esta situación, 156 informantes dos 345 que habitan no interior (o 45,22%) afirman falar só ou máis galego, fronte aos 133 (o 38,55%) que se declaran castelanfalantes. Os datos porcentuais completos recóllense no gráfico 15, no que se pode visualizar con claridade o contraste entre unha e outra situación xeográfica.

Táboa 8. Distribución dos informantes por situación do centro educativo e lingua habitual

<i>l. habitual</i> ⇨ <i>situación</i> ↓	<i>só gal</i>	<i>máis gal</i>	<i>ambas</i>	<i>máis cast</i>	<i>só cast</i>	total
costeira	31	42	57	208	117	455
interior	79	77	56	108	25	345

Gráfico 15. Distribución dos informantes por situación do centro educativo e lingua habitual



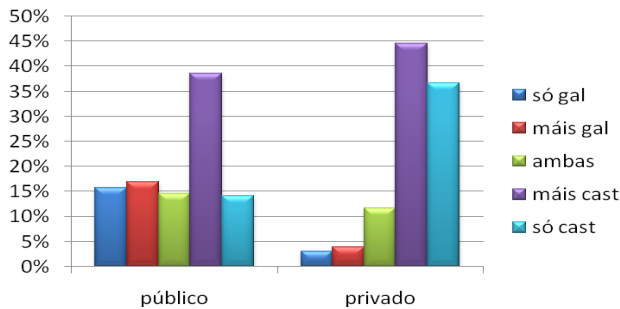
Como podemos comprobar tanto na táboa 9 coma no gráfico 16, a variable ‘Tipo de centro educativo’ (ou tipo de ensino) relaciónase directamente coa lingua escollida polos nosos informantes, no sentido de que os alumnos dos centros privados practicamente desbotan o galego como lingua de uso habitual en favor do castelán: deles, só 9 de 128 din falar só ou máis galego (apenas o 7,03%), mentres que 104 (o 81,25%) se declara castelanfalante; fronte a esta situación, o 32,74% dos estudantes dos centros públicos (220 dun total de 672) prefiren empregar o galego, porcentaxe esta bastante máis próxima á dos que se decantan polo castelán (o 52,68%, 354 rapaces).

Táboa 9. Distribución dos informantes por tipo de centro e lingua habitual

<i>l. habitual</i> ⇨ <i>tipo centro</i> ↓	<i>só gal</i>	<i>máis gal</i>	<i>ambas</i>	<i>máis cast</i>	<i>só cast</i>	total
público	106	114	98	259	95	672
privado	4	5	15	57	47	128

O gráfico 16, que recolle os datos porcentuais de cada un dos valores, amosa con claridade o contraste entre as dúas poboacións; nel podemos visualizar non só o baixo emprego do galego entre os informantes que se educan nos centros privados, xa comentado, senón tamén a elevada proporción, dentro deste grupo, dos que se declaran falantes exclusivos de castelán.

Gráfico 16. Distribución dos informantes por tipo de centro e lingua habitual



Non hai que esquecer, a este respecto, que todos os centros privados da mostra están situados en poboacións urbanas, nas que é maioritaria a presenza da lingua castelá; así mesmo, unha boa parte das familias destes alumnos corresponde ao NSC alto¹² e, como veremos máis adiante, este valor tamén se relaciona directamente co emprego do castelán. Con todo, estes datos tamén se poderían relacionar coa propia presenza da lingua galega

¹² Case o 27%, fronte ao 9% dos rapaces que estudan en centros de titularidade pública.

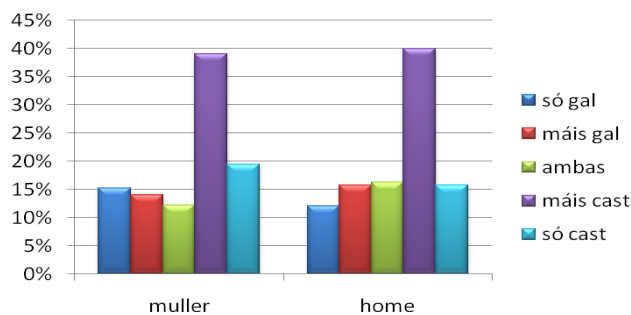
nos centros privados, que semella irrelevante, así como, talvez, co menor compromiso, en termos xerais, deste tipo de ensino de cara á potenciación do emprego desta lingua.

No que se refire ao sexo, aínda que as disimilitudes nas preferencias lingüísticas son mínimas, constatamos que as mulleres se declaran máis monolingües ca os homes, tanto en lingua galega coma en lingua castelá (máis de tres puntos porcentuais en ambos os dous casos), mentres que hai unha porcentaxe maior de homes que din falar as dúas linguas indistintamente, como se pode ver no gráfico 17. Por iso non é de estrañar que, segundo se desprende dos datos recollidos na táboa 10, a cantidade de mulleres que se decantan por falar só ou maioritariamente tanto galego coma castelán sexa lixeiramente superior ca a dos homes: as que prefiren o galego suman 125 (o 29,27% das rapazas), fronte aos 104 homes (o 27,88%), e as que empregan o castelán son 250 (o 58,55%), fronte aos 208 rapaces (o 55,76%).

Táboa 10. Distribución dos informantes por sexo e lingua habitual

<i>l. habitual</i> ⇨ <i>sexo</i> ↓	<i>só gal</i>	<i>máis gal</i>	<i>ambas</i>	<i>máis cast</i>	<i>só cast</i>	total
mulleres	65	60	52	167	83	427
homes	45	59	61	149	59	373

Gráfico 17. Distribución dos informantes por sexo e lingua habitual



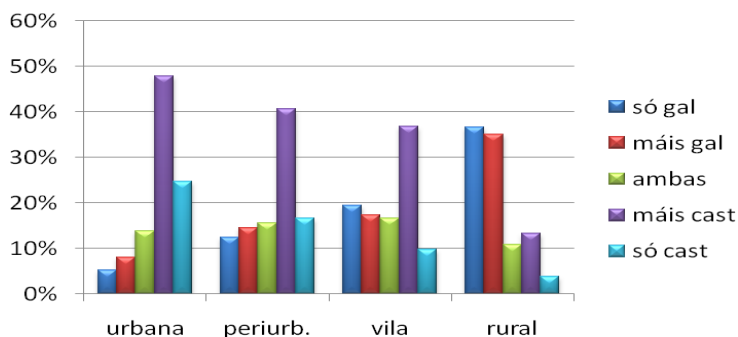
O tipo de poboación en que se sitúa o centro educativo incide significativamente na escolla da lingua habitual por parte dos estudantes de 2º curso de bacharelato, sobre todo se consideramos os dous tipos máis afastados entre si, a poboación urbana e a poboación rural; de feito, as porcentaxes dos que se decantan polo emprego exclusivo ou maioritario dunha das dúas linguas son practicamente idénticas, pero en sentido inverso: así, o 72,69% dos que estudan nas cidades (314 informantes de 432) prefiren o castelán, mentres que o 71,88% dos que o fan no medio rural (92 informantes de 128) son galegofalantes.

Táboa 11. Distribución dos informantes por poboación do centro educativo e lingua habitual

<i>l. habitual</i> ⇨ <i>pob. centro</i> ↓	<i>só gal</i>	<i>máis gal</i>	<i>ambas</i>	<i>máis cast</i>	<i>só cast</i>	total
urbana	23	35	60	207	107	432
periurbana	12	14	15	39	16	96
vila	28	25	24	53	14	144
rural	47	45	14	17	5	128

No gráfico 18 observamos tamén que nos centros educativos das vilas e das poboacións periurbanas os distintos valores, exceptuando o relativo ao uso maioritario do castelán, acadan porcentaxes similares, nun abano entre o 10% e o 20%, aínda que se observa unha lixeira preferencia polo uso do galego no primeiro destes dous tipos de poboación.

Gráfico 18. Distribución dos informantes por poboación do centro educativo e lingua habitual

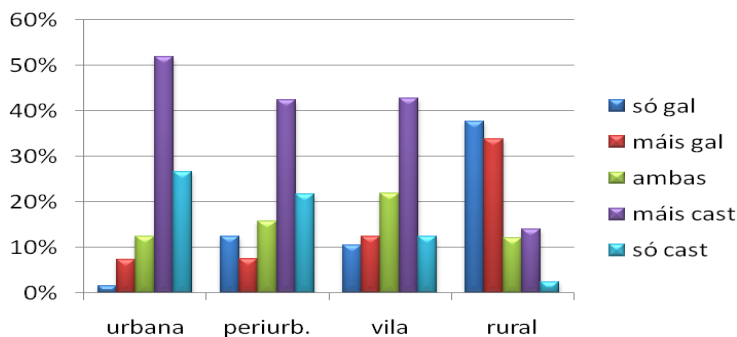


A variable ‘Tipo de poboación do domicilio familiar’ achega resultados moi similares á anterior, aínda que os contrastes son máis marcados neste caso, basicamente polo que atinxe aos informantes que habitan nas cidades, entre os que a porcentaxe de castelanfalantes sobe ao 78,53%, en claro contraste coa dos que falan só ou maioritariamente galego, que só chega ao 8,97%.

Táboa 12. Distribución dos informantes por poboación do domicilio familiar e lingua habitual

<i>l. habitual</i> ⇨ <i>pob. domic.</i> ↓	<i>só gal</i>	<i>máis gal</i>	<i>ambas</i>	<i>máis cast</i>	<i>só cast</i>	total
urbana	6	27	46	191	98	368
periurbana	15	9	19	51	26	120
vila	11	13	23	45	13	105
rural	78	70	25	29	5	207

Gráfico 19. Distribución dos informantes por poboación do domicilio familiar e lingua habitual



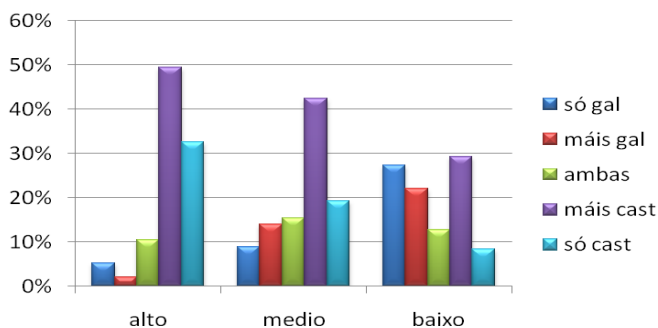
Finalmente, o nivel educativo das familias semella relacionarse directamente coa selección da lingua habitual, que é maioritariamente o castelán nas de nivel medio e, de xeito máis marcado, nas de nivel alto, e maioritariamente o galego nas de nivel baixo, aínda que neste caso as porcentaxes están moito máis equilibradas, como se pode comprobar na táboa 13 e mais no gráfico 20: observamos que só 7 dos 95 informantes do NSC alto din falar exclusiva ou maioritariamente galego (o 7,37%), fronte aos 85 de 226 informantes do NSC baixo que se declaran castelanfalantes (o 37,61%), mentres que o 82,11% dos primeiros prefiren expresarse en castelán, fronte ao 49,56% dos segundos que escollen o galego coma lingua de uso habitual.

Táboa 13. Distribución dos informantes por NSC e lingua habitual

<i>l. habitual</i> ⇔ NSC ↓	<i>só gal</i>	<i>máis gal</i>	<i>ambas</i>	<i>máis cast</i>	<i>só cast</i>	total
alto	5	2	10	47	31	95
medio	43	67	74	203	92	479
baixo	62	50	29	66	19	226

O gráfico 20 permite visualizar que as diferenzas se establecen, sobre todo, no NSC baixo, en particular no que se refire ao emprego da lingua galega, dado que só neste NSC o galego é o idioma de uso maioritario; por contra, o desequilibrio nas porcentaxes é moi evidente nos outros NSC, sobre todo no alto, no cal o emprego do galego non pasa de ser anecdótico.

Gráfico 20. Distribución dos informantes por NSC e lingua habitual



Como recapitulación de todos estes datos cabe dicir que, salvo a relativa ao sexo dos informantes, as variables de estratificación social da nosa mostra se relacionan coa escolla de lingua habitual, de xeito que un retrato do estudante prototípico de segundo de bacharelato que fala galego podería ser o do rapaz ou rapaza cuxos pais teñen un nivel formativo baixo, que vive no interior de Galicia, en particular nas provincias de Lugo ou Ourense, que habita e estuda, nun centro de ensino público, no rural.

2.3 Criterios de edición dos materiais

Como xa indicamos, este traballo segue moi de preto a metodoloxía empregada na elaboración do LDEG, o cal permite a realización de cotexos directos e significativos entre o léxico dispoñible do galego e o léxico dispoñible do castelán de Galicia¹³. Da mesma maneira que sucedía neste, dous son os principios fundamentais que rexen a edición dos materiais no LDG: un principio ao tempo unificador e estandarizador, que permite a agrupación de variantes así como a eliminación das formas que non se adecúan ao solicitado, e un segundo principio de respecto ou fidelidade ás respostas ofrecidas polos informantes. Procuramos un equilibrio entre ambos os principios, que tentaremos describir nas liñas que seguen.

¹³ De súa vez, o LDEG segue os criterios de edición do proxecto panhispanico, empregados igualmente na elaboración de diversos léxicos dispoñibles do español (véxanse, por exemplo, González Martínez 2002, Galloso Camacho 2003 ou Hernández Muñoz 2004). En calquera caso, non sempre existe un criterio común de edición (compárese, por exemplo, Mateo García 1998, Galloso Camacho 2002 e Gómez Molina / Gómez Devís 2004 en relación co mantemento ou supresión das marcas comerciais). Remitimos ao LDEG para máis detalles e referencias bibliográficas sobre a xustificación da adopción dos diferentes criterios (López Meirama 2008: 17-91). Para unha presentación dalgúns dos criterios máis relevantes empregados na edición dos léxicos dispoñibles pode consultarse Samper Padilla (1998) ou Samper Padilla / Bellón Fernández / Samper Hernández (2003).

2.3.1 Formas eliminadas e formas conservadas

Aínda que, de maneira xeral, conservamos as respostas ofrecidas polos informantes, en todos os CI levamos a cabo a eliminación dalgunhas palabras. De calquera xeito, fomos bastante moderadas á hora de levar a cabo estas supresións, en aras da fidelidade a que acabamos de facer referencia. A continuación sinalaremos cales foron os criterios xerais empregados para realizar as eliminacións, aínda que no comentario individual dos CI realizado en § 2.3.4 daremos conta dos criterios particulares usados en cada caso:

- i. En ocasións un informante repite unha palabra nun mesmo CI ou ofrece dúas formas que, de acordo cos criterios que se comentarán en § 2.3.2 e § 2.3.3, deben ser asignadas ao mesmo lema: variantes flexivas, forma galega e forma castelá cognada, forma acurtada e forma plena, variantes galegas que comparten a mesma raíz... En todos estes casos mantemos en xeral a primeira mención e suprimimos a seguinte ou, no seu caso, as seguintes. No entanto, cando a primeira das palabras amosa claramente unha errata, fronte á segunda, conservamos esta última.

Dado o interese que pode ter coñecer a coexistencia nunha mesma listaxe de variantes galegas e dos seus cognados casteláns (por exemplo *culler* e *cuchara*) ou ben de dúas variantes fónicas ou morfolóxicas galegas (por exemplo, *igrexas* e *eirexas*), en § 2.3.3.4 ofrecemos a nómina completa de parellas deste tipo rexistradas nas listaxes iniciais do LDG.

- ii. Debido ao método empregado para a obtención das listaxes, enquisas nas que os informantes van escribindo as diferentes respostas a medida que lles van xurdindo na mente, é natural que moitas de tales respostas se obteñan por asociación inmediata dunhas con outras. A modo de exemplo, vexamos esta listaxe ofrecida para o CI 18 ‘O mar’:

(...) conchas auga algas praias surf body-surf toalla sandalias crema xeados frotador burbulla socorrista peixes nadadores bañistas bañadores bikinis (...)

Como pode observarse, existe un nexos de conexión evidente entre gran parte das palabras fornecidas: a voz *praias* evoca as actividades que se realizan nela e todo o que ten que ver con tales actividades —*surf*, *toalla*, *socorrista*...—, e dentro desta asociación, *frotador* vincúlase con *burbulla*, *bañadores* con *bikinis* etc.

Agora ben, en ocasións as sucesivas asociacións que se realizan conducen a respostas que só posúen un vínculo moi feble co CI sobre o que se está a facer a enquisa ou mesmo que quedan absolutamente fóra del. En casos deste tipo, procedemos á eliminación das formas en cuestión. Así, por exemplo, no CI 14 ‘Os animais’ eliminamos, entre outros rexistros, *o efecto mariposa* (que figura despois da voz *mariposa*) ou no CI 15 ‘Os xogos e as distraccións’ suprimimos o rexistro *amarelo* (proporcionado despois de *parchís e ficha*). Por suposto, tamén eliminamos aquelas formas nas que mesmo é difícil rastrexar o vínculo asociativo que conduciu ata elas (por exemplo, *fotosíntese* no CI 14 ‘Os animais’).

No entanto, cando as asociacións teñen un certo vínculo coa temática do CI e son establecidas por varios informantes, tendemos a conservar as formas ofrecidas. Así, por exemplo, é considerable o número de estudantes que estenden o contido do CI 02 ‘A roupa’ á súa adquisición e ao gasto que implica. Deste xeito, mantemos rexistros como *gastar, comprar, diñeiro, vendedor...*¹⁴. De igual xeito, mantemos no CI 01 ‘O corpo humano’ denominacións de diferentes enfermidades ou voces relacionadas con elas (*alerxia, colesterol, hepatite, miopía, neuralxia...*), posto que esta asociación é relativamente frecuente nas respostas dos informantes. Porén, se as formas ofrecidas se corresponden claramente con outros CI do LDG, tendemos a eliminar as palabras: así, por exemplo, suprimimos *calcetín, zapatos* ou *xersei* do CI 01 ‘O corpo humano’, non só polo feito de que o vínculo con este CI non é demasiado forte, senón tamén porque tales formas xa teñen o seu lugar no CI 02 ‘A roupa’. Da mesma maneira, suprimimos do CI 07 ‘A cociña e os seus utensilios’ a maior parte das respostas que designan alimentos (*chocolate, laranxas, pastas, grelos...*), ás que lles corresponde o CI 05 ‘Alimentos e bebidas’.

En calquera caso, debe terse en conta que, como xa tivemos ocasión de comprobar, os diferentes rexistros ofrecidos polos informantes posúen un contexto fundamental: as respostas anteriores e posteriores. Moitas veces é tan só a través do coñecemento deste contexto como se lle pode dar sentido á presenza dun rexistro concreto nun CI determinado. Vexamos un exemplo: é moi probable que se percorremos a listaxe de lemas correspondentes ao CI 11, ‘O

¹⁴ De maneira xeral, ao longo desta introdución exemplificaremos con lemas normalizados e non con variantes concretas presentes nas listaxes. No entanto, nalgunha ocasión faremos referencia ás variantes. Se estes casos excepcionais non quedan suficientemente claros polo contexto, ofreceremos unha indicación ao respecto.

campo’, nos chame a atención o lema *tirarse*, e mesmo que pensemos que tal lema debería ser eliminado por non ser tematicamente pertinente. No entanto, se acudimos á listaxe de formas ofrecidas polo informante, podemos entender a presenza deste verbo no CI en cuestión: *vexetación, flor, correr, brincar, tirarse, andar a rolos*. Como pode observarse, o estudante ofrece diferentes actividades que se poden realizar nun día de campo e dentro deste marco, o verbo *tirarse* é perfectamente acaído ao CI. Así pois, a presenza nas listaxes dalgún dos rexistros aparentemente menos pertinentes pode explicarse precisamente porque no proceso de revisión e preparación dos materiais foi consultado o seu contexto de aparición e, a partir do coñecemento deste, concluímos a conveniencia de mantelo.

- iii. Outro grupo de formas eliminadas está constituído por aquelas que posúen un carácter moi xenérico, de tal xeito que, sen poder consideralas alleas ao CI de que se trate, non parecen fornecer información relevante sobre este. Por poñer un exemplo, no CI 13 ‘Os traballos do campo e do xardín’, eliminamos voces como *falar, home, muller, mirar, señor...*
- iv. Tamén prescindimos dalgunhas series das que botan man os informantes para ir cubrindo as celas do cuestionario con pouco esforzo, moitas veces coa intención de enchelas de vez. Así, por exemplo, no CI 02 ‘A roupa’ atopámonos cunha listaxe de talles que eliminamos (*s, x, s, m, l, xl*). No CI 15 ‘Os xogos e as distraccións’ descubrimos no medio das respostas cinco versións do videoxogo FIFA (*Fifa 00, Fifa 01, Fifa 02, Fifa 03, Fifa 04*), das que nós só mantemos o primeiro rexistro lematizado baixo a voz xenérica FIFA, ou tres da *PlayStation* (*PlayStation 1, Play station 2, Play station 3*), das que mantemos tamén só o primeiro rexistro lematizado baixo *PlayStation*. Outro exemplo deste tipo, entre algúns outros que poderíamos citar, atopámolo no CI 12 ‘Os medios de transporte’, no que un alumno ofrece este grupo de rexistros: *moto 50 cc, moto 125 cc, moto 250 cc, moto 600 cc, moto 1000 cc, moto 1200 cc, moto 1600 cc*. Só conservamos o primeiro rexistro lematizado baixo *moto*.
- v. Noutras ocasións, suprimimos formas nas que é palpable o desexo do informante de transgredir a tarefa encomendada. Son relativamente frecuentes en case todos os CI (nuns en maior medida ca noutros) voces con claras referencias sexuais (*cipote* ou *zoofilia* no CI 13 ‘Os traballos do campo e xardín’, *sexo* ou *tetorras* no CI 09

‘A calefacción e a iluminación?..’), escatolóxicas (*cagar* no CI 15 ‘Os xogos e as distraccións’, *mexar* no CI 08 ‘A escola: mobles e materiais’...) ou, en xeral, formas coas que o informante procura un efecto lúdico e/ou dá conta das súas ideas e opinións (*Babe o cerdito valiente*, *Garfield* ou *George W. Bush* no CI 14 ‘Os animais’, *algúns só decoran* —despois de profesor— no CI 08 ‘A escola: mobles e materiais’, *pirómano no cárcere* no CI 09 ‘A calefacción e a iluminación’, *mar-celo*, *mar-mota*, *mar-tín* no CI 18 ‘O mar’...). Todas as formas citadas, escollidas a modo de ilustración entre outras varias, foron eliminadas.

- vi. Retomando un dos exemplos que se acaban de ofrecer, eliminamos os rexistros que conteñen predicacións, na medida en que o que nos interesa é recoller *léxico*, non enunciados creados a partir de unidades léxicas. Deste modo, quedan fóra expresións (fixadas ou non) rexistradas en diferentes CI, como a xa mencionada *algúns só decoran* no CI 08, *acende o candil Marica chus chus*¹⁵ e *fai un calor de carallo* (no CI 09 ‘A calefacción e a iluminación’) ou *calma os problemas de cada un* (no CI 18 ‘O mar’).
- vii. Si conservamos, pola contra, expresións complexas non lexicalizadas que se empregan para designar determinadas realidades concibidas como unitarias: así, a carón de *paneira*, no CI 06 ‘Os obxectos colocados na mesa para a comida’, poderemos atopar *bandexa do pan*, *recipiente para o pan*, *fonte do pan*, *prato do pan*... En ocasións, estas expresións poden dar conta do desexo do informante de nomear determinadas realidades das que descoñece, ou non lembra, a denominación habitual. Neste sentido, non debe estrañar tampouco a presenza nas listaxes de compostos morfolóxicos e derivados non recollidos nos dicionarios como *pousafontes* ou *pousapotas* (equivalentes seguramente de *gardamanteis*, obxecto empregado para pousar os recipientes quentes sobre a mesa), *colgatoallas*, *cubridor*, *suxeitotoallas*, *secalouzas*... Coi-damos que pode ser de especial interese para o investigador descubrir os distintos recursos de que botan man os informantes para verbalizar un mesmo concepto e, por tal motivo, mantemos todos estes rexistros, aínda que carezan de lexicalización e non estean referendados polos dicionarios. Porén, eliminamos algunhas pe-rífrases complexas que pouco achegan ao investigador: por poñer dous exemplos, no CI 08 ‘A escola: mobles e materiais’ suprimi-

¹⁵ Modificación da parte inicial do retrouso dunha canción popular galega (*Apaga o candil, Marica, chus, chus*).

mos *couso de madeira para poñer o cadro mentres pintas e cousa sobre a que se proxecta*.

- viii. Tamén mantemos algunhas creacións dos informantes que resultan de peculiares adaptacións ao galego de palabras ou expresións do castelán, como *vaia* (adaptación de *valla*) ou *petivermello* (adaptación de *petirrojo*). Aínda que varias destas formas son creacións *ad hoc*, sen vitalidade máis alá das propias listaxes, coidamos que resultaba interesante mantelas, entre outros motivos, para deixar constancia dos diferentes recursos dos que botan man os informantes á hora de “galeguizar” voces ou expresións do castelán.
- ix. Conservamos igualmente algunhas respostas cuxa forma, na nosa opinión, resulta dun cruzamento con outras palabras ou dun deficiente coñecemento do informante, sempre e cando teñamos unha intuición relativamente clara do lema co que debemos vinculalas como variantes. Así, por exemplo, no CI 01 ‘O corpo humano’ mantemos *incisores* (por *incisivos*: o contexto é *dentes, moas, incisores*) ou *ceballas* (por *cellas*: o contexto é *ollos, ceballas, pestanas*).
- x. Como xa sinalamos, nas instrucións que cumpría seguir para realizar as diferentes enquisas, púñase fincapé no feito de que os informantes non deberían incluír marcas comerciais nin nomes propios nas respostas. Malia esta indicación, nalgúns CI, como por exemplo no CI 02 ‘A roupa’ (*Nike, Zara, Adidas...*), no CI 15 ‘Os xogos e as distraccións’ (*PlayStation, Monopoly, Trivial...*), no CI 05 ‘Os alimentos e as bebidas’ (*Coca-Cola, Fanta, Kas...*) ou no CI 12 ‘Os medios de transporte’ (*Vespa, Zodiac, Vespino...*), é notoria a presenza de nomes de diversas marcas, ocupando algúns deles postos elevados nos niveis de dispoñibilidade léxica. Tendo en conta que, malia as instrucións en contra, as palabras deste tipo xorden con moita frecuencia nas respostas dalgúns CI, consideramos que debiamos conservarlas en todos eles (sempre que fosen tematicamente pertinentes), deixando constancia da súa representatividade nas listaxes¹⁶. Por suposto, tamén conservamos os substantivos lexicalizados procedentes de marcas: *ioió, plastilina, tabasco...*
- xi. Fomos algo máis restritivas cos antropónimos e topónimos, menos frecuentes nas respostas, mantendo só aqueles que aparecían con

¹⁶ Empregamos un criterio diferente ao utilizado no LDEG, pois neste suprimíronse todos os nomes de marcas que non atinxían unha mínima representatividade numérica. O número inicial de nomes comerciais era considerablemente maior no LDEG porque nas instrucións sobre a realización das enquisas non se advertía da conveniencia de non introducir tal tipo de palabras nas listaxes. *Vid.* § 3.1.4 a este respecto.

certa recorrencia nas listaxes. Por poñer un exemplo, no CI 18 ‘O mar’ conservamos as denominacións de mares e océanos, pero desbotamos algúns topónimos designadores de poboacións costeiras galegas (probablemente poboacións coas que os informantes tiñan unha vinculación persoal: *Baiona, Muros, Muxía, Noia, Portonovo* etc.). Nos comentarios específicos sobre os diferentes CI poden atoparse máis exemplos de nomes propios suprimidos. Si, conservamos, por suposto, as expresións complexas que conteñen nomes propios, sempre que tales expresións sexan acaídas ao CI en que figuran (*azul Klein, azul Prusia, coche de san Fernando, pipeta Pasteur* etc.), así como os substantivos propios que sufriron un proceso de lexicalización (*bikini, rebeca, sándwich* etc.).

2.3.2 Criterios xerais de lematización

Co obxectivo de unificar variantes (sobre todo fónicas, morfolóxicas, ortográficas e flexivas) e de ofrecer, na medida do posible, formas normalizadas que faciliten a busca do usuario, vinculamos a un *lema* (noutros léxicos dispoñibles, *vocabulo*) todas e cada unha das palabras conservadas nas listaxes. Deste xeito, e por poñer un exemplo, respostas como *abductor, abdutores, abductor* ou *adutor* aparecerán recollidas baixo o mesmo lema *abductor*, forma normalizada sobre a que se realizan os cálculos de dispoñibilidade léxica, frecuencia etc. Aínda que ao longo das páxinas que seguen ampliaremos información con respecto á selección do lema, de maneira xeral podemos agora sinalar que escollemos como tal a forma estándar, coa ortografía normativa e na forma de cita habitual en función da súa categoría gramatical. Agora ben, fronte a outros léxicos dispoñibles, o LDG (ao igual ca o LDEG) ten a peculiar vantaxe de permitir coñecer as variantes concretas que son realización de cada lema, o cal, no caso do léxico dispoñible do galego, nos parece especialmente importante. Estas variantes, que se poden consultar no apéndice 1, recóllense na forma na que foron ofrecidas polos informantes, incluíndo, daquela, as faltas de ortografía e as erratas (*oreyas, patatas cocicas...*)¹⁷.

Antes de expoñer os diferentes criterios empregados no proceso de lematización das variantes, debemos indicar que este se realiza de maneira independente para cada CI. Deste xeito, cando para a selección do lema empreguemos, por exemplo, criterios de frecuencia (tal e como se comentará nas vindeiras páxinas), debe terse en conta que a frecuencia se mide dentro de

¹⁷ Así pois, e continuando co mesmo exemplo ofrecido previamente, o **lema** *abductor* realízase a través de sete **variantes** distintas (*abductor, abductores, abductor, abductores, aductor, adutores*), que se corresponden cun total de 10 **palabras, rexistros** ou **respostas** (posto que hai dous exemplos de *abductor, abdutores* e *aductor* e un das restantes variantes).

cada CI e non no conxunto do LDG, e máis en concreto, dentro do conxunto de palabras que se agrupan baixo un mesmo lema. Isto implica que en diferentes CI poderemos atopar lemas diferentes para expresar a mesma unidade léxica: así, por exemplo, no CI 06 ‘Os obxectos colocados na mesa para a comida’ atoparemos o lema *tesoira*, porque, no conxunto das respostas dos informantes, as formas en singular son máis frecuentes ca as formas en plural, mentres que no CI 07 ‘A cociña e os seus utensilios’ atoparemos o lema *tesoiras* porque, contrariamente, neste CI son máis frecuentes os rexistros en plural ca en singular. De igual xeito, o criterio empregado determina que mesmo dentro dun único CI unha palabra pertencente a varias expresións complexas de estrutura similar poida aparecer lematizada de dúas formas distintas segundo a expresión: así, no CI 2 ‘A roupa’ atopamos o lema (*pantalón*) *pitillo*, porque as formas en singular recollidas baixo tal lema son máis frecuentes ca as formas en plural, pero o lema (*pantalóns*) *piratas*, porque no conxunto dos rexistros agrupados baixo esta forma predominan as respostas en plural sobre as respostas en singular.

Optamos por traballar coa frecuencia dentro de cada un dos CI e nestes, dentro de cada grupo de variantes (e non no conxunto de exemplos do LDG), porque consideramos que existía a posibilidade de que ben a acepción, determinada pola temática do CI, ben a expresión complexa na que se realiza unha palabra, puidese condicionar nalgúns casos a selección dunha ou outra variante. Entendemos, pois, que era ben respectar a idiosincrasia de cada CI e de cada expresión pluriverbal.

2.3.2.1 O xénero

Como xa sinalamos, o lema permite o agrupamento de variantes de diferente tipo. No que atange á variación de xénero, nos adxectivos escollemos como lema o masculino, pois esta é a forma de cita habitual nos dicionarios e a máis frecuente nas respostas dos informantes. Deste modo, no CI 17 ‘As cores’ *bonito* e *bonita* lematízanse baixo *bonito* ou no CI 10 ‘A cidade’ *aburridas* baixo *aburrido*. Non debe estrañar a presenza nas listaxes elaboradas polos informantes dalgúns adxectivos en feminino, na medida en que os estudantes con frecuencia establecen un vínculo entre tales adxectivos e ben as respostas precedentes, ben o tema do CI: así, por exemplo, a forma *aburridas* aparece ao pé de *grandes* e *iluminadas*, tres adxectivos cos que o informante achega características xerais que el atribúe ás poboacións urbanas.

O criterio de lematización é diferente nos substantivos designadores de seres sexuados. Neste caso lematizamos sempre co xénero que proporcionan

os informantes. Se unicamente hai casos da forma masculina ou da forma feminina, o lema aparece en masculino ou en feminino, respectivamente. Así, no campo 13, ‘Os traballos do campo e xardín’, *porco* e *porcos* lematízanse baixo *porco*, posto que non hai exemplos de *porca*, mentres que no CI 16, ‘As profesións e os oficios’, o único exemplo que se rexistra de *depiladora* aparece lematizado tamén baixo *depiladora* porque non hai exemplos do masculino.

Se nas listaxes aparecen os dous xéneros, recollemos as dúas terminacións separadas por guión, colocando en primeiro lugar a máis frecuente. Así, para os 66 exemplos de *administrativo* e os 5 de *administrativa* o lema será *administrativo-a*, mentres que para os 4 exemplos de *maquilladora* e os 2 de *maquillador* o lema será *maquilladora-o*. Estes casos son especialmente frecuentes no CI 16, onde se recollen numerosas denominacións de oficios e profesións, como as que acabamos de comentar. Coidamos que pode resultar interesante para os investigadores coñecer se se rexistran ou non exemplos designadores dos dous sexos e no seu caso, cales son as respostas maioritarias. Aínda que, por suposto, a esta información sempre se pode acceder a través das variantes, o feito de recollela no lema facilita a consulta de quen estea interesado en estudar aspectos relacionados con esta cuestión.

Non unificamos os substantivos masculinos e femininos designadores de seres sexuados cando algún dos xéneros remite a acepcións específicas non compartidas polo outro xénero. Así, por exemplo, no CI 16 atoparemos os lemas diferenciados *modista* e *modisto*, na medida en que, segundo sinala o *Diccionario da Real Academia Galega* —Real Academia Galega (1997)¹⁸— (en adiante DRAG), a forma feminina *modista* posúe unha acepción non compartida por *modisto*: ‘muller que fai pezas de roupa ou as amaña’. Tampouco figurará o dobre xénero no lema cando os dicionarios interpreten que o substantivo é epiceno, por moito que os informantes ofrezan exemplos de variación xenérica: así os exemplos de *cangura* lematízanse baixo *canguro*. Por outra banda, non se unifican os masculinos cos femininos irregulares. Deste xeito, formas como *actor* e *actriz*, *alcalde* e *alcaldesa* ou *galo* e *galiña* terán entradas independentes.

No que atangue aos seres non sexuados, e tal e como cabía esperar, non unificamos os substantivos relacionados semántica e formalmente, pero que posúen significados parcialmente distintos: *xerra* e *xerro*, *sacha* e *sacho*,

¹⁸ Cando editamos os materiais, a nova edición do DRAG (Real Academia Galega 2012, en liña en <http://www.realacademiagalega.org/diccionario#inicio.do>) non estaba aínda a disposición dos usuarios. Por tal motivo, nomearemos con frecuencia nestas liñas a primeira edición do Diccionario da Real Academia Galega (DRAG), que contaba con 25.000 entradas. No entanto, tamén faremos referencia nalgunha ocasión á nova versión electrónica, que denominaremos DRAG2 e que conta con 50.000 artigos lexicográficos.

dedo e *deda* etc. Si unificamos, pola contra, as formas deste tipo que teñen o mesmo significado: *azucreira* e *azucreiro* (baixo a forma estándar *azucreiro*), *licuador* e *licuadora* (baixo *licuadora*...) etc.

No caso de que as diferenzas de xénero se establezan entre substantivos ou adxectivos que funcionan como complementos en sintagmas máis amplos, levamos a cabo a unificación baixo o masculino, mesmo se tratamos con seres sexuados. Así, *despacho do director* e *despacho da directora* unifícanse baixo *despacho do director*.

2.3.2.2 O número

De maneira xeral, e seguindo o mesmo criterio empregado nos dicionarios, os substantivos e adxectivos aparecen lematizados en singular. Este tamén é o número habitual nas respostas dos informantes, o que non obsta para atoparmos con certa frecuencia algúns rexistros en plural (*vid.* en 2.3.2.1 o comentario sobre a presenza do adxectivo *aburridas* no CI 10). Nas liñas que seguen, comentaremos algunhas excepcións a este tratamento.

Os substantivos que posúen unha acepción específica en plural son lematizados neste número: así, por exemplo, no CI 1 ‘O corpo humano’, podemos atopar o lema *miolos*, pois a forma plural ten un valor particular, tal e como rexistran os dicionarios (o plural de *miolo* é definido como ‘encéfalo’ no DRAG). O mesmo vale, por exemplo, para *rebaixas* ou *lentes*.

Por outra banda, e co obxectivo de ser o máis fieis posible ás respostas dos informantes e ao tempo, de evitar lemas algo artificiosos, tamén ofrecemos o lema en plural cando este número é maioritario nas respostas e algún dos dicionarios consultados (DRAG, *Gran Dicionario Xerais da Lingua*, en adiante GDXL —Carballeira Anllo 2009²— ou *Diccionario de galego Digalego*, en adiante *Digalego*) sinala que é especialmente frecuente o uso en plural da palabra en cuestión. Así, por exemplo, os dous rexistros *ovas* que se recollen no CI 18 ‘O mar’ lematízanse baixo *ovas*, porque, malia que esta forma se pode considerar o plural semanticamente regular de *ova*, o DRAG sinala o seu uso frecuente en plural (“OBS. Emprégase xeralmente en plural”) e este número é maioritario (neste caso exclusivo) nas respostas dos informantes. Como se pode observar, aplicamos un criterio lexicográfico, de tal xeito que o lema só aparecerá en plural se algún dos dicionarios indicados informa da preeminencia cuantitativa deste número no uso da palabra. Isto leva sen dúbida a algunhas incongruencias, como o feito de que no CI 05 ‘Os alimentos e as bebidas’, atopemos o lema *macarróns* fronte ao lema *fideo*, sendo o plural moito máis frecuente nas respostas dos informantes en ambos os dous casos. No entanto, desde o punto de vista práctico precisabamos dalgún criterio

sistemático ao que aternos (máis alá da mera intuición sobre os niveis de frecuencia de singular e plural) e entendemos que o referendo lexicográfico podía axudarnos nesta tarefa.

Cando o dicionario informa da presenza habitual do plural, pero este número é minoritario nas respostas ou coincide co número de respostas en singular, lematizamos en singular: deste modo, no CI 03 'As partes da casa', atoparemos o lema *alicerce*, posto que nas listaxes hai un exemplo de *alicerces* e outro de *alicerce*.

Utilizamos igualmente o criterio de frecuencia naqueles casos nos que se emprega o singular e o plural coa mesma designación, facendo referencia a un único elemento constituído por dúas partes simétricas: *tesoira(s)*, *pantalón(s)*, *alicate(s)*... Neste caso, e sempre que o *Vocabulario Ortográfico da Lingua Galega*, en adiante VOLG (González González / Santamarina Fernández 2004), e/ou os dicionarios arriba citados recollan dunha ou outra maneira o uso en plural para designar un único elemento, apelamos ás respostas dos informantes para seleccionar o lema: lematizamos en singular cando os rexistros do CI son máis frecuentes en singular e lematizamos en plural cando son maioritarias as respostas neste número. En caso de coincidencia, o lema escollido é o singular. Se as obras de referencia non rexistran o uso en plural, lematizamos tamén en singular.

De por parte, unificamos singulares e plurais cando as diferenzas de número se establecen entre substantivos e adxectivos que funcionan como complementos en sintagmas máis amplos. Neste caso, empregamos tamén un criterio de frecuencia á hora de escoller entre o singular e o plural. *Limpar o monte* e *limpar os montes* lematízanse baixo *limpar o monte* porque hai tres exemplos do primeiro rexistro e só un do segundo, pero *recollida das patacas* e *recollida da pataca* unifícanse baixo a forma plural, porque hai máis casos co complemento en plural ca en singular.

Con respecto ao segundo compoñente de compostos morfolóxicos, lematizamos co número que figura nos dicionarios (*abrebotellas* ou *tirarrollas* pero *lavalouza*). Agora ben, son bastante numerosas as formas recollidas nas listaxes que non se encontran nas obras lexicográficas: así, atoparemos rexistros como *limpasolos*, *cortapizza*, *apoia vasos*, *apoia vasos*, *pousa-fontes*... Nestes casos, lematizamos en principio de acordo co rexistro máis frecuente nas respostas dos informantes dentro de cada CI (por exemplo, *cortapizza* no CI 06, pero *cortapizzas* no CI 07), agás que o dito rexistro presente unha forma bastante antiintuitiva (por exemplo, *abrecorcho* lematízase como *abrecorchos*).

2.3.2.3 As palabras con morfemas avaliativos

As palabras que posúen morfemas avaliativos son lematizadas independentemente cando entendemos que posúen ao menos certo grao de lexicalización. Deste modo, teñen lemas independentes, por suposto, formas como *camisón*, *fouciño* ou *corpiño* (este último no CI 02 ‘A roupa’), pero tamén outras sen presenza no dicionario pero que aparecen con certa frecuencia nas listaxes a carón das formas non sufixadas, e que parecen designar realidades con certa especificidade: *armariño*, *lampadiña*, *lampadiña de estudio*, *pratiño*... Aínda que é verdade que estes substantivos denotan armarios, lámpadas e pratos de pequeno tamaño, non é menos certo que poden designar ao tempo obxectos con características específicas (armarios pequenos que se colgan nas paredes, pratos que se empregan para a sobremesa ou determinadas lámpadas que se colocan sobre as mesas), o que xustificaría a súa presenza relativamente frecuente nas respostas, así como a coexistencia de formas afixadas e non afixadas nunha mesma listaxe.

Pola contra, lematízanse baixo a forma non sufixada aqueles rexistros nos que consideramos que non hai lexicalización, de tal xeito que o morfema avaliativo só indica tamaño, intensidade ou unha valoración cualitativa: *caloriño* lematízase baixo *calor*, *casíña* baixo *casa* (no CI 11 ‘O campo’) ou *quentiño* baixo *quente*. Tamén lematizamos sen diminutivo os substantivos en *-iño* que designan ou poden designar crías de animais: *gatiños*, *xureliños*, *pitiños*...

Finalmente, cando nas respostas dos distintos informantes atopamos formas sufixadas claramente lexicalizadas e formas non sufixadas, ambas con entradas no dicionario e coa mesma designación, escollemos como lema a máis frecuente. Así, por exemplo, as locucións *cabaliño do demo* e *cabalo do demo* lematízanse baixo a forma con diminutivo por ser máis numerosa nos rexistros dos informantes.

2.3.2.4 Os verbos

Os poucos verbos que non figuran en infinitivo nas listaxes son lematizados con esta forma verbal: *sulfata* baixo *sulfatar* ou *vota o abono* baixo *botar o abono*. No entanto, este tratamento non é aplicable, como cabía esperar, aos casos nos que se produciu unha substantivación —*pillá (pillá)*, *corre que te pilló* etc.—, nin tampouco ás formas de participio —*abarrotado*, *contaminado*, *acendido* etc.—.

En segundo lugar, os verbos que nas listaxes aparecen sempre e exclusivamente en forma pronominal lematízanse co clítico, con independencia do seu valor (reflexivo, recíproco, valor resultativo ou un simple formante léxico): *abafarse*, *anicarse*, *lavarse*, *levantarse*, *mancarse*... Se conviven as formas

con clítico e sen el, empregamos as parénteses para unificar —*agochar(se)*, *alimentar(se)*, *distraer(se)*...—, agás no caso de que unhas e outras impliquen diferentes acepcións do verbo non vinculables regularmente: así, por exemplo, no CI 15 ‘Os xogos e as distraccións’ mantemos separados *tocar* (nótese a presenza no mesmo CI de rexistros como *tocar o baixo*, *tocar a guitarra*, *tocar música*) e *tocarse* (que convive con varios verbos que fan referencia a actividades de carácter sexual) ou no CI 02 *vestir* e *vestirse*, posto que a primeira forma pode significar, ademais de ‘cubrir con roupa’, acepción á que lle correspondería o uso pronominal reflexivo, ‘levar posto’ ou ‘proporcionar roupa’.

No caso de rexistrarmos nas listaxes verbos sen clítico que, segundo os dicionarios, posúen sempre carácter pronominal, recuperamos a partícula *se* no lema, sen parénteses. Así, *accidentar* figura lematizado como *accidentarse*. No sentido contrario, cando no rexistro aparece un uso pronominal que non é propio da sintaxe normativa, lematizamos sen clítico. Deste xeito, *secarse o pelo* lematízase como *secar o pelo*.

2.3.2.5 Os acurtamentos, as abreviaturas gráficas, as siglas e acrónimos e os símbolos

É ben sabido que son relativamente frecuentes os acurtamentos de palabras en galego, dando lugar a formas que adoito levan unha marca de rexistro coloquial. Á hora de levar a cabo o proceso de lematización, decidimos agrupar as formas plenas e as formas acurtadas botando man das parénteses. Deste xeito, atoparemos lemas como *profe(sor-a)*, *contra(ventá)*, *tele(visión)*, *cole(xio)*, *(auto)busero*... Por outra banda, lematizamos coa forma plena, sen parénteses, cando só hai rexistros deste tipo e lematizamos co acurtamento cando só hai rexistros acurtados (*metro*). Isto explica que, por exemplo, no CI 03 teñamos como lema *contra(ventá)*, pero no CI 08 *contraventá*, posto que neste último centro non se rexistra a forma abreviada. Estes criterios de lematización aplícanse tamén cando as formas acurtadas figuran no interior de sintagmas: así temos, por exemplo, *mesa do profe(sor)*.

Empregamos tamén un criterio unificador cando nos rexistros atopamos formas acurtadas e frecuentemente sufixadas do tipo de *pantacas*, *furgona* ou *bocata*, que se lematizan baixo as correspondentes formas plenas e máis frecuentes (respectivamente, *pantalóns*, *furgoneta* e *bocadillo*).

Contrariamente, mantemos separados algúns acurtamentos cando entendemos que poden designar realidades distintas ás denotadas polas formas plenas. É o caso de *bus* / *autobús* e *moto* / *motocicleta*, que serán discutidos en § 2.3.4.12.

Un caso diferente ao comentado constitúeno as abreviaturas gráficas, moitas delas empregadas polos informantes para ganar tempo á hora de encher as enquisas. Así, por exemplo, temos rexistros como *c. comercial, sist. nervioso, transp. urbano...* En todos estes casos, desenvolvemos a abreviatura no lema (*centro comercial, sistema nervioso, transporte urbano*).

Por outra banda, as siglas e acrónimos lematízanse como tales, con letras maiúsculas e sen puntos de separación entre os compoñentes, independentemente de que aparezan só ou no interior de sintagmas: *AVE, PC, LED, bombillas LED...* Constitúen excepción aquelas palabras que naceron como siglas ou acrónimos, pero que os dicionarios xa recollen con grafía en minúsculas (e, no seu caso, con acento gráfico) e que nós tamén lematizamos desta maneira (*ovni, radar, sida, láser...*). Así mesmo, poden constituír excepción algunhas marcas comerciais, comentadas no apartado seguinte.

Contrariamente, os símbolos lematízanse sempre baixo a forma plena que representan, unificados con esta se figura nos rexistros: así, por exemplo, CO_2 lematízase baixo *dióxido de carbono* ou H_2O , que encontramos, por exemplo, no CI 05 ‘Os alimentos e as bebidas’, lematízase baixo *auga*.

2.3.2.6 Os nomes de marcas, empresas e produtos comerciais

No que atange ás marcas, empresas e produtos comerciais, tentamos empregar no lema o nome rexistrado de maneira exacta e completa, aínda que nas respostas figuren denominacións abreviadas ou populares. Así, por exemplo, lematizamos *Ballantine's** (recollido nas listaxes como *Ballantains, Ballantines...*), *Cocodrilo Sacamuelas** (fronte ao rexistro *Cocodrilo quita muelas*), *Budokai Tenkaichi** (presente nas respostas dos alumnos como *Bodecci ta-caichi*), *Rotavator** (*retovato* nas listaxes), *Springfield** (representado por un informante como *SPF*), *MINI** (*mini* nas respostas)... No caso de que o nome comercial sexa o resultado dunha sigla ou acrónimo, respectamos sempre a grafía empregada pola empresa ou compañía. Así, por exemplo, escribimos *Seat**, e non *SEAT**, porque, malia tratarse dunha sigla, o nome comercial só ten maiúscula na letra inicial. Como se pode observar, os nomes de marcas e produtos comerciais son facilmente identificables mediante a presenza dun asterisco no lema. Este símbolo tamén se emprega cando o nome da marca ou produto comercial aparece no interior dun sintagma: por exemplo, *botella de Coca-Cola**.

Marcamos con asterisco realidades de moi diferente tipo: son especialmente numerosas as denominacións de produtos comerciais, sobre todo alimentos, bebidas, xogos e xoguetes (*100 Pippers**, *Nesquik**, *Monopoly**, *Nenuco**), os nomes de consolas e de videoxogos (*Nintendo**, *Mario Bros**,

*Medal of Honor**), as empresas de confección e marcas de roupa (*Pronovias**, *Zara**, *Rei-Zentolo**), así como as compañías de automoción e marcas de vehículos (*Maclaren**, *Seat**, *MINI**). Pero tamén podemos atopar empresas de transporte (*Monbus**, *ATSA**, *RENFE**), programas de televisión (*Cifras e letras**, *Pasapalabra**)¹⁹, produtos de limpeza (*Mistol**), marcas de electrodomésticos (*Thermomix**, *Minipimer**) e un longo etcétera de produtos, empresas e marcas rexistradas.

Cando se produciu un proceso de lexicalización, de tal xeito que o nome do produto figura como entrada nos dicionarios consultados, non incluímos o asterisco e empregamos minúscula inicial. É o caso de *ioiío*, *tabasco*, *tergal*, *termo*... Somos plenamente conscientes de que os dicionarios sempre van un pouco por detrás da sociedade, de tal xeito que existen nomes plenamente lexicalizados neste momento e que, no entanto, non figuran nos dicionarios (por tal motivo, nas nosas listaxes irán seguidos de asterisco: por exemplo, *Gusanitos** ou *Hula Hoop**). A pesar deste feito, botamos man do criterio lexicográfico por razóns de carácter práctico. Resulta imposible establecer un límite discreto entre o que está lexicalizado e o que non o está, posto que é evidente que a lexicalización é un proceso gradual.

2.3.2.7 Os sintagmas

En relación coas respostas dos alumnos que constitúen sintagmas, lexicalizados ou non, o criterio xeral é lematizar tamén mediante un sintagma. Deste xeito, atoparemos no LDG lemas como *gama cromática* ou *ganar diñeiro*.

Agora ben, con certa frecuencia conviven na lingua, e nas listaxes dos distintos informantes, formas complexas, de carácter sintagmático, con outras que posúen a mesma designación e que resultan da simplificación das anteriores, constituídas normalmente só polo elemento final. Así, por exemplo, atopamos nun mesmo CI parellas como *pantalón pitillo* e *pitillo*, *pavillón polideportivo* e *polideportivo*, *pluma estilográfica* e *estilográfica*, *forno microondas* e *microondas*, *cuarto de baño* e *baño*... En casos como estes, agrupamos os rexistros mediante o emprego de parénteses no lema: (*pantalón*) *pitillo*, (*pavillón*) *poliderportivo*, (*pluma*) *estilográfica*, (*forno*) *microondas*, (*cuarto de*) *baño*... Se a única forma que figura nas listaxes é a simplificada, esta é a que se adopta como lema: así, por exemplo, os catorce casos de *descapotable* que figuran no CI 12 lematízanse baixo *descapotable*. Se só se rexistra a forma complexa non incorporamos parénteses: o lema é *tren de cercanías* no CI 10.

¹⁹ En realidade, nos dous casos o informante pode estar facendo referencia aos xogos de mesa vinculados con cadanseu programa.

Por outra banda, tamén levamos a cabo a unificación de variantes que, posuíndo a mesma designación e presentando diferentes compoñentes, comparten no entanto as mesmas raíces morfolóxicas. Así, unificamos sintagmas que difiren no tipo de complemento (un adxectivo fronte a un substantivo introducido por preposición): *culler de sopa* e *culler sopeira*, *auga salgada* e *auga con sal*, *cadea de música* e *cadea musical* etc. Da mesma maneira, unificamos formas que alternan no modificador entre substantivos e infinitivos (*sala de prancha* e *sala de pranchar*) así como, en certas ocasións, compostos morfolóxicos e sintagmas (*cortapizza* e *cortador de pizza*, *cortacésped* e *cortador de céspede...*)²⁰. En todos estes casos, escollemos como lema a forma máis frecuente nas listaxes, agás no caso de que a resposta con menor número de rexistros estea, no entanto, claramente máis estandarizada.

Contrariamente, mantemos separadas formas como *recipiente para o aceite* e *aceiteira*, *pousavasos* e *apoiavasos* ou *bañador* e *traxe de baño*, porque nos tres casos as formas só comparten unha das raíces (respectivamente, *acet-*, *vaso* e *bañ-*). Como se observa, este criterio aplícase mesmo cando as expresións non teñen referendo lexicográfico e resultan, moi probablemente, de creacións *ad hoc* dos informantes (*apoiavasos*). Tampouco unificamos as combinacións de verbo soporte e substantivo cos verbos da mesma raíz ca o substantivo: *botar esterco* e *estercar*, *facer debuxos* e *debu-xar* presentarán lemas independentes.

Si unificamos os rexistros constituídos por un adxectivo cos rexistros constituídos por un sintagma que inclúe o dito adxectivo se o ámbito de aplicación deste último coincide coa clase designada polo núcleo do sintagma. Deste xeito, *domésticos* e *animais domésticos* lematízanse conxuntamente baixo (*animal*) *doméstico* ou *prehistóricos* e *animais prehistóricos* baixo (*animal*) *prehistórico*, dado que resulta evidente que no CI no que nos situamos, ‘Os animais’, os adxectivos *doméstico* e *prehistórico* se aplican a estes (non habendo no CI ningún outro substantivo distinto de *animal* que apareza modificado por tales adxectivos). Outros exemplos serían (*cor*) *cálida* no CI 17 ‘As cores’ ou (*mar*) *calmo* no CI 18 ‘O mar’. Se só hai exemplos do adxectivo, lematizamos con este último: *domesticado* ten como lema *domesticado*.

Por outra banda, empregamos parénteses para encerrar outros elementos opcionais, pero frecuentes, sen os cales é posible manter a mesma interpre-

²⁰ Nalgúns casos, moi poucos, non unificamos compostos morfolóxicos e sintagmas. Ocorre así con *abrebotellas* e *abridor de botellas* e *pelapatacas* e *pelador de patacas*. Nótese que tanto *abridor* como *pelador* son voces recollidas nos dicionarios co mesmo significado dos compostos morfolóxicos e consideramos que eran merecentes de posuír entrada propia, en pé de igualdade con estes últimos.

tación do sintagma: *malas en arrancar (as) (malas) herbas, quente en bolsa de auga (quente), de comer en darlles (de comer) ás galiñas...* Por suposto, colocamos parénteses sempre e cando atopemos nas listaxes exemplos da forma máis simple (verbigracia, exemplos de *darlles ás galiñas*). En caso contrario, a expresión aparecerá sen parénteses e en toda a súa extensión.

Como pode observarse en *arrancar (as) (malas) herbas*, a aplicación dos diferentes criterios pode conducir ao emprego de varias parénteses nun único lema. Outros exemplos do mesmo tipo atopámoslos en *carril (de) bici(cle-tas)* ou en *mando (a distancia) da tele(visión)*.

Finalmente, cómpre sinalar que algúns agrupamentos se realizaron só unha vez consultado o contexto no que figuraban os rexistros, pois o coñecemento das respostas anteriores e posteriores permitiu determinar en que sentido se debían interpretar as palabras. Así, por exemplo, no CI 02 ‘A roupa’ unificamos o único exemplo de *gastar* e o único exemplo de *gastar cartos* baixo *gastar (cartos)*, malia que *gastar* tamén se podería interpretar neste CI no sentido de ‘estragar polo uso’. No entanto, o contexto no que figura o verbo —*modas, comprar, gastar, marcas*— parece dar a entender que *gastar* se debe interpretar como ‘empregar [diñeiro]’. No entanto, debe sinalarse que estas consultas só se realizaron puntualmente e para voces pouco frecuentes nas listaxes.

Faremos agora algunhas outras precisións relativas ao proceso de lematización dos sintagmas, a maiores das que xa foron mencionadas ata o de agora.

2.3.2.7.1 Determinantes e cuantificadores

Algúns dos rexistros das listaxes son frases nominais introducidas por artigo. Tales frases nominais son lematizadas eliminando o determinante. Así, por exemplo, *o parchís* ou *o campo de fútbol* teñen como lema, respectivamente, *parchís* e *campo de fútbol*. Tamén se exclúen do lema os posesivos nos poucos casos en que aparecen (*a miña habitación* ou *habitación dos meus pais* lematízanse, respectivamente, baixo *habitación* e *habitación dos pais*). No que atangue ao tratamento dos cuantificadores, cando estes indican cantidade elevada ou maior, o cuantificador desaparece do lema. Deste xeito, lematizamos *moito traballo* como *traballo* ou *máis osíxeno* como *osíxeno*, unificando as formas con e sen cuantificador. O mesmo vale cando o cuantificador ten uso adverbial: *ilumina moito* lematízase baixo *iluminar*. No entanto, mantemos no lema o cuantificador cando este designa unha cantidade pequena ou menor: así, atoparemos lemas como *poucos coches* ou *menos*

contaminación. Neste caso, a unificación cos rexistros *coches* ou *contaminación* (no caso de existiren), non tería moito sentido ao estar dando conta de ideas absolutamente contrarias²¹.

Outros rexistros das listaxes son frases verbais que conteñen algún complemento no seu interior. Cando estes complementos aparecen introducidos por artigo, tentamos ser o máis fieis posible ás respostas dos informantes. Deste xeito, se todos os exemplos figuran con artigo, mantemos este no lema (é o caso de *muxir as vacas*), pero se todos os exemplos carecen de determinante, tamén carecerá del o lema (*vender vacas*). No caso de atoparmos tanto exemplos con artigo como sen el, unificamos os rexistros, pero, sempre que sexa posible, botamos man do uso das parénteses para dar conta da variación: así, *apañar as patacas* e *apañar patacas* lematízanse baixo *apañar (as) patacas* ou *facer as maletas* e *facer maletas* baixo *facер (as) maletas*. No entanto, nalgúns exemplos resulta imposible unificar mediante parénteses. Nestes casos, lematizamos baixo a forma máis frecuente. Así, por exemplo, *recoller froita* (tres exemplos) e *recoller as froitas* (un exemplo) lematízanse baixo *recoller froita*.

Tamén atopamos nas listaxes con certa frecuencia frases substantivas con complementos nominais, dándose o caso de que moitos destes complementos poden aparecer en galego con ou sen determinante. De feito, nos nosos rexistros figuran parellas como *coitelo da carne* e *coitelo de carne*, *culler do café* e *culler de café*, *cesta do pan* e *cesta de pan...* En realidade, o significado das expresións que conforman estas parellas non é estritamente o mesmo ou pode non selo. Así, por exemplo, *cunca de viño* e *cunca do viño* son susceptibles de designar realidades algo diferentes. *Cunca de viño* pode designar unha *cunca* con viño ou ben un *tipo* específico de *cunca* (para viño). Pola súa banda, con *cunca do viño* designamos a *cunca* que empregamos habitualmente na casa para tomar o viño (que pode ou non ter a forma específica das *cuncas* de viño). Dado que calquera dos valores indicados pode ter sentido nas listaxes, pero parecía demasiado drástico establecer unha lematización separada dos rexistros con artigo e dos rexistros sen el, optamos por escoller como lema a forma máis frecuente (nótese a imposibilidade de empregar parénteses pola presenza de contraccións). Esta é a explicación de que teñamos como lemas,

²¹ Os cuantificadores aparecen esencialmente nos CI 10 'A cidade' e 11 'O campo'. A presenza dos comparativos (*máis* e *menos*) explícase polo establecemento por parte dos informantes dunha comparación entre o campo e a cidade: *menos coches*, *menos contaminación*, *máis osíxeno* (no CI 11), *máis liberdade*, *máis ruído...* (no CI 10). En todos os casos, os rexistros con cuantificador comparativo de superioridade conviven con rexistros sen el (*osíxeno*, *ruído...*), e consideramos que a unificación reflectiría máis fielmente o nivel de dispoñibilidade léxica dos substantivos en cuestión. O mesmo sucede co cuantificador *moito*. Só mantemos o adverbio *máis* no lema *casas máis grandes* (CI 11), pois non atopamos exemplos de *casas grandes*.

por exemplo, *coitelo do pan* pero *coitelo de carne* ou *xerra do leite* pero *xerra de zume*. Por outra banda, e en caso de termos o mesmo número de rexistros con artigo e sen el, optamos por lematizar coa forma sen artigo. Así temos *pocillo de té* ou *bota de viño*. Aínda sendo conscientes de que o galego é unha lingua cunha forte tendencia ao emprego do artigo (en contextos nos que outras linguas próximas non o empregarían), tomamos esta decisión polo feito de que a forma sen artigo parece posuír un contido máis xenérico e abstracto que a converte nunha mellor candidata para funcionar como lema²². De xeito parello, cando os rexistros con que contamos son rexistros abreviados (sen preposición), do tipo de *garrafa auga*, repoñemos no lema a preposición sen artigo (*garrafa de auga*) (vid. § 2.3.2.7.2 a este respecto). E finalmente, cando os informantes ofrecen rexistros con outras preposicións seguidas de artigo (nomeadamente *para*), existindo denominacións estándares, coa preposición *de* e sen artigo (vid. *infra* a este respecto), substituímos a primeira preposición pola segunda, prescindindo do determinante (así, por exemplo, lematizamos *servizos para o café* como *servizo de café*).

Por outra banda, cómpre sinalar que, independentemente das formas usadas nos rexistros dos informantes, nós nunca empregamos a segunda forma do artigo nos lemas e utilizamos sempre a contracción *ao*. Deste xeito, temos lemas como *agrandar o marco* ou *botarlle cal ao prado*.

2.3.2.7.2 Preposicións

En xeral mantemos as preposicións ofrecidas polos informantes que introducen os complementos do nome (*moble para as chaves*), aínda que en certos casos, se existe unha forma normalizada, lematizamos con esta (tal e como acabamos de indicar, lematizamos *servicios para o café* como *servizo de café*). Neste mesmo sentido, unificamos os sintagmas que parecen ter as mesmas posibilidades designativas, presentan os mesmos compoñentes léxicos e difiren no emprego de preposición. Así, por exemplo, *bote con sal*, *bote de sal* e *bote do sal* lematízanse baixo *bote de sal*. En todos estes casos, escollemos para o lema a forma máis frecuente nas listaxes, agás que consideremos que o rexistro menos repetido ten un carácter claramente máis estandarizado. En caso de termos o mesmo número de rexistros en todos os exemplos, escollemos a forma que consideramos máis estandarizada ou máis frecuente nos discursos en galego.

²² Nótese que a presenza do artigo indefinido determinando o conxunto bloquea ou fai estraña a presenza do artigo definido no complemento. Dicimos “dáme unha copa de champaña” pero máis dificilmente “dáme unha copa do champaña”. E xustamente ese trazo de indefinición cadra perfectamente coas características que debe posuír un lema.

Como xa indicamos, en certos casos os informantes presentan as súas respostas de maneira abreviada. Unha das formas máis habituais de abreviar para ganar tempo á hora de facer as listaxes é eliminar as preposicións dos rexistros. Así, atopamos con certa frecuencia respostas do tipo de *copa fresco*, *coitelo carne* ou *botella viño*. Nestes casos, recuperamos no lema a preposición que consideramos máis axeitada para o contexto en cuestión e non empregamos parénteses. Nos exemplos indicados, trátase sempre da preposición *de*: *copa de fresco*, *coitelo de carne*, *botella de viño*. Si empregamos parénteses cando na lingua habitual conviven a forma con e sen preposición e atopamos exemplos dos dous tipos nas listaxes: *licor (de) café*.

Un problema doutro tipo é o que ofrecen rexistros que aparecen directamente introducidos por preposición. En función da preposición de que se trate, esta é mantida ou eliminada nos lemas. Así, as formas introducidas pola preposición *con* son lematizadas suprimindo esta e unificando, no seu caso, cos substantivos correspondentes: *con capucha*, *con rodas* ou *con plumas* son lematizados, respectivamente, baixo *capucha*, *roda* e *pluma*. No entanto, esta supresión non se pode levar a cabo cando a preposición que figura é a de significado oposto: nas listaxes atoparemos lemas como *sen tráfico* ou *sen plumas* (así como lemas introducidos por *non*: *non atascos* ou *non coches*). Tamén conservamos a preposición noutros casos: *de madeira* ou *de cores*.

Por suposto, mantemos a preposición cando está fixada constituíndo unha locución: *a patas* ou *a pé* no CI 12 ‘Os medios de transporte’. Pero no caso de que a carón das formas fixadas con preposición existan nas listaxes outras sen ela, unificamos con parénteses: deste xeito, *cuadros* e *a cuadros* lematízanse no CI 2 ‘A roupa’ baixo (*a*) *cuadros*. Finalmente, se os informantes ofrecen un rexistro sen preposición, pero a voz só se emprega nunha construción preposicional, repoñemos a preposición no lema e non usamos parénteses. Así, por exemplo, no CI 15 ‘Os xogos e as distraccións’ lematizamos *revolos* como *a rebolos*.

2.3.2.7.3 Orde

En xeral, a orde dos constituíntes dos sintagmas recollidos nas listaxes responde á ordenación habitual do galego, co núcleo na primeira posición e o modificador ou o complemento a seguir. No entanto, nalgún caso os rexistros, por diferentes motivos, amosan outra organización. Na procura dunha lematización o máis estandarizada e regular posible, en boa parte destes exemplos recuperamos no lema a orde núcleo – complemento, aínda sendo

conscientes de que o tipo de modificación é diferente en función do lugar que ocupa o adxectivo. Deste xeito, lematizamos *pequenos animais* como *animal pequeno*, *grandes cubertos* como *cuberto grande* ou *diferente estilo* como *estilo diferente*. Por suposto, a inversión da orde non se leva a cabo cando a expresión presenta cando menos certo grao de fixación ou cando dá lugar a un novo sentido claramente diferenciado: deste xeito, atoparemos lemas como *gran superficie* ou *gran empresa*. Menos frecuentemente, a inversión faise no sentido contrario: lematizamos *dunha cor soa* baixo *dunha soa cor*.

Un caso doutro tipo é o constituído por rexistros nos que o modificador vai seguido dun elemento clarificador entre parénteses. Así, atopamos respostas como *urbano (autobús)*, ou *pitillo (pantalón)*. Consideramos estas variantes como equivalentes aos correspondentes compostos sintagmáticos (*autobús urbano* e *pantalón pitillo*) e lematizamos coa orde habitual.

2.3.2.7.4 Supresión e mantemento de constituíntes

En xeral, mantemos nos lemas os modificadores e complementos ofrecidos polos informantes. No entanto, suprimímolos en calquera das circunstancias que comentaremos a seguir. En primeiro lugar, cando a información que achegan é redundante, no sentido de deducible a partir do coñecemento do significado dos restantes membros do sintagma, no marco do CI en que nos situemos. Así, por exemplo, *limpar as cortes dos animais* lematízase baixo *limpar as cortes* (con supresión do complemento *dos animais*), polo feito de que no sentido de *corte* xa vai implícito o contido ‘animal’. O dito complemento non restrinxe en absoluto a extensión de *cortes*. De xeito parello, lematizamos *biberón do bebé* baixo *biberón*, *xerra de bebida* baixo *xerra*, *exprecedor do zume* baixo *exprecedor* etc. En ocasións, o que encontramos nas listaxes son aclaracións dos informantes entre parénteses, que tamén suprimimos se non son necesarias para a interpretación do rexistro: *caza (avión combate)* no CI 12 ‘Os medios de transporte’, *camellos (non animais)* no CI 10 ‘A cidade’...

No entanto, mantemos no lema o modificador se este, malia ser semanticamente prescindible no marco do CI en cuestión, constitúe co núcleo unha unidade designativa con relativa fixación e difusión: *fauna mariña*, *patio de recreo*, *balanza de cociña*... No caso de que no mesmo CI haxa rexistros sen complemento coa mesma designación, e de acordo co indicado en § 2.3.2.7, unificámolos todos mediante o emprego de parénteses. Deste xeito, no CI 18 ‘O mar’ *brisa mariña* e *brisa* lematízanse baixo *brisa (mariña)*. Téñase en conta ao respecto do indicado que á hora de levar a cabo a lematización sem-

pre tivemos en conta o CI concreto co que estabamos traballando. Deste xeito, no CI 12 ‘Os medios de transporte’ unificamos *cinto* e *cinto de seguridade* baixo *cinto (de seguridade)* porque, dada a temática do CI en cuestión, consideramos que ambos os rexistros fan referencia á mesma realidade. De igual xeito, no CI 07 ‘A cociña e os seus utensilios’ unificamos *armario* e *armario de cociña* baixo *armario (de cociña)*. Pola contra, noutros CI, como no CI 04 ‘Os mobles da casa’, esta última unificación non se pode levar a cabo posto que a coincidencia designativa non ten por que darse. Así pois, a aplicación deste criterio pode dar lugar, unha vez máis, á presenza de diferentes lemas en distintos CI. Vexamos un último exemplo: no CI 5 ‘Os alimentos e as bebidas’ ou no CI 14 ‘Os animais’, non podemos unificar *boi* e *boi de mar*, mentres que no CI 18 ‘O mar’ o agrupamento *boi (de mar)* si se pode levar a cabo, dado que a única interpretación posible para *boi* neste CI é a do crustáceo.

Tamén mantemos os modificadores, pero desta volta sen parénteses, cando, no marco dun CI concreto, permiten desambiguar as palabras e saber cal é a realidade a que está aludindo o informante. Así, por exemplo, no CI 10 ‘A cidade’ atopamos tres exemplos lematizados baixo *banco de sentar(se)* (*banco de sentarse, banco de sentar, banco para sentarse*). Dado que neste CI tamén ten pleno sentido a presenza de *banco* como ‘entidade bancaria’ (de feito tamén hai dous exemplos de *banco de cartos*), non consideramos conveniente unificar as variantes *banco* e *bancos* coas anteriormente mencionadas, posto que descoñecemos cal é o sentido que o informante lles está atribuíndo a estas últimas palabras. Así pois, mantemos separados eses sintagmas nos que o informante deixou ben ás claras cal é o valor que lle quería outorgar ao núcleo, feito que, neste caso concreto, dá lugar a tres lemas distintos: *banco, banco de sentar(se)* e *banco de cartos*. Polas mesmas razóns, no CI 18 ‘O mar’, mantemos separadas as entradas *banco, banco de peixes, banco de pesca* e *banco de area*. Cando o elemento desambiguador aparece entre parénteses, como aclaración do informante, mantémolo sempre que sexa posible extraer a partir del un lema estándar: por exemplo, *chave (de acender la luz)* lematízase baixo *chave da luz* e *vasos (medir)* (no CI 08 ‘A cociña e os seus utensilios’) baixo *vaso de medir*.

Por outra banda, suprimimos do lema aqueles complementos ou modificadores que achegan restricións pouco relevantes, por seren ou ben demasiado xerais, ou ben demasiado específicas e, en definitiva, perfectamente prescindibles. Compárese, por exemplo, o valor de *grande* en *culler grande* (empregado para designar un tipo específico de *culler*, oposta á pequena que se emprega para a sobremesa) co valor que posúe en *grandes xeadas*, onde se utiliza exclusivamente cun valor intensificador. Neste segundo caso, lema-

tizamos *grandes xeadas* baixo *xeadas*, mentres que no caso anterior debemos manter necesariamente o modificador no lema. Outros exemplos de modificadores que desaparecen no lema son os que figuran en *deportes en xeral* (lematizado baixo *deporte*), en *prato normal* (lematizado baixo *prato*), en *bufanda amarela* (lematizado baixo *bufanda*), en *buque de Pescanova* (lematizado baixo *buque*), en *coches rápidos* e *coches lentos* (lematizados baixo *coche*), en *condutor novel* (lematizado baixo *condutor*)...

En liña cos criterios que estamos comentando, propoñemos outras lematizacións cunha casuística difícil de sistematizar: lematizamos *aumento do efecto invernadoiro* baixo *efecto invernadoiro*, *cuarto comedor* baixo *comedor*, *traballar e traballar* baixo *traballar* etc.

2.3.2.7.5 Recuperación de constituíntes

Xa vimos previamente como os informantes abrevian as respostas con certa frecuencia, suprimindo, por exemplo, as preposicións, que se recuperan no lema. Outra forma de ganar tempo é eliminar os núcleos de sintagmas cando tales núcleos se comparten cos rexistros anteriores ou foron mencionados previamente. Así, por exemplo, un informante ofrece esta listaxe no CI 09 ‘A calefacción e a iluminación’: *estufa, apagado, encendido, de auga, de leña, de butano*. Parece claro que o estudante fai referencia aos diferentes tipos de estufas, pero escusa a presenza do núcleo porque xa foi mencionado con anterioridade na listaxe. Nestes casos, recuperamos no lema o elemento elidido, sen situalo entre parénteses (daquela, nun caso como o indicado teriamos os lemas *estufa de auga, estufa de leña* e *estufa de butano*). Encontramos outro exemplo na listaxe *rosado, ribeiro, rioja, barrantes, catalán*. No caso do último rexistro, recuperamos *viño* para que o usuario teña claro cal é a referencia da palabra ofrecida polo informante.

Menos frecuentemente, falta algún elemento que non está presente no contexto das respostas próximas, pero que parece necesario para completar o rexistro. Así ocorre, por exemplo, coa resposta *cabaliño*, que figura no CI 18 ‘O mar’ e que lematizamos baixo *cabaliño de mar*. En casos coma este, tampouco botamos man das parénteses.

Contrariamente, e como xa comentamos en § 2.3.2.7, non recuperamos elemento ningún cando os únicos exemplos que figuran nas listaxes son formas resultantes dun proceso de elisión dalgún constituínte, pero que se empregan frecuentemente nesa forma abreviada (*vid. supra* o caso de *descaptable*).

No caso de que nos rexistros figuren exemplos de Complementos Indirectos, incluímos o clítico de dativo (*lle* ou *lles*) no lema, con independencia

de que este estea presente ou non nos rexistros dos informantes e de que posúa ou non nestes variación de número. Así, *botarlle de comer os animais* lematízase como *botarlles de comer aos animais* e *dar de comer os animais* lematízase como *darlles de comer aos animais*.

2.3.2.8 A separación / fusión gráfica

Á hora de representar graficamente algunhas unidades, sobre todo compostos morfolóxicos e lexías complexas, os informantes amosan diferentes posibilidades relativas ao grao de fusión dos compoñentes. Algúns dos compostos máis problemáticos desde o punto de vista gráfico son aqueles que posúen a estrutura “verbo + substantivo” (do tipo de *afialapis*, *abrebotellas*...). Por poñer un exemplo, temos rexistros de *afia lapis*, de *afia-lapis* e de *afialapis*. Pero non son estas as únicas expresións que amosan variabilidade: *mesa camilla*, *mesa-camilla* e *mesacamilla*; *alta mar*, *alta-mar* e *altamar*, *verde azulado*, *verde-azulado* e *verdeazulado*, ou mesmo *auto bus*, *auto-bus* e *autobús* ou *bambán* e *bam-bam*. Por suposto, en todos estes casos lematizamos coa forma que corresponde no código normativo, sexa do galego (*afialapis*, *alta mar*, *verde azulado*, *autobús*, *bambán*) ou doutras linguas (*mesa camilla*).

No caso de que a expresión que ofrece o informante non se recolla nos dicionarios, adaptámola ao sistema gráfico característico da súa estrutura formal. Así, por exemplo, o único exemplo rexistrado de *apoiavaos*, lematízase como *apoiavaos*, coa fusión gráfica característica dos compostos de esquema “verbo + substantivo”.

2.3.2.9 O emprego de maiúsculas

En relación co emprego de maiúsculas, seguimos as normas xerais do galego, de tal xeito que atoparemos maiúscula inicial en topónimos (*Galicia*), nomes de empresas, marcas ou produtos (*Zara**, *Absolut**, *Trivial**), institucións (*Garda Civil*) etc. Tamén iniciamos con maiúscula os substantivos que designan estudos que conducen á obtención dun título, presentes esencialmente no CI 16 (*Maxisterio*, *Odontoloxía*...), mentres que empregamos minúscula nos substantivos que denotan materias. Deste modo, poderemos atopar un mesmo lema escrito con maiúscula ou con minúscula inicial: no CI 08, dedicado á escola, atoparemos *bioloxía* porque esta palabra designa unha materia de estudo, mentres que no CI 16, no que o estímulo son os oficios e profesións, o lema será *Bioloxía*, posto que a voz designa a correspondente carreira. Tamén empregamos maiúsculas nalgunhas siglas, como *LED* ou *PVC* (*vid.* a epígrafe correspondente, § 2.3.2.5).

Mantemos o uso de maiúsculas aínda cando a palabra xa non designe un individuo ou un lugar (*Manhattan* como denominación dun cóctel), sempre e cando os dicionarios non recollan a voz con minúscula inicial (por exemplo, *tabasco*). O mesmo sucede cando o nome propio se integra nunha expresión complexa (*coche de san Fernando, azul Prusia...*).

2.3.2.10 As variantes ortográficas, os lapsus cálami e as creacións dos informantes

De acordo cos criterios comúns empregados na edición dos léxicos dispoñibles, lematizamos baixo a forma estándar galega variantes con diverxencias exclusivamente gráficas con respecto a aquela²³. Así, por exemplo, *adesivo* lematízase baixo *adhesivo*, *azibeché* baixo *acibeché* ou *automóvil* baixo *automóbil*. Este criterio é aplicable incluso nos casos en que as formas con erros ortográficos son maioritarias ou mesmo exclusivas. Así, por exemplo, *avafarse* lematízase baixo *abafarse*, por moito que a forma grafada con <v> sexa a única realización deste verbo no LDG.

Por outra banda, resúltanos irrelevante á hora de lematizar o feito de que a diverxencia poida ter ou non unha explicación na influencia do cognado castelán: deste modo, os noventa e catro exemplos de *abogado* / *abogada* unifícanse, baixo o lema *avogado-a*, cos douscentos setenta e oito restantes nos que figura o grafema <v>, seguindo o mesmo criterio polo que os seis exemplos de *bestido* se sitúan conxuntamente, baixo *vestido*, coas duascentas corenta e unha formas restantes escritas correctamente.

No caso de que no código normativo se admitan en pé de igualdade dúas ortografías, lematizamos coa forma máis frecuente nas listaxes. Deste xeito, no CI 05 ‘Os alimentos e as bebidas’, o lema é *alcohol*, posto que é máis frecuente a grafía con <h> nas respostas dos informantes, mentres que no CI 08 ‘A escola: mobles e materiais’, o lema, contrariamente, é *alcol*, posto que o único rexistro que figura neste CI aparece escrito sen o dito grafema.

De por parte, as variantes ortográficas de palabras do castelán sen cognado galego lematízanse baixo a forma normativa do castelán. Así, no CI 13 *avonar* figura baixo *abonar*. Este último criterio sería de aplicación incluso se as formas con <v> fosen maioritarias ou mesmo exclusivas. No caso de que no conxunto das respostas prevalecese a forma adaptada ao galego, esta tamén se axeitaría á ortografía do castelán (deste xeito, se só houberse como resposta *avoar*, o lema sería *aboar*). Sobre os criterios de lematización das palabras do castelán sen cognado galego, vid. § 2.3.2.12 e § 2.3.3.

²³ En § 2.3.3 indicamos as fontes que empregamos para determinar a forma normativa galega.

Por outra banda, en certas ocasións atopamos algunhas respostas cuxa forma parece obedecer a un lapsus cáلامي. Os criterios de lematización son os mesmos sinalados para os erros ortográficos: *patnete* lematízase baixo *patinete* ou *figorífico* baixo *frigorífico*.

Pero ademais, de cando en vez, encontramos nas respostas formas que non figuran nos dicionarios, que nós nunca lemos nin escoitamos e que non se poden explicar como simples lapsus cáلامي nin como adaptacións ao galego de voces do castelán. Nalgúns destes casos, tendo en mente o CI no que se insire a palabra así como as respostas que acompañan a tales formas na listaxe, podemos intuír cal é a voz que está na orixe da cando menos aparente creación do informante. Se esta intuición é relativamente clara, mantemos a resposta, lematizándoa coa forma estándar correspondente. Así, por exemplo, lematizamos *incisores* como *incisivo* (o contexto é *dentes, moas, incisores*) ou *labregar* como *labrar* (os tres exemplos de *labregar* foron proporcionados por alumnos do mesmo centro, o CPR Grande Obra de Atocha, da Coruña). Nalgún caso, a forma que ofrece o informante existe en galego, pero o seu contido non é acaído ao CI en cuestión nin encaixa coas palabras que a acompañan nas listaxes. Así, por exemplo, varios alumnos dun mesmo centro inclúen a voz *borralleiro* no CI o8 ‘A escola: mobles e materiais’, que nós, dado o CI e tendo en conta os contextos en que aparece, intepretamos como *borrador* e lematizamos baixo esta voz.

2.3.2.11 Os estranxeirismos

Analizamos nesta epígrafe o tratamento de formas presentes nas listaxes que non teñen orixe galega nin castelá (estas últimas serán analizadas en § 2.3.2.12 e § 2.3.3). No caso de que a palabra en cuestión figure nos dicionarios, o lema coincide coa variante que recollen estes. Como se sabe, as obras lexicográficas poden incluír tanto expresións perfectamente integradas no galego desde o punto de vista formal (*pixama* ou *chatear*), as cales normalmente non ofrecen problemas ao informante canto á súa representación gráfica, como outras formas con menor grao de integración formal (*whisky, jet, kaiak, iogur*²⁴...), que adoitan supoñer maiores dificultades na escrita.

Agora ben, como se sabe, os dicionarios son bastante cautos á hora de introducir estranxeirismos, de tal xeito que unha boa parte das voces de orixe foránea que figuran nas listaxes está ausente das obras lexicográficas (e malia

²⁴ Estas son as variantes que atopamos nas listaxes: *güisquy, visky, whiskey, whiskey, whisky, whisky, whisky, whyskey, whyski, whysky, wiscky, wishky, wiskey, wiski, wiskie e wisky; jet e yet; caiak, callac, cayac e kayak; iogourt, iogues, iogur, iogures, iogurs, iogurt, iogurth, iogurts, iogur, xogures, xogurt, yoghurt, yogurt, yogur, yogures, yogurt, yogurts e yogurt*.

que o *Digalego*, como recolle expresamente na súa Introducción, está especialmente aberto á súa incorporación²⁵). Nestes casos, cando a forma non se recolle en ningunha das obras de referencia, empregamos no lema a grafía da lingua de orixe ou, se isto non é posible, a grafía máis habitual á hora de representar esa palabra na escrita, independentemente de que tal grafía coincida ou non coa empregada polos informantes. Así, temos *balonkorf*, *aye-aye*, *dropknee*, *moto naked*...

No que atangue ao tratamento de marcas comerciais e nomes de empresas, moitos dos cales son estranxeirismos, véxase § 2.3.2.6.

2.3.2.12 As variantes galegas e as voces castelás

Ao pé das variantes flexivas e gráficas, que xa comentamos, son especialmente frecuentes e relevantes no LDG outros tipos de variantes, que en parte imos tratar de maneira conxunta, aínda que respondan a explicacións de diferente clase. Consideramos que a complexidade da casuística atopada xustifica que lle dediquemos un apartado específico (§ 2.3.3) a esta cuestión, de tal xeito que agora simplemente presentaremos de maneira esquemática a tipoloxía rexistrada e os criterios básicos de lematización.

Táboa 14. Criterios empregados para lematizar as variantes galegas fónicas e morfolóxicas e as voces procedentes do castelán

Tipo de variantes	Exemplos de variantes	Criterio lematizador	Lemas correspondentes
VARIANTES FÓNICAS E MORFOLÓXICAS DE ORIXE GALEGA	<i>abelleiro</i> <i>silar</i> <i>choiva</i> , <i>chuvia</i>	Forma normativa do galego co mesmo significado, designación e raíz. No caso de haber varias formas normativas igualmente recomendables (por exemplo, <i>chovia</i> e <i>chuvia</i>), escóllese como lema a variante máis frecuente, ou única, nas listaxes de cada CI.	<i>abelleiro</i> <i>ensilar</i> <i>choiva</i> ou <i>chuvia</i> (a única ou a máis frecuente no CI)
FORMAS DO CASTELÁN (ADAPTADAS OU NON AO GALEGO) COGNADAS DE VOCES GALEGAS	<i>sombreiro</i> <i>naranzo</i>	Forma normativa do galego cognada e co mesmo significado, designación e raíz.	<i>sombreiro</i> <i>laranxeira</i>
FORMAS DO CASTELÁN (ADAPTADAS OU NON AO GALEGO) SEN COGNADO GALEGO	<i>cerdo</i> <i>acha</i> ¹ <i>abonar</i> , <i>aboar</i> ²	Forma normativa do castelán ou forma adaptada. Escóllese como lema a forma máis frecuente, ou única, nas listaxes de cada CI.	<i>cerdo</i> <i>hacha</i> <i>abonar</i> , <i>aboar</i> (a única ou a máis frecuente no CI)

¹ O contexto deixa ás claras que se trata da ferramenta.
² En ambos os casos co significado de 'fertilizar'.

²⁵ De feito, algunhas formas só as atopamos nesta obra, como é o caso de *yonqui*.

2.3.3 A lematización das variantes galegas fónicas e morfolóxicas e das voces procedentes do castelán

Como indicamos en § 1, as utilidades dun léxico dispoñible poden ser de moi distinto tipo (ensino de linguas, psicolingüística, lexicoloxía, lexicografía...) e, en función da análise que se queira levar a cabo, o interese do investigador pode centrarse nos conceptos ou nas unidades léxicas que vehiculan tales conceptos. Isto é, haberá investigadores, interesados na análise conceptual, para os que sería útil que nós unificásemos, por exemplo, as voces *silla* e *cadeira*, as cales posúen exactamente o mesmo significado e designación ('asento individual con respaldo, provisto ou non de brazos', DRAG²), pero outros, pola contra, centrarían a súa análise en cada unha das unidades léxicas tomadas de forma individualizada, ou mesmo no contraste entre o nivel de dispoñibilidade dunha e outra. En calquera caso, parece claro que unha obra que se autodenomina "léxico dispoñible" non pode levar a cabo unificacións do tipo da indicada, na medida en que *silla* e *cadeira* son, sen ningún xénero de dúbidas, dúas unidades léxicas claramente diferentes, por moito que só a segunda forme parte do código normativo galego ao considerarse a primeira unha interferencia do castelán non desexable na variedade estándar.

Contrariamente, noutros casos, como nos de *harado* / *arado*, *biquini* / *bikini* ou *abogado* / *avogado*, en que nos atopamos con variacións meramente gráficas, sen correlato fónico, parece claro que a unificación se impón como necesaria, seguindo, por outro lado, o mesmo criterio empregado na edición doutros léxicos dispoñibles, tal e como indicamos en § 2.3.2.10.

Pero entre estas dúas situacións extremas, nas que o criterio lematizador parece abondo claro, existe unha complexa casuística de variantes que ofrece bastantes máis problemas e incertidumes, máxime no caso dunha lingua coma o galego. Téñase en conta que este é un idioma altamente interferido polo castelán, cun código normativo de recente creación, e aínda non demasiado ben coñecido por parte dunha boa parte dos galegos (mesmo por aqueles que están inseridos nun sistema de ensino que se imparte parcialmente en galego e no que se estuda a lingua galega, como é o caso dos nosos informantes) e cun claro retroceso no uso, sobre todo por parte da xente máis nova. Todo isto conduce a un considerable número de galegos non só ao deficiente dominio do código galego normativo e da variedade estándar na que se plasma, senón, en xeral, da lingua histórica galega, en calquera das súas variedades. Esta situación dá lugar a unha enorme variación nas respostas dos informantes, que queremos ilustrar cun par de exemplos.

No código normativo do galego, a única forma aceptada para designar a ‘máquina para lavar os pratos, vasos, tarteiras etc., que funciona por medio de electricidade’ (DRAG²) é *lavalouza*. Pois ben, nas listaxes do CI 07 ‘A cociña e os seus utensilios’ atopamos vinte variantes distintas para designar este aparello (unificando os rexistros que presentan diferenzas na maneira de establecer o vínculo entre os dous compoñentes do composto —unión directa, trazo ou espazo separador— e unificando tamén as formas con e sen faltas de ortografía ou lapsus cáلامي, pois se considerásemos o total de formas distintas chegaríamos a corenta e dúas variantes). Velaquí as vinte realizacións ordenadas alfabeticamente (entre parénteses indicamos o número de rexistros): *lavalousas* (1), *lavalouza* (36), *lavalouzas* (75)²⁶, *lavalouza* (1), *lavalouzas* (3), *lavaplatos* (35), *lavapratos* (23), *lavavailletas* (1), *lavavaillexas* (1), *lavavaixelas* (2), *lavavaixela* (9), *lavavaixelas* (183), *lavavaixelas* (1), *lavavajillas* (27), *lavavaxela* (1), *lavavaxelas* (3), *lavavaxilla* (1), *lavavaxillas* (20), *lavavixela* (1) e *lavavixelas* (1). Nótese, por outra banda, que a forma normativa ocupa tan só o terceiro posto por orde de frecuencia, precedida de *lavavaixelas* e *lavalouzas*.

Vexamos agora outro exemplo: para designar a tan coñecida raíz comestible de cor laranxa e forma de cono o galego normativo propón a voz *cenoria*. No CI 05 ‘Os alimentos e as bebidas’ encontramos sete variantes distintas (unificando singulares e plurais, formas con e sen <h>, formas con <c> e <z> ante <e> e formas con e sen acento gráfico, pois sen levar a cabo tal unificación o número total de variantes sería de vinte e unha): *celoura* (1), *cenouira* (1), *cenoria* (62), *cenoura* (11), *zanahoria* (91), *zanoira* (4) e *zanoria* (27). Nótese unha vez máis que a forma máis frecuente non coincide coa normativa galega, senón, neste caso, coa propia do castelán.

Da nosa insistencia sobre o código normativo do galego, non queremos que se extraia a conclusión de que, á hora de facer as enquisas, nós solicitamos dos informantes que se limitasen a incluír nas listaxes palabras deste código ou que nós tivéssemos especial interese en que así se fixera. Como xa sinalamos en § 2.1, nas instrucións deixábaselles claro aos estudantes que o propósito da enquisa era “recoller as palabras que se lles ocorren poñéndose

²⁶ Poderíase interpretar que o <s> final de *lavalouzas* marca o plural de *lavalouza* e, daquela, que *lavalouza* e *lavalouzas* non son dúas variantes distintas, formadas con dous compoñentes distintos, *louza* e *louzas*, senón o singular e o plural dunha mesma forma (o mesmo podería dicirse de *lavalouza* e *lavalouzas*). Pero resulta un pouco estraña esta interpretación, e cando menos non parece poder aplicarse a todos os casos. Nótese, por exemplo, que neste mesmo CI, das catrocentas sete formas lematizadas baixo *neveira* só dúas son plurais e que os catrocentos oitenta e catro exemplos do lema *forno* están en singular. E aínda sería máis cuestionable no caso das formas cuxo segundo elemento é *vaixela(s)* (ou variantes), pois debe terse en conta que a forma estándar do castelán, na que moi probablemente se inspiran as variantes galegas, é precisamente *lavavajillas*, un substantivo formalmente invariable en número, cuxo segundo elemento se incorporou na forma plural.

na situación de falar ou escribir en galego sobre o tema que representa cada centro de interese”. Non hai, pois, nin sequera unha mínima referencia ao código normativo nin a ningunha variedade en particular, pero dado o contexto situacional (unha tarefa realizada no marco dunha clase) é esperable que os informantes empreguen o galego normativo ou, cando menos, que queiran demostrar que o coñecen. Daquela, a ausencia das formas normativas nas listaxes poderase explicar moitas veces simplemente polo seu descoñecemento ou, en calquera caso, pola súa escasa dispoñibilidade na mente dos falantes. E este feito, cando menos en parte, conduce á variación a que antes faciamos referencia, con presenza de dialectalismos, de castellanismos, de cognados do castelán, de adaptacións ao galego feitas *ad hoc* etc.

Pois ben, ante esta complexa situación era necesario analizar os pros e os contras dos posibles criterios de lematización e adoptar un sistema o máis útil posible para o usuario e que ao tempo fose axeitado para o lematizador. Nas seguintes epígrafes tentamos resumir esta análise, e asemade explicamos máis polo miúdo a aplicación de cada un dos criterios lematizadores presentados de forma esquemática en § 2.3.2.12, con todas as matizacións e excepcións precisas. Nesa explicación, faremos referencia con frecuencia á forma estándar galega como forma lematizadora. Son as nosas referencias fundamentais para determinar a variante lematizadora o VOLG, o DRAG e o *Diccionario castelán-galego* da Real Academia Galega, en adiante DCG—Real Academia Galega (2004)—. No caso de non atoparmos resposta no VOLG nin nos dicionarios indicados, acudimos a outras obras de referencia, esencialmente o GDXL e o dicionario en liña *Digalego*²⁷.

2.3.3.1 Variantes fónicas e morfolóxicas de orixe galega

Comecemos polas formas galegas procedentes dun mesmo étimo e que con frecuencia teñen un reparto dialectal (puidendo ter unha ou varias delas, asemade, carácter normativo): cumpría lematizar conxuntamente palabras como *cenoria*, *cenaira*, *cenoura* e *zanoria* ou sería preferible a lematización separada?

É evidente que esas catro formas poden considerarse catro unidades léxicas distintas e, neste sentido, poderían manterse individualizadas, pero, dada a enorme variabilidade existente, este sistema de lematización daría lugar a unha gran dispersión das respostas. Xa sinalamos previamente que

²⁷ Como xa sinalamos na nota 18, cando levamos a cabo a lematización das diferentes formas, non puidemos empregar a segunda edición do dicionario da Real Academia Galega (DRAG²), que duplica en rexistros a primeira, posto que non estaba dispoñible. Por outra banda, nunhas poucas ocasións encontramos no DCG algunhas formas ausentes no VOLG e no DRAG, que empregamos de todos os xeitos como lemas. É o caso de *mandarino* ou de *manteigado*.

traballamos en principio con dispoñibilidade léxica e non con dispoñibilidade conceptual, pero non é menos certo que os léxicos dispoñibles tamén se empregan para estudar esta última. Lematizar de maneira separada esas catro formas dificultaría considerablemente a análise da dispoñibilidade conceptual de ‘cenoria’, complicaría a busca realizada por investigadores interesados en coñecer as formas que os alumnos de segundo de bacharelato empregan para designar a raíz comestible en cuestión e entorpecería os contextos co nivel de dispoñibilidade léxica da voz equivalente noutras linguas.

Como xa indicamos, mantemos separadas formas sinónimas como *silla* e *cadeira* ou, engadindo máis exemplos, *oculista* e *oftalmólogo*, *saia* e *faldra* (entorpecendo as análises de dispoñibilidade conceptual en prol da diferenciación léxica), pero levar este criterio a casos como o indicado (ou a *aixada* e *eixada*, *bicho* e *becho*, *tabán* e *tabao*...) non nos parecía útil nin acertado. Nótese que, se no caso de sinónimos como *oculista* e *oftalmólogo*, ou mesmo *silla* e *cadeira*, é frecuente que as diferentes respostas convivan nunha mesma listaxe, esta situación é moito menos habitual nos exemplos do outro tipo, os cales, tal e como xa sinalamos, obedecen adoito a diferente distribución dialectal. Coidamos que na intuición de todos está a posibilidade de considerar *bicho* e *becho*, *tabán* e *tabao*... como variantes (fronte ao que sucede con *silla* e *cadeira*, por exemplo) e entendemos que resultaba máis útil para o usuario a lematización conxunta, que foi, en definitiva, a que aplicamos.

Pero ademais cumpría tomar outra decisión relativa ao conxunto de formas encaixables neste grupo. Cando a diferenza entre as voces tiña unha explicación de carácter fónico e había unha clara proximidade entre elas, non cabía dúbida da súa consideración como variantes: *bicho* e *becho*, *xirelo* e *xurelo*, *legón* e *leghón*... Máis dificultades se nos presentaban cando, compartindo todas as formas unha raíz etimolóxica común, había unha diferenza morfolóxica entre elas ou, máis en xeral, cando as voces estaban bastante afastadas formalmente: debiamos lematizar baixo a mesma entrada *cerdeira*, *cereixeira* e *cereixo*; *castiñeiro* e *castiro* ou *silar* e *ensilar*? Dadas as dificultades de establecer un corte claro, por unha banda, entre o que está próximo ou non formalmente, e, por outra banda, entre as diferenzas morfolóxicas e as meramente fónicas (pénsese, sen ir máis lonxe, no valor do *a-* inicial como prefixo ou como mero elemento protético), optamos por lematizar conxuntamente todas as formas de igual significado que compartisen unha raíz etimolóxica común, sempre e cando fosen consideradas como variantes por cando menos algunha das obras lexicográficas consultadas²⁸. Deste xeito,

²⁸ Entendemos que dúas voces son consideradas variantes nos dicionarios, ou que poden tratarse como tales, cando neles se remite dunha a outra forma, sexa mediante unha remisión directa ou mediante

tamén contribuíamos a evitar a dispersión nas respostas e favoreciamos os cotexos con outros léxicos dispoñibles, de acordo co arriba indicado.

Escollemos como lema para todas estas variantes que comparten designación e raíz a variante normativa, independentemente de que figure ou non nas listaxes. Deste xeito, lematizamos *xirelo* baixo *xurelo* ou *silar* baixo *ensilar*.

A aplicación do criterio que acabamos de mencionar pode conducir a algunhas lematizacións chocantes polo afastamento entre lema e variantes, como, por exemplo, o feito de que a voz *denosiña* figure baixo *donicela* ou *cereixo* baixo *cerdeira*. Repárese en que apelamos a un criterio etimolóxico moi amplo, que dá cabida a palabras de idéntico significado que proveñen exactamente do mesmo étimo (*choiva* / *chuvia*); que derivan dunha mesma raíz, pero que teñen diferenzas morfolóxicas, pola presenza de diferentes afixos (*bebedeiro* / *bebedeiro*) ou pola presenza / ausencia de afixos (*silar* / *ensilar*) e, de maneira xeral, a formas de igual significado encaixables nunha mesma familia léxica, con diferentes relacións entre si (*cerdeira* / *cereixeira*). Aínda que, como acabamos de mencionar, o criterio empregado pode conducir a algunhas lematizacións pouco intuitivas, resultaba moi difícil, como xa indicamos, establecer unha separación de doada aplicación entre as diferentes formas que comparten unha mesma raíz etimolóxica, de tal xeito que consideramos que as vantaxes do criterio acordado eran superiores ás desvantaxes, máxime tendo en conta que o usuario podería prever doadamente os lemas, que coinciden cos que se propoñen no código normativo.

No caso de que haxa dúas ou máis variantes normativas, todas elas consideradas co mesmo nivel de corrección, seguimos un criterio de frecuencia para determinar o lema. Así, no CI 01 ‘O corpo humano’, lematizamos as formas *unlla* e *uña* conxuntamente baixo *uña*, dado que, aínda que as dúas son voces normativas e recomendadas, *uña* é máis frecuente nas respostas dos informantes. O mesmo pasa con *castiro* e *castiñeiro* (lema *castiñeiro*) ou con *chopo* e *choco* (lema *choco*). Por suposto, e de acordo co explicado en § 2.3.2, nalgún caso o criterio adoptado pode conducir a diferentes lematizacións en distintos CI. Así, por exemplo, nos CI 10 e 18 o lema é *chuvia*, pero é *choiva*

definición sinonímica. Deste xeito, por exemplo, aínda que *cantil* e *acantilado* pertencen á mesma familia léxica e teñan unha designación moi próxima, son formas lematizadas separadamente porque ningún dos dicionarios consultados as trata como variantes (o DRAG sinala que a voz *acantilado* se emprega particularmente para designar os terreos escarpados de fonte pendente que se sitúan na costa; esta restrición non figura na definición de *cantil*). Un caso diferente é o daquelas variantes que non se recollen no dicionario, pero que se poden vincular claramente con formas que si figuran neles. É o caso de *lerme*, variante de *lesme* pronunciada con rotacismo, que á súa vez é variante recollida nos dicionarios da forma normativa *lesma*, voz que tomamos como lema. Outro exemplo, de diferente tipo, é *cociñante*, que figura no CI 16 ‘As profesións e os oficios’, e que unificamos con *cociñeiro*. Nalgúns destes casos, pode resultar difícil distinguir entre variantes propiamente ditas, creacións / adaptacións dos informantes e lapsus cálemi.

en 11 e 13. Incluso pode haber diverxencias nun mesmo CI: así, por exemplo, no CI O1 encontraremos *aparelo reprodutor*, pero *aparato circulatorio* ou *aparato dixestivo*, porque no primeiro composto sintagmático os exemplos con *aparelo* son máis frecuentes ca os exemplos con *aparato*, mentres que nos restantes sucede ao revés²⁹.

Somos conscientes dos problemas que pode supoñer o recurso á forma máis frecuente (imprevisibilidade do lema), pero consideramos que este procedemento é o máis axeitado. Outra opción obrigaría a escoller unha de entre as varias formas normativas e necesariamente habería que acudir a un criterio relativamente arbitrario como, por exemplo, o de seleccionar como lema a primeira forma de acordo coa ordenación alfabética. Agora ben, considerar que este criterio trae consigo a previsibilidade do lema é supoñer que o usuario coñece todas as formas aceptadas no código normativo, o cal, na nosa opinión, dista bastante de ser certo. Dado que, en principio, o léxico dispoñible é reflexo da fala dos informantes, entendemos que, nos casos nos que coexisten varias palabras da mesma raíz e do mesmo significado no código normativo, o usuario poderá intuír en boa medida cales van ser as formas máis frecuentes e, daquela, os lemas (así, *amorodo* e non *morote*, *coello* e non *coenllo*, *gruñir* e non *grñir*, *poñer* e non *pór...*). Evidentemente, non sempre ocorrerá así (*vid. supra* o caso de *chuvia* e *choiva*), pero isto sobre todo sucederá nas voces máis marxinais, con menor número de rexistros e daquela, menos relevantes para quen analiza o léxico dispoñible³⁰.

No caso dos sufixos *-ble* / *-bel* e *-ería* / *-aría*, aceptados en pé de igualdade no VOLG, optamos por lematizar sempre, respectivamente, con *-ble* e con *-ería*, con moito os sufixos máis frecuentes nas listaxes dos alumnos³¹.

Se no código normativo se aceptan varias formas, pero hai unha diferente valoración da súa corrección, escollemos como lema a forma recomendada (con independencia de que figure ou non nas respostas). Así, *morango* ou *cuidar*, voces aceptadas pero non recomendadas, lematízanse, respectiva-

²⁹ En caso de igual número de exemplos de *aparato* e *aparelo* dentro dun composto sintagmático concreto (por exemplo, hai un exemplo de *aparelo auditivo* e outro de *aparato auditivo*), escollemos *aparato* (lema *aparato auditivo*), por ser *aparato* máis frecuente no CI, globalmente, ca *aparelo*. Isto é aplicable a outros casos similares. No caso de que a análise do CI non nos axudase, acudimos ao conxunto de exemplos do LDG.

³⁰ Outra opción sería a de escoller como lema a variante que os dicionarios consideran como principal (aquela na que figura unha definición perifrástica e á que a outra forma remite). No entanto, isto implica apelar ao criterio de frecuencia (aínda que nin o DRAG nin o *Digalego* o sinalan explicitamente, supoñemos que a forma na que figura a definición perifrástica é a que se considera máis habitual). E entendemos que, de apelar a este criterio, era preferible acudirmos ás respostas dos falantes e ofrecer, deste xeito, un lemario máis fiel ás listaxes.

³¹ No caso de *-ería* / *-aría*, por suposto só nos casos, maioritarios, en que as dúas formas alternan no código normativo. *Romaría* e *romería*, por exemplo, lematízanse baixo *romaría*, a única forma aceptada no citado código.

mente baixo *amorodo* e *coidar*, que son as variantes consideradas con maior grao de corrección no VOLG³².

2.3.3.2 Voces do castelán cognadas de palabras galegas

A toma de decisión con respecto ás formas galegas procedentes dun mesmo étimo determina os criterios de lematización empregados noutros casos. É frecuente, como vimos no caso de *cenoria*, que nas listaxes convivan formas de orixe galega e formas cognadas de orixe castelá (sen distinguirmos agora entre os castelanismos, isto é, as formas de orixe castelá que se empregan habitualmente nalgunha variedade do galego, e as palabras do castelán que figuran nas listaxes, pero que non son habituais nos discursos en galego)³³. Así, ademais de *zanahoria* atopamos palabras como *ratón*, *queso*, *avellana*, *murciélagos*, *tejado*..., por citar só algúns exemplos de entre os moitísimos que se poden encontrar nas listaxes. Unha opción sería a de lematizar estas palabras de forma independente con respecto aos cognados galegos: *rato* e *ratón*, *queso* e *queixo*, *avellana* e *abelá*... Pero este criterio carrexaría moitos problemas, porque existe unha gran cantidade de palabras a medio camiño entre as formas galegas e as formas castelás. Que facer, por exemplo, con *morcélagos*, *teiado*, *herreiro*, *suxetador*, *amarelo*, *amarillo*...?, por non citar os casos das formas galegas con morfoloxía castelá (*mexillones*, *cullarones*...) ou das formas castelás con morfoloxía galega (*cañóns*, *señais*...). Parece claro que tales palabras non deben ter un lema propio, pero baixo que lema debería aparecer *morcélagos*, baixo *morcego* ou baixo *murciélagos*, baixo que lema debería aparecer *amarelo*, baixo *amarelo* ou baixo *amarillo*? Resulta imposible comparar e medir o nivel de influencia da forma galega e da forma castelá nestes casos híbridos, nestas formacións que, sendo en

³² O lema é *amorodo*, e non *amorogo* ou *morote*, tamén voces recomendadas, porque nas respostas dos informantes só hai exemplos da primeira forma. Aplícase, pois, o criterio de frecuencia consonte o explicado un pouco máis arriba.

³³ No LDEG identificouse o léxico específico de Galicia (esencialmente palabras de orixe galega) mediante a etiqueta [g]. Non consideramos axeitado levar a cabo unha marcación parella para as palabras do castelán rexistradas no LDG, por varios motivos. En primeiro lugar, por razóns de carácter práctico, pois non sempre é doado saber con certeza se estamos ante un dialectalismo do galego ou ante un castelanismo (*pimento* pode ser unha adaptación ao galego de *pimiento* ou unha variante dialectal, *agua* pode ser a palabra do castelán ou un dialectalismo...); por outra banda, son frecuentes as formas híbridas como *pavellón*, *prato ondo*...). Ademais, se no caso do castelán de Galicia a etiqueta galeguismo ten un valor xeográfico e de orixe (en xeral, marca voces de orixe galega específicas do castelán de Galicia), no caso do galego a dimensión xeográfica desaparece (non se trata de sinalar palabras específicas do galego de Galicia, fronte a outros galegos doutros lugares) e unicamente resta a dimensión da orixe, á que en realidade se engadiría unha dimensión prescritiva: a etiqueta [c] aparecería en voces de orixe castelá que son rexeitadas ou non recomendadas para a variedade estándar e non aparecería en voces aceptadas no código estándar galego, por exemplo, *bocadillo* ou *tortilla*. E coidamos que esta información non é apropiada nin relevante para este traballo, o que non obsta para que a análise das formas do castelán presentes nas listaxes sexa de gran interese —véxase ao respecto Álvarez de la Granja (no prelo)—.

xeral residuais tomadas individualmente (hai un exemplo de *morcélago*, un de *herreiro*, dous de *amarello*, aínda que corenta e nove de *suxetador*), no conxunto das listaxes non son en absoluto desdeñables.

Por outra banda, algunhas formas dialectais galegas coinciden coas voces castelás, de tal xeito que en moitas ocasións resultaría imposible atribuír-lles unha orixe ás respostas. Por exemplo, dos trinta e un informantes que ofrecen a variante *agua* no CI 05 ‘Os alimentos e as bebidas’, en que casos estamos ante a palabra castelá que se introduce directamente nas listaxes (como se introduce, por exemplo, *perro* en setenta e sete ocasións no CI 14 ‘Os animais’) e en que casos estamos ante a variante dialectal galega? É realmente imposible sabelo agás que fagamos unha comparativa da distribución dialectal das variantes e da orixe dos alumnos, pero isto, á parte de delongar e complicar enormemente o proceso de lematización, non conduciría necesariamente a conclusións acertadas.

En vista das dificultades prácticas de manter separadas formas galegas e formas castelás cognadas optamos por unificar todas elas baixo un mesmo lema. Deste xeito, as voces galegas e as voces castelás cognadas (así como as formas híbridas) do mesmo significado e extensión aparecerán lematizadas de maneira conxunta baixo a palabra normativa galega, independentemente de que esta última forma sexa minoritaria ou non figure nas listaxes: *rato* e *ratón* baixo *rato*, *queixo* e *queso* baixo *queixo*, *abelá* e *avellana* baixo *abelá* ou *morcego*, *morcélago* e *murciélago* baixo *morcego*. Deste modo, unha vez máis, evitamos a dispersión nas respostas e favorecemos a realización de cotexos con outros léxicos dispoñibles. Como indicamos, aplicamos este criterio de lematización mesmo cando nas respostas dos alumnos é maioritaria a palabra do castelán ou mesmo cando non hai exemplos da voz galega. Así, lematizamos como *adianto* a forma *adelanto* que aparece no CI 10 ‘A cidade’, malia que a variante *adianto* non figura nas listaxes. O mesmo vale, claro está, cando o único exemplo que aparece é unha forma híbrida. Lematizamos a voz *avellás* que figura no CI 11 ‘O campo’ baixo *abelá*. Os criterios para a selección do lema galego estándar, caso de existir varias posibilidades, son os explicados no apartado anterior. Nos poucos casos en que a unha palabra do castelán lle corresponde unha palabra cognada do galego aceptada no estándar, aínda que pouco recomendada, lematizamos a través desta. Deste xeito, *disfrutar* lematízase baixo *desfrutar*.

O mesmo criterio laso empregado para a consideración de dúas formas galegas como variantes, agrupables baixo un mesmo lema, é aplicable con respecto ás formas galegas e castelás (e adaptacións destas). Deste xeito, unificamos baixo a variante galega estándar as formas que comparten a raíz

etimolóxica, isto é, as palabras que proceden exactamente do mesmo étimo (*ferreiro* e *herrero*), as palabras que derivan dunha mesma raíz, pero presentan diferenzas morfolóxicas pola existencia de diferentes afixos (*concelleiro / concejal, concellal, concexal*) ou pola presenza / ausencia de afixos (*castiñeiro / castaño*) e as que, compartindo raíz, manteñen outro tipo de relacións (*ril / riñón; sebe / seto*). No entanto, lematizamos baixo a forma castelá (e non unificamos variantes galegas e castelás, caso de atoparmos ambas as dúas) cando algunha das voces (a galega ou a castelá), podendo remontarse ao mesmo étimo ca a outra, presenta influencias doutras palabras na súa etimoloxía, o que determina certo afastamento formal entre galego e castelán. É o caso de *cohete* (fronte a *foguete*) ou de *cerrojo*, rexistrada na forma adaptada *cerroxo* (fronte a *ferrollo*).

A aplicación deste criterio conduce a algunhas lematizacións chocantes, como por exemplo, o feito de que *calabacín* figure baixo *cabaciña* ou *ferreteria* baixo *ferraxería*, máxime cando non hai exemplos nas respostas dos alumnos das formas galegas estándares. Neste sentido, a listaxe de lemas debe verse en boa medida como unha relación de formas galegas normativas (cando estas existen), as cales permiten unificar voces coa mesma designación e coa mesma raíz etimolóxica e facilitan as buscas do usuario, pero non podemos esquecer que ás veces está bastante afastada das respostas reais dos alumnos. Estas poden consultarse no apéndice 1.

Por outra banda, se nas respostas atopamos expresións complexas con elementos do castelán, de tal xeito que só algúns deles posúen cognados galegos, prescindimos en xeral de lematizar con estes, para evitar formas cunha mestura potencialmente chocante de voces galegas e castelás. Así, por exemplo, no CI 13 hai un caso de *perro guardián*³⁴. Aínda que *guardián* posúe o cognado galego *gardián*, optamos por lematizar directamente mediante *perro guardián* e non mediante *perro gardián* polas razóns que acabamos de comentar. O mesmo vale para *rojo teja* ou *rojo pasión*³⁵. No entanto, cando a voz sen cognado en galego é habitual nesta lingua (un castelanismo ben implantado, non “etiquetado” por boa parte dos falantes como forma doutro idioma e sen claras marcas identificadoras da súa orixe foránea³⁶) non facemos excepción (o único exemplo de *silla de mimbre* lematízase baixo *silla de vime*). Nótese, en calquera caso, que tratamos sempre con formas minoritarias, exemplos únicos, situados en posicións marxinais nas listaxes.

³⁴ É unha expresión marxinal no CI, cun único rexistro. Ao seu pé non hai exemplos de *perro gardián*.

³⁵ Nótese que o cognado galego *roxo*, no código normativo, non é equivalente semanticamente de *rojo*, co cal os únicos elementos potencialmente substituíbles por un cognado galego serían *teja* (*tella*) e *pasión* (*paixón*). Non atopamos, por outro lado, exemplos de *rojo tella* nin de *rojo paixón*. Tampouco hai exemplos de *silla de vime / vimio* no caso que comentamos máis adiante.

³⁶ *Rojo* é un castelanismo moi estendido, pero a súa grafía mostra claramente a súa orixe.

No caso de voces aparentemente creadas polos alumnos, sen referendo normativo pero cunha estrutura transparente, procuramos adaptalas ao código normativo galego: así, por exemplo, *posa manteles* lematízase baixo *pou-samanteis*, *pasaplatos* como *pasapratos* ou *quitahumo* como *quitafume*.

No mesmo sentido, cando encontramos unha expresión derivada ou complexa en castelán, sen equivalente nos dicionarios galegos, pero que presenta noutras linguas un esquema formativo semellante ao do castelán, optamos por propoñer un lema con compoñentes galegos (por exemplo, tal e como se explica no comentario ao CI 14, lematizamos *escarabajo bombardero* baixo *escaravello bombardeiro* ou *ciervo-mula* baixo *cervo mula*, malia os dicionarios galegos non recoller ningunha das dúas expresións, polo feito de que noutras linguas as designacións dos devanditos animais seguen o mesmo esquema formativo ca en castelán). Outro criterio empregado para levar a cabo unha lematización deste tipo é o feito de que as voces con elementos galegos se recollan en discursos en galego, nomeadamente se estes emanan de organismos oficiais, institucións educativas etc. (por tal motivo, lematizamos, por exemplo, *pizzero* baixo *pizzeiro* no CI 16; *vid.* o comentario a este CI para unha xustificación máis detallada).

Por outra banda, a lematización baixo a forma galega cognada só se produce, tal e como indicamos, cando a voz de orixe galega e a voz de orixe castelá, ademais de posuíren raíces cognadas, teñen o mesmo significado e a mesma designación ou extensión. É evidente, por exemplo, que no CI 18 ‘As cores’ non podemos lematizar *rojo* baixo *roxoxo*, na medida en que ambas as palabras, aínda que etimoloxicamente emparentadas, posúen significados claramente distintos³⁷.

Noutras ocasións, a palabra galega e a cognada castelá só comparten parcialmente a súa extensión. A diferenza entre ambas pode deberse a que a forma galega é máis extensa ca a castelá ou ao revés. A primeira circunstancia ocorre, por exemplo, coas palabras *percha* e *perchero* (adaptada frecuentemente como *percheiro*), que non se unifican baixo *percha*, a única forma aceptada no código normativo galego. Isto é debido a que *percha* pode designar tanto a peza ou moble con varios soportes para a roupa como o soporte individual

³⁷ Dito isto, estamos convencidas de que moitos dos cento sesenta e sete exemplos de *roxoxo* que figuran no CI ‘As cores’ son, sen ningunha dúbida, adaptacións fónicas do castelán *rojo* (sinónimos, logo, de ‘vermello’) e, neste sentido, poderían unificarse con esta última forma. No entanto, non é menos certo que nalgúns deses exemplos se podería estar usando a voz *roxoxo* no seu sentido propiamente galego, para designar unha cor de pelo (de feito, no mesmo CI hai exemplos de *moreno* ou de *louro* e un exemplo de *roxoxo do pelo*). Dado que resulta imposible descubrir o sentido que quería atribuír cada informante aos diferentes exemplos de *roxoxo*, imponse a necesidade de manter separados *roxoxo* e *rojo*, a non ser que o contexto lingüístico nos permita interpretar o sentido da voz. É o caso de *roxoxo inglés* e de *boli roxo*, que se unifican con *rojo inglés* e *boli roxo*.

con gancho. No entanto, *perchero* / *percheiro* só posúe o primeiro valor. Se situamos *perchero* (ou *percheiro*) baixo a voz máis xenérica *percha* estamos perdendo información relevante sobre o tipo de realidade concreta a que os infomantes están facendo referencia ao empregar a primeira forma. Outro exemplo encaixable neste grupo é o de *fajín* e *faixa*, pois a voz galega posúe un significado moito máis xeral ca a palabra do castelán, de tal xeito que lematizando *fajín* baixo *faixa* a especificade da primeira voz queda diluída.

Outras veces é a voz castelá a que posúe unha maior extensión. Así ocorre, por exemplo, con *buhardilla*, que, seguindo o *Diccionario de la Real Academia Española* (en adiante DRAE —Real Academia Española 2001²²—), designa tanto unha abertura que deixa pasar a luz como unha estancia, mentres que a voz cognada do galego, *bufarda*, só remite a unha ‘pequena abertura na parede ou no teito dunha casa para dar luz, para ventilar ou para saír ó tellado’. A lematización de *buhardilla* baixo *bufarda* supón tomar partido por un dos sentidos do vocábulo castelán, o cal non parece conveniente.

Agora ben, debe terse en conta que, como xa indicamos, a lematización é independente para cada CI e pode considerar, ademais, o contexto imposto pola temática deste. Isto determina que os lemas poden variar de CI a CI. Así, por exemplo, no CI 03 ‘As partes da casa (sen mobles)’ non podemos lematizar de maneira unificada a voz de orixe castelá *pileta* e a voz *pía*, posto que *pileta* fai referencia en Galicia ás pías pequenas da cociña ou do cuarto de baño, mentres que *pía* designa un ‘recipiente cóncavo de pedra ou doutro material, para conter líquidos, de diversas formas e tamaños segundo o uso que se lle dá’ (a pía das cociñas ou dos baños, pero tamén, por exemplo, a dos porcos)³⁸. No entanto, no CI ‘07 ‘A cociña e os seus utensilios’ *pía* e *pileta* si se unifican baixo a forma normativa *pía*, posto que a designación de ambas as palabras neste CI é necesariamente a mesma (a pía da cociña). Outras veces, é o contexto lingüístico o que permite lematizar baixo a forma estándar galega voces que noutras circunstancias son lematizadas baixo o vocábulo do castelán. É o caso de *bote do pimentón*, lematizada baixo *bote do pemento* (no CI 06 ‘Os obxectos colocados na mesa para a comida’), ou de *tarta de Santiago*, lematizada baixo *torta de Santiago* (no CI 05 ‘Os alimentos e as bebidas’). Pola contra, as variantes *pimentón* por unha banda e *tarta*, pola outra, cando aparecen illadamente, fóra de expresións como as arriba sinaladas, lematízanse respectivamente baixo *pimentón* e *tarta*, posto que a súa designación non ten por que coincidir necesariamente coa de *pemento* e *torta* (en ambos os casos, a forma galega ten maior extensión ca a castelá).

³⁸ Téñase en conta que neste CI o concepto ‘casa’ é entendido de maneira ampla, dando cabida a todo tipo de realidades que se poden atopar ao redor da vivenda unifamiliar, como *cortello*, *canastro*, *corte*, *xardín*...

2.3.3.3 Voces do castelán sen cognado galego

Un problema doutro tipo ofrécenos aquelas respostas coincidentes con palabras do castelán, pero que non posúen un cognado galego no código normativo (sexan de novo castelanismos propiamente ditos, isto é, interferencias presentes no galego ou nalgunha variedade deste, ou palabras tiradas directamente do castelán que o alumno introduce nas listaxes). Non son problemáticas as voces nas que figura a forma castelá sen ningún tipo de adaptación (ou simplemente con algunha falta de ortografía), pois neste caso o lema coincide coa forma normativa do castelán: así, lematizamos baixo *cerdo* os corenta e oito exemplos desta voz que figuran no CI 14 ‘Os animais’, e seguimos o mesmo criterio para lematizar os catrocentos sesenta e dous exemplos de *silla* que atopamos nos CI 4 ‘A casa: mobles e materiais’³⁹ ou o único exemplo de *acha* que figura no CI 11 ‘O campo’, lematizado baixo *hacha* (*vid.* nota a pé 1 da táboa 14).

A dificultade xorde naqueles casos en que nas respostas figuran adaptacións ao galego de voces do castelán sen cognado galego no código normativo. Estas adaptacións poden empregarse realmente na fala (é o caso, por exemplo, de *xudía*) ou ser moi probablemente creacións *ad hoc* cuxa vitalidade non vai máis aló da propia listaxe (*renacuallo*). Por outra banda, pode tratarse de adaptacións cunha base fónica, como as indicadas, ou ben de calcos morfoléxicos, como o que temos en *lavavaixelas*, en que os compoñentes dunha unidade complexa se traducen ao galego, a través dos correspondentes equivalentes, moitas veces cognados dos elementos do castelán.

A disxuntiva que se nos ofrece é a de lematizar coa forma castelá normativa ou ben a de manter no lema a adaptación ao galego feita polos informantes. Probablemente nuns casos é máis apropiada a lematización coa forma galega (cando realmente esta existe na fala, caso de *xudía*), mentres que noutros, nas creacións *ad hoc*, que poden presentar formas moi diversas (*renacuallo* convive con *renacuaxo* e mesmo con *ranacuaxo*), o máis aconsellable é a lematización a través da forma do castelán (neste caso, *renacuajo*, palabra que, por outra banda, tamén figura nas listaxes). No entanto, non sempre é posible discernir entre uns casos e outros, de tal xeito que se impón como necesario desde un punto de vista práctico un tratamento conxunto de todos os exemplos.

Ante esta disxuntiva, optamos por escoller unha solución a medio camiño entre as dúas indicadas, lematizando a través da forma máis frecuente nas listaxes. Así, por exemplo, no CI 13 ‘Os traballos do campo e do xardín’ atopamos as respostas *abonar* e *aboar*. Escollemos como lema a forma con <n>, sen adaptación ao galego, dado que as respostas *aboar* son minoritarias. En caso de que haxa o mesmo número de rexistros da palabra adaptada e sen

³⁹ Neste mesmo CI hai douscentos setenta e un exemplos da forma do código normativo *cadeira*.

adaptar, lematizamos mediante esta última. Aplicamos o mesmo criterio nas palabras do castelán que posúen un cognado galego, pero cunha extensión ou significado diferentes e non apropiados para o CI no que nos estamos situando. Así, por exemplo, no CI 08 ‘A escola: mobles e materiais’, os dous exemplos de *suelo* e os tres de *solo* (que resultan dunha adaptación ao galego de *suelo* e non do emprego da voz galega estándar *solo*⁴⁰) lematízanse baixo *solo* por ser máis frecuente a forma adaptada ca a forma sen adaptar.

As vantaxes deste tratamento son que resulta coherente, que se mostra fiel ás respostas ofrecidas polos informantes e que, daquela, cando menos nas voces con maior grao de frecuencia, que son as máis representativas, dá lugar a un leuario máis axustado á realidade (por exemplo, no caso de *xudía*, no CI 05 ‘Os alimentos e as bebidas’ hai setenta e catro exemplos con <x> e tan só seis con <j>). Ao mesmo tempo, dado que estamos tratando cun léxico dispoñible do galego consideramos que a fuxida cara ao castelán, que por outro lado, pode ser moi doada desde o punto de vista práctico, agacha de vez e de maneira pouco conveniente esa tendencia á adaptación ao galego que se extrae das respostas dos informantes.

Non se nos escapa, de por parte, que o sistema adoptado implica tamén unha serie de problemas. Talvez o máis importante é que o criterio de frecuencia conduce a un lema menos previsible que o criterio da forma castelá normativa. Isto é, o usuario que desexa buscar un lema determinado non sabe se debe procurar as formas *xudía* ou *judía*, *acera* ou *aceira*, *carretera* ou *carreteira*. Por suposto, se se lematiza sempre a través da forma castelá, este problema desaparece. Outra cuestión, mostra palpable do anterior, ten que ver co feito de que, ao aplicar un criterio de frecuencia, os lemas, unha vez máis, poden variar dependendo do CI: atopamos o lema *aceira* no CI 12, pero *acera* no CI 03; no CI 10 lematizamos o único exemplo de *canerías* como *canería*, mentres que no CI 09 o lema é *cañería*, por ser máis frecuente entre as respostas dos informantes a variante con nasal palatal etc. Finalmente, aplicando o criterio comentado, as listaxes de lemas poden verse salpicadas moi de cando en vez de formas estrañas, rexistros únicos resultado de procesos de adaptación *ad hoc* e sen vitalidade na fala (por exemplo, o xa mencionado *renacuallo* ou *vaia* no CI 03 ‘As partes da casa’, adaptación de *valla*). No entanto, son escasos os exemplos (lémbrese que estamos falando só de formas sen cognado galego) e moi pouco representativos (precisamente polo seu carácter *ad hoc*), co cal ocuparán sempre lugares absolutamente marxinais nas listaxes.

⁴⁰ Nótese que a forma equivalente de *suelo* no sentido no que se debe interpretar a voz neste CI (‘superficie sobre a que se anda’) é *chan*. *Solo* posúe outros significados (‘parte superficial da codia da Terra’, ‘terreo onde se pode edificar’ etc.).

Valorados os pros e os contras, optamos, como xa indicamos, por concederlles aos informantes o rol decisorio á hora de lematizar as formas de orixe castelá sen cognado galego.

O criterio é de aplicación tamén no caso dunha unidade complexa (por exemplo, *lavaplatos*) que estea constituída por elementos que posúen cognados galegos (neste caso, *lava* e *pratos*), pero que non teña un correlato estrutural no galego estándar (a forma estándar galega é *lavalouza*, non *lavapratos*), pois consideramos as unidades deste tipo, no seu conxunto, como carentes de cognado. Así, no CI 04 lematizamos os dous exemplos de *lavapratos* e os catro de *lavaplatos* baixo *lavaplatos*, a forma máis frecuente. Para excepcións a este tratamento, véxase o indicado no apartado anterior ao respecto de formas como *escaravello bombardeiro* ou *cervo mula*.

Por outra banda, cando os falantes ofrecen un calco morfolóxico dunha expresión do castelán e o resultado é unha palabra con cando menos unha raíz non cognada da orixinal (por exemplo, *petivermello*, creada sobre *petirrojo*), a voz terá o seu propio lema, e non se unificará coa forma castelá (caso de existir nas listaxes) nin coas correspondentes adaptacións “fónicas” desta (deste xeito, *petivermello* non se unifica cos dous exemplos de *petirroxo* que figuran no mesmo CI).

Para as expresións complexas formadas ao tempo con palabras do castelán que posúen cognado galego e con outras sen cognado, véxase a sección anterior. Se o falante ofrece directamente unha forma híbrida, con formas galegas e formas do castelán sen cognado, mantemos a mestura no lema: *baixar o perro* lematízase como *baixar o perro*.

2.3.3.4 Supresión de variantes galegas e de cognados

O criterio empregado para a unificación de variantes galegas e de cognados galegos e casteláns, en combinación co feito de que non pode haber dous lemas iguais na mesma listaxe, pode supoñer unha pequena perda de información para o investigador. Naqueles casos nos que unha persoa ofrece para o mesmo CI, isto é, nunha mesma listaxe, dúas variantes diferentes, pero etimoloxicamente relacionadas, ou unha forma galega e o seu cognado castelán, eliminamos a segunda mención, para evitar a reduplicación de lemas. No entanto, pode ser relevante coñecer precisamente a coexistencia destas formas no léxico dun único individuo. Cando esta coexistencia se dá, non sería descartable a especialización semántica dalgún dos vocábulos⁴¹ ou a súa

⁴¹ Isto é probablemente o que sucede con *batedor* (manual) e *batedora* (eléctrica). No entanto, dado que a única forma do código normativo é *batedor*, que os dicionarios de galego lle atribúen a esta forma calquera das dúas designacións (aparelo manual ou aparelo eléctrico) e que o DRAE considera *batidor* e *batidora* como variantes co mesmo significado (nótese que *batedora* é probablemente unha formación creada a partir do castelán *batidora*), optamos por lematizar tanto as formas masculinas como as

diferente marcación diafásica, por moito que estas non aparezan recollidas nos dicionarios. Por tal motivo, decidimos tomar nota de todas as formas deste tipo eliminadas nas respostas dos informantes e ofrecelas neste traballo. No cadro que segue, as formas en cursiva (á dereita de cada parella) son as formas eliminadas (que non figuran, daquela, na listaxe de variantes e que non se computan como realizacións do lema) e as formas en redonda (á esquerda) son as que se conservaron, por apareceren antes nas listaxes (nótese que non son lemas, senón formas; de aí as parellas *tesoira / tisoira* e *xerra / serra*). Entre parénteses indicamos o número de respostas eliminadas (isto é, o número de veces que se produce a coincidencia)⁴².

Táboa 15. Variantes e formas cognadas repetidas nunha mesma listaxe

abuelo– <i>avoa</i> (1)	amarelo– <i>marelo</i> (1)	amarillo– <i>amarelo</i> (1)
amizade– <i>amistade</i> (1)	anzol– <i>anzó</i> (1)	batedor– <i>batedora</i> (2)
becada– <i>becaciña</i> (2)	branco– <i>blanca</i> (1)	castaño– <i>castiñeiro</i> (1)
chocolate– <i>chiculate</i> (1)	codo– <i>cóbado</i> (1)	coidar– <i>cuidar</i> (1)
coitelo– <i>cuchillo</i> (2)	cuchara– <i>culler</i> (5)	cucharón– <i>cullerón</i> (1)
cuchillo– <i>coitelo</i> (2)	culler– <i>cuchara</i> ¹ (6)	culleriña– <i>cuchariña</i> (1)
cultivar a horta – <i>cultivar o horto</i> (1)	doctor– <i>doutora</i> (1)	ensilar– <i>silar</i> (2)
espalladeira– <i>espalladora</i> (1)	espumadeira– <i>espumadera</i> (1)	farol– <i>farola</i> (1) ²
farola– <i>farol</i> (1)	foz– <i>hoz</i> (1)	fritir– <i>frixir</i> (1)
garabata– <i>grabata</i> (1)	gasoil– <i>gasóleo</i> (1)	igreja– <i>eirexa</i> (1)
jallina– <i>galiña</i> (1)	laranxa ³ – <i>naranxa</i> (1)	linterna– <i>lanterna</i> ⁴ (1)
maestro– <i>mestre</i> (1)	paloma– <i>pombiña</i> (1)	peatones– <i>peatóns</i> (1) ⁵
peixe– <i>pez</i> (1)	piñeiro– <i>pino</i> (1)	rato– <i>ratonciño</i> (1)
ratón– <i>rato</i> (1)	regla– <i>regra</i> (1)	riñón– <i>ril</i> (3)
rosado– <i>arrosado</i> ⁶ (1)	sembrar– <i>sementeira</i> (1)	sembrar (1)– <i>sementar</i> (2)
semillar– <i>sementar</i> (1)	sótano– <i>soto</i> (1)	soto– <i>sótano</i> (2)
sudadera– <i>sudadeira</i> (1)	tendero– <i>tendeiro</i> (1)	tesoira/tisoira– <i>tixeira</i> (2)
tixeira– <i>tixoira</i> (1)	topo– <i>toupa</i> (1)	touro– <i>toro</i> (1)
water– <i>váter</i> (1)	xabalín– <i>xabarín</i> (1)	xarra– <i>xerra</i> (2)
xavali– <i>xavarín</i> (1)	xerra / serra– <i>xarra</i> (5)	

¹ Nun caso dos 6 computados, a forma eliminada é *chuchara*.
² Tanto este caso como o seguinte son rexistrados no CI 10 'A cidade', onde a interpretación máis plausible de *farol* é á correspondente ao cognado *farola*.
³ A resposta que figura nas listaxes é *laranza*.
⁴ A resposta que figura nas listaxes é *lantena*.
⁵ Como se ve, incluímos tamén esta parella onde a variación afecta exclusivamente ao morfema de plural.
⁶ Esta variante non figura en ningún dicionario dos consultados e é posible que se trate dunha creación do informante.

femininas baixo o lema normativo *batedor*.

⁴² Reproducimos as voces sempre en singular, aínda que nas listaxes figuren en plural, e repoñemos acentos gráficos.

2.3.4 Criterios específicos de edición. Os centros de interese

2.3.4.1 CI 01. O corpo humano

As voces que integran este centro de interese non se limitan aos substantivos que designan os diversos órganos (*corazón, fígado, ril...*), partes (*brazo, cabeza, perna...*) substancias e elementos en xeral (*auga, enzima, sangue...*) que constitúen o corpo humano; os informantes interpretan o estímulo de maneira moi xeral e achegan con bastante frecuencia palabras relacionadas con distintos aspectos vinculados máis ou menos directamente con aquel.

Así, encontramos respostas que fan referencia ao funcionamento dos órganos e elementos arriba mencionados e ás diferentes funcións vitais dos seres humanos (*alimentar, anabolismo, bombear, comunicación, crecemento, dor, excreción, expiración, metabolismo, mover, pensar, reprodución, respirar...*), entre as que se inclúen bastantes palabras referentes a accións e procesos frecuentes nos seres humanos (*andar, beber, correr, durmir, escoitar, falar, mastigar, sentar, tragar, tusir, ver...*), algunhas delas con clara incidencia na saúde (*deporte, fumar*). Precisamente, tamén encontramos numerosas respostas que se relacionan coa (a)normalidade do funcionamento do organismo, coa saúde e a enfermidade (*alergia, astigmatismo, hepatite, medicación, taquicardia, transplante...*). Non faltan neste grupo as designacións de profesións enmarcables na sanidade (*dentista, enfermeiro, médico de familia, oftalmólogo...*), así como de disciplinas relacionadas (*bioloxía, medicina, química, xenética*). Cómpre mencionar tamén a presenza nas listaxes de varias voces que fan referencia á evolución do ser humano (*feto, neno, puberdade, home, muller; nacer, morrer...*).

Por outra banda, son relativamente frecuentes os vocábulos vinculables co estado ou o aspecto físico das persoas (*alto, atractivo, bonito, calvo, cansazo, delgadeza, pálido, rubio, sobrepeso...*), así como os que se relacionan coa psicoloxía humana (*alegría, emoción, erotismo, intelixencia, odio, personalidade, pracer...*).

Como se pode extraer do anteriormente mencionado, unha boa parte dos informantes amplían o contido do CI, que pasa de 'O corpo humano' a 'O ser humano'. Nós respectamos esta ampliación e mantemos nas listaxes as palabras encadrables nos grupos anteriores. Excluímos voces que non se axustan a eles ou que o fan de maneira moi indirecta, así como aquelas cun contido moi xeral (*abrir, cefalotórax, diferente, fitoplasma, increíble, lente, poder, ximnasio...*). Tamén deixamos fóra uns algúns exemplos de pezas de roupa, que teñen o seu lugar natural no CI 02 (*calcetín, pantalón, xersei, zapato...*).

No que atangue ás particularidades da lematización, cómpre xustificar algunhas das decisións tomadas:

Lematizamos a variante *beixos* baixo *beizo* e non baixo *beixo* ('bico, demostración de afecto'), porque o contexto (*beixos-queixo*), o feito de que a palabra apareza en plural, así como a maior centralidade do concepto 'labio' neste CI parecen conducir á interpretación do rexistro como unha forma errónea que en realidade quere transmitir este último sentido (nas listaxes hai tamén tres exemplos de *bico*, pero debe terse en conta que esta palabra designa en galego tanto a demostración de afecto como a boca das persoas).

Por outra banda, é posible que chame a atención nas listaxes a presenza baixo *man* da variante *man esquerda* (e *mano izquierda*), sen que ao seu carón se rexistren variantes co adxectivo *dereito*. Os alumnos que ofreceron estas respostas escribiron a continuación *man dereita* (e *mano derecha*). No proceso de lematización consideramos que o modificador era prescindible, de tal xeito que lematizamos ambas as respostas con *man*. Isto deu lugar a lemas repetidos nunha mesma listaxe e á necesidade de suprimir o segundo exemplo (*man dereita / mano derecha*), de aí que estas variantes non figuren nas listaxes finais. Problemas similares aconteceron con *brazo* e *pé*.

No que atangue a cuestións de número, cómpre sinalar que, aínda que nos dicionarios consultados non atopamos recollida unha acepción específica para o plural de *cervical*, decidimos lematizar separadamente *cervical* e *cervicais*, entendendo que a primeira palabra é un adxectivo e o segundo un substantivo plural que designa as vértebras cervicais.

Con respecto aos agrupamentos de rexistros, mantemos separados *papila* e *papila gustativa* porque a palabra *papila* ten un carácter máis xenérico, tal e como se pode comprobar nesta definición extraída do DRAG²: 'cada unha das pequenas protuberancias cónicas que se forman na superficie da derme, e nalgúns mucosas, especialmente da lingua'. Noutros casos si levamos a cabo o agrupamento porque entendemos que a designación da forma simple e da complexa coinciden: *columna (vertebral)*, *planta (do pé)*.

2.3.4.2 CI 02. A roupa

Neste CI encontramos, como cabía esperar, nomes de pezas de roupa para vestir (*camisa, saia, xersei...*), de distintos tipos de calzado (*tenis, zapatilla, zapato...*) e algúns representantes da roupa do fogar (*edredón, manta, saba...*).

Entre os primeiros, é merecente de atención a presenza dos vocábulos *fórmula* e *ringo*, con un e tres exemplos respectivamente, e que non atopamos

en ningún dicionario como designadores de pezas de vestir (ambas as dúas palabras foron citadas por alumnos da provincia de Lugo). Tras interrogar a varias persoas da zona dos informantes, descubrimos que as dúas voces fan referencia aos xerseis de colo alto, finos, que adoitan colocarse baixo outras pezas de vestir, concretamente baixo vestidos sen manga ou baixo outros xerseis. Un falante que posúe no seu idiolecto o termo *fórmula* vinculou a orixe desta palabra co grupo Fórmula V, cuxos compoñentes, famosos nos anos 70, adoitaban empregar tal tipo de xerseis. A palabra *ringo* podería vincularse co compoñente dos Beatles, Ringo Starr, que tamén utilizaba con frecuencia esa peza de roupa. En calquera caso, lematizamos en minúscula as dúas palabras pois parece claro que, se esas foron as orixes, estamos ante exemplos evidentes de lexicalización.

Pero ademais este centro recolle tamén numerosos elementos que, sen constituíren a cerna do ámbito temático, posúen unha clara vinculación con el. De maneira xeral, mantéñense todos estes vocábulos na presentación final do centro.

Por unha banda atopamos un numeroso grupo de nomes de accesorios e complementos, entendidos en sentido amplo: *abano, bolso, goma do pelo, leontina, paraugas, pinza, pulseira, reloxo, tatuaxe...* Outro grupo numeroso de voces fai referencia a actividades e entidades que teñen que ver directamente coa confección de pezas de roupa (*agulla, calcetar, coser, costureira, dedal, máquina de coser, meter baixos, tecer, xastre, zurcir...*), así como co mundo empresarial que rodea a citada actividade (*empresa, empresa téxtil, grande empresa...*) e co ámbito da moda en xeral (*deseñador, fashion, fotografía, modelo, pasado de moda, pasarela, Nova York, tendencia...*). Como cabía esperar, o centro acolle tamén unha gran cantidade de respostas que se vinculan de forma máis ou menos directa co acto de comprar roupa (*boutique, cadea, cliente, escaparate, liquidación, probador, rebaixas, vender...*) e, aínda que se pedira expresamente que non se incorporasen nomes propios, atopamos tamén unha elevada cantidade de marcas e establecementos: *Adidas*, Adolfo Domínguez*, H&M*, Inditex*, Levi's*, Reebok*, Rei Zentolo*, Umbro*, Zara*...* Como xa indicamos (§ 2.2.1), mantemos estas respostas porque o feito de figuraren nun número elevado, malia a súa prohibición expresa, é un claro indicio da súa importancia na conformación mental deste ámbito temático.

Rexistramos tamén algúns exemplos de usuarios de roupa, dos que eliminamos os máis xenéricos (*chico, home, muller, neno, persoa...*) e mantemos os que designan grupos de individuos que en parte se caracterizan polo emprego dun tipo de roupa específico (*pijo, punky...*). A carón dos usuarios,

atopamos os contextos marcados precisamente pola utilización dun tipo de roupa en particular ou pola importancia que se lle concede á vestimenta: *ca-lor, entretempo, festa, inverno, saída nocturna, voda...*

Dado o papel fundamental que a roupa ten na caracterización das persoas, rexistramos tamén varios vocábulos que fan referencia ao arraxo persoal, á aparencia física e a outros aspectos relacionados: *arreglarse, atractivo, conxuntado, elegancia, gordura, ligar, look, prepararse, presumir, sinxeleza...* Atopamos tamén varias respostas que amosan a directa relación entre moda e aparencia física e aspectos psicolóxicos e sociolóxicos: *anorexia, bulimia, manipulación, personalidade, poder de elección, publicidade, sociedade, televisión...*

No que atangue á caracterización da roupa propiamente dita, encontramos materiais (*algodón, flanela, poliéster, satén, seda, terciopelo, tergal...*), partes de pezas de roupa e de calzado (*cámara de aire, escote, manga, puño, revés, solapa...*), e respostas concernentes sobre todo ás cores e vistosidade (*azul, gris, rechamante, vermello, vistoso...*), ao prezo e calidade (*barato, bo, caro...*), aos estampados, formas, adornos e complementos que pode ter a roupa (*con volantes, plisado, raia, rombo...*), á fermosura (*bonito, feo, lindo...*), á textura e sensacións, sobre todo térmicas, que produce (*abrigar, caloroso, fresco, suave, textura...*), á época, ao contexto ou ao estilo (*diferente, formal, friki, orixinal, outonal, veraniego...*), á adecuación ao corpo e disposición neste (*apertado, cinguido, cómodo, folgado, interior...*) e á dimensión (*ancho, grande, longo, talle...*). Ao respecto deste último subgrupo eliminamos as designacións específicas de talles (*m, s, xxl...*). Ademais, hai rexistros que aluden tamén aos estados nos que se poden atopar as pezas de roupa (*con engurras, descosido, esfiañado, raído, roto, vello...*).

Outro grupo de respostas vincúlase coas actividades que se fan coa roupa. As máis evidentes son aquelas que teñen que ver co feito de poñela e quitala: *abotoar, descalzar, mudarse, poñer enriba, quitar, vestir(se)...* Tamén atopamos un grupo numeroso de palabras que se vinculan coa súa limpeza, mantemento e conservación: *deterxente, encoller, lavadora, planchar, secadora, suavizante, tender...* Finalmente, algúns rexistros atanguen á súa colocación ou ordenación: *armario, colgar, gardarroupa, maleta, percha...*

Á parte dos xa mencionados, eliminamos outros rexistros que non encaixan en ningún dos grupos indicados ou que o fan de maneira moi indirecta: *collóns, crecer, enfermidade, estar mollado, tempo, viaxe...*

Cómpre agora facer algunhas precisións relativas á lematización das formas rexistradas nas respostas dos informantes:

En primeiro lugar, cómpre sinalar que lematizamos as diferentes variantes *sudadera, suadeira, sudadeira...* baixo *suadoiro*, posto que o DCG ofrece *suadoiro* como proposta de tradución de *sudadera* (con este exemplo: “ao terminar a carreira puxo un pantalón de chándal e un suadoiro”). No entanto o DRAG², que, como indicamos, non estaba dispoñible no momento de editar os materiais, non lle asigna a *suadoiro* máis ca o valor de ‘manta que se lles pon ás cabalarías debaixo da sela’.

Nalgunha ocasión a lematización está condicionada claramente polo contexto. É o que sucede, por exemplo, co lema asignado a *cortos* na secuencia: *pantalón-vaqueiros-piratas-cortos*. Aínda que nas respostas dos alumnos figuran diversos adxectivos designadores de dimensións (de xeito que poderíamos lematizar *cortos* baixo *curto*), coidamos que o alumno se está a referir nese contexto a un tipo de pantalón, e deste xeito lematizamos *cortos* baixo *pantalón curto* (escollemos como lema *pantalón curto*, en singular, por ser máis frecuente o rexistro *pantalón curto* ca o rexistro *pantalóns curtos*). No caso dos dous exemplos de *pinzas*, lematizamos baixo o singular *pinza* porque en ningún caso o contexto dá a entender que a palabra faga referencia ao ‘instrumento de metal formado por dúas pezas unidas, que serve para coller cousas miúdas’ (DRAG²), senón máis ben aos obxectos que se empregan para suxeitar o pelo.

Por outro lado, dous dos exemplos de *nike* figuran lematizados baixo *Nike**, como designación da marca comercial, e os restantes baixo *nique*. Esta diferente lematización pode xustificarse igualmente a través do contexto, pois nos dous primeiros casos a palabra vai acompañada doutras marcas comerciais (nos dous casos, o seguinte rexistro é *Adidas*). Tamén lematizamos baixo *Nike** os tres exemplos escritos con maiúscula inicial.

Un novo exemplo da influencia do contexto témolo nun caso que xa foi mencionado ao abordar os criterios de edición dos materiais: lematizamos conxuntamente *gastar cartos* e *gastar*, baixo *gastar (cartos)*, porque o contexto no que figura *gastar* (inmediatamente despois de *comprar*) nos fai entender que o sentido do verbo é o económico, desbotando outras acepcións que tamén serían posibles neste centro. De xeito parello, desvinculamos a forma *tendeira* do concepto ‘muller que traballa nunha tenda’ e entendemos que se trata dunha errata por *tendedeiro* (lematizando a palabra baixo *tendal*). A voz aparece despois dos verbos *lavar, pranchar* e *coser*. Finalmente, lematizamos *pasada* baixo *pasada (de moda)* posto que a palabra figura no seguinte contexto: *fermosa, pasada, moda, seria*.

Tamén pode chamar a atención a distinta lematización de *náuticas* e *náutica*. *Náutica* pode facer referencia a un tipo de camisa, pero tamén pode ser

sinónima de *náutico* e designar un tipo de calzado. O contexto no que figura *náuticas* (antes de *zapatillas deportivas*) e o feito de que se empregue o plural parece conducir á segunda interpretación, e por tal motivo lematizamos baixo a variante máis frecuente *náutico*. Esta interpretación é tamén plausible, por suposto, para o único caso de *náutica*, pero dado que o contexto no que aparece non é revelador e que se recolle en singular, optamos por esco-ller *náutica* como lema.

De por parte, e como xa indicamos, lematizamos en plural as formas constituídas por dúas pezas, que admiten tanto o singular como o plural para designar unha única entidade, e das que temos máis rexistros en plural nas nosas respostas, pero facémolo exclusivamente se algún dicionario menciona a posibilidade do emprego do plural co mesmo valor do singular. Deste modo, atoparemos lemas como *bragas*, *calzas* ou *leotardos*, pero *party*, dado que o único dicionario galego no que atopamos esta forma (o *Digalego*) rexístraa en singular e non fai referencia ao seu uso en plural para designar unha entidade única.

No que atangue aos lemas con sufixos apreciativos, cómpre sinalar que atoparemos o lema *chaquetiña* como forma independente de *chaqueta*, recollendo as variantes *chaquetiña* e *chaquetita*. A pesar de que a voz non figura nos dicionarios consultados, entendemos que é moi probable que os falantes empreguen tal palabra para designar un tipo de *chaqueta* e non simplemente unha chaqueta de pequenas dimensións. Así mesmo, a voz *chaquetilla* ten o seu propio lema, porque esta palabra do castelán se recolle nos dicionarios desta lingua cunha acepción específica (DRAE: ‘chaqueta corta y ajustada, generalmente con adornos’). Por suposto, amais dos citados atoparemos outras formas lexicalizadas e recollidas nos dicionarios: *camiseta*, *camisón*, *chaquetón*...

2.3.4.3 CI 03. As partes da casa (sen mobles)

Este CI amosa un gran número de rexistros que van máis alá do que en principio entenderíamos estritamente por “partes de casa” (ou sexa, *cociña*, *cuarto*, *sala*, *salón*...). No entanto, mantemos a maioría destes vocábulos por formaren parte de ámbitos amplamente representados nas respostas dos alumnos.

En primeiro lugar, debe terse en conta a frecuencia das casas unifamiliares en Galicia, con terreo e construcións anexas ao seu redor. ‘Casa’ identifícase con ‘propiedade’ e isto dá lugar a que sexa moi frecuente atopar nas respostas dos alumnos palabras do tipo de *almacén*, *alpendre*, *canizo*, *capela*, *casa do can*, *coelleira*, *corte*, *cuberto*, *galiñeiro*, *galpón*, *piscina*, *pozo*... Da mesma maneira, rexistramos palabras que designan diferentes tipos de terreo, así

como elementos da natureza, ou outros relacionados cos cultivos ou a gandería: *árbore, camiño, céspede, eido, finca, herba, leira, palleiro, seto, terreo, viña...*

Por outro lado, o concepto 'partes' pode entenderse de maneira moi lasa e así o fixeron os alumnos, que ofreceron como respostas numerosos materiais de construción (*area, argamasa, cemento, escaiola, madeira, mármore, PVC...*), voces relacionadas coas conducións, cos elementos de saneamento e cos sistemas de calefacción e eléctricos (*bañeira, cableado, calefacción, chave da luz, conduto da electricidade, estufa, jacuzzi, sistema eléctrico...*), así como elementos estruturais e arquitectónicos (*armazón, cimentos, estrutura, falso teito, placa, varanda...*). Moi probablemente por asociación con vocábulos como os que se acaban de mencionar, tamén xurdiron nas respostas outras palabras relacionadas coa construción, como as que designan ferramentas (*brocha, cepillo, trade*), traballadores (*albanel, escaiolista, pintor*) ou actividades vinculadas con ela (*arquitectura, construción, edificar, pintar, reformar...*). Deixamos fóra, no entanto, elementos que posúen unha vinculación menos directa coa casa, como *amperio, intensidade* ou *voltio*.

Tamén teñen cabida partes de partes, esencialmente de portas ou ventás, que incluímos na súa totalidade: *bisagra, cadeado, cristal, cristaleira, manilla, marco, pechadura, pomo, xamba...*

A xa mencionada consideración ampla do concepto 'partes da casa' esténdese á presenza de nomes de mobles, malia que na denominación do centro se fai explícita a exclusión destes. Os estudantes amosan un amplo repertorio de tal tipo de entidades e é este o motivo de que os conservemos no CI, malia que exista un centro de interese específico para este ámbito temático (*aparador, armario, banco, estante, estantería, gardarroupa, lacena, librería, mesa, roupeiro, silla, sofá, zapateiro...*). O mesmo ocorre con outros complementos e elementos ornamentais (*alfombra, cenefa, cortina, escobilla, espello, felpudo, maceta, moqueta, póster...*). Nesta mesma liña de inclusión dos elementos que con frecuencia se atopan dentro da casa, recollemos tamén numerosos electrodomésticos (*forno, frigorífico, lavadora, neveira, ordenador, secadora...*). Deixamos fóra, no entanto, algúns utensilios de coíña, moito menos abundantes nas listaxes (*cazarola, cazola, coitelo, culler, garfo...*). Damos tamén cabida á roupa do fogar (*almofada, colcha, toalla...*), pero excluimos o nome de pezas de roupa vestidas polas persoas (*albornoz, bufanda, pantalón, vestido...*).

A casa e as súas partes aparecen recorrentemente caracterizadas nas respostas dos alumnos. Pola súa frecuencia, mantemos aqueles elementos que denotan calidades do ambiente (*acolledor, agradable, comodidade, comfortable,*

familiar, hostil, incómodo, soidade, tranquilidade...), o aspecto da casa ou das súas partes (bonito, feo, fermoso, maxestoso, vistoso), a súa dimensión (alto, ancho, angosto, baixo, enorme...), a súa temperatura (calor, frío, quente), ou a súa luz (claridade, luminoso...). Tamén incluímos adxectivos que designan o estado no que se atopa a casa (*amoblado, baleiro, cheo, desamoblado, desfeito, limpo, ocupado...*), así como o seu tempo de existencia (*novo, vello...*). Deixamos fóra algúns outros adxectivos que se apartan dos grupos mencionados (*branco, coloreado, duro, liso, negro, perezoso, poroso, sentado...*).

Da mesma maneira que se describen as características da casa, tamén se sinalan actividades que se fan prototipicamente nela. Por un lado, atopamos accións que teñen por obxectivo o mantemento, limpeza ou ornamentación da vivenda (*amoblar, aspirar, decorar, decoración, enmoquetar, fregar, limpar, limpeza, ordenar, varrer, ventilar...*), así como algúns substantivos relacionados directamente con tales accións (*lixo, vasoira...*). Por outro lado, rexistramos nas respostas dos alumnos accións que as persoas realizan normalmente dentro das casas: *bañarse, cea, comer, defecar, deitarse, ducharse, durmir, estudar, lavarse, ler, merenda, relaxarse, ver a televisión, xogar...* Finalmente, atopamos algúns exemplos de verbos que teñen a *casa* como complemento habitual: *abrir, entrar, habitar, mudarse, pechar, recoller, vivir...* Pola frecuencia deste tipo de respostas, damos cabida a todas as que encaixan nalgún dos grupos mencionados, deixando fóra outros verbos como *arder, correr, nadar, podar, solagar, vixiar...*

Tamén podemos atopar varias palabras que designan ou cualifican tipos de construción. Mantemos estes vocábulos sempre que o seu referente se poida utilizar como vivenda: *casa, caserío, chabolo, chalé, chamizo, edificio, mansión, pazo, subterráneo, tecnolóxico, zulo...* Excluimos outras construcións que non se empregan para vivir (*merendeiro, museo, teatro*), así como partes de edificios que tampouco están destinados a ese fin (*conserxería, probador, sala de cine...*). Mantemos, no entanto, *recepción*, por amosar sete rexistros e designar, moi probablemente, o recibidor (de feito, no CI 04 hai un exemplo de *mesíña para a recepción*).

Finalmente, excluimos outras palabras cun significado máis dificilmente encaixable nalgún dos grupos mencionados: *aforas, albufeira, can, cidade, coche, libro, merda, mordomo, necesidade, pasta de dentes, pirola, revista...*

Canto a cuestións particulares de lematización, optamos por lematizar *lareiro* baixo *lareira* porque o contexto (*garaxe, habitacion, dormitorio, comedor, cociña, lareiro, soto*) non facía pensar noutra posible interpretación.

No que atangue aos sufixos apreciativos, lematizamos as formas *cuartucho* e *saliña* como formas independentes, posto que entendemos que é posible que

os falantes empreguen tales palabras para designar un tipo de estancia diferente da que designan *sala* e *cuarto* (de feito, no caso de *sala* e *saliña*, varios informantes ofrecen as dúas respostas). Coherentemente, mantemos tamén o diminutivo nas formas complexas *saliña de estar* e *saliña de visitas*. Por suposto, facemos o mesmo con formas complexas cuxo núcleo é un diminutivo claramente lexicalizado, de modo que distinguimos, por exemplo, entre *caseta do can* e *casa do can*, malia que moi probablemente o referente sexa o mesmo.

2.3.4.4 CI 04. Os mobles da casa

Estamos ante un CI que presenta un grao de heteroxeneidade bastante elevado, o que nos levou a facer unha interpretación ampla do estímulo e a incluír, xa que logo, voces de diverso tipo. Entre estas, mantemos algunhas palabras encaixables en grupos que se corresponden con outros CI (electrodomésticos, partes da casa...), pero que están altamente representados nas respostas. En calquera caso, neste CI fomos máis restritivas á hora de conservar voces ca no anterior (CI 03 ‘As partes da casa’). O concepto ‘parte’ pode interpretarse de maneira bastante xenérica, mentres que ‘moble’ ten unha restrición temática moito maior. Isto levounos a aceptar no CI 03 algúns grupos temáticos, concibibles cunha óptica ampla como partes (como pode ser o da roupa do fogar), pero a desbotalos aquí.

Así, ao pé das denominacións de mobles propiamente ditos (*armario, cadeira, mesa, sofá...*), mantemos os nomes de electrodomésticos (*batedor, conxelador, forno, fritideira, lavadora, neveira, tostador, vitrocerámica...*), obxectos de iluminación (*bombilla, candil, flexo, lámpada, lamparilla...*), sanitarios (*bañeira, bidé, ducha, prato de ducha, taza de váter...*) e audiovisuais (*minicadea, radiocasete, tele(visión), vídeo...*). Tamén conservamos palabras que designan elementos decorativos (*cadro, figura, floreiro, fotografía, xarrón...*) e obxectos diversos que se atopan normalmente dentro das casas (*alfombra, arquivador, atril, caixa forte, felpudo...*), pero deixamos fóra as voces que designan roupa de fogar e utensilios de cociña, propios dos CI 02 e 07 respectivamente (*coitelo, colcha, cortina, culler, edredón, manta, mantel, toalla, vaixela, vaso...*).

Por outra banda, incluímos neste CI voces que denotan calidades e características aplicables aos mobles, relativos, por exemplo, ao seu tamaño, estado ou forma (*antigo, astillado, cómodo, con rodas, de madeira, dúas prazas, grande, lacado, moderno, sobrio...*) e palabras relacionadas co seu coidado, limpeza e mantemento (*abrillantar, limpar...*) ou cos usos que lles outorgamos (*decorar, durmir, sentar, tombarse...*). Eliminamos, no entan-

to, outras voces que designan accións vinculadas con entidades que non son propiamente mobles, aínda que tales entidades si se manteñan no CI: *bañarse, ducharse, cociñar...*

Incluimos tamén voces que fan referencia a partes dos mobles (*andel, balda, cabeceira, cabeceiro, cabezal, dosel, módulo, pata* etc.) ou a realidades vinculadas coa súa elaboración (*aluminio, carpinteiro, cravo, cristal, madeira, metal, tornillo, verniz, vidro...*), entre elas varios substantivos designadores de tipo de madeira (*faia, piñeiro, roble...*).

Por outra banda, tendo en conta que as casas unifamiliares son un tipo de vivenda moi frecuente en Galicia, conservamos palabras que designan obxectos que normalmente están fóra delas, como *arrandeadoiro, balancín, churrasqueiro, hamaca* etc. Tamén mantemos algunhas designacións de partes da casa que poden conter mobles como *baño, biblioteca, estudio, recibidor, retrete...*

Finalmente, deixamos fóra algunhas voces dificilmente encaixables nos grupos anteriores ou demasiado xerais: *barco, diñeiro, mirar, quente, romper, vagancia...*

En relación con cuestións concretas de lematización, cómpre sinalar que no caso de *minicadea*, ofrecemos esta forma como lema, posto que aínda que non figura nos dicionarios consultados, si se vén empregando nos discursos en galego, mesmo emanados de institucións oficiais⁴³. O mesmo pode afirmarse da forma *sandwicheira*⁴⁴.

Por outra banda, ao non atopar axuda nos dicionarios, decidimos ofrecer como lemas *lámpada de estudio, lampadiña de estudio, mesa de estudio e silla de estudio* (malia que nas respostas se podían encontrar tanto formas con *estudo* como con *estudio*), porque entendemos que o modificador fai referencia ao lugar no que se atopan os mobles (*estudio*) e non á actividade (*estudo*). No entanto, somos conscientes de que calquera das dúas interpretacións é plausible.

No que atange aos sufixos apreciativos, decidimos ofrecer como lemas separados *lámpada* e *lampadiña*, entendendo que coa forma sufixada se pode facer referencia a un tipo especial de lámpada, unha lámpada de mesa. Aínda que este último valor é moi probablemente o que lle corresponde ao único exemplo de *lamparilla* que figura no CI, decidimos lematizar a dita voz separadamente porque, segundo o DRAE, o substantivo pode facer referencia tamén ao recipiente no que se colocan as mechas que sobre todo

⁴³ Como neste da Xunta de Galicia [consulta realizada o 15 / 04 / 2013]: http://www.xunta.es/dog/Publicados/1997/19970210/Anuncio354A_gl.html

⁴⁴ Véxase, por exemplo, este documento [consulta realizada o 15 / 04 / 2013]: http://traballo.xunta.es/export/sites/default/Biblioteca/Documentos/Publicacions/Perfis_profesionais_camareiro_barra.pdf

antigamente se empregaban para ter luz de noite. Por outra banda, tamén son frecuentes neste CI as formas creadas sobre *mesa* con diferentes sufixos diminutivos. Optamos por lematizar da seguinte maneira:

- i. *Mesa de noite, mesiña de noite, mesilla de noite e mesita de noite* (nalgún rexistro figura *noche* en vez de *noite*) unifícanse baixo *mesa de noite*, dado que en todos estes casos a designación parece ser idéntica. Canto á selección do lema, escollemos *mesa de noite*, aínda que sexa menos frecuente nas respostas ca *mesiña de noite*, dado que ningún dicionario galego recolle este último sintagma como forma lexicalizada, pero si o primeiro.
- ii. Unificamos *mesita* e *mesiña* baixo *mesiña*, e non baixo *mesa*, entendendo que os informantes que ofrecen estas formas designan un tipo de mesa en particular, esencialmente de cuarto ou de sala, e non só unha mesa de pequenas dimensións.
- iii. En coherencia co tratamento anterior, mantemos o diminutivo en lemas como *mesiña dos nenos* e *mesiña da salíña*. Facemos o mesmo en *mesiña da cama* e *mesiña de cabeceira*, aínda que neste caso a designación sería a mesma que tería a forma sen diminutivo. A razón é que non hai rexistros similares sen diminutivo. No entanto, unificamos *mesa auxiliar* e *mesiña auxiliar* (baixo *mesa auxiliar*), pois cremos que o referente de ambas as formas coincide e que o diminutivo unicamente engade, se o fai, matices de tamaño. É o adxectivo *auxiliar* a palabra que determina o tipo de mesa de que se trata.
- iv. Mantemos separada a forma *mesilla*, que segundo os dicionarios de castelán é, na súa primeira acepción, sinónimo de *mesa de noite*. Non o unificamos con *mesita* e *mesiña* posto que estes substantivos non teñen por que designar necesariamente o tipo de mesa que se coloca ao pé das camas. Por outra banda, *mesilla* tamén designa unha parte das escaleiras (sinónimo de *meseta*), feito que determina necesariamente a súa lematización separada.
- v. No entanto, si lematizamos *mesilla da entrada* como *mesiña da entrada*, porque aquí o sentido das palabras co afixo castelán e galego parece ser o mesmo. *Mesilla* non se está empregando co sentido con que figura recollido nos dicionarios de castelán (equivalente, como acabamos de indicar, de *mesa de noite*)⁴⁵.

45 Empregamos o mesmo criterio utilizado no LDEG: neste, *mesita* e *mesilla* figuran como lemas diferentes cando se rexistran como elementos illados, pero unifícanse, baixo as formas con *mesita*, en construcións como *mesilla / mesita auxiliar* ou *mesilla / mesita de teléfono*.

Por outra banda, ofrecemos *secador* como lema separado de *secador do pelo*, posto que entendemos que é moi probable que a palabra *secador* poida tamén designar a máquina que se utiliza para secar a roupa, chamada habitualmente *secadora*. Concluimos isto do feito de que a voz *secador* aparece nunha ocasión precedida de *lavadora* e tamén do feito de que hai exemplos de *secador* no CI 07 ‘A cociña e os seus utensilios’, CI onde a interpretación máis plausible da palabra é a de ‘máquina para secar a roupa’.

Finalmente, cómpre sinalar que é relativamente habitual neste CI atopar formas compostas que non están rexistradas nos dicionarios e que en moitos casos parecen creacións individuais, como *colgatoallas*, *gardalouza*, *gardapratos*, *gardazapatos*, *secalouzas* ou *suxeitatoallas*. De acordo co indicado nos criterios xerais de lematización (§ 2.3.1), todas estas palabras son conservadas nas listaxes.

2.3.4.5 CI 05. Os alimentos e as bebidas

Neste CI, ademais de moitos substantivos e expresións complexas que fan referencia a bebidas (*auga*, *gasosa*, *viño...*), condimentos (*ourego*, *pementa*, *sal...*), alimentos (*carne*, *pan*, *pasta*, *peixe...*) e pratos que se preparan con eles (*carne ao caldeiro*, *cocido*, *patacas con polo...*), encontramos secundariamente algunhas outras palabras que teñen unha vinculación máis ou menos directa co CI e que en xeral mantemos nas listaxes.

Deste xeito, conservamos aquelas palabras e expresións que remiten a calidades, propiedades ou estados dos alimentos ou das bebidas e voces relacionadas, como por exemplo as que aluden á temperatura, o sabor, o estado ou forma de conservación etc. Daquela, no CI poderán atoparse vocábulos como *caducidade*, *delicioso*, *enlatado*, *mofo*, *podre*, *quente*, *refrescante*, *rico*, *saber mal*, *saudable...*

Outro grupo de palabras que se manteñen nas respostas é o das que fan referencia á forma de preparar os alimentos (*aliñar*, *cocer*, *fritir*, *quentar*, *refogado...*), así como aos utensilios empregados nese proceso (*cacharro*, *cazarola*, *tixola...*).

Non podían faltar as voces que remiten ao proceso da alimentación. Deste xeito, encontramos palabras que se vinculan máis ou menos directamente coa actividade de inxerir alimentos e bebidas (*bébedo*, *beber*, *carnívoro*, *comer*, *dixestión*, *empacho*, *morder*, *saborear*, *ter fame*, *ter sede*, *tragar*, *xaxún...*). Inclúense neste grupo os verbos e substantivos que designan as distintas comidas do día (*almorzar*, *cear*, *merenda*, *xantar...*), así como diversos utensilios empregados para beber ou comer (*coitelo*, *cuberto*, *culler*, *prato*, *servilleta*, *vaso...*).

De por parte, cómpre mencionar a existencia dalgunhas palabras que designan lugares e establecementos nos que se atopan alimentos e bebidas (*bar, hipermercado, neveira, panadería, restaurante...*) ou profesionais relacionados coa alimentación (*carniceiro, peixeiro...*), así como unhas poucas formas que se vinculan coa adquisición de comida (*comprar, oferta...*).

Dentro dos vocábulos que fan referencia directa a alimentos e bebidas, cómpre sinalar a elevada frecuencia de nomes de marcas, sobre todo no eido das bebidas. Deste xeito, podemos encontrar nas listaxes *Aquarius**, *Bitter Kas**, *Donuts**, *Gusanitos**, *Licor 43**, *Nestea**, *Tío Pepe**, *Trina**... De acordo cos criterios xerais expostos previamente, mantemos todas estas denominacións nas listaxes.

Por outra banda, chamamos a atención sobre o feito de que nas listaxes se atopan de cando en vez substantivos que designan alimentos que podemos cualificar, desde o noso punto de vista, de “exóticos”. Así, podemos encontrar, entre outras, voces como *golfinho*, *insecto* ou *serpe*. Por outra banda, e dado que no epígrafe do CI non se restrinxe a temática á alimentación humana, decidimos manter palabras, como *alfalfa* ou *bebedoiro*, que en principio se relacionan co mundo animal.

Finalmente, eliminamos algunhas palabras que tiñan unha vinculación moi escasa ou nula coa temática do centro: *bailar, cansazo, excursión, lama, noite, pai, sofá, traballo...*

En relación con cuestións concretas de lematización, cómpre sinalar que optamos por manter *corto* (que figura despois de *café*) na forma do castelán, pois non atopamos recollido en ningún dicionario a substantivación do adxectivo galego *curto* con ese valor. Por outra banda, lematizamos *Xeada* como *Xiada**, entendendo que a informante quería designar a marca de gaseosa (moi consumida no sur de Galicia, de onde é orixinaria a alumna). A palabra precedía a *limonada* e *zume* (a inicial non serve de referencia porque esta informante encabezou todas as palabras con maiúscula). Por último, optamos por lematizar *sapo* e *peixe sapo* separadamente porque o primeiro vocábulo pode designar nalgunhas zonas de Galicia peixes comestibles diferentes á rabada (de acordo coa información ofrecida por Ríos Panisse s.v. *sapo* [consulta realizada a través do *Diccionario de diccionarios* —Santamariña Fernández (2003)]).

No que atangue ao número, cómpre sinalar que son moi frecuentes neste CI as denominacións de pratos que presentan unha forma plural ao estar constituídos moi frecuentemente por varios elementos do mesmo tipo (*canelóns, lentellas, patacas fritas...*). No entanto, tamén é certo que moi frecuentemente a forma singular é perfectamente acaída ao CI, ao designar

directamente o alimento (e non o prato: cf. *macarróns* e *macarrón*) ou ao designar unha porción individual do preparado (de feito, hai exemplos tanto de *canelón* como de *lentella*). Nestes casos, nos que cabía a dobre posibilidade e o plural era maioritario nas respostas, optamos por lematizar con este número só cando algún dos dicionarios consultados introducía unha acepción específica para o plural ou facía referencia á frecuencia do seu uso neste número. Isto é o que explica que, por exemplo, teñamos como lema *lentella*, pero *canelóns*. Como indicamos na introdución (§ 2.3.2.2), somos conscientes de que este tratamento conduce a algunhas incoherencias na lematización das formas.

Por suposto, mantemos o plural cando a única posibilidade de designación é un prato e este se formula necesariamente en tal número: *macarróns á boloñesa* ou *patacas con polo*.

Manténdonos aínda dentro da cuestión do número, gustaríanos facer unha precisión sobre a unificación de determinadas formas singulares e plurais. Nas respostas dos informantes atopamos *orella* e *orellas*, e *pasta* e *pastas*. Poderíamos pensar que o informante que ofrece a voz *pastas* está pensando nas galletas que frecuentemente se toman co té ou co café, mentres que quen escribe *pasta* ten en mente os produtos de diversas formas e tamaños que se fan cunha masa de fariña diluída en auga (*macarróns*, *fideos*...). No entanto, isto é unha simple suposición que non parece cadrar sempre coa realidade. Ao repasar os listaxes comprobamos que moitos dos exemplos da forma *pastas* van precedidos da voz *arroz* ou de substantivos que designan tipos de pasta (*canelóns*, por exemplo). Botando man do outro exemplo, encontramos exemplos de *orellas* en contextos similares a este que segue, que parecen evocar unha parte do porco e non o doce típico do entroido: *grellos*, *repolo*, *orellas*, *uñas de porco*, *carne de vaca*... En ambos os dous casos, unificamos singulares e plurais baixo un lema en singular. O usuario que desexe coñecer o número de rexistros en plural debe acudir ás listaxes de lemas e variantes e debe ter en conta que, de acordo co indicado, parece que non se pode establecer unha relación biunívoca entre acepcións e número.

Si establecemos dous lemas, por suposto, cando os dicionarios recollen a diferenza de significado entre singular e plural. Así, por exemplo, distinguimos entre *cereal* e *cereais*, pois o GDXL asigna esta acepción para o uso plural: ‘alimento preparado coa semente destas plantas [os cereais] que adoita estar enriquecida con vitaminas e outras substancias’.

No que atange ao xénero, e a diferenza do criterio empregado noutros CI, como ‘Os animais’ ou ‘O mar’, decidimos distinguir entre o masculino e o feminino de seres sexuais cando o macho e a femia se valoran de diferente

maneira desde o punto de vista da alimentación. Así, aínda que é probable que a forma *centolo* sexa unha adaptación ao galego do castelán *centollo*, non é menos certo que nalgunhas zonas de Galicia, coa voz *centolo* se fai referencia ao macho da *centola* (tal e como recolle o GDXL), sendo a súa carne peor valorada ca a da femia. Por tal motivo, nestes CI lematizamos separadamente masculino e feminino (mentres que no CI 14 e no CI 18, atoparemos o lema *centolo-a*). Tamén lematizamos separadamente *pita* e *pito*, pero unificamos o único exemplo de *tenreiro* (*terneiro* nas listaxes) baixo *tenreira*.

Por outra banda, outorgámoslles lemas propios aos diminutivos que implican un prato ou alimento (que se considera) distinto ou que se prepara de maneira diferente do que designa a forma sen apreciativo, se esta existe. Velái a razón de que atopemos como lemas diferenciados *choco* e *choquiño*, *pescada* e *pescadilla* ou *pepino* e *pepinillo*. Noutros casos, a forma sen diminutivo nin sequera existe como denominación dun alimento: *bocadillo*, *perrito caliente* etc. Sobre a lematización separada de *culler* e *culleriña*, vid. o comentario ao CI 06.

2.3.4.6 CI 06. Os obxectos colocados na mesa para a comida

Dada a relativa concreción implícita na denominación deste CI, as respostas dos alumnos non alcanzan a diversidade temática que se pode encontrar noutros centros. Malia isto, e ao pé das respostas máis esperadas, isto é, elementos de cubertería (*culler, garfo, coitelo...*) e mantelería (*mantel, pano de mesa, servilleta...*), así como algúns outros utensilios frecuentes —ou posibles— na mesa á hora de comer (*azucreiro, cesta do pan, fonte, saleiro...*), encontramos algúns rexistros que escapan do prototípico, sen perder a súa vinculación co CI, co cal os mantemos nas listaxes.

Deste xeito, conservamos algunhas palabras, nomeadamente verbos, que fan referencia ao acto de comer ou beber (*beber, comer, comida, dixerir, fame, tragar, xantar...*). O acto de comer concíbese en toda a súa extensión, como un acto social que ten lugar ao redor da mesa e que implica unha serie de accións anteriores e posteriores (*sentar, quitamanchas, recollemigas...*). Por tal motivo, conservamos tamén algunhas palabras que, sen estar directamente vinculadas co acto fisiolóxico de comer, si o están co social: *móbil, periódico, televisión...* Non faltan tampouco os nomes de utensilios empregados para cociñar, aínda que non é menos certo que, con frecuencia, estes poden atoparse sobre a mesa á hora de comer: *pota, tarteira, tixola...* Encontramos igualmente designacións de obxectos que se utilizan ou se poden utilizar para servir a comida: *cullerón, espumadeira...*

Tamén conservamos algunhas denominacións de alimentos e condimentos que moi habitualmente se empregan para acompañar ou aderezar certos pratos e que, de maneira xeral, non constitúen unha comida (*aceite, galleta, manteiga, pan, pementa, queixo en po, sacarina, sal, tabasco...*). Do mesmo xeito, mantemos denominacións xenéricas (*acompañamento, condimento, doce, entrante, froita, postre...*), así como os nomes de bebidas (*auga, café, gasosa, viño, zume...*). Pero eliminamos voces máis específicas e algunhas outras palabras que non encaixan nos grupos anteriores (*carne, chocolate, perdiz, plátano, tortilla, turrón, xamón...*) e que teñen o seu lugar natural no CI 05 ‘Os alimentos e as bebidas’.

De igual modo, poderán atoparse nas listaxes algúns adxectivos moi xerais que fan referencia a calidades dos alimentos ou bebidas: *frío, picante, quente, rico, saboroso...*

Suprimimos finalmente outras voces que non manteñen vinculación co CI ou que manteñen unha vinculación moi feble: *carteira, divertido, esvarar, fregar, lavar, limpo, mamá, reunión, sombreiro, tontear...*

No que atangue a aspectos concretos de lematización, cómpre sinalar algunhas cuestións:

O rexistro *coller* encóntrase tras unha serie de verbos e nunha listaxe na que previamente figuraba o termo *culler*. Este é o motivo de que non o consideremos unha variante deste último substantivo e que o lematicemos baixo a forma verbal *coller* (como verbo designador dunha acción característica do acto de comer ou beber). Por outra banda, unificamos *pano de cociña, paño* e *pañe de cociña* baixo *pano de cociña*, posto que en castelán a forma *pañe* só fai referencia ao trapo. Pero deixamos separado *pano* porque con esta voz o informante podería estar designando o pano de mesa. En terceiro lugar, lematizamos *tartera* baixo *tarteira* porque aínda que en castelán a palabra designa un ‘recipiente cerrado herméticamente, que sirve para llevar los guisos fuera de casa o conservarlos en el frigorífico’ (DRAE), a palabra do castelán de Galicia *tartera* é empregada esencialmente para designar un tipo de recipiente para cociñar, isto é, constitúe un calco semántico da voz galega *tarteira*, que posúe tal significado.

É este, por outra banda, un CI con gran número de compostos sintagmáticos e morfolóxicos, algúns deles creados polos informantes (como indicamos en § 2.3.1, conservamos todas estas creacións nas listaxes finais). Con respecto aos compostos morfolóxicos, figura nas listaxes a voz *cortafriames* lematizando a variante *corta fiambres*. Aínda que *cortafriames* non aparece recollido nos dicionarios, optamos por tal lematización porque esta voz vén sendo de uso no galego culto, como o demostra o seu emprego en documen-

tos emanados de organismos oficiais ou no ensino⁴⁶. O mesmo sucede con *pousapratos*, que empregamos como forma lematizadora de *posa platos*⁴⁷.

Por outra banda, neste centro de interese (fronte ao que facemos noutros) unificamos *cesta* e *cesto* baixo o mesmo lema (*cesta*, o máis frecuente) porque, aínda sendo conscientes de que ambas as formas poden convivir nun mesmo falante, designando obxectos de distinto tamaño, o feito de situarnos no contexto ‘Os obxectos colocados na mesa para a comida’ restrinxe as posibilidades designativas das voces a recipientes de tamaño reducido. Ademais, ningún dos informantes ofrece as dúas palabras na súa listaxe. Como apoio á lematización conxunta realizada, nas respostas dos informantes atopamos tanto *cesto do pan* como *cesta do pan*, feito que amosa a posibilidade de que ambos os vocábulos designen a mesma realidade. Tamén unificamos baixo a variante estándar outras formas de igual significado como *aceiteira* e *aceiteiro*, *azucreiro* e *azucreira*, *salseira* e *salseiro*...

Finalmente, e en relación coas palabras con afixos apreciativos, á parte das que están claramente lexicalizadas e figuran xa no dicionario, optamos por ofrecer como lemas separados *culleriña*, *garfíño*, *pratiño*, *taciña* e *tenedorciño*, entendendo que tales formas non só indican un obxecto de tamaño pequeno, senón un tipo especial de obxecto (os empregados na sobremesa, por exemplo). No entanto, unificamos *taciña de café* e *taza de café* porque ambas as formas designan en principio un mesmo tipo de cunca. Por outra banda, ofrecemos como lema *toalliña de limón* e non *toalla de limón* ao considerar que a forma *toalliña* (e non *toalla*) é a habitual para designar o produto empregado para a limpeza das mans.

Con respecto ao número, mantemos separadas as formas *pinza* e *pinzas*, posto que a forma plural pode posuír unha acepción específica (‘instrumento de metal formado por dúas pezas unidas, que serve para coller cousas miúdas’, DRAG²).

2.3.4.7 CI 07. A cociña e os seus utensilios

Neste CI atopamos respostas bastante diversas, na medida en que a cociña é un lugar onde se realizan actividades de diferente tipo.

Así, por un lado, encontramos denominacións de utensilios para cociñar e preparar os alimentos (*cazola*, *pota*, *rodillo*, *sartén*, *tesoiras*...), así como de

⁴⁶ Por exemplo [consulta realizada o 15/04/2013]: http://www.xunta.es/dog/Publicados/2006/20061221/Anuncio253BE_gl.html ou [http://www.vicensvives.es/DIDACLIC/descarga_recursos/docs_comuns/docs_cd/es/primaria/SegundoCiclo/Medio/Medio 3-4 \(Galicia\)/Progr. Aula Medio 4/09 G Medio 4 Gal PA.doc](http://www.vicensvives.es/DIDACLIC/descarga_recursos/docs_comuns/docs_cd/es/primaria/SegundoCiclo/Medio/Medio 3-4 (Galicia)/Progr. Aula Medio 4/09 G Medio 4 Gal PA.doc)

⁴⁷ A voz *pousapratos* recóllese, por exemplo, en [consulta realizada o 15/04/2013]: http://emigracion.xunta.es/files/manual_programa_castellano.pdf

electrodomésticos que empregamos coa mesma finalidade (*batedor, cociña, forno...*). Mantemos tamén partes ou elementos necesarios para o funcionamento destes, como *botón* ou *mando*. Aínda que escapan da cerna do CI, tamén mantemos algunhas palabras, nomeadamente verbos, que designan accións relacionadas coa preparación de alimentos ou outras realidades vinculadas con esta actividade: *aliñar, amasar, cociñeiro, cortar o queixo...*

Outras respostas están vinculadas máis directamente co acto de comer, de tal modo que rexistramos palabras que designan elementos de cubertería (*coitelo, culler, tenedor...*) ou de mantelería (*mantel, pano de mesa...*), en clara confluencia co CI 06. Non faltan tampouco neste grupo as denominacións do mobiliario empregado para comer (*cadeira, mesa, silla...*). Paralelamente ao que fixemos no grupo anterior, mantemos palabras que designan accións ou procesos relacionados co acto de comer ou beber: *beber, comer, merendar...*

En terceiro lugar, a cociña está directamente vinculada, por diversos motivos, coa limpeza e a sucidade, de tal xeito que tampouco faltan denominacións de realidades que se relacionan directamente con este aspecto, desde electrodomésticos (*lavadora, lavalouza...*), utensilios e produtos de limpeza (*estropallo, fregona, Mistol[®], trapo, xabón...*), elementos empregados para non lixarnos (*babeiro, mandil...*) etc.

Moitos dos obxectos mencionados nos grupos anteriores son gardados frecuentemente na cociña, de tal xeito que un cuarto grupo de repostas ten que ver co mobiliario característico deste lugar da casa: *armario de cociña, caixón, estantería...* Non faltan tampouco as denominacións doutros elementos que conforman ou poden conformar a fisionomía das cociñas galegas: *baldosa, cadro, calendario, chaveiro, cheminea, cinceiro, encimeira, flor, granito, lareira, televisión...*

Tamén atopamos un considerable número de palabras que designan alimentos. Optamos por manter as denominacións xenéricas (*bebida, conserva, froita...*) así como os substantivos máis frecuentes nas listaxes, con máis de cinco respostas en total (*aceite, azucre, pan...*). Seguindo este criterio, suprimimos un considerable número de palabras que facían referencia a alimentos, condimentos etc., pero que figuraban moi poucas veces nas respostas (*carne, cebola, garavanzo, kiwi, macela...*). Tamén eliminamos algunhas palabras que designaban estados ou calidades dos alimentos, como *adobado, cociñado, doce, frito...* e outras que escapaban en demasía dos intereses do CI ou eran demasiado xerais: *atrapar, barato, esixir, espada, falar, home, pedir...*

No que atangue ás particularidades de lematización, podemos atopar neste CI o composto sintagmático *cortafriames*, que xa foi comentado ao tratar

o CI 06 (*vid.* tamén o indicado alí para *pano*, *pañó* e *pañó de cociña*, así como o sinalado para *tartera*). A este último respecto, cómpre sinalar que neste CI a forma *tartera* en todos os casos vai acompañada nas listaxes doutras denominacións de utensilios para cociñar. Por outra banda, lematizamos *secador* baixo *secadora*, posto que neste CI a interpretación da primeira palabra como ‘electrodoméstico para secar o pelo’ non parece apropiada. En relación con *minicadea* e *sandwicheira*, *vid.* o indicado no comentario ao CI 04.

Como sucede no CI 06, no 07 tamén se evidencia con claridade a vacilación na selección do xénero de determinados substantivos. Unificamos baixo a forma recollida nos dicionarios *azucreiro* e *azucreira* (baixo *azucreiro*), *filloeira* e *filloeiro* (*filloeira*); *froiteira* e *froiteiro* (*froiteiro*); *licuador* e *licuadora* (*licuadora*)... Mesmo no caso de que só haxa exemplos dunha variante non estándar, seguindo o criterio xeral, escollemos a forma estándar como lema: así, por exemplo, *amasador* lematízase baixo *amasadora*. Por suposto, mantemos separadas as formas cando designan obxectos diferentes: *caldeira* e *caldeiro*, *cazola* e *cazolo*, *cunco* e *cunca*, *maza* e *mazo*...

No que atangue aos afixos apreciativos, tal e como fixemos no CI 06, lematizamos as formas *culleriña* e *pratiño* como formas independentes de *culler* e *prato*, entendendo que designan obxectos específicos.

2.3.4.8 CI 08. A escola: mobles e materiais

Neste CI, ademais dos mobles (*mesa*, *cadeira*, *pupitre*, *encerado*...) e materiais de moi diferente tipo (*balón*, *bolígrafo*, *frauta*, *goma*, *lapis*, *libro*, *microscopio*...) aos que se alude especificamente no título, podemos atopar bastantes expresións vinculadas co mundo da escola, pero non encaixables en ningún dos dous subgrupos anteriores.

Polo considerable número de respostas rexistradas, decidimos manter os nomes de materias ou disciplinas que os alumnos cursan ao longo da súa vida académica (*debuxo*, *educación física*, *galego*, *xeografía*...). Tamén conservamos unha serie de voces específicas de certas materias, como as que designan realidades lingüísticas (*frase*, *oración*, *palabra*...) ou matemáticas (*ecuación*, *multiplicar*, *resta*...).

Son altamente frecuentes os substantivos que fan referencia a dependencias e anexos dos centros escolares. Todos eles son conservados: *comedor*, *sala de ordenadores*, *urinario*, *vestíbulo*, *ximnasio*... Mesmo mantemos tamén a especificación das aulas por materias (*aula de electrónica*, *aula de mecánica*, *aula de música*...). De igual modo, respectamos as voces que designan os elementos constitutivos das dependencias citadas, entre eles os ornamentais: *alarma*, *azulexo*, *cadro*, *ladrillo*, *ventá*, *xanela*...

Un grupo moi numeroso, que mantemos, é o conformado polas voces que fan referencia aos individuos que constitúen a comunidade escolar desde o punto de vista da súa función nela: *alumno, director, profesor, titor, xefe de estudos...* Cómpre citar tamén aquí os colectivos correspondentes (*alumna-do, profesorado, xunta*) e os substantivos designadores das tarefas realizadas polos distintos membros (*orientación, tutoría...*). Tamén se inclúen as palabras que designan persoas encargadas dos servizos no colexio (*camareiro, señora da limpeza...*), voces que fan referencia ás relacións que se establecen entre os alumnos (*amigo, amizade, compañeiro...*) ou aos distintos “tipos” de alumnos (*chapón, gamberro, macarra...*).

Tamén son relativamente frecuentes, e tamén conservamos, as palabras que designan diferentes actividades propias da vida no centro escolar ou determinadas por esta: *amoestación, atender, castigar, conversar, madrugar, ordenar, pensar, traballar* etc. Finalmente, como é habitual nas listaxes de léxico dispoñible, os informantes mostran os seus sentimentos e valoracións, desta volta sobre o traballo ou a vida no centro escolar. Seguindo a pauta xeral nos outros CI mantemos estas formas, entre as que podemos atopar *aburrimiento, agradable, cansazo, chorar, difícil, insoportable, pasalo mal, risa* ou *tristeza*.

En relación coas palabras que se sitúan baixo un dos subgrupos polos que se preguntaba, os materiais, cómpre facer algunha matización. No caso da roupa, mantemos aquelas voces que designan pezas que se poden empregar especificamente para realizar determinada actividade escolar (*bañador, bata, chándal...*), ademais, por suposto, de *uniforme*. Excluímos, no entanto, as que teñen un valor xenérico (*abrigo, falda, xersei...*) e que teñen o seu lugar propio no CI 02 ‘A roupa’.

Por outra banda, mantemos diferenciados os dicionarios especificados por linguas (*dicionario de castelán, dicionario de galego, dicionario de inglés...*) e os libros por materias (*libro de artes, libro de ciencias...*). No que atange aos bolígrafos tamén conservamos as diferenzas en función da cor con que escriben (*bolígrafo negro, bolígrafo vermello...*), pero eliminamos outras mencións dos adxectivos de cor nas que estes se empregan para encher as listaxes (*tinta, azul, negro, vermello; escribir, azul, vermello, negro, verde*).

Do mesmo xeito, eliminamos outras formas que moi dificilmente encaixan na temática do CI ou que teñen un valor moi xenérico (*cu, desayuno, gordo, grande, noite, roto, vello, xente...*). Neste grupo, poden incluírse palabras que designan ou están vinculadas con actividades que, realizándose con certeza nos centros escolares, non son específicas delas (*golosina, ligar, mexar, tabaco...*). Probablemente a presenza na listaxe dalgunha das formas eliminadas responda á idiosincrasia dos distintos centros educativos dos in-

formantes ou do momento no que se fixeron as enquisas (*árbore de Nadal, cadela, faiado, garaxe...*).

En referencia á lematización, cómpre xustificar algunhas das decisións tomadas:

Optamos por lematizar *cañón* baixo *canón*, *cañón de luz* baixo *canón de luz* e *cañón de proxectar* baixo *canón de proxectar*, posto que, aínda que a palabra *canón* non se recolle especificamente nos dicionarios xerais consultados co valor implícito nesas expresións, si figura co dito valor en vocabularios específicos (por exemplo, en Pazó Fernández 2005). Ademais, noutras linguas como o francés e o portugués atopamos no mesmo contexto o cognado correspondente. No que respecta a *minicadea*, *vid.* o comentado no CI 04.

No caso do rexistro *libro de Lengua* optamos por lematizalo como *libro de lingua* (e non *libro de lingua*) entendendo que se está facendo referencia ao libro empregado na materia “lingua castellana”. Nótese que o rexistro figura despois de *libro de galego* e que é habitual chamar xenericamente *lingua* á materia lingua castellana e *galego* ou *lingua* (aínda falando en castelán) á materia lingua galega. Deste xeito, o lema *libro de lingua* sería chocante e pouco transparente. Polas mesmas razóns que acabamos de aducir, lematizamos *lingua* baixo *lingua* e non baixo *lingua*.

Por outra banda, ás veces son as formas do contexto circundante as que xustifican a escolla de lema: así, lematizamos *cutre* baixo *cúter* porque o contexto non dá pé á interpretación da palabra como adxectivo: *pintura, sillas, plastilina, papeleira, cutre, cera*. De igual xeito, lematizamos *lentes* baixo *lentes*, en plural, porque o contexto tampouco dá a entender que o informante se refira ás lentes usadas no laboratorio, senón que probablemente estea a designar os anteollos: *bolígrafos, estuches, lentes, lápices*.

Por outra banda, lematizamos separadamente *televisor* e *televisión* porque neste CI (fronte ao que sucede nos CI 03, 04, 06 e 07), a palabra *televisión* pode estar facendo referencia ao aparello eléctrico, pero tamén ao medio como un instrumento para o ensino / aprendizaxe. Sobre a lematización separada de *bus* e *autobús*, *vid.* o indicado no comentario ao CI 12.

Canto á representación gráfica das denominacións de materias, estas formas son lematizadas con minúscula inicial, pois, tal e como indicamos en § 2.3.2.9, restrinximos o emprego de maiúscula inicial aos substantivos que designan estudos que conducen á obtención dun título. Deste xeito, lematizaremos *Química* no CI 16 ‘Os oficios e as profesións’, pois está designando a carreira, pero *química* no CI 8, pois está designando a materia.

No que atange ao número, mantemos como lemas diferentes *estudo* (a actividade) e *estudos* (‘conxunto de materias que se aprenden, e actividades

que se realizan durante a formación nun centro de ensino', DRAG²). En relación cos afixos apreciativos, conservámoslos no lema en casos como *arma-riño*, *caderniño*, *carríño* e *mesíña*, entendendo que nas catro palabras hai cando menos certo grao de lexicalización.

2.3.4.9 CI 09. A calefacción e a iluminación

Ademais dos substantivos que designan obxectos que xeran luz ou calor (*bombilla*, *calefactor*, *candil*, *estufa*, *farol*, *fluorescente*, *foco*, *halóxeno*, *lámpada*, *lanterna*, *radiador* etc.), conservamos outros vocábulos que se poden situar en diferentes grupos.

En primeiro lugar, cómpre destacar as voces que designan substancias, materias etc. empregadas para iluminar e quentar, como *aceite*, *aire*, *cera*, *electricidade*, *enerxía*, *fuel*, *gasolina*, *gas*, *leña*, *lume*, *luz natural*, *luz solar*, *petróleo*, *sol* etc. (incluídos os medios de uso puramente individual: *calor humana*, *cama*, *ducha quente*, *fregas*). Conservamos igualmente as voces que designan pezas de vestir ou roupa de cama (*bata*, *cazadora*, *chaqueta*, *guante*, *roupa*, *roupa de abrigo*; *manta eléctrica*, *saba de flanela*...), así como espazos físicos asociados á calor (*casa*, *fogar*, *sauna*, *vivenda*...).

Por outra banda, tamén conservamos nomes de aparellos, de pezas, de ferramentas etc. que forman parte dos diferentes utensilios empregados para quentarnos ou iluminarnos, que se utilizan na súa instalación ou que dunha maneira máis ou menos directa están relacionados con eles: *abrazadeira*, *alicate*, *arame*, *botón*, *buscapolos*, *cable*, *cargador*, *chave inglesa*, *cronómetro*, *falso teito*, *fusible*, *interruptor*, *manguito*, *portalámpada*, *regulador*, *roda*, *válvula* etc. Rexistramos e mantemos tamén outro grupo de palabras que fan referencia ao funcionamento ou regulación dos aparellos (*acender*, *apagar*, *aumentar*, *funcionar*, *estalido*, *regular*...) ou aos efectos derivados da luz ou da calor (*abrigar*, *cegar*, *iluminar*, *ver*...). Tamén podemos atopar na listaxe final algúns vocábulos xenéricos relacionados coa calor e a luz, como *calor*, *claridade*, *cor*, *frío*, *luz*, *temperatura* etc., así como certos nomes comerciais (esencialmente empresas vinculadas coa calefacción ou a iluminación): *Cep-sa**, *Gusy Luz**, *Iberdrola**, *Petromás**, *Repsol**, *Unión Fenosa**... Asemade, poden atoparse expresións valorativas, vinculadas moi probablemente á sensación de calor. Así, conservamos formas como *acolledor*, *benestar* ou *pracer* no sentido positivo ou *agobiante* e *agobio* no negativo.

Son igualmente bastante frecuentes as voces relacionadas de maneira xeral coa enerxía, tanto aquelas que fan referencia a unidades de medida (*ampere*, *volt*, *watt*...) como ao seu consumo. Neste último grupo temáti-

co, cómpre destacar a presenza dun bo número de expresións valorativas (*aforrar, consumismo, consumo, dano, dependencia, derroche, despilfarro, destrución, factura, investimento, prexudicial, progreso, seguridade...*). Hai que salientar finalmente que algúns informantes amosan un amplo dominio do vocabulario específico relacionado coa enerxía, e moi en particular coa electricidade (*circuíto paralelo, condensador, conmutador, convector, disipador, fotocélula, panel fotovoltaico etc.*).

Fronte ás formas anteriores, desbotamos outras que escapan en exceso do núcleo temático ou que nada teñen que ver con el. Velaquí unha mostra: *altavoz, avó, cliente, coche, colexio, discoteca, fábrica, fame, gripe, ler, soidade, soldo, tranquilidade, vello, xuntanza...* Entre elas encontramos algunhas ocorrencias dos informantes como *libro de filosofía* ou a xa comentada (*vid. § 2.3.1) acende o candil, Marica chus chus*.

No que afecta a aspectos concretos de lematización neste CI, cómpre xustificar algunha das decisións tomadas (para *canón* e *secador / secadora*, véxase o indicado no comentario ao CI 08 e ao CI 04, respectivamente).

Con respecto ás formas en plural, conservamos neste número *lucos de Nadal*, entendendo que os informantes se están a referir á iluminación que se pode atopar nesta época nas vilas e cidades de Galicia. De igual modo, mantemos *lucos de festa*. Por outra banda, se se consultan as listaxes de lemas e variantes, atopárase un exemplo de *plomo*, en singular, lematizado como *chumbo* e dous casos do mesmo rexistro lematizados como *plomos*. A explicación atópase no feito de que só no primeiro exemplo parece que o falante se está a referir ao material, mentres que nos outros dous máis probablemente está aludindo aos fusibles. O contexto para o primeiro é *gas, CO₂, plomo, tubería* e para os dous restantes, estoutros: *plomo, bombilla, fundir, apagado; electricidade, chave, enchufe, plomo, candeia*.

En relación cos afixos apreciativos, consideramos formas lexicalizadas *farolillo, lampadiña, lampadiña de estudio* ou *lamparilla* (sobre *lampadiña* e *lamparilla vid. o indicado no comentario ao CI 04*). Por outra banda, *lámpada de mesita* lematízase baixo *lámpada de mesiña* e *lámpada de mesilla* baixo *lámpada de mesilla*. Véxase tamén o comentario ao CI 04 para unha xustificación deste tratamento, así como para a explicación da escolla de *estudio* e non *estudo* nos compostos *lámpada de estudio* e *lampadiña de estudio*.

Por outra banda, son relativamente frecuentes neste CI as siglas (dentro da escaseza xeral no conxunto das listaxes): *LDR, NTC, PTC, transistor PNP...* A presenza destas formas mostra o coñecemento de vocabulario específico de electrónica por parte dalgúns informantes ao que faciamos referencia máis arriba.

2.3.4.10 CI 10. A cidade

A causa do alto grao de heteroxeneidade que presenta este CI, fixemos unha interpretación ampla do estímulo e nel incluímos voces de moi diverso tipo. Poden atoparse nas listaxes denominacións de elementos característicos das cidades, como *barrio, beirarrúa, catedral, edificio, facultade, hospital, estación, pensión, rañaceos...* No entanto, tamén demos entrada a outras moitas formas que, sen designaren realidades específicas das urbes, si se poden atopar frecuentemente nelas: *balcón, basura, gaivota, igrexa...*

En conxunto, son frecuentes as palabras que designan establecementos e edificios (*adosado, cafetería, casa do concello, colexio, librería, restaurante, taller, xoguetería...*), así como as que fan referencia ás persoas que traballan neses lugares. Deste xeito, encontramos e conservamos denominacións de profesións e oficios (*bombeiro, dentista, dependente, médico, panadeiro, taxista, xardineiro...*), así como de cargos públicos (*alcalde, concelleiro, conselleiro, ministro, presidente...*). Tamén atopamos e conservamos algúns nomes de establecementos comerciais: *Alcampo**, *Carrefour**, *Zara**... Non faltan tampouco as palabras, sobre todo verbos, que designan accións vinculadas co traballo que se desenvolve nas cidades, desde a perspectiva do traballador ou do cliente ou beneficiario: *comercialización, comprar, construír, consumir, edificar, traballar* etc.

Outro grupo numeroso de palabras é o constituído polas que se vinculan cos medios de transporte e a circulación de persoas e vehículos. Deste xeito, e malia a confluencia co CI 12 ‘Os medios de transporte’, conservamos as numerosas palabras que designan vehículos (*ambulancia, autobús, bicicleta, camión, furgoneta, metro, taxi, tranvía...*). Por suposto, facemos outro tanto coas voces que fan referencia á circulación, desde varias perspectivas (*atasco, beirarrúa, peón, rúa, semáforo, sinal...*). A este respecto, cómpre citar a presenza dalgunhas palabras que designan accións relacionadas con este subgrupo temático (*aparcar, atropelo, camiñar, conducir...*).

Como cabía esperar, son tamén frecuentes as voces que designan as diferentes realidades, máis alá dos edificios, que conforman a fisionomía das cidades, como o mobiliario urbano (*banco, cabina, fonte, papeleira...*) ou as vinculadas coa distribución e organización do territorio ou cos diversos espazos que conforman as urbes (*barrio, parque, paseo marítimo, praza...*). Como pode observarse na listaxe anterior, algunhas das voces rexistradas amosan o carácter marítimo dalgunhas das cidades galegas. Deste xeito, encontramos tamén, e conservamos, formas como *barca, peirao, porto* ou *praia*, pero eliminamos outras como *acantilado, chalana, nasa* ou *rocha*.

Non faltan tampouco certas palabras que designan substancias ou materiais relacionados con algunhas das facetas da cidade arriba mencionadas: *alcátrán, asfalto, butano, hormigón, gas, gasolina, pedra...*

Centrándonos agora nas palabras que fan referencia a seres humanos, ademais das xa citadas, cómpre sinalar que encontramos e mantemos denominacións de xentilicios (*árabe, chinés, marroquí...*) e de tribos urbanas ou tipos de individuos que poboan as cidades (*hippy, macarra, pijo-a...*). Tamén conservamos as denominacións dalgúns animais que son característicos das poboacións urbanas, como poden ser *can, paxaro* ou *pomba*, pero eliminamos outros de difícil encaixe, como *ovella*.

Non faltan tampouco adxectivos que denotan calidades aplicables ás urbes (*antigo, gris, industrial, turístico, vello* etc.). Cómpre sinalar, unha vez máis, a presenza de termos valorativos, moitos deles substantivos que designan realidades abstractas que os informantes atribúen ás cidades. Mantemos todos eles: *aburrido, agobio, corrupción, crispación, delincuencia, droga, estrés, liberdade, perigoso, problema, progreso, violencia...*

Por outra banda, son relativamente frecuentes as voces que aluden á contaminación: *contaminar* (e derivados), *dano ecolóxico, merda subterránea, monóxido de carbono, polución...* Neste aspecto, e noutros, os informantes establecen unha comparación con outro tipo de vida, a desenvolvida no campo. Así atopamos rexistros como *máis contaminación, máis estrés, máis liberdade, máis ruído, máis traballo* (sobre a lematización destas expresións *vid.* § 2.3.2.7.1). Como se pode observar, aquí e no parágrafo precedente, a valoración é tanto positiva como negativa, aínda que probablemente predomine a segunda sobre a primeira.

Finalmente, os nosos informantes proporcionan algunhas palabras que non encaixan neste CI e que, polo tanto, decidimos eliminar: ademais das que xa mencionamos, suprimimos *calzado, traxe, zapatos* (máis propias do CI 'A roupa'); *gandería, prado, tractor* (máis propias do CI 'O campo'), así como algunhas outras palabras cun vínculo moi feble coa temática da cidade: *bailar, estafador, hípica, inundación, reunión, tumbona, verán* etc. Por outra banda, eliminamos nomes propios con poucas ocorrencias (cidades, vilas, rúas, monumentos...): *Barcelona, Madrid, río Lérez, Vigo...*

En relación con aspectos concretos de lematización, cómpre sinalar tres casos nos que botamos man do contexto para asegurar a correcta interpretación das formas rexistradas polos informantes. En primeiro lugar, lematizamos *basoira* e *vasoiras* baixo *basura* (e *vasoiras de reciclaxe* baixo *basura de reciclaxe*). Ningún dos contextos nos que aparecen estas formas é propicio para a aparición de *vasoira* no seu sentido habitual ('instrumento

para varrer’), sentido que é bastante alleo á temática do CI. Por outra banda, algúns dos contextos parecen claramente indicar que se está a pensar no valor de ‘lixo’ (ademais da expresión *vasoiras de reciclaxe*, atopamos unha serie como esta: *basoira, contenedores, basureiros*). Finalmente, fóra das listaxes, o contido ‘lixo’ é atribuído con frecuencia á palabra *vasoira*⁴⁸ (o feito de que, como *basura*, a palabra se escriba adoito con , responde tamén probablemente á confusión establecida)⁴⁹.

En segundo lugar, lematizamos *tendedeiro* baixo *tendeiro*, pois a forma figura no contexto *comercio, tenda, tendedeiro*. Por último, incluímos *venta* baixo *ventá* en dous casos e baixo *venda* noutros dous. Os exemplos figuran nas series *mesa, venta, edificio // edificio, venta* (lema *ventá*) e nestas outras: *librerías, quiscos, ventas // ventas, tendas* (lema *venda*).

Para a xustificación da lematización separada de *bus* e *autobús* e de *moto* e *motocicleta*, *vid.* o comentario ao CI 12 ‘Os medios de transporte’.

Con respecto ás variantes flexivas, cómpre xustificar algunha das decisións tomadas. En relación coa diferenciación entre singular e plural, ademais de *galería / galerías* ou *festa / festas*, con clara distinción semántica, tamén separamos as formas *gas* e *gases*, aínda que os dicionarios non recollen unha acepción específica para o plural. Consideramos que *gas* pode facer referencia ao elemento enerxético, mentres que con *gases* se alude á presenza de contaminación na cidade (de feito, tamén se recollen *gases tóxicos* e *gases contaminantes*). Por outro lado, mantemos separados *buzón de correo* e *buzón de correos*, porque entendemos que no primeiro caso o informante pode estar referíndose á caixa do correo que hai nas casas e no segundo á que podemos atopar na rúa.

Canto ao xénero, non unificamos *xardineira* e *xardineiro* mediante *xardineiro-a*, dado que coa forma feminina probablemente o informante estea facendo referencia ao recipiente no que se cultivan as plantas. Por razóns semellantes mantemos separados *caixeira* e *caixeiro*, pois esta última voz pode estar designando ao *caixeiro automático*. Para *mula / mulo* *vid.* o comentario ao CI 14.

⁴⁸ Non nos foi difícil atopar nos buscadores de internet algúns exemplos de “recoller a vasoira”. Velaquí simplemente dous casos (os destacados son nosos): “Eu penso que sería mellor que *recolleran a basoira* epacial dalgunha maneira porque senon acabara destruindo o espazo” (*sic*) (<http://oblogde6.blogspot.com.es/2011/09/basura-espacial.html>); “Non podo dicir o que pago porque me veñan a *recoller a basoira* (é máis, teño a obriga de levala ós contedores estratéxicamente sitiados, ou situados,...) (<http://pontedeume-fala.blogspot.com.es/2007/05/relaxamiento-mental-antes-de-ir-durmir.html>) [consultas realizadas o 15 / 04 / 2013].

⁴⁹ Os exemplos aparecen nas seguintes series: *marquesinas, peóns, vasoiras, pasos de peóns // iluminación, basoira, andamios // choiva, alcantarillado, basoira // humo, basoira, contenedores, basureiros // farolas, vasoiras, vasoiras de reciclaxe, coches // carteis, basoira, casas*.

2.3.4.11 CI 11. O campo

Dada a amplitude temática deste CI, non debe estrañar que nel se recollan palabras designadoras de realidades de moi diferente tipo. Pero máis alá deste feito, o certo é que podemos atopar distintas visións do ámbito temático, con tres eixos principais. Por unha banda, rexistramos unidades relacionadas directamente co campo como elemento de explotación humana, onde a agricultura e a gandería son os focos fundamentais. Neste sentido, podemos atopar elementos comúns co CI 13 ‘Os traballos do campo e xardín’: animais de compañía e domésticos (*can, boi, galiña, mula, touro...*) e vocábulos relacionados (*bosta, bovino, cacarexo, carne, celo, la...*), plantas, árbores, flores ou froitos cultivados polo home, con algunhas das súas partes (*abelá, hortensia, millo, pétalo, tomate...*), tarefas e o seu resultado (*abonar, acequia, alimentar as vacas, aramado, paca, recoller, regar...*) ou máquinas (mesmo marcas: *Bertolini, John Deere...*), ferramentas, aparellos, instrumentos, materiais, medidas, recipientes, roupa etc. empregados en tales tarefas (*abonadora, abono, aixada, albarda, arame, barril, bota, botadora, cavón, cortasebes, empacadora, esterco, sombreiro de palla, tanque de leite, tractor...*). Tamén podemos atopar elementos designadores de espazos, vías e lugares característicos do traballo, sobre todo terreos e construcións (*agro, aira, alboio, bodega, cabaceiro, carreiro, leira, terreo, veiga...*), así como sistemas de cultivo ou de explotación agrícola ou gandeira (*minifundio, transhumancia...*).

Son tamén frecuentes as voces que fan referencia ao tipo de actividade desenvolvido no campo (*agrario, agrícola, agricultura, gandería, reforestación...*) e á explotación comercial dos produtos (*autoconsumo, cooperativa, explotación, produción...*), así como ás persoas que levan a cabo as correspondentes tarefas (*agricultor, apicultor, granxeiro, pastor...*) ou que se vinculan dalgunha maneira con elas (*agrónomo*).

A visión do campo, entendida no sentido indicado, esténdese á contemplación das poboacións, de carácter rural, nas que normalmente viven as persoas que se dedican ás tarefas agrícolas e gandeiras. Deste modo, encontramos elementos que nos permiten describir as *aldeas* ou *vilas* (dous vocábulos que tamén figuran rexistrados nas respostas, onda *concello, pobo, poboación...*) e a vida que se desenvolve nelas, isto é, habitantes, comidas, construcións, espazos, vías de comunicación, ocio ou comercio e servizos: *alcalde, aldeán, andar a rolos, artesán, avó, basureiro, caldo, cacique, campo da festa, canteiro, capela, casa, casa da saúde, cemiterio, comer ben, cruceiro, festa, igrexa, estrada comarcal, tenda, verbena, xente vella, xogar...*

En segundo lugar, o campo pode entenderse nun sentido máis amplo, case como equivalente de natureza, de modo que atopamos nomes de ani-

mais de moi diferente tipo, os característicos do hábitat galego (*abella, cacacol, londra, lucecú, rato de campo, saltón, xílgaro...*), pero tamén algúns propios de ecosistemas diferentes ao noso (*bisonte, cebra, xirafa...*). O mesmo sucede cos nomes de árbores, plantas, flores e colectivos correspondentes (*alecrín, arboredo, bidueiro, bidueiral, landra, piñeiro, piorno...*, pero tamén *banana* ou *pistacho*). Por suposto, aparecen ademais voces xenéricas como *animal, bicho, árbore* ou *planta*. Dentro da mesma interpretación xeral, podemos situar algúns vocábulos designadores de formacións xeolóxicas, vías ou zonas naturais, cursos de auga etc. (*barranco, beira do río, berrocal, cavorco, cerro testemuña, río, montaña, val...*) e outras expresións que fan referencia a fenómenos ou elementos da natureza, ou calidades relacionadas con eles, que frecuentemente son máis puros ou “naturais” no campo, ou que nel se perciben de forma máis clara ou directa: *aire* (adoito modificado: *aire puro, aire fresco...*), *arco iris, auga, brisa, calor, ceo, xeadada...* Non faltan tampouco as voces que sinalan as cores da natureza (*amarelo, azul, colorido, cor, marrón, verde...*) ou as que indican as súas dimensións (*amplitude*).

En terceiro lugar, atopamos o campo enfocado como un lugar de descanso, retiro ou ocio, de modo que podemos encontrar diversas palabras que fan referencia ás actividades que se realizan nel con algúns dos fins indicados, así como diversos vocábulos relacionados: *acampada, albergue, bañador, bañarse, bar, barbacoa, bocadillo, cacería, campamento, camping, casa de turismo rural, casa de vacacións, merendeiro, motocrós, pínic...*

Secundariamente rexistramos unha cuarta interpretación. Debido ao feito de que moitos deportes se desenvolven en *campos*, atopamos tamén palabras como *baloncesto, campo de baloncesto, campo de béisbol, campo de fútbol* (cinco ocorrencias), *fútbol* (tamén con cinco rexistros), *terreo de xogo...*

Non podían faltar, por outra banda, as valoracións, directas ou indirectas e tanto positivas como negativas, as cales poden relacionarse con varios factores, estéticos, anímicos, relativos á saúde, á hixiene e ao benestar, económicos etc.: *acolledor, agarimoso, agradable, alegría, amabilidade, antigüidade, beleza, benestar, bo olor, boa respiración, bonito, bucólico, descanso, saúde, tranquilidade...* fronte a *aburrimento, desesperanza, malestar, malvivir, pobreza, pouca limpeza, subdesenrolo, xente non civilizada...*

Por outro lado, e en relación co que acabamos de sinalar, en varias das respostas pode talvez apreciarse o tipo de relación que os informantes man-

teñen ou mantiveron co campo. Non escasean aqueles rexistros que amosan a presenza de vínculos afectivos como *amigos, amigos da infancia, amizade, amor, compañeiro...* Outros rexistros inciden na relación de proximidade ou afastamento espacial (*afastado, alonxado, aforas, apartado, proximidade...*). O mesmo que sucedía no CI 10 ‘A cidade’, é tamén frecuente a comparación implícita con esta, habitualmente a través de elementos como *máis, menos, mellor...*: *menos contaminación, mellor relación, máis comunicación cos veciños, non semáforos...*

Quedan fóra algúns rexistros demasiado xenéricos, ou que non encaixan neste ámbito temático ou que o fan moi indirectamente: *apelido, coche fúnebre, comunión, conxelar, pantalón, parede, xoguete, zapatilla...* Tamén suprimimos algúns nomes propios (moito menos frecuentes ca no CI 10 ‘A cidade’): *Miño* ou *Vilarchán*.

No que atangue a aspectos concretos de lematización, cómpre facer algunhas precisións:

En cinco casos, apelamos ao contexto para xustificar a decisión tomada: en primeiro lugar, lematizamos *acha* como *hacha*, posto que aínda que a palabra galega *acha* posúe significados acaído ao centro (por exemplo, ‘anaco pequeno e irregular de madeira que salta cando esta se parte ou rompe violentamente’, DRAG²), o contexto no que figura o rexistro, ao pé de nomes de ferramentas, varias delas na forma castelá ou adaptadas —*foz, fouciño, acha, rastrillo*—, parece indicar que estamos ante un exemplo da palabra do castelán *hacha*, que tamén rexistramos noutro informante (desta volta, xa con <h>). En segundo lugar, agrupamos *carballal* con *carballeira* (baixo a forma máis frecuente, *carballeira*), porque o contexto (*finca, chouso, carballal, souto, veiga*) fainos pensar nun lugar plantado de carballos e non nun tipo de pera, a outra acepción da palabra, tamén acaída a este CI. Con respecto a *maceiro*, lematizamos esta palabra baixo *maceira*, porque aínda que a voz masculina pode designar tamén un paxaro, o contexto (nun caso tras *mazá* e noutro antes dunha serie de nomes de árbores), parece dar a entender que os informantes estaban pensando nas froiteiras antes ca na ave. Por outro lado, non podemos esquecer que *maceiro* é variante dialectal empregada para designar a especie arbórea. No caso de *pola*, esta forma figura nunha listaxe despois de *abelá* e noutro antes de *viña*, de tal xeito que interpretamos que o informante quere facer referencia ás ramas das árbores. Finalmente, aínda que a palabra *toro* designa en galego unha parte das árbores, lematizamos *toro* como *touro* porque o contexto no que aparece a voz (en once ocasións), ao pé doutros nomes de animais, dá a entender que os informantes teñen en mente o macho da vaca e non o tronco das árbores.

Por outra banda, aínda que non atopamos a forma *cerro testemuña* rexistrada en ningún dicionario nin enciclopedia galega, decidimos lematizar *cerro testigo* baixo *cerro testemuña* posto que esta é unha forma que se vén empregando no ensino medio⁵⁰.

Sobre a lematización de *cerro mula* e *mula / mulo*, *vid.* o comentado no CI 14 e sobre a de *colleitar*, *cortasetos*, *forcada*, *forquita*, *horquilla*⁵¹, *rastrelo*⁵² e *recolleitar* consúltase o CI 13, dedicado aos traballos do campo e do xardín.

Por outra banda, representamos no lema a variación xenérica dos seres animados que se rexistrou nas respostas, seguindo o criterio xeral comentado en § 2.3.2.1. Deste modo temos casos como *coello-a*, *raposo-a* ou *vello-a*. Constitúen aparente excepción *pito/pita*, *polo/pola* e *tenreiro/tenreira*, que se comentarán ao describir o CI 14, dedicado aos animais. De acordo coa norma xeral, cando a variación formal non se corresponde (ou pode non corresponderse) cunha variación de sexo, non unificamos as formas: así *cultivador* e *cultivadora* non se lematizan conxuntamente porque a palabra *cultivadora* pode designar unha máquina, nin tampouco *agrario* e *agraria*, porque esta última forma se emprega para facer referencia a unha asignación económica para os traballadores do campo.

No que atangue aos afixos apreciativos, conservamos no lema algúns sufixos, malia non atoparmos nos dicionarios, cun significado acaído ao CI, as formas afixadas. É o caso de *maquinillo* ou *tractorillo*. Ao pé destas formas encontramos, por suposto, outras con referendo lexicográfico, como *forquita* ou *fouciño*.

Como xa se sinalou, son frecuentes neste campo os sintagmas complexos con cuantificadores, comparativos ou non, ou con elementos de negación, como *sen* ou *non*: *sen tráfico*, *menos coches*, *pouca contaminación...* Sobre a lematización destas expresións *vid.* § 2.3.2.7.1.

2.3.4.12 CI 12. Os medios de transporte

Ademais das voces máis claramente previsibles neste CI, como son as que designan os vehículos prototípicos (*avión*, *barco*, *coche*, *moto*, *tren...*), encontramos (e conservamos) designacións doutros medios de transporte menos centrais ou doutras realidades cuxa consideración como tales pode resultar

⁵⁰ Como se reflicte, por exemplo, no feito de ser o termo escollido pola Comisión Interuniversitaria de Galicia nas directrices xerais sobre Xeografía para o curso 2009-2010 [consulta realizada o 15/04/2013]: <http://ciug.cesga.es/grupos/xeografiaorient.html>

⁵¹ Polos mesmos motivos esgrimidos no comentario ao CI 13, tamén lematizamos de maneira separada *horquillo*, que se recolle exclusivamente no CI 11.

⁵² Tamén no CI 11, como no 13, os contextos conducen á interpretación da palabra *rastrelo* como designadora dun aveño.

cando menos polémica (*ascensor, cadeira de rodas, monopatín, submarino, táboa de surf, teleférico* etc.), así como algunhas voces que fan referencia aos diferentes tipos de desprazamento (*andar, correr, montar, navegar, planar, sucar, voar...*).

Ademais, tal e como cabía esperar, rexistramos igualmente formas directamente relacionadas cos diferentes tipos de vehículos, como nomes de materiais e pezas que os constitúen (*buxía, caucho, cinto, eixe, ferro, manillar, pedal, turbina* etc.), voces vinculadas coa circulación, en todas as súas dimensións: vías (*asfaltar, autoestrada, estrada, socavón, túnel, vía...*); combustibles (*electricidade, gasolina, gasoleneira, petróleo...*); regulación do tráfico (*Dirección Xeral de Tráfico, semáforo, sinal, stop...*); sancións e permisos (*carné, multa, permiso de moto, radar...*); accidentes (*accidente, choque, estrelarse...*); condución, pilotaxe ou manexo do vehículo (*acelerar, condutor, frear, guiar, maquinista, navegante, pedalada, pilotaxe, taxista...*); instalacións (*aeroporto, estación, marquesiña...*); entradas e saídas (*baixar, desembarcar, terminal, transbordar...*); adquisición (*concesionario, prezo, vendedor, vender...*) etc. Por suposto, conservamos todas elas nas listaxes.

Algunhas das voces que mantemos teñen carácter xenérico e fan referencia ás viaxes, ao transporte ou ao desprazamento de maneira xeral: *billete, chegar, destino, ir de viaxe, levar, moverse, pagar, partir, pasaporte, percorrer, regresar, ruta, transportar, traslado...*

Non faltan tampouco as formas que se centran na finalidade da viaxe, esencialmente ocio (*excursión, turismo, visitar...*), aínda que tamén comercio (*comerciar, mercadoría*).

Algunhas palabras, sobre todo adxectivos, son empregadas para establecer unha tipoloxía dos medios de transporte desde diferentes puntos de vista (*aéreo, barato, comfortable, gratuíto, perigoso, privado, rápido...*). Non faltan referencias á contaminación, un tema moi presente en varios dos CI do LDG (*contaminador, ecolóxico, fumes, polución...*), así como ás consecuencias que pode ter o desprazamento nos usuarios (*agobio, estrés, intranquilidade, mareado...*) etc.

Por outra banda, tamén conservamos nomes comerciais (*Airbus**, *Cas-tromil**, *Concorde**, *RENFE**, *Vespino** etc.). Entre eles son especialmente frecuentes os designadores de marcas e modelos automobilísticos: *Corsa**, *Golf**, *Mercedes**, *Renault**, *Seat**, *Suzuki**...

Por outra banda, eliminamos os nomes de máquinas (*cosechadora, moto-niveladora*), excepto aquelas que presentan un número de aparicións superior a cinco (*escavadora, segadora, tractor*); tamén suprimimos os nomes

de animais (*cabra, cervo, ovella*), salvo os que se utilizan como un medio de transporte ou son parte esencial del (*boi, caballo, dromedario, elefante*). Finalmente, deixamos fóra outras formas con escasa ou nula vinculación co CI que nos ocupa, ou que posúen un valor demasiado xeral (*calor, corpo, cousa, filosofía, mexillón, nabo, Nadal, roupa, sentir, ver, zapato...*).

No que atangue aos lemas, cómpre facer referencia á lematización independente de *bus* e *autobús* e *moto* e *motocicleta*. As formas plenas e acurtadas conviven nas listaxes dalgúns informantes, o que pode ser un indicio de que designan realidades diferentes. Aínda que a distinción non se recolle en ningún dos dicionarios consultados, é moi probable que para algúns falantes a *moto* designe un vehículo de maior cilindrada ca a *motocicleta*, e que certos informantes empreguen especificamente a palabra *bus* para facer referencia ao medio de transporte de carácter urbano⁵³.

Sobre a lematización de *mula / mulo*, vid. o comentario ao CI 14.

Con respecto ás palabras con afixos apreciativos, consideramos lexicalizadas formas como *barcaza, camioneta, carretilla, carretillo* ou *carriño*, pero non outras como *barquiña* e *trenciño*, que tratamos, respectivamente, como variantes de *barca* e *tren*. No caso de *cochecito / cocheciño* interpretamos que os informantes poden estar facendo referencia ao carro empregado para transportar os nenos de pouco tempo (a variante *cocheciño* precede a *carro de bebé*), de tal xeito que lematizamos baixo a forma sufixada. Finalmente, lematizamos *ventanilla* baixo *ventaíña*, porque esta palabra é un dos equivalentes de tradución que se ofrecen para a voz do castelán no DCG.

2.3.4.13 CI 13. Os traballos do campo e do xardín

Neste CI, que se solapa en boa medida co CI 11, dedicado de modo xenérico ao campo, é especialmente frecuente a presenza de verbos, por seren estes os elementos que prototipicamente designan accións. É, daquela, o centro de interese con máis representantes da categoría verbal (*arar, cavar, coidar o gando, esfollar, muxir, partir leña, podar, segar...*), sendo algúns de tales verbos especialmente produtivos na formación de sintagmas: *coidar (a facenda, a horta, as flores, as ovellas...)*; *limpar (a finca, as cortes, as malas herbas...)*; *plantar (cebolas, centeo, coles...)*. Pero a carón dos verbos, das frases verbais e dalgúns substantivos (*pastoreo, sementeira...*) que

⁵³ En calquera caso, esta última distinción non sería xeral, posto que atopamos exemplos de *bus de liña* ou de *bus de cercanías*. En coherencia coa lematización establecida para *bus* e *autobús*, tampouco agrupamos as formas cando van modificadas (*autobús escolar* e *bus escolar* ou *autobús urbano* e *bus urbano* non se unifican), aínda que nalgún caso, como a do último exemplo, a interpretación é unívoca. Deste xeito, pode atoparse a listaxe de hiperónimos e hipónimos cun só golpe de vista. Si agrupamos, no entanto, cando o medio de transporte funciona como modificador: *estación de (auto)buses* ou *conductor de (auto)bús*.

fan referencia directa a traballos ou actividades do campo e do xardín, encontramos outros elementos que tamén parecen encaixar neste ámbito temático. Por un lado, atopamos algunhas voces que, sen designaren tarefas prototípicas do campo ou do xardín, amosan un vínculo coas actividades indicadas (*comprar sementes, contratar un xardineiro, ensuciar, madrugar, mal olor, uñas sucias, vacinar...*). Tamén rexistramos algúns vocábulos designadores, non de tarefas no sentido habitual da palabra, pero si de accións que se fan no campo: *agrandar o marco, camiñar, pasear, roubar laranxas...*

Por outro lado rexistramos un considerable número de instrumentos (coas súas partes e mesmo con marcas), construcións, recipientes, maquinaria, medidas, roupa, vehículos etc. empregados directa ou indirectamente nas distintas tarefas (*alpendre, arado, bocoí, camiión, capacho, cortasebes, chimpín, ferrado, fouce, galleta, hectárea, John Deere*, legón, luva, mango, muíño, pesticida, poleiro, roupa vella, sombreiro...*), así como materiais de traballo (*abono, alimento, area, auga, esterco, pedra, pesticida...*). Eliminamos no entanto algunhas voces moi pouco específicas como *coche* ou *automóvil*. Atopamos así mesmo designacións de diferentes produtos e elementos da natureza propios do campo, sexan froitos, verduras, árbores, arbustos... (*acelga, dente de león, enredadeira, espárrago, fresa, froita, hortensia, piñeiro...*), produtos obtidos destes (*aceite, bagazo, pole, viño...*) e voces relativas ao seu aproveitamento ou comercialización (*autoconsumo, comercio, cooperativa, negocio, pequeno comercio, producir, serradoiro, subsistencia...*). Tamén atopamos partes de plantas (*espiña, esqueje, flor, folla, talo...*), procesos sufridos por elas e vocábulos relacionados (*agromar, brotar, crecer, enraizar, florecer, praga...*) ou mesmo características (*oloroso*). Chama a atención así mesmo a presenza de voces vencelladas coa meteoroloxía e as épocas do ano, debido con toda seguridade á directa vinculación entre tales realidades e as tarefas agrícolas: *calor, chuvia, fresco, sarabiar, seca, sol, vento...*

A carón das plantas figuran, por suposto, os animais: nomes xenéricos (*animal, bicho*), tipos (*bovino*), animais criados para a súa explotación comercial ou doméstica (*coello, galiña, porco, vaca...*), animais de compañía e empregados nos traballos de campo (*can, burro, mula...*), ou outros que se encontran facilmente en campos ou xardíns (*avespa, caracol, formiga, grilo, paxaro...*). Da mesma maneira que sucedía coas plantas, atopamos aquí tamén nomes de partes de animais, sobre todo as vinculadas con certas tarefas (*ubre*), produtos que se obteñen deles (*la, leite, ovo*) e actividades feitas por eles (*empollar*).

Certos rexistros fan referencia a alimentos así como a actividades ou elementos que teñen que ver coa súa preparación ou conservación. Trátase esencialmente de alimentos extraídos do campo ou dos animais domésticos: *adobar, amasar, cocer pan, macerar, pan, queixo, touciño...* Eliminamos no entanto os adxectivos empregados para caracterizar eses produtos: *rico, saboroso...*

Rexistramos por outro lado algunhas palabras que fan referencia aos diferentes lugares e vías nos que se levan a cabo as tarefas ou que interveñen nelas dunha ou outra maneira: ao pé das propias palabras *campo* e *xardín* atopamos *bosque, carreiro, eira, finca, galiñeiro, horta, leira, monte, río...* e voces máis xerais como *aldea* e *pobo*. Encontramos igualmente algúns vocábulos que fan referencia a realidades habituais nos xardíns, como *casa do can, fonte, mesa, parrilla* ou *barbacoa*. Tamén figuran nas listaxes algúns rexistros que designan as persoas que realizan as tarefas do campo e do xardín (*agricultor, aldeán, apicultor, criador, esquilador, inmigrante, leñador, pastor, traballador, xardineiro...*, pero tamén *ancián, avó, vello...*), os distintos ámbitos, disciplinas e saberes cos que se vinculan (*agrícola, agricultura, apicultura, botánica, floricultura, gandería, xardinería...*) ou os diferentes sistemas de cultivo ou de pastoreo (*transhumancia, rotación...*). Atopamos tamén certo número de rexistros que designan oficios ou profesións asociadas fortemente coa vida rural, algunhas delas polo seu carácter orixinariamente artesanal: *artesán, canteiro, carpinteiro, costureira, curandeiro, ferraxeiro, panadeiro, taberneiro, zapateiro...* Teñen tamén cabida outras palabras que remiten a actividades que se desenvolven con frecuencia no campo: *arqueólogo, bombeiro forestal, adestrador de mascotas, garda civil...* Pero eliminamos algunhas formas que non amosan esa vinculación específica co mundo rural: *albanel, fontaneiro, soldador...*

Finalmente, non son escasas as palabras que implican unha valoración das tarefas que nos ocupan, ou do campo en xeral, ben sexa positiva (*agrimo, alegría, beleza, bonito, gratificante, harmonía, pracer, relax, saúde, silencioso, tranquilidade...*), ben negativa. Neste último sentido, abundan as palabras que inciden na dureza do traballo: *aburrimento, cansazo, custoso, dificultade, duro, esforzo, esgotador, laborioso, romper o lombo, sufrir...*

Quedan fóra rexistros demasiado xenéricos ou de máis difícil encaixe no ámbito temático que nos ocupa: *estudar, falar, fumar, home, instaurar, mirar, muller, neno, PlayStation, prostitución, señor, sobar, zoofilia...*

No que se refire á lematización de casos concretos, cómpre facer unha serie de precisións e xustificacións:

A forma *colleitar*, rexistrada en vinte ocasións nas respostas dos alumnos, non figura no VOLG nin no DRAG. No DCG da Academia tampouco se menciona como posible equivalente do seu cognado castelán *cosechar*. No entanto, o GDXL dálle entrada con formato de palabra estándar e o DCG da Academia ofrece *colleitadora* como equivalente de tradución de *cosechadora* (nótese que *colleitadora* supón como base o verbo *colleitar*). Por tal motivo, ofrecemos *colleitar* como lema e situamos a palabra *cosechar* (tamén vinte exemplos) baixo *colleitar*. Debe terse en conta, por outro lado, que a palabra *colleitar* se rexistra en respostas de alumnos de moi distinta procedencia xeográfica, entre eles varios que teñen o galego como lingua nai. En coherencia con este tratamento, lematizamos as frases verbais que posúen o verbo *cosechar* (*cosechar o trigo*, por exemplo) baixo frases verbais con *colleitar* (*colleitar o trigo*).

No que atangue á forma *recolleitar*, atopamos dezanove rexistros, aos que debemos engadir once formas *recolectar*. A forma do castelán *recolectar* figura en oitenta e sete ocasións. O verbo *recolleitar* non se rexistra nos dicionarios máis modernos (nin tampouco no VOLG), agás no caso do GDXL, que lle dá entrada no formato correspondente ás palabras galegas non estándares. Ante esta situación, optamos por lematizar *recolleitar*, *recolectar* e *recolectar* baixo o verbo máis frecuente, *recolectar*. No caso das frases verbais con complemento, seguimos tamén o criterio da frecuencia. Deste modo temos *recolleitar uvas* pero *recolectar alimentos*.

Por outra banda, optamos por lematizar separadamente *forquita*, *horquilla* e *forcada* porque algunhas das obras consultadas diferencian na forma ou funcionalidade entre uns e outros avehos (*vid.* por exemplo o GDXL).

A palabra *retrobatear*, que se recolle nas listaxes, é un verbo formado sobre o nome comercial *Rotavator*, marca lexicalizada para facer referencia a un rotocultivador. Dado que non hai unha forma estándar para este verbo (tamén podemos atopar na lingua, entre outras, as formas *rotovatear* e *rotabatear*) lematizamos coa forma que ofrece o alumno.

Por outra banda, lematizamos *cortasetos* baixo *cortasebes*, malia esta última palabra non aparecer recollida nos dicionarios consultados, porque se trata dunha palabra difundida na lingua escrita e mesmo recollida en documentos oficiais⁵⁴. Ademais, o proceso de formación para denominar a mesma máquina noutras linguas coincide coa do castelán e coa que seleccionamos para o galego: así, por exemplo, *cortasebes* en portugués ou *coupe-haies* en francés. No que atangue a *hidrosiembra*, escollemos *hidrosementeira* como

⁵⁴ Vid. por exemplo [consulta realizada o 15/04/2013]: http://www.edu.xunta.es/fp/sites/fp/files/fp/Curriculos/LOE/Agraria/cs_paisaxismo_medio_rural.pdf

lema porque aínda que esta forma non se recolle nos dicionarios galegos si é de uso na lingua, e mesmo podemos recollela en documentos oficiais⁵⁵.

Noutras ocasións o contexto condiciona o lema seleccionado: a forma *fresal* lematízase baixo *fresar*, posto que o feito de que os restantes rexistros fornecidos polo alumno (agás *xornaleiro*) sexan verbos designadores de accións propias do campo leva a interpretar a palabra como unha errata por *fresar* e non como a voz do castelán que fai referencia ao terreo no que medran as amordeiras. En segundo lugar, lematizamos *polas* baixo *póla* porque o contexto dá a entender que o informante está designando unha parte das árbores (a palabra aparece ao pé de *ramas*). Por outra banda, lematizamos *rastrelo* coa forma recomendada *restrelo*, porque, aínda que *restrelo* tamén pode ser unha castaña que se desprende soa do ourizo, os contextos nos que figura a voz (ao pé doutras denominacións de ferramentas e máquinas agrícolas) conducen á súa interpretación como designadora dun instrumento de labranza. Por último, aínda que lematizamos a forma *sembra* (e *siembra*) baixo *sementeira*, considerándoa habitualmente como substantivo, nun caso interpretámola como verbo e lematizámola baixo *sementar*, polo feito de ir acompañada doutras voces nas que se impón a consideración como verbos en 3ª persoa de singular: *sulfata*, *vota* [sic] o *abono*. Por idéntico motivo, lematizamos a forma *ara* que ofrece o mesmo alumno baixo *arar*, intepretrándoo como verbo e non como substantivo.

En relación coa unificación de formas femininas e masculinas, non agrupamos nun mesmo lema as formas *recolector* e *recolectora* porque esta última palabra parece estar designando unha máquina, aínda que non a encontremos nos dicionarios consultados con ese valor. Velaquí o contexto no que figura: *máquina de cortar leña*, *sementadora*, *recolectora*.

No que atangue ás formas complexas, unificamos mediante parénteses algunhas formas nominais que aparecen con e sen complemento, como *rotación (de cultivos)*, pero non facemos así no caso de *matanza* e *matanza do porco*, posto que a primeira unidade, ademais doutras acepcións compartidas coa forma complexa, posúe os significados de ‘conxunto de porcos que van ser sacrificados’ ou ‘os produtos que se obteñen do sacrificio dos porcos’ (GDXL). Tales acepcións non parecen atribuíbles ao emprego habitual que se fai de *matanza do porco*.

En relación cos afixos apreciativos, consideramos formas lexicalizadas *maquinillo* ou *tractorillo*, que aparecen tanto neste CI como no CI 11, malia non figurar como máquinas propias do campo nos dicionarios consultados.

⁵⁵ Por exemplo en [consulta realizada o 15/04/2013]: http://www.xunta.es/dog/Publicados/2011/20110615/AnuncioCA04-090511-1362_gl.html. Figura ademais na listaxe de formas da *Neoteca. Banco de Datos de Neoloxismos da Universidade de Vigo* (Seminario de Lingüística Informática 2006-2008).

2.3.4.14 CI 14. Os animais

A carón dun elevado grupo de nomes de animais, no seu nivel básico de categorización (*can, gato, paspallás, reno, vaca...*), atopamos tamén vocábulos máis xenéricos (filos, clases, ordes...): (*anfibio, crustáceo, equinodermoo, insecto, molusco, roedor* etc.) e outros máis específicos, designadores nomeadamente de razas (*yorkshire, rubia galega* etc.). Outras voces xenéricas baséanse en taxonomías de moi distinto tipo, relativas á alimentación dos animais (*herbívoro, omnívoro...*), á reprodución (*ovíparo, vivíparo...*), ao medio no que viven (*acuático, terrestre...*), á cronoloxía (*prehistórico*), a características formais (*invertebrado, paquidermo, vertebrado*) etc. Finalmente, tamén encontramos algunhas denominacións xenéricas do tipo de *animal, bicho* ou *fera* (xunto con *raza* ou *especie*). Mantemos dentro da listaxe os nomes de figuras mitolóxicas con características animais: *centauro, harpía, lobishome, minotauro, serea...* Excluímos, pola contra, o grupo de nomes propios designadores sobre todo de individuos concretos, na súa gran maioría imaxinarios: *Babe, Garfield, Piolín, Puercoaraña, Ratatouille...* Como pode comprobarse, varios destes animais son protagonistas de *debuxos animados*, e o propio termo composto, tamén eliminado, figura rexistrado nas respostas dos alumnos.

A morfoloxía animal aparece representada a través de varias palabras designadoras de partes do corpo: *á, cachola, dente, espiña, fociño, garra, orella, pata, pel, pelaxe, pouta, rabo, sangue, trompa...* Tamén atopamos varias respostas que serven para caracterizar os animais desde o punto de vista formal: *bonito, peludo, pequeno, pesado, suave...*

A fisioloxía e o comportamento animal tamén están amplamente representados. Rexistramos respostas relativas á reprodución (*camada, cría, poñer ovos, preñado, seme...*), á alimentación, incluíndo tipos de alimentos habituais para certos animais (*amamantar, bebida, comida, espiga, fariña, gran, millarada, millo, pacer, pastar, pendón...*), ás voces características (*balar, bramar, ladrar, ladrido, muar, ouvear, rinchar...*), ao desprazamento (*andar, correr, galopar, lento, reptar, saltar, veloz, voar...*), á excreción (*bosta, cagada, excremento...*), ao proceso evolutivo e á saúde (*crecemento, enfermarse, morte, nacer, rabioso, san, sarnoso...*), ao cheiro que desprenden (*cheirente, olor*) e, de maneira xeral, ao comportamento animal (*agresividade, arañar, babeco, bo, bondade, bravo, bruto, cariñoso, manso, marca de territorio, mordedor, pelexa, picar, sucio, saltareiro, salvaxe, tímido, tranquilo...*).

Dada a estreita vinculación establecida entre os seres humanos e os animais ao longo de moitos anos de convivencia, non debe estrañar tampou-

co a presenza de numerosos vocábulos que amosan claramente a existencia dese vínculo. Se no parágrafo anterior aludiamos á alimentación animal, en moitas das respostas encadrables neste grupo queda patente que é o ser humano o que alimenta os animais: *alimentar o gando, ben mantido, dar de comer, alpiste, bebedoiro, folgada, lavadura, penso...* De súa vez, os animais convértense en fonte de recursos para o ser humano, sobre todo, aínda que non só, de alimentación, o que queda de manifesto nalgunhas respostas dos alumnos: *abrigo, carne, caza, chourizo, churrasco, leite, mel, pescar, raxo, touciño, xamón...* O interese humano tamén ten un compoñente estético. Isto reflíctese en respostas como *adorno* ou *disecado*.

De maneira xeral, a actividade de cría e coidado dos animais (para a súa explotación comercial ou para o seu aproveitamento doméstico) está representada nas listaxes obtidas: *arre, caxato, espicho, estrar, gandería, matar, recoller, sacrificio, traballo, zurro...* Pero ao pé dos animais domésticos criados para o sustento das familias, os animais de compañía, cada vez máis numerosos, están tamén altamente representados. Atopamos varios rexistros que poderían encadrarse doadamente neste ámbito temático: *abandonar, adestrar, adoptar, caricia, coidador, coidar, gasto, lavar, limpar, mimar, pasear, peitear, recompensa, xogar...* É especialmente significativa a presenza de rexistros que introducen un compoñente afectivo e que implican unha valoración da relación que se establece entre o home e o animal na sociedade actual: *acompañar, amigo, animar, compañeiro, compromiso, confianza, dedicación, fidelidade, inseparable, lealdade, tenrura...* A valoración positiva da relación vai máis alá dos animais de compañía. Nas respostas dos alumnos reflíctese tamén unha preocupación ecolóxica: *comercio, extinción, extinguir, explotación, protexer*. Isto non obsta para que certos animais teñan unha consideración negativa, por razóns de distinto tipo. Este feito reflíctese na presenza de formas como *fumigar, medo, molesto, noxo, perigo* ou *Ratibrón**.

Outras respostas aluden ao dominio do ser humano sobre o animal, que queda reflectido en formas como *amansar, domesticado* ou *montar*. Poden mencionarse tamén aquí aqueles instrumentos ou aparellos que o home adoita colocar nos animais, con diferentes funcionalidades: *cabezada, cadea, cascabel, correa, solta...* Contrasta con estas respostas a visión característica do animal como ser libre: *liberdade, libre*.

Finalmente, hai numerosas respostas que fan referencia ao lugar habitado polos animais, sexan construcións, hábitats ou elementos naturais: *aire, alpendre, bosque, campo, canceira, cortello, establo, estepa, gaiola, natureza, niño, piscifactoría, pocilga, pombal, refuxio, sabana, selva, terra...* Algunhas voces aluden a colectividades animais: *camada* ou *enxame*.

Eliminamos outras formas que non encaixan en ningún dos grupos indicados ou que o fan moi indirectamente: *branco, cateca, espido, fotosíntese, George W. Bush, mamá, negro, papá, o efecto mariposa, sofá, verde...*

No que atangue á lematización, algúns casos deben ser comentados:

Optamos por lematizar o rexistro *ciervo-mula* como *cervo mula* porque, aínda non atopando unha forma estándar galega para o dito animal, nas linguas próximas (inglés, portugués, francés...) o composto segue o mesmo esquema de formación que se rexistra no castelán. A mesma xustificación empregada no caso de *ciervo-mula* condúcenos a lematizar *escarabajo bombardeiro* como *escaravello bombardeiro*.

Con respecto á voz *cancodrilo*, que figura dúas veces nas respostas, atopamos rexistrado o seu uso no galego tanto co significado de animal a medio camiño entre o can e o crocodilo como co significado de crocodilo. Dado que os alumnos que ofrecen a palabra non inclúen ao tempo o nome habitual do reptil, optamos por lematizar a voz baixo *crocodilo*.

Por outra banda, lematizamos separadamente *mantis* e *mantis relixiosa* porque segundo indican os dicionarios, *mantis* designa o xénero e *mantis relixiosa* unha especie. Tamén separamos *mel* e *mel das abellas* porque o mel non ten por que extraerse necesariamente daquelas, pois tamén pode proceder doutros insectos.

Noutras ocasións é o contexto o que determina a selección do lema: en todos os casos lematizamos a variante *óso* como *oso*, posto que as palabras circundantes parecen indicar claramente que se trata do ser vivo e non da parte do esqueleto. Sempre figura entre nomes de animais e nunca a carón doutras denominacións de partes do corpo. Como proba da existencia de confusión, pode citarse a presenza dun *óso formigueiro*. En segundo lugar, aínda que a forma *reo* designa en galego un peixe que nace no río, lematizamos un dos nosos rexistros *reo* baixo *reno* pola súa situación: a carón de *cervo, alce* e *veado*. Nótase que é habitual nas listaxes esta “adaptación” ao galego suprimindo o [n] intervocálico.

Finalmente, optamos por lematizar conxuntamente *silla* e *silla de montar* baixo *sela*. En principio, se o núcleo fose unha palabra galega, as dúas formas deberían unificarse baixo *silla (de montar)*. Agora ben, *silla* é un castelanismo que posúe un cognado galego, *sela*, co significado específico de aparello de montar (designa ademais o asento das bicicletas). Así pois, *silla* debería lematizarse baixo *sela*. Para non separar a forma simple e a complexa do castelán, *silla* e *silla de montar*, optamos por agrupar as dúas baixo *sela*, aínda en contra do criterio de non lematizar nunca baixo formas que non compartan todas as raíces das variantes rexistradas (pois neste caso, no lema *sela* non aparece a raíz de *montar*).

No que atangue ao xénero, seguindo o criterio xeral, lematizamos en masculino os femininos de formas que o dicionario considera masculinos epicenos (*cangura* baixo *canguro*) e en feminino os masculinos de formas consideradas nas obras lexicográficas como femininos epicenos (*foco* baixo *foca*), pero facemos excepción coa voz *mulo* (lematizado baixo *mula-o*). A voz *mulo* probablemente apareza por castelanismo (en galego o macho da *mula* chámase precisamente *macho* e *mula* é invariable).

Por outra banda, seguindo o criterio xeral situamos baixo un mesmo lema masculinos e femininos correspondentes a machos e femias (*becerro-a, burro-a...*), pero non agrupamos o masculino e o feminino cando unha das palabras posúe algún significado sen correspondencia no outro xénero. Así, mantemos separados *pata* e *pato*, posto que o substantivo feminino pode facer referencia a unha parte dos animais. Tampouco unificamos *tenreiro* e *tenreira*, posto que, amais de designaren a cría da vaca macho ou femia, a voz feminina pode facer referencia a unha vaca que acaba de parir, non tendo neste caso, por suposto, forma masculina asociada. O mesmo pasa con *cocha* e *cocho*, posto que este último vocábulo, amais de denotar o macho da cocha, tamén significa ‘cortello’. Por suposto, tamén mantemos separadas outras voces que formalmente poderían considerarse como masculino e feminino, pero que non designan, sen máis (ou non necesariamente), o macho e a femia dun mesmo animal: *pito* e *pita*, *polo* e *pola*, *touro* e *toura*... Ademais, lematizamos independentemente os femininos irregulares (*tigre* e *tigresa*, *galo* e *galiña*) e, por suposto, os que non comparten raíz (*boi* e *vaca*, *cabalo* e *egua*...).

Con respecto ás formas con sufixos apreciativos, lematizamos con diminutivo expresións lexicalizadas como *cabaliño do demo* e *cabaliño de mar* (máis frecuentes ca as variantes sen diminutivo), *pescadilla* ou *vaquilla*, pero prescindimos do sufixo noutros casos como *cabuxiño* ou *paxariño*. Si consideramos que *galiño* pode facer referencia a un tipo especial de galo (o *quico*), co cal o mantemos como lema diferenciado.

2.3.4.15 CI 15. Os xogos e as distraccións

Neste CI, ademais dos agardados nomes de xogos ou de tipos de xogos, en sentido amplo (*acción, adiviña, backgammon, balón prisioneiro, deportivo, diana, didáctico, parchís, policías e ladróns, xogo en equipo, zapatiño inglés*...), de xoguetes (*Action Man**, *boneca, cochiño*...), de actividades deportivas ou de mantemento físico (*aeróbic, alpinismo, andar, cazar, 110 metros vallas, pescar, taichí, voleibol, ximnasia*) ou de pasatempos (*autodefinido, charada, encrucillado, sopa de letras*...) atopamos outras actividades

clasificables en diferentes grupos, que mantemos nas listaxes finais: amor e sexo (*bicar, coito anal, facer as beiras, facer unha palla, sesenta e nove*), natureza (*acampada, bañarse no río, buscar caracois, sendeirismo...*), artes plásticas (*actividade plástica, escultura, fotografía...*), artes escénicas e musicais (*actuar, bailar, ballet, canción, coral, obra teatral...*), comer e beber (*beber, botellón, cociñar, saír de bares...*), cine e televisión (*animación, audiovisual, cine, película...*), lectura e escritura (*banda deseñada, escribir, novela, tebeo...*), viaxes, paseos e desprazamentos en xeral (*andar en moto, conducir, excursión, pasear, viaxar...*), animais (*apicultura, baixar o perro, xogar cos animais...*), aprendizaxe (*aprender, aprender idiomas, conferencia, estudar, pasantía...*), actividades agrícolas e gandeiras (*muxir, sachar...*), festas, feiras e actividades circenses (*atracción, circo, coche de choque, esmorga, festas...*), descanso e relaxación (*descansar, durmir, estar nas verzas, mirar para as moscas, relaxarse...*), informática (*estar no ordenador, formatear discos duros, navegar por internet...*), manualidades e decoración (*bolillos, bordar, calcetar, colorear, coser, decorar, facer manualidades...*), relación e comunicación (*charlar, chatear, conversar, escoitar, maruxear...*), coleccionismo (*coleccionar cromos, selos...*), compras (*comprar, ir de tendas, mirar tendas...*), ciencias e técnicas (*electricidade, electrónica, mecánica, tecnoloxía...*) e outras actividades máis dificilmente clasificables (*coidar un bebé, comer as unllas, escaravellar no fociño, manifestarse...*), ademais de vocábulos xenéricos como *actividade extraescolar, enredo, facer algo, xogo...*

Ámais das actividades encadrables nalgúns dos grupos anteriores, atopamos numerosas voces relacionadas máis ou menos directamente con elas, que tamén conservamos. Así, por exemplo, rexistramos palabras que designan consecuencias dos xogos ou distraccións indicados (*adicción, borracheira, desestrés, distraerse, endebedarse, enfadarse, ludopatía, mancarse...*), accións ou procesos vinculados co seu desenvolvemento (*agochar, animar, apoiar, arbitrar, atopar, baza, brilar, eliminar, empatar, encestar, perder, recortar, saltar, trampa...*), lugares, espazos ou momentos relacionados coas actividades (*aire libre, bar, biblioteca, cafetería, centro comercial, discoteca, época de exames, escaparate, escola, mar, parque, recreo, teatro...*), sistemas, mecanismos, substancias, pezas de roupa, animais etc. empregados na realización das actividades (*acordeón, ala delta, alcohol, amplificador, animal, batería, cabalo, consola, cadea de música, caderno, camiseta deportiva, carro, cartucho, cromo, cronómetro, discman, droga, piano, plastilina...*) ou persoas e colectividade que participan nas mesmas (*amigo, anano, banda...*). Tamén atopamos algúns elementos caracterizadores dos xogos e

distraccións. Deixando a un lado os que serven para establecer unha tipoloxía de xogos, encontramos adxectivos que os cualifican desde o punto de vista da súa capacidade de entretemento (*aburrido, divertido...*), da súa época (*antigo, moderno, tradicional...*) ou do seu valor económico (*barato, caro*).

Non lles damos entrada, no entanto, a algúns vocábulos que se vinculan con actividades que non parecen centrais ao ámbito temático, como partes do corpo ou outros elementos relacionadas co campo sexual (*lingua, lefa, ollo, pé, preservativo, teta...*), ou outras palabras cun vínculo demasiado indirecto co centro (*amarelo, calar, calor, cemento, contaminación, frío, respirar, vestimenta...*) ou cunha designación demasiado xenérica (*xente*).

No que atange á lematización, este centro ofrece bastantes problemas derivados da complexidade das vocés ás que lles dá cabida. Por unha banda, é este o CI que rexistra maior número de nomes de marcas, marcas de xogos e xoguetes, que recollemos baixo o nome comercial empregado en Galicia, independentemente de que algúns informantes adapten a denominación ao galego (*La herencia de la tía Ágata**, *Mariquita Pérez**, *¿Quién es quién?**, *Scalextric**, *Tabú**, *Trivial**...). Este tratamento é válido para todas aquelas marcas que non aparecen recollidas no dicionario, mesmo aquelas que con seguridade alcanzaron a lexicalización (*Hula Hoop**). Cómpre destacar, por outra banda, a presenza dun elevado número de videoxogos e xogos de ordenador, como *Fifa**, *Medal of Honor**, *Pro Evolution Soccer**⁵⁶... Varios destes xogos presentan diferentes versións, recollidas nas respostas dos alumnos, que nós unificamos baixo a voz xenérica. Así, por exemplo, *Fifa 00*, *Fifa 2008* ou *Fifa Pro* lematízanse baixo *FIFA**. No caso de que un mesmo alumno ofrezca varios rexistros da mesma serie, só mantemos a primeira forma (lematizada baixo a pabra xenérica, segundo acabamos de indicar). Tamén atopamos algúns nomes coincidentes con programas de televisión (*Alla tú**, *Cifras e Letras**...) (*vid. infra*). Atopamos igualmente nomes de xornais (*El Mundo**, *El País**), que mantemos tamén na forma orixinal.

Deixando a un lado as marcas rexistradas, son moi frecuentes os nomes de xogos de diferente tipo (on line, de cartas, de campamento...) que non aparecen recollidos nos dicionarios. Nestes casos (*cuadrado, churra montada la burra, Simón dice...*), a non ser que por un ou outro motivo teñamos constancia do emprego habitual da forma galega, apelamos ao criterio de frecuencia á hora da lematización. Así, por exemplo, lematizamos *cuadrado* e non *cadrado*, porque os catro rexistros deste xogo de cartas figuran coa se-

⁵⁶ Algúns dos xogos que naceron como marcas rexistradas alcanzaron unha gran extensión no seu uso a través do ordenador, de tal xeito que actualmente se empregan para nomear xogos on line de factura similar ao orixinal. É o caso de *Hundir la flota** ou *Conecta 4**, que lematizamos no entanto como marcas rexistradas.

cuencia inicial <cu>, pero lematizamos *galiña cega* (fronte a *gallinita cega/ciega*) por ser a forma máis frecuente. A aplicación rigorosa do criterio de frecuencia leva á presenza de formas híbridas entre castelán e galego, como *beso, verdade e condición*⁵⁷.

Cómpre agora facer algunha precisión relativa á lematización dalgunhas das voces rexistradas no CI:

Ofrecemos *subhastado* como lema porque esta palabra figura como forma estándar no GDXL, aínda que non está recollida no DRAG nin no VOLG. Sobre *minicadea, vid.* o indicado no comentario ao CI 04.

En bastantes casos botamos man do contexto para levar a cabo a selección do lema: así, unificamos *agacharse* cos rexistros de *agochar(se)*, baixo a forma máis frecuente *agochar(se)*, porque o contexto no que aparece a palabra dá a entender que o informante se está a referir ao acto de esconderse e non a unha acción específica dun xogo que se estea a describir (*amigos, xuntanza, agacharse, correr, pillar*). En segundo lugar, optamos por lematizar *Celda* como *The Legend of Zelda** porque o contexto presupón que se trata dun videoxogo (sitúase antes de *Colin McRae 2.0* e despois doutros xogos do mesmo tipo) e non da casa dun taboleiro (*celda* en castelán). Por outra banda, o citado xogo abréviase frecuentemente en *Zelda*. En terceiro lugar, conservamos a forma galega en *Cifras e Letras**, posto que os informantes poden estar facendo referencia ao programa da Televisión de Galicia do mesmo nome. É tamén probable, no entanto, que se trate dunha adaptación ao galego do nome do xogo de mesa *Cifras y Letras*, comercializado en castelán⁵⁸. Por outra banda, lematizamos *dama* como *damas* porque o contexto dá a entender que o informante se refire ao xogo, ao non aparecer ao seu pé o nome de ningunha outra peza de xadrez e si de varios xogos de mesa. Con respecto ao rexistro *debuxos*, optamos por lematizalo baixo *debuxos (animados)*, aínda que a palabra podería facer referencia ao acto de debuxar. No entanto, os contextos nos que figura a voz e a aparición do plural parece conducir cara á interpretación como ‘debuxos animados’. En sexto lugar, lematizamos *Kart* como *Mario Kart** porque a palabra figura despois de *Mario Bros* e con maiúscula inicial. Finalmente, lematizamos *Moto GP**, con asterisco, entendendo que o informante fai referencia a un xogo de ordenador, pois o rexistro aparece despois da denominación doutras realidades do mesmo tipo.

Son moi habituais neste CI as lematizacións en plural. En certos casos, o dicionario recolle as formas neste número (*bólas, damas, construcións, ouros,*

⁵⁷ Se o rexistro presenta algún elemento anómalo, sen representatividade na fala, modificamos o lema para evitar a forma estraña: *ti a chevas* aparece lematizado como *ti a levas*.

⁵⁸ As mesmas dúbidas xorden cos dous rexistros *Allá tú**, que se poden interpretar como designadores do programa de televisión ou do xogo de mesa.

palmas, reflexos...), pero con moita frecuencia son formas non rexistradas nas obras lexicográficas. No entanto, en varias ocasións lematizamos en plural porque só o plural ten sentido neste CI ou porque o nome do xogo parece lexicalizado en tal número. Temos así sobre todo nomes de xogos (*catro esquinas, polis e cacos, barcos, submarinos, relevos, tazos...*, frecuentemente xogos de roles: *médicos, bandidos, piratas...*), ou designacións abreviadas de actividades que só se conciben como tales cunha forma plural (*estrelas* [ver as estrelas], *selos* [coleccionar selos]...). En certos casos conviven nas listaxes finais o lema plural e o lema singular, nomeadamente porque a voz en plural designa o xogo e o vocábulo en singular un elemento que se emprega na realización deste (*bólas e bóla, canicas e canica, dardos e dardo, naipes e naipe...*).

Un caso particular atopámolo nas palabras *mozas* (catro ocorrencias), *mozo* (unha ocorrencia) e *mozos* (unha ocorrencia). Todos os informantes que ofrecen *mozas* son rapaces, e ademais fano en contextos de festa ou sexo, como se pode comprobar nas seguintes series: *mozas, festa, movida, música; amigos, falar, liar, mozas; sexo, mozas, festas*; só hai un informante que coloca esta palabra en primeiro lugar, seguida de *ordenador, computadora...*, pero incluso nesta serie parece que a denotación é a mesma ca nos outros informantes. Con respecto aos casos en masculino, sucede que a forma singular é proporcionada por unha rapaza na serie *música – banda – mozo – internet*, e a forma plural, por un rapaz, no contexto *saír – mozos – vellos*, no que posiblemente o informante asocie o termo *saír* coa idea de xuventude. Tendo en conta estes datos, optamos por lematizar separadamente *mozas* en plural e *mozo* en singular. O singular representa ben o ‘noivo’, ben ‘unha persoa nova’, mentres que co plural se tenta denotar o contido ‘persoas do sexo contrario’ como elemento de distracción (neste caso, rapazas, sendo os informantes todos homes; a palabra *mulleres*, con sete ocorrencias, tamén se lematiza en plural). Pensamos que este contido (como máis arriba indicabamos para *estrelas* ou *selos*) non quedaría recollido a través da lematización en singular.

En relación cos rexistros en singular *ouro, copa, basto, espada, diamante, corazón e trévol*, que ofrece un único informante, optamos por lematizalos baixo a forma plural, que representa o pau da baralla, para así unificar estas formas con outros rexistros en plural (que atopamos nos casos de *corazóns e diamantes*). No que se refire á forma *espada*, mantemos baixo o singular un rexistro que, polo contexto, nada ten que ver coas cartas da baralla.

No que atangue ás voces con sufixos apreciativos, ofrecemos como lemas *casiñas, chupito, maquinaña, mariquitilla, abuelita...* Menos frecuentemente, as palabras sufixadas recóllense baixo a voz sen afixo apreciativo: *casitas de muñecas*, baixo *casa de bonecas* ou *palmitas* baixo *palmas*.

Con respecto aos agrupamentos de formas, e igual que fixemos no CI 08 e fronte ao criterio empregado noutros CI, lematizamos separadamente *televisor* e *televisión*, porque esta última palabra pode entenderse de dúas maneiras: como aparello receptor e como o ‘conxunto de servizos encargados da elaboración e difusión por ese medio dunha serie de programas destinados á formación, á información ou ao ocio’ (DRAG²). Por outra banda, convén destacar que na frecuente construción *xogar a...* as unificacións se limitan a aqueles casos nos que a forma simple é tamén un verbo, como en *xogar a pillar* e *pillar* lematizados baixo (*xogar a*) *pillar* (no entanto, *xogar ás casiñas* e *casiñas*, por exemplo, lematízanse separadamente). Polo que respecta ao segmento tamén moi frecuente *xogo de...*, as fusións establécense unicamente cando se dá unha clara equivalencia denotativa entre a forma simple e a complexa, como en (*xogo da*) *botella*, (*xogo da*) *oca* ou (*xogo da*) (*pill*) *pill*, pero non, por exemplo, en *pelota* e *xogo da pelota*. Para unha xustificación máis detallada deste tratamento, *vid.* López Meirama (2008: 81).

2.3.4.16 CI 16. As profesións e os oficios

Neste CI, como cabía esperar, as palabras máis frecuentes son tanto as que fan referencia ás persoas que teñen determinados oficios ou profesións (*adestrador, astrónomo, carpinteiro, médico, periodista, químico, socorrista, xoieiro*⁵⁹...), como, de maneira máis ou menos xenérica, ás propias actividades (*albanelería, educación, Enxeñería, hostalería, Medicina, política, transporte...*). Algunhas das actividades que se citan só se poden considerar como oficios ou profesións outorgándolles un sentido amplo a estes vocábulos: *buscador de mitos, camello, ladrón...*

Aínda que en menor medida, tamén podemos atopar algúns outros grupos de palabras, que mantemos.

Así, por exemplo, rexistramos algúns substantivos que designan cargos, sexan de carácter eclesiástico ou laico (*bispo, concelleiro, conselleiro, director, ministro, xefe de estudos...*). Non faltan os rangos militares (*almirante, comandante, sarxento...*) ou mesmo algún tratamento asociado aos cargos (*monseñor*). Como acabamos de ver, están presentes varios substantivos que fan referencia ás vocacións relixiosas: *cura, monxa, monxe, sacerdote...*

Tamén podemos atopar un grupo de palabras que designan lugares nos que se realizan os diferentes traballos: *ambulancia, avión, axencia, camiión, clínica, cociña, colexio, despacho, escola...* Eliminamos, non obstante, al-

⁵⁹ Aínda que algúns deste lemas recollen a variación xenérica (*xoieiro-a*), nesta exposición ofrecemos só a forma maioritaria, agás cando conveña para a explicación.

gúns rexistros que fan referencia a ferramentas, obxectos, substancias etc. que se vinculan con determinadas profesións ou oficios (*anzol, area, chave inglesa, formigoneira, micrófono, recibos, serra...*), así como as palabras que denotan o resultado dun traballo (*casas, cestas, comida, libro, novela...*). Conservamos, no entanto, algúns vocábulos encadrables neste grupo, como *ferramenta*, polo seu carácter xeral, ou *ladrillo*, por facer referencia xenericamente ao traballo da construción.

Constitúen un grupo numeroso as palabras, sobre todo verbos, que designan accións que se realizan nos distintos traballos, teñan carácter máis ou menos xeral. Conservamos todas elas: *amasar, arranxar, asfaltar, calcular, cortar, crear, curar, embalar, fregar, navegar, operación, peitear, pensar, sachar, traballar...*

Outras voces mantidas nas listaxes fan referencia ao mundo laboral de maneira xenérica: *ascender, contrato, despedir, emprego, sindicato, soldo...* Entre elas, algunhas que designan aspectos negativos da relación laboral: *explotación, explotado, imponerse...* Non están tampouco ausentes as expresións relacionadas coa falta de traballo (*desemprego, non traballar, parado, paro...*) ou mesmo nalgún caso, coa desgana (*folgazán, vago, zangonear...*) ou co requirido descanso (*descansar*).

Non podían estar ausentes as palabras que se centran na importancia dos oficios e profesións para vivir (*comer, manter, sobrevivir...*) e na dureza do traballo (*aburrido, cansado, custoso, duro, esforzo, estrés, madrugar, sufrir...*). No entanto, tamén se destaca de cando en vez o lado positivo (*amizade, convivencia, divertirse, independencia...*), con algunha referencia ao que se pode conseguir mediante o traballo (*cartos, fama, millonario...*).

Algunhas expresións serven para designar características que en principio deberían estar presentes no traballo: *eficiente, formalidade, responsabilidade...*

Finalmente, e como cabía esperar tendo en conta as características dos informantes que realizaron a enquisa, encontramos así mesmo algunhas palabras que fan referencia ao proceso que conduce ao coñecemento dun oficio ou profesión: *carreira, doutorando, estudos universitarios...* Por outra banda, a presenza de setenta e nove exemplos de *estudiante* dá a entender que os informantes consideran a súa actividade como unha verdadeira profesión.

En canto ás palabras eliminadas, ademais daquelas ás que xa fixemos referencia, suprimimos outras demasiado xenéricas ou con difícil encaixe no CI. Algúns exemplos son *apagar, falar, familia, lambón, pelexa, pena de morte, romper, sida, ventá, verán, xente...* Cómpre sinalar, por outra banda, que foron bastante numerosas as eliminacións de variantes flexivas, pois con

certa frecuencia os informantes encheron as listaxes incluíndo o masculino e o feminino dun mesmo vocábulo. Así ocorre, por exemplo, nunha listaxe na que *camareira* figura a continuación de *camareiro* (elimínase *camareira*) ou noutra na que *ama de casa*, que se suprime, figura despois de *amo de casa*.

No que atangue á lematización de formas concretas, cómpre facer algunhas precisións:

Non sendo *pizzeiro* unha voz recollida nos dicionarios, trátase dunha palabra bastante frecuente, que se pode atopar mesmo en documentos emanados de institucións oficiais⁶⁰. Por tal motivo aparece como lema nas listaxes finais do LDG. Por outra banda, aínda que o DRAG non recolle a voz *floreiro* co sentido de persoa que vende ou prepara flores, esta palabra si aparece no GDXL con este valor e no DCG como un dos equivalentes de *florista*. De por parte, segundo o GDXL a voz *florista* en galego (fronte a *floreiro*) designa tamén a un botánico especializado no estudo da flora. Por tal motivo, e dado que rexistramos ambas as dúas variantes no LDG, decidimos ofrecer dous lemas diferenciados, *florista* e *floreiro*.

Noutras ocasións botamos man do contexto para a selección do lema: lematizamos un caso de *enfermeria* baixo *enfermeira*, porque entendemos que estamos ante un lapsus cálamí. O rexistro sitúase entre *médico* (*medico*) e *farmacéutica* (*farmaceutrica*) e non hai na listaxe do informante ningún rexistro que designe un estudo, pois todos fan referencia a persoas que desempeñan un oficio ou profesión. No caso de *labores*, o feito de que este rexistro se sitúe detrás de *ama de casa* lévanos a lematizalo baixo *labores do fogar*.

É frecuente neste CI o emprego de maiúscula inicial nos lemas debido á presenza de numerosos substantivos que designan estudos conducentes á obtención dun título. Tal e como comentamos en § 2.3.2.9, este tipo de palabras van introducidas por maiúscula: *Arquitectura*, *Bioloxía*, *Dereito*, *Enfermería*, *Enxeñería*... Nalgún caso, o contexto permitiunos saber se a palabra estaba facendo referencia ao individuo ou aos estudos. Así, por exemplo, lematizamos un exemplo de *química* baixo *Química* e dous baixo *químico-a*⁶¹ ou interpretamos *óptica* como a designación da profesional e non do establecemento⁶². Nalgũa ocasión a disxuntiva establécese entre os estudos e

⁶⁰ Por exemplo [consulta realizada o 15/04/2013]: http://traballo.xunta.es/export/sites/default/Biblioteca/Documentos/Publicacions/Perfis_profesionais_panadeiro.pdf

⁶¹ Lematizamos baixo *químico-a* nas seguintes series:

- *enfermeira, médico, bióloga, química, físico.*
- *operadora, bióloga, química, investigadora.*

E lematizamos baixo *Química* nestoutra, onde en ningún caso (agás *enfermeira*) o informante ofrece a profesión en feminino:

- *neurólogo, física, química, filólogo.*

⁶² No contexto *actriz, óptica, médico.*

outras realidades vinculadas coa profesión. É o caso de *farmacia*, que pode designar tanto a carreira como o establecemento. En casos coma este, a non ser que o contexto fose moi claro, optamos sempre por lematizar de acordo coa acepción máis central ao CI (neste caso *Farmacia*).

Por outro lado, cómpre sinalar que con algunha frecuencia os informantes, para designar unha profesión, combinan a raíz cun sufixo que non é o lexicalizado. Neste caso, o lema retoma a forma habitual: así *administrante* lematízase baixo *administrador-a* ou *camionista* recóllese baixo *camioneiro-a*.

Con respecto á flexión, seguindo a regra xeral agrupamos masculinos e femininos (*cociñeiro-a*, *limpador-a...*). No entanto, separamos *modista* e *modisto* porque os dicionarios sinalan unha acepción especial para a forma feminina, como sinónima de *costureira* ('muller que fai pezas de roupa ou as amaña', DRAG²).

En relación cos agrupamentos, unificamos mediante parénteses as formas que levan o modificador *profesional* e aqueloutras que non o fan: *xogador (profesional)*, *nadador (profesional)*... Mantemos separados, no entanto, *corredor profesional* ('atleta') e *corredor*, posto que este substantivo designa varias profesións amais da pertencente ao ámbito deportivo.

Sobre a lematización separada de *estudo* e *estudos*, *vid.* o comentado en relación co CI 08.

2.3.4.17 CI 17. As cores

Este CI ten unha clara restrición temática que non dá lugar a demasiada dispersión nas respostas dos informantes. Atopamos os agardados nomes de cor, desde os máis frecuentes e xerais (*azul*, *vermello*, *negro*, *verde...*) ata os máis específicos (*maxenta*, *rosa bebé*, *verde caqui...*). Entre estes últimos cabe destacar a presenza de denominacións de tonalidades de carácter especializado: *azul Klein*, *azul Prusia*, *vermello Sudán*... Por outra banda, somos bastante laxas na introdución de nomes de cor, dado que mantemos algunhas palabras que non parecen estar lexicalizadas na expresión de determinadas tonalidades: *acelga*, (*amarelo*) *plátano*, *perexil*... No entanto, eliminamos algunhas combinacións que teñen unha lectura absurda ou moi difícil: *verde fucsia*, *azul fucsia*, *verde branco*; *azul tirando a roxo*, *blanco tirando a negro*... e que son proporcionadas por dous informantes en concreto. Si mantemos os rexistros, por suposto, cando se trata de compostos que designan a presenza simultánea de varias cores, aínda que non figuren no dicionario: *branquiazul*, *verdiamarelo*... Igualmente, prescindimos dalgúns formas "orixinais" como *vermello suspenso* ou *amarelo festa*, posto que en-

tendemos que o segundo elemento non achega ningún matiz cromático á cor expresada no núcleo (o informante previamente xa ofreceu na listaxe as formas *vermello* e *amarelo*). Se, contrariamente, o modificador introduce unha matización nas características da cor mantemos o dito modificador no lema, aínda que sexa peculiar: *granate a lo cura*. Esporadicamente atopamos nas listaxes nomes de cores en inglés, que eliminamos: *black, blue, green, red...*

Encontramos outras palabras que encaixan claramente no CI. Inclúense neste grupo, por exemplo, os adxectivos que fan referencia á tonalidade, brillo ou a outras propiedades das cores (*apagado, brillante, (cor) escura, (color) rechamante...*) ou dunha cor en particular (*impoluto, voz ofrecida despois de branco*), as expresións que sinalan características de determinadas realidades con cor, ou sen ela (*contraste, destacar, efecto, incoloro, luminosidade...*) ou aquelas palabras vinculadas coa cor en si mesma, coa súa clasificación ou cos seus trazos físicos: (*cor*) *fría, (cor) primaria, (cor) secundaria, cromático, espectroscopia, intensidade, frecuencia, gama cromática, luz...*

Tamén mantemos outras palabras que se vinculan de forma máis ou menos directa coa temática do estímulo proposto. Por unha banda, atopamos vocábulos que se relacionan coa pintura e o debuxo: *artista, boceto, cubismo, debuxo, paleta, pintar, pintor...* Entre estes, encontramos algúns substantivos que designan obxectos que se empregan ou se poden empregar para dar cor: (*bolí*)*grafo, lapis, pincel, rotulador...*

Conservamos algunhas palabras relacionadas coa percepción das cores e coa influencia que poden ter en nós: *colorterapia, daltónico, ver, vista...* Precisamente, outro subconxunto de voces está constituído por palabras que designan sentimentos que teñen un vínculo coas cores (ou coa ausencia das mesmas): *alegría, ánimo, esperanza, paixón, paz, tristeza...* O carácter simbólico e cultural dalgúns destes vencellos queda reflectido na presenza de rexistros como *símbolo* ou *cultura*. Situándonos xa no ámbito das opinións e percepcións persoais, podemos citar outro grupo de palabras que caracterizan as cores deste esta perspectiva: *bonito, (cor) favorita, feo, hortera, triste...*

Tamén conservamos algunhas palabras que designan realidades nas que a cor, por un ou outro motivo, é especialmente caracterizadora, sobranceira ou relevante: *bandeira, bolboreta, camaleón, flor, maruxiña, moda, outono, plastilina, roupa...*

De por parte, eliminamos algunhas palabras que se afastan en exceso da temática do CI ou son demasiado xerais: *adxectivo, aprender, bo, cova, encerado, escultura, guerra, interno, lavar, país, parto, racismo, rendición, saúde, sorte, tenda, vaidade, vidro...*

No que fai referencia a aspectos concretos de lematización, cómpre sinalar algunhas cuestións:

Lematizamos *anegrado* e *negruzco* baixo *anegrazado* porque esta voz figura no GDXL en pé de igualdade con outras formas normativas, aínda que en realidade non se recolle no VOLG nin no DRAG. En calquera caso, non atopamos outra opción máis axeitada. En segundo lugar, nas listaxes figura un exemplo da voz *bengé*, que designa un tipo de madeira, así como unha cor asociada a ela. Dado que non atopamos unha proposta estándar para o galego e dado o feito de que existe fóra das listaxes moita variedade na ortografía da palabra (*bengué*, *bengé*, *wengué* e *wengé*...), decidimos lematizar a palabra coa mesma forma ofrecida polo informante. En terceiro lugar, lematizamos *burdeos* como *bordeos*, porque aínda que esta palabra non figura no DRAG nin noutros dicionarios galegos consultados, é o equivalente de tradución que se ofrece para *burdeos* ('cor') no DCG.

De por parte, son habituais neste CI os sintagmas (fixados ou non) encabezados polos substantivos *cor* ou *color*, que van sempre enmarcados entre parénteses, pois en todos os casos o complemento aparece tamén de forma illada: (*cor*) *café*, (*cor*) *mel*, (*cor*) *laranxa*; (*color*) *alegre*, (*color*) *rechamante*, (*color*) *perla*... Nalgún dos sintagmas introducidos por *cor* / *color*, os informantes incorporaron a preposición *de* (*cor da pel*, *cor-de-rosa*...). Nós eliminamos esta preposición dos lemas: (*cor*) *pel*, (*cor*) *rosa*... Por outra banda, tal e como acabamos de ver, ás veces figura *cor* no lema e noutras ocasións atopamos *color*. Este feito responde á aplicación do criterio comentado en § 2.3.3.1, segundo o cal cando dúas formas da mesma raíz teñen o mesmo nivel de aceptación no código estándar, escollemos como lema a forma máis frecuente entre as respostas dos informantes agrupables baixo aquel.

De por parte, lematizamos separadamente formas como *limón* e *amarelo limón* / *verde limón*, *melocotón* e *laranxa melocotón*, *mar* e *azul mar* / *verde mar* porque consideramos que poden designar matices de cor distintos, ou porque nos resulta imposible escoller o sintagma co cal se deberían agrupar (*limón* debería xuntarse con *verde limón* ou con *amarelo limón*?). Si unificamos noutros casos nos que parece que a designación da forma simple e da complexa é a mesma: (*amarelo*) *plátano*, (*azul*) *celeste*, (*branco*) *marfil*, (*verde*) *aceituna*... Tamén agrupamos os adxectivos *metalizado* e *metálico* cando modifican un termo de cor (*verde metalizado* e *verde metálico* figuran baixo *verde metalizado*), pero mantemos separados os dous adxectivos cando aparecen sós. Sobre o tratamento dos sintagmas con *roxo*, *vid.* § 2.3.3.2.

Ademais, cómpre sinalar que mantemos separado *plástico* e *plástica*, porque mentres que a primeira palabra é un adxectivo, a segunda probablemente

sexa o substantivo que designa a actividade artística. Por outra banda, aínda que son frecuentes os cualificativos en feminino (*apagada, bonita...*), porque aluden á palabra *cor / color*, nós ofrecemos o lema sempre en masculino.

Finalmente, conservamos *luces de discoteca* en plural entendendo que o informante se está a referir ao conxunto de luces (de diferentes cores) propio deste establecemento. A lematización en singular non daría conta desa variedade cromática característica do referente indicado.

2.3.4.18 CI 18. O mar

Como outros CI previamente comentados, e fronte ao anterior, este presenta unha grande amplitude temática, pois o mar pode contemplarse desde perspectivas realmente moi diversas. Así, por un lado, atopamos respostas relacionadas co mar como ecosistema, onde teñen cabida os seres vivos, presentados como especies, pero tamén como clases, grupos etc. (*ameixa, castañeta, cefalópodo, molusco, nécora, peixe, rapante...*), as aves (*corvo mariño, gaivota, pingüín...*), os mamíferos mariños (*balea, foca, león mariño...*), as algas (*alga, argazo...*) etc. Como se pode observar, a fauna que se menciona nas listaxes pode ou non corresponder coa propia de Galicia. Mantémola en todos os casos, evidentemente, pero eliminamos as designacións de animais que non corresponden en absoluto ao medio mariño ou que non son específicos deste (*cisne, lagarto, ra...*). A carón dos seres vivos, encontramos o medio físico, en toda a súa amplitude (*auga salgada, baía, costa, cova, falla, farallón, fosa abisal, illa, rocha...*). Podemos mencionar aquí léxico relacionado co movemento do mar: *corrente (mariña), baixamar, marea, onda...* A influencia do tempo no mar é evidente, de aí que encontremos tamén varias palabras que fan referencia ao estado do mar ou á meteoroloxía: *embravecido, mar bravo, (mar) calmo, mar de fondo, mar (picado), tempestade, temporal, vento...* Finalmente, algúns dos informantes ofrecen nomes concretos de mares e océanos: *Atlántico, mar do Norte, (mar) Exeo, mar Morto...* Eliminamos das listaxes outros nomes propios que aparecen moi esporadicamente: *Baía Cochinos, Hawai, illas Cíes...*

Por outro lado, o mar é un medio de vida para moitas persoas, e isto reflíctese na presenza nas listaxes de moito léxico relacionado sobre todo coa actividade pesqueira, aínda que tamén con outras industrias ou actividades lucrativas que dunha ou outra maneira se vinculan co mar: (*barco*) *mercante, conserveira, estaleiro, narcotraficante, plataforma petrolífera, Pescanova**, *petroleiro, pirata...* A actividade pesqueira reflíctese na presenza de embarcacións (*barco de pesca, bou, dorna, gamela...*), de artes e utensilios

de pesca (*anzol, cana de pescar, rapeta, rede (de pesca), volanta...*), de tipos de pesca (*pesca de altura, pesca de arrastre, pesca furtiva...*) etc. O vínculo da pesca coa alimentación é evidente, de aí que atopemos tamén algunhas palabras que se vencellan con estoutro ámbito, desde a venda do peixe ou marisco para o consumo ata o momento no que se serve no prato: *alimento, cociñar o peixe, escamar, lonxa, mariscada, praza do peixe...*

O léxico da navegación en xeral (*atracar, calado, boia, embarcadoiro, navegar, porto...*) ou das partes das embarcacións (*popa, proa, quilla...*), pode poñerse en relación coa pesca ou con outras actividades que teñen o mar como protagonista. Así tamén encontramos embarcacións como *barco (veleiro)* ou *transatlántico*, que nos sitúan no terreo do mar como un lugar para o ocio e o recreo. Nesta mesma interpretación encaixan *cruceiro, pasaxeiro, turista...* Esta vertente lúdica do mar pode ir máis alá do ámbito da navegación, como nos casos en que o mar evoca a praia con todo o que isto implica (*bañista, bikini, parasol, pedaleta...*). No entanto, decidimos eliminar algunhas expresións residuais que atopaban a súa xustificación por ter só un vínculo feble coa praia (*atasco, bocadillo, balón, bolsa, refresco, xullo...*), así como algúns nomes de praias ou de vilas e cidades con praia que aparecían moi esporadicamente (*Benidorm, Samil, Sanxenxo...*). Podemos mencionar tamén neste parágrafo outras actividades ou deportes que se realizan no mar ou os elementos requiridos para levalos a cabo, os individuos que os realizan, as accións asociadas etc.: *invertido*⁶³, *moto de auga, regata, surf, surfista, táboa de surf, windsurf...* Tamén podemos atopar designacións para o equipamento e servizos que adoita haber nos lugares máis turísticos ao pé das praias: *chiringuito (de praia), paseo marítimo, piscina...* De maneira xeral, esta vertente lúdica do mar e da praia recóllese en palabras como *alegría, disfrute, diversión, pasalo ben* ou *vacacións*.

En xeral, o mar, a costa e a praia tamén se observan desde unha perspectiva valorativa, frecuentemente positiva, como lugares de liberdade e de pureza: *amplitude, grandioso, infinito, liberdade, osíxeno, respirar aire puro...* A beleza, a paisaxe, é outro valor que se destaca: *atardecer, belo, fermosura, mirador, paisaxe, posta de sol...*

No sentido negativo ponse de relevo o mal cheiro e, claro está, o perigo que supón o mar e o temor que pode chegar a infundir: *afogado, ameazador, desgraza, escuridade, fedor, mal olor, medo, perigoso, peste, temeroso...* Non faltan, por suposto, as alusións á contaminación: *contaminación, desperdicios, marea negra, residuos, sucidade...* Neste sentido, cómpre sinalar que, transcorridos seis anos entre o afundimento do petroleiro Prestige fronte ás

⁶³ Movemento do bodyboard.

costas galegas e a realización das enquisas, a catástrofe segue viva no recordo de varios dos informantes. Así, vinte persoas ofrecen a palabra *Prestige* nas súas respostas, vinculada a outras como *chapapote*, *fuel* ou *petróleo*⁶⁴.

Finalmente, cómpre sinalar que non faltan as voces que relacionan o mar coa emigración, vinculada frecuentemente coa idea do naufraxio: *emigrante*, *emigrar*, *caiuco*, *naufraxio*, *morte* etc.

En canto a outras formas eliminadas que aínda non foron mencionadas, prescindimos de topónimos que designan portos de mar (*Ferrol*, *Muros*, *Muxía*, *Noia*, *Portonovo*, *Vigo*...), aínda que mantemos os nove exemplos de *Galicia*, que dan conta da estreita vinculación existente entre esta e o mar. Tamén prescindimos de nomes propios que designan individuos, reais ou de ficción, que por un ou outro motivo se vinculan co mar (*Calamardo*, *Dory*, *Flipper*, *Javier Bardem*⁶⁵, *Moby Dick*, *Poseidón*...) ou nomes de embarcacións (*Titanic*). Finalmente, eliminamos outras formas cun vínculo co mar moi débil ou difícil de percibir, ou cun contido moi xeral: *cachondo*, *casa*, *cemento*, *consumidor*, *creador*, *historia*, *montaña*, *perdón*, *premio*, *principio*, *ver*, *vivir*...

Canto a aspectos concretos de lematización, cómpre xustificar algunhas das propostas:

Lematizamos conxuntamente e baixo *galo* os dous casos de *galo* e o único exemplo rexistrado de *gallo*. A voz *galo* pode facer referencia a diversas especies mariñas, entre elas a designada comunmente como *gallo* en castelán (*rapante* ou *meiga* en galego). Por outra banda, os dous exemplos de *galo* son proporcionados por informantes castelanfalantes e entendemos que é posible que sexan galeguizacións da voz *gallo* (nun dos casos, a palabra aparece despois de *meiga*).

En relación con *culler*, escollemos como lema *culleriña*, que se recolle no GDXL co significado de ‘instrumento de pesca cun anzol de tres puntas’. O contexto é *pescador*, *cana*, *culler*.

En certos casos, o contexto condiciona a decisión tomada. Así, un dos exemplos de *bater* que se recolle neste CI figura lematizado baixo *batea*, porque as palabras circundantes parecen indicar que o informante está pensando neste substantivo e non no golpear das ondas: *escamar*, *marisco*, *bater*, *corda*, *percebeiro*. Outro informante, despois de *batea*, escribe *bater mexillóns*, que lematizamos como *batea de mexillóns*. En segundo lugar, lematizamos *cortada* como *cortar*, entendendo que o informante fai referencia á acepción de *cortar* ‘producir unha sensación de frío moi forte a [alguén]’

⁶⁴ No entanto, debe terse en conta que nas enquisas realizadas tres anos antes para o LDEG, encontramos máis do dobre de rexistros da voz *Prestige* ca no LDG.

⁶⁵ Protagonista da película *Mar adentro*, na que o actor dá vida ao mariñeiro tetrapléxico Ramón Sampedro. O título da película aparece despois do nome do actor e tamén o eliminamos.

(DRAG²) (“esta auga corta”). A palabra precede a *fría* na listaxe. En terceiro lugar, optamos por lematizar *marisqueiro/a* baixo *mariscador-a* e non baixo o adxectivo *marisqueiro*, entendendo que os informantes fan referencia á persoa que se dedica a coller marisco (en practicamente todos os casos a palabra vai acompañada da denominación doutros oficios vinculados co mar: *pescador, navegante, mariñeiro...*). Téñase en conta ademais que no CI 16, dedicado aos oficios e profesións, hai precisamente varios exemplos de *marisqueira* e *marisqueiro*. Por outra banda, lematizamos *mar roxa* como *mareta roxa*, porque o contexto (*maruxía, resaca, mar roxa, alta mar*) non dá a entender que o informante estea pensando no nome do *mar Vermello* (*mar Rojo* en castelán, do que tamén hai exemplos nas listaxes). Por último, nas listaxes pode atoparse un exemplo de *pacífico* lematizado baixo *pacífico* e os seis restantes baixo *Pacífico*. O contexto xustifica este diferente tratamento. No segundo caso, a palabra sempre vai acompañada doutras denominacións de océanos, mentres que no primeiro nos atopamos ante unha enunciación de adxectivos cualificativos aplicables ao mar. O contexto é este: *cálido, grande, pequeno, salgado, pacífico, alto, baixo*.

No caso das denominacións de mares e océanos, se os informantes ofrecen tanto o composto sintagmático (*mar Mediterráneo*), como o segundo elemento presentado de forma illada (*Mediterráneo*) empregamos parénteses cando esta segunda opción resulta natural: (*mar*) *Mediterráneo*, (*mar*) *Exeo...* No entanto, se a supresión do substantivo *mar* resulta estraña, prescindimos das parénteses por moito que os informantes ofrecen o modificador illadamente: *Rojo* lematízase como *mar Rojo*. Outros agrupamentos son os que establecemos entre a denominación simple e complexa de embarcacións: *mercante* e *barco mercante* baixo (*barco*) *mercante*; *veleiro*, *barco veleiro* e *barco de vela* baixo *barco (veleiro)* ou *remolcador* e *barco remolcador* baixo (*barco*) *remolcador*. Por outra banda, non agrupamos *pesqueiro* e *barco pesqueiro* entendendo que a primeira palabra pode ser un adxectivo cun valor xeral (“relativo á pesca”).

Tamén levamos a cabo o agrupamento con parénteses en varios casos nos que a expresión complexa porta o modificador *mariño/a* ou *de mar*, entendendo que o ámbito temático do CI restrinxe as posibles interpretacións da voz sen especificador: *brisa* e *brisa mariña*, *corrente* e *corrente mariña*, *fauna* e *fauna mariña*, *estrela* e *estrela de mar / estrelamar*, *serpe* e *serpe mariña...* Pola contra, non empregamos a paréntese cando a forma sen modificador resulta estraña: *león mariño* e *león* lematízanse baixo *león mariño*. Outros agrupamentos son: *castelo* e *castelo de area* (pois o informante que proporciona *castelo* elabora unha listaxe centrada na súa totalidade na praia) ou *gafas* e

gafas de buceo (no contexto *submarinista*, *gafas*, *osíxeno*). Pola contra, non unificamos *cana* e *cana de pescar*, porque o substantivo ten outras acepcións que encaixan neste CI (a *cana* é, por exemplo, unha parte da áncora ou unha parte do remo) nin *nó* e *nó mariñeiro*, posto que a primeira palabra pode facer referencia a unha unidade empregada para medir a velocidade dos barcos.

Dentro do grupo dos peixes, cómpre sinalar que lematizamos *sapo* e *peixe sapo* separadamente porque, tal e como indicamos en § 2.3.4.5, a primeira palabra pode designar nalgunhas zonas de Galicia peixes diferentes á rabada. Si agrupamos *espada* e *peixe espada* por unha banda e *porco*⁶⁶ e *peixe porco* por outra, sen utilizar parénteses, porque consideramos que os informantes nun e noutro caso se están a referir ao mesmo animal e porque a denominación estándar inclúe o substantivo *peixe*. No sentido contrario, lematizamos *tiburón* e *peixe tiburón* baixo a forma simple e estándar *tiburón*. Por outra banda, o contexto (*peixe espada*, *peixe tiburón*, *balea*) parece desbotar por completo que o informante estea designando con *peixe tiburón* unha especie propia de acuario, de pequeno tamaño.

No que atange a variacións formais de xénero, cómpre sinalar que lematizamos separadamente formas como *patrulleira* e *patrulleiro*, posto que a forma feminina probablemente fai referencia a unha lancha e a masculina pode indicar unha embarcación de maior tamaño. En relación co plural, lematizamos separadamente *aparells* e *aparello (de pesca)*. A primeira expresión, en plural, e segundo recollen os dicionarios, pode facer referencia ao ‘conxunto de instrumentos e utensilios necesarios para realizar unha actividade’ (DRAG²), neste caso para pescar.

Canto ás expresións con afixos apreciativos, prescindimos destes nos lemas de formas como *barcucho*, *conchiñas*, *ondiña*, *pedriñas*, *pedrols...* pero si o mantemos en expresións que consideramos lexicalizadas como *cabaliño de mar* (máis frecuente nas respostas ca *cabalo de mar*), *chiringuito (de praia)*, *culleriña*, *gambón*, *illote* ou *pescadilla*.

2.4 Tratamento informático dos datos

2.4.1 O proxecto panhispánico

Como indicamos en § 1, os estudos sobre dispoñibilidade léxica no ámbito hispánico inspíranse nos elaborados en volta da lingua e a cultura francesas que se remontan aos anos cincuenta do século XX, e en esencia empregan a súa

⁶⁶ Non o interpretamos como adxectivo (‘sucio’) porque a palabra aparece nunha listaxe na que non hai ningunha referencia á sucidade ou á contaminación e, pola contra, si figuran moitas denominacións de especies mariñas.

metodoloxía. Non obstante, esta metodoloxía tarda décadas en ofrecer unha análise afinada da dispoñibilidade, posto que, para o seu cálculo, en principio utiliza exclusivamente o factor da frecuencia, de xeito que as palabras con maior grao de dispoñibilidade son simplemente aquelas que aparecen máis veces nas listaxes elaboradas polos informantes. Con todo, cómpre ter en conta que o concepto de dispoñibilidade está indisolublemente unido ao de espontaneidade, no sentido de que as unidades léxicas máis dispoñibles son as que acoden á memoria máis rapidamente trala presentación do estímulo (o CI), polo que as que ocupan os primeiros postos nesas listaxes son as que resultan dispoñibles en máis alto grao. Isto quere dicir que a orde de aparición dos vocábulos debería constituír un factor tan relevante como o da frecuencia.

López Morales (1999) sinala que, aínda que desde as primeiras publicacións son varios os autores que chaman a atención sobre a necesidade de considerar ambos os dous factores, non é ata a década dos oitenta cando se empezan a deseñar fórmulas matemáticas capaces de ponderar, xunto coa frecuencia, o lugar de aparición de cada unidade léxica nas listaxes, isto é, fórmulas que non se limitan a contabilizar as ocorrencias de cada unidade, senón que o fan en cada posición que esta ocupa.

Seguindo na mesma liña, no proxecto panhispánico calcúlase o índice de dispoñibilidade léxica empregando unha fórmula, ideada por López Chávez / Strassburger Frías (1991), que combina varios factores: a frecuencia absoluta de cada palabra en cada posición; a frecuencia absoluta de cada palabra, resultado de sumar as diferentes frecuencias nas distintas posicións; o número de informantes que participan na enquisa; o número de posicións acadadas en cada centro de interese (n) e as posicións en que aparece cada palabra (que poden ir de 1 a n). Esta fórmula é a que se aplica nos dous programas informáticos utilizados no proxecto panhispánico: *LexiDisp*, patrocinado polo Instituto Cervantes, a ALFAL e a Universidade de Alcalá de Henares, e *Dispoplex*, deseñado para a Universidade de Salamanca. Ambos coñecen varias versións, elaboradas para mellorar as carencias que o seu uso pon en evidencia e para aumentar as posibilidades de acceso e manipulación dos datos.

O segundo, *Dispoplex*, é o maioritariamente empregado nos últimos anos e constitúe o banco de datos do proxecto panhispánico, do que sairá o futuro dicionario do léxico dispoñible do español. Como se indica na páxina de presentación (www.dispoplex.com), proporciona as ferramentas necesarias para realizar os cálculos máis habituais en dispoñibilidade léxica, entre os que podemos destacar o índice de dispoñibilidade, a frecuencia dunha unidade léxica nun CI, a porcentaxe de aparición, o número de palabras por CI, o promedio de produción de palabras por informante e CI ou o índice de cohesión (*vid.* § 3.1.2).

No subproxecto galego realizamos a parte relativa á lingua castelá empregando este programa, aínda que a partir do ano 2010 contamos cun programa propio, *Dispogal* (<http://www.gramatica.usc.es/disvogal/>), no que envorcamos os datos desta lingua e introducimos os da lingua galega.

2.4.2 O proxecto galego

Dispogal é unha adaptación de *Dispoplex* ás necesidades do proxecto galego, de xeito que, aínda que en esencia emprega a mesma estrutura, resulta, na nosa opinión, máis versátil, xa que facilita e acelera as buscas e ofrece máis posibilidades na extracción da información.

No que atinxe á dispoñibilidade xeral, as buscas realizadas sobre cada CI dan como resultado unhas táboas que recollen esquematicamente os datos máis relevantes de cada unidade léxica. Como exemplo, reproducimos a seguir a parte alta da táboa a través da cal a aplicación amosa os datos relativos ao CI 14 ‘Os animais’:

Táboa 16. Datos sobre dispoñibilidade xeral do CI 14 (fragmento)

Nº	Palabra	Dispoñib.	Aparición	Frec. absoluta	Frec. relativa	Frec. acumulada
1	gato-a	0,74731	89,625%	717	3,404%	3,404%
2	can	0,73275	84,250%	674	3,200%	6,604%
3	vaca	0,57272	75,875%	607	2,882%	9,486%
4	cabalo	0,41692	59,250%	474	2,250%	11,736%
5	porco-a	0,38341	56,000%	448	2,127%	13,863%
6	galiña	0,33520	52,625%	421	1,999%	15,862%
7	león-oa	0,31724	54,500%	436	2,070%	17,932%
8	ovella	0,31504	49,750%	398	1,889%	19,821%
9	coello	0,29566	47,375%	379	1,799%	21,620%
10	tigre	0,23586	43,750%	350	1,662%	23,282%

Previamente a esta información, o usuario accede a datos máis xerais: sabe que a media de respostas por informante no CI 14 é 26,33, que o total de informantes é 800 e que o total de respostas, é dicir, o total de palabras do CI, é 21.064. Obviamente, na derradeira fila tamén coñece o número de lemas (neste caso, 1010)⁶⁷.

A primeira columna recolle o rango que ocupa cada unidade léxica na listaxe de dispoñibilidade; a segunda, o lema que representa cada unidade

⁶⁷ Se seleccionase, por exemplo, os informantes de sexo masculino cuxa lingua habitual é exclusivamente o galego, descubriría que a media de respostas é de 27,29, que o total de informantes é 45 e que o total de respostas é 1.228. Indo á parte final da táboa sabería que a listaxe está composta por 350 lemas.

léxica; a terceira, o índice de dispoñibilidade; a cuarta, a porcentaxe de aparición da unidade en cuestión (ou a porcentaxe de enquisas en que aparece); a quinta, a frecuencia absoluta⁶⁸; a sexta, a frecuencia relativa de cada palabra sobre o total (por exemplo, as 717 ocorrencias de *gato-a* constitúen o 3,404% de 21.064; as 674 de *can* o 3,2%; etc.); a sétima, a frecuencia acumulada, que como é obvio aumenta a medida que se engaden palabras (por exemplo, as cinco primeiras unidades léxicas, *gato-a*, *can*, *vaca*, *cabalo* e *porco-a*, acumulan 2.920 ocorrencias (717 + 674 + 607 + 474 + 448), o que constitúe o 13,863% do total).

Toda esta información –tamén a forma de ofrecela por parte da aplicación e a forma de seleccionala por parte do usuario– non varía, en esencia, respecto á que podemos atopar en *Dispoplex*, programa do que *Dispogal* só se diferencia, como dixemos, na súa maior versatilidade. Porén, o subproxecto galego conta con outro programa informático, elaborado e manipulado por Guillermo Rojo, ideado para facilitar o proceso de transcrición dos datos a soporte informático e, sobre todo, para favorecer e mellorar o traballo de lematización.

En primeiro lugar, o programa está deseñado para detectar erros de diverso tipo e repeticións, o cal alixeira e mellora a tarefa da edición; en segundo lugar, permite seleccionar o material empregando diferentes criterios, o cal facilita a súa manipulación; en terceiro lugar –e máis importante–, permite integrar todas as formas conectadas co lema normalizado correspondente, o cal é fundamental na tarefa de lematización.

O seguinte exemplo representa un pequeno fragmento dunha enquisa tal e como foi transcrita para ser sometida á primeira revisión. Como é fácil observar, cada parágrafo recolle as formas léxicas dun CI, e na súa parte inicial aparecen tres díxitos. Estes rexistran os seguintes datos de identificación: 019 representa un centro de ensino (neste caso, o IES Punta Candieira de Cedeira; *vid.* táboa 1); 01 representa un alumno concreto dese centro; o dígito final representa un CI (03 ‘As partes da casa’, 04 ‘Os mobles da casa’ etc.)⁶⁹. Por outra banda, sempre que a palabra proporcionada polo informante non coincide co lema, este inclúese tras unha barra oblicua; faise así, por exemplo, coas formas plurais (*portas / porta*, *ventás / ventá*, *sardiñas / sardiña...*), coas abreviaturas (*tv / televisión*), coas formas castelás cogna-

⁶⁸ Ofrece a mesma información da columna anterior en cifras absolutas; por exemplo, *gato-a* recóllese en 717 das 800 enquisas da mostra, o cal equivale ao 89,625%. Se, como na nota a pé anterior, seleccionásemos aos informantes masculinos que só falan galego, observaríamos que a primeira palabra, *can*, foi anotada por 41 dos 45 rapaces, isto é, polo 91,111%.

⁶⁹ Noutra parte do programa almacénanse os datos de estratificación social dos informantes.

das (*tuberías / tubaxe, nevera / neveira, costilletas / costeleta*), coas formas castelás galeguizadas (*desaugues / desaugadoiro, naranxas / laranxa*) etc. Máis adiante, en sucesivas fases de revisión, van propoñéndose novos lemas; por exemplo, a forma *vertedeiro* que aparece entre as últimas palabras do primeiro parágrafo será substituída por *vertedeiro / vertedoiro*).

019 - 01 - 03 - ladrillo - habitación - portas/porta - ventás/ventá - cemento - estrutura/estrutura - columnas/columna - zapatas/zapata - placa - desaugues/desaugadoiro - tuberías/tubaxe - perlita - falso teito - ducha - lavabo - váter - pedra - pintura - plaqueta - madeira - tarima - vertedeiro - caldeira - cristais/cristal

019 - 01 - 04 - nevera/neveira - microondas - forno - cociña - mesa - sillas/silla - estanterías/estantería - cama - armario - moble do salón - tv/televisión - sofá - televisión/televisión - coqueta - piano

019 - 01 - 05 - auga - leite - viño - guiso - caldo - croquetas/croqueta - costilletas/costeleta - chourizos/chourizo - churrasco - sardiñas/sardiña - grelos/grelo - leituga - verzas/verza - rabizas/rabiza - nabos/nabo - mazás/mazá - melocotóns/melocotón - plátanos/plátano - naranxas/laranxa - ovos/ovo - arroz - pan - fabada - lentellas/lentella

019 - 01 - 06 - culler - coitelo - tenedor - abridor - prato fondo - prato chan - fonte - pota - vasos/vaso - tazóns/tazón - pocillos/pocillo

A identificación que se realiza cos primeiros díxitos permite organizar o material de maneira diferente, de modo que en fases posteriores a revisión pode facerse por CI, o cal propicia unha lematización máis axeitada (para iso, xúntanse todos os parágrafos que teñen o mesmo díxito no terceiro lugar). Ao respecto, convén lembrar que, como xa sinalamos en § 2.3.2, neste proxecto de investigación léxica os lemas son unidades asociadas aos diferentes estímulos que serviron para xeralas, polo que podelos agrupar en volta de tales estímulos, os CI, favorece a coherencia da lematización. Así, en fases posteriores manexamos listaxes de variantes e lemas xeradas polo programa, das que a continuación ofrecemos, como exemplo, un fragmento da listaxe correspondente ao CI 18 ‘O mar’ ordenada alfabeticamente:

18 bogavante	15
18 bogabante.....	8
18 bogavante.....	6
18 vogabante.....	1
18 boi de Francia	2
18 boi de Francia	2

18 boi (de mar)	44
18 boi	21
18 boi de mar	2
18 bois	6
18 boy	3
18 buei	8
18 bueis	1
18 buey	2
18 buey de mar	1
18 boia	56
18 boia	18
18 boias	12
18 bolla	6
18 bollas	14
18 boya	1
18 boyas	5
18 boliche	1
18 boliche.....	1

A listaxe recolle, debaixo de cada lema, as distintas variantes vinculadas con el, e informa sobre a frecuencia absoluta tanto do lema como de cada unha das formas. Por exemplo, sabemos que só seis das quince ocorrencias de *bogavante* teñen a ortografía normalizada, ou que seis das cincuenta e seis ocorrencias de *boia* se escribiron con <y> e vinte con <ll>, ou que hai dúas ocorrencias de *boi de Francia* e unha de *boliche*.

É fácil inferir que o emprego destas listaxes axuda enormemente na lematización, non só porque facilita a detección de erros, senón tamén porque a información sobre frecuencias das distintas variantes pode empregarse, nos casos nos que se estime oportuno, na selección do lema (*vid.* § 2.3.2). Pero máis alá disto, cremos que a agrupación das variantes baixo un mesmo lema ofrece unha información ben interesante que pode abrir o camiño a investigacións de diverso tipo, polo que, como xa indicamos, decidimos incluír esta información en apéndice. Lembremos a este respecto que baixo un mesmo lema non só agrupamos variantes flexivas ou ortográficas, senón tamén variantes fónicas, morfolóxicas e mesmo cognados do castelán.

3 ANÁLISE DE RESULTADOS

As aplicacións informáticas empregadas para recompilar os datos das enquisas das diferentes sintopías do proxecto panhispánico (nomeadamente, *LexiDisp* e *Dispolex*) non se limitan á edición dos materiais, senón que, como acabamos de comentar, ofrecen tamén a posibilidade de realizar unha serie de cálculos estándares nos estudos de dispoñibilidade léxica, como son o reconto de palabras totais en cada CI, o de palabras diferentes ou lemas, os promedios de produción de palabras, o índice de cohesión, a compatibilidade entre as listaxes de lemas producidos por distintas poboacións establecidas a través das variables de estratificación social etc.

Para manter a necesaria homoxeneidade metodolóxica no marco do proxecto panhispánico, segundo acabamos de dicir o programa *Dispogal* da Universidade de Santiago de Compostela está elaborado tomando como modelo *Dispolex*, aínda que, ademais de engadir ás xa establecidas as variables específicas do subproxecto galego, este programa permite realizar máis combinacións nas buscas, o que redunda nunha maior flexibilidade na selección dos distintos valores e por tanto na posibilidade de obter máis información.

Neste capítulo presentaremos brevemente os datos globais correspondentes ao léxico galego dispoñible, comezando pola información relativa aos centros de interese e continuando coa análise da incidencia das variables de estratificación social na produción de lemas e coa avaliación do grao de compatibilidade léxica entre distintas poboacións da mostra. No que atinxe aos CI, aínda que a información será de carácter basicamente cuantitativo, cando menos poderemos ofrecer unhas pinceladas a través das cales se caractericen como ámbitos temáticos máis ou menos compactos; no que se refire ás variables, constataremos se hai trazos sociolingüísticos que incidan de xeito relevante na produción das palabras.

3.1 Os centros de interese. Índices xerais de cuantificación

Comezaremos a análise ofrecendo información sobre a produción de unidades léxicas dos preuniversitarios galegos en todos e cada un dos dezaoto CI, para o que contabilizaremos as palabras totais (en adiante, palabras), as palabras distintas (os lemas) e os promedios de produción de palabras por estudante e CI; a estas cifras engadiremos as relativas ás palabras proporcionadas por un único informante e as das palabras que acadan índices de dispoñibilidade superiores a 0,1, e toda esta información permitirá caracterizar minimamente as distintas áreas de significado.

Aínda que non é a nosa intención ofrecer un estudo contrastivo co léxico dispoñible do español de Galicia, haberá unha referencia ao LDEG, basicamente para comprobar se é posible detectar diferenzas significativas entre as dúas linguas nos distintos dominios conceptuais (vid. 3.1.4).

3.1.1 Número de palabras

Os oitocentos informantes seleccionados proporcionaron centos de miles de palabras que, tralo proceso de revisión das enquisas (vid. § 2.3), quedaron cifradas en 272.758; noutros termos, podemos dicir que cada estudante produciu unha media de case que 341 palabras. Así mesmo, o número de lemas da nosa mostra é de 18.279, pero iso non significa que a mostra estea constituída por preto de 18.300 elementos léxicos diferentes, xa que, como indicamos no seu momento, neste traballo empregamos o termo *lema* para facer referencia a unha unidade adscrita a un CI determinado, e non para identificar unha unidade do lexicón global. Naturalmente, este feito trae consigo inevitables repeticións, sobre todo nos CI que constitúen áreas temáticas moi próximas entre si, como por exemplo os dedicados ao campo (CI 11) e aos traballos do campo e do xardín (CI 13), que amosan unha coincidencia do 18,17% dos lemas (entre eles, *abeto, abonar, barbeito, boi, campesiño, cultivar, finca, maceira, ovella, pa, recolleita* etc.).

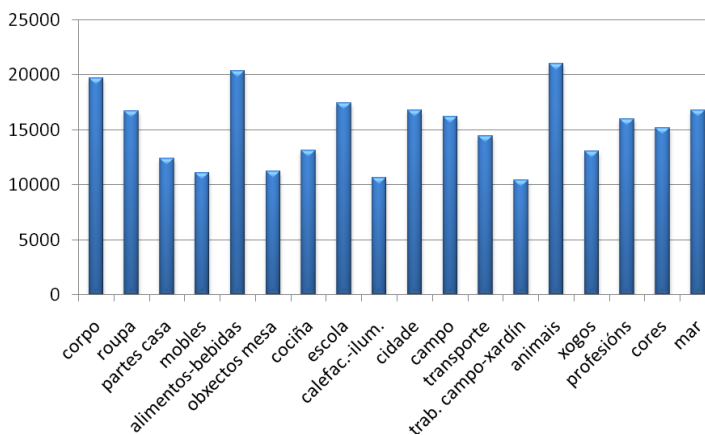
Na táboa 17 ofrecemos os datos cuantitativos de produción de palabras nos dezaoto CI: a segunda columna recolle o número de palabras contabilizadas en cada CI; a terceira, o promedio de palabras emitidas por informante (ou o resultado de dividir o total de palabras en cada CI entre oitocentos, o número de informantes); a cuarta, o número de palabras distintas ou lemas de cada CI. Se contrastamos estas cifras cos promedios ofrecidos na derradeira fila, comprobaremos que o reparto de unidades léxicas é moi desigual nas diferentes áreas temáticas:

Táboa 17. Datos xerais sobre produción de palabras e lemas

Centros de interese	Total de palabras	Promedio de palabras	Lemas
01 O corpo humano	19753	24,69	858
02 A roupa	16707	20,88	790
03 As partes da casa (sen mobles)	12380	15,48	677
04 Os mobles da casa	11109	13,89	451
05 Os alimentos e as bebidas	20342	25,43	1078
06 Os obxectos colocados na mesa...	11248	14,06	505
07 A cociña e os seus utensilios	13132	16,42	720
08 A escola: mobles e materiais	17477	21,85	959
09 A calefacción e a iluminación	10640	13,30	945
10 A cidade	16803	21,00	1330
11 O campo	16214	20,27	1626
12 Os medios de transporte	14449	18,06	963
13 Os traballos do campo e do xardín	10410	13,01	1503
14 Os animais	21064	26,33	1010
15 Os xogos e as distraccións	13041	16,30	1594
16 As profesións e os oficios	15967	19,96	1284
17 As cores	15192	18,99	667
18 O mar	16830	21,04	1319
Promedios	15153,22	18,94	1015,5

Efectivamente, observamos que hai dez CI que superan o promedio de produción de palabras (que se cifra en algo máis de 15.150 por CI) e oito que están por debaixo del; ademais, a diferenza entre o máis voluminoso —o 14, dedicado aos animais, de 21.064 palabras— e o máis pequeno —o 13, dedicado aos traballos do campo e do xardín, de 10.410 palabras— é de algo máis do dobre. O seguinte gráfico permite visualizar con claridade as diferenzas:

Gráfico 21. Número total de palabras por CI



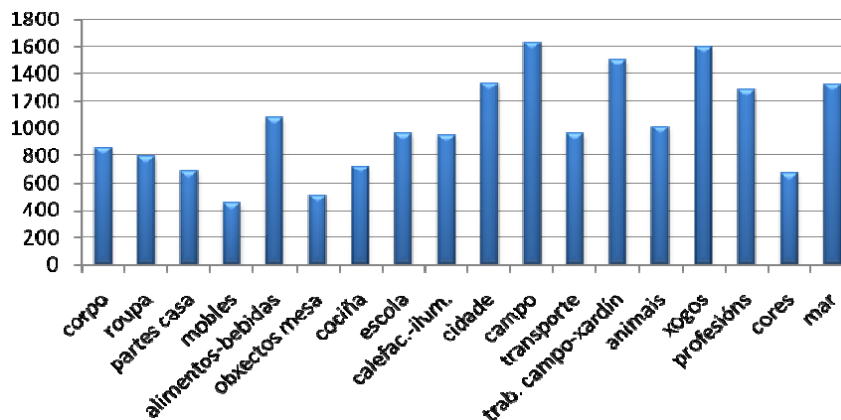
Na táboa 18 ofrecemos os CI ordenados segundo a produción de palabras: a primeira columna indica o rango que ocupa cada CI e a cuarta, a porcentaxe de palabras que achega cada área temática ao total da mostra.

Táboa 18. Nivel dos CI por produción de palabras

Rango	Centro de interese	Nº palab.	Distr. %
1	14 Os animais	21064	7,72
2	05 Os alimentos e as bebidas	20342	7,46
3	01 O corpo humano	19753	7,24
4	08 A escola: mobles e materiais	17477	6,41
5	18 O mar	16830	6,17
6	10 A cidade	16803	6,16
7	02 A roupa	16707	6,13
8	11 O campo	16214	5,95
9	16 As profesións e os oficios	15967	5,85
10	17 As cores	15192	5,57
11	12 Os medios de transporte	14449	5,30
12	07 A cociña e os seus utensilios	13132	4,81
13	15 Os xogos e as distraccións	13041	4,78
14	03 As partes da casa (sen mobles)	12380	4,54
15	06 Os obxectos colocados na mesa...	11248	4,12
16	04 Os mobles da casa	11109	4,07
17	09 A calefacción e a iluminación	10640	3,90
18	13 Os traballos do campo e do xardín	10410	3,82
Totais		272758	100

Pero as disimilitudes entre os CI son aínda máis marcadas se comparamos o número de lemas, xa que neste caso o máis voluminoso —o 11, no que se contabilizan 1.626 lemas relativos ao campo— é 3,6 veces maior ca o máis pequeno —o 04, ‘Os mobles da casa’, que suma 451 lemas—. A este dato hai que engadir que só superan a media sete CI, tres dos cales se sitúan moi por riba dela, con máis de 1.500 lemas: o antedito CI 11 e os CI 15 ‘Os xogos e as distraccións’ e 13 ‘Os traballos do campo e do xardín’.

Gráfico 22. Número de lemas por CI



A táboa 19 recolle a mesma información ca a 18, pero agora referida ao número de lemas por CI: na columna da dereita vese con claridade que a distribución porcentual de lemas é, como acabamos de indicar, máis desigual ca a das palabras totais, xa que oscila entre o 8,90% e o 2,47% (fronte ao 7,72% e o 3,82% destas).

Táboa 19. Nivel dos CI por produción de lemas

Rango	Centro de interese	Nº lemas	Distr. %
1	11 O campo	1626	8,90
2	15 Os xogos e as distraccións	1594	8,72
3	13 Os traballos do campo e do xardín	1503	8,22
4	10 A cidade	1330	7,28
5	18 O mar	1319	7,22
6	16 As profesións e os oficios	1284	7,02
7	05 Os alimentos e as bebidas	1078	5,90
8	14 Os animais	1010	5,52
9	12 Os medios de transporte	963	5,27
10	08 A escola: mobles e materiais	959	5,25
11	09 A calefacción e a iluminación	945	5,17
12	01 O corpo humano	858	4,69
13	02 A roupa	790	4,32
14	07 A cociña e os seus utensilios	720	3,94
15	03 As partes da casa (sen mobles)	677	3,70
16	17 As cores	667	3,65
17	06 Os obxectos colocados na mesa...	505	2,76
18	04 Os mobles da casa	451	2,47
Totais		18279	100

Por outra banda, contrastando os datos da primeira columna das táboas 18 e 19 comprobamos que en boa parte das áreas temáticas as cantidades das palabras e dos lemas están equilibradas, xa que a maior número de palabras, maior número de lemas no mesmo CI, e viceversa: dunha banda, nos CI 05 ‘Os alimentos e as bebidas’, 18 ‘O mar’, 10 ‘A cidade’, 11 ‘O campo’ e 16 ‘As profesións e os oficios’ contabilizamos o número máis elevado tanto de palabras como de lemas (por iso ocupan os primeiros rangos en ambas as dúas táboas); e por outra, nos CI 09 ‘A calefacción e a iluminación’, 04 ‘Os mobles da casa’, 06 ‘Os obxectos colocados na mesa para a comida’, 03 ‘As partes da casa’, 07 ‘A cociña e os seus utensilios’ e 12 ‘Os medios de transporte’ detectamos o número máis pequeno de palabras e de lemas (de aí que se sitúen na parte baixa das táboas). No entanto, tal correlación non se dá no resto dos CI, xa que os hai que exhiben cifras elevadas de produción de palabras fronte a un número relativamente pequeno de lemas (os CI 14 ‘Os animais’, 01 ‘O corpo humano’, 08 ‘A escola’, 02 ‘A roupa’ e 17 ‘As cores’) e, polo contrario, algúns son parcos en produción de palabras e extensos en número de lemas (os CI 15 ‘Os xogos e as distraccións’ e 13 ‘Os traballos do campo e do xardín’), razón pola cal presentan rangos moi diferentes nas táboas 18 e 19. Ofrecemos a continuación outra que repite esta información, para facilitar a visualización das diferenzas:

Táboa 20. Rango dos CI por produción de palabras e de lemas

Centro de interese	Palabras	Lemas
01 O corpo humano	3	12
02 A roupa	7	13
03 As partes da casa (sen mobles)	14	15
04 Os mobles da casa	16	18
05 Os alimentos e as bebidas	2	7
06 Os obxectos colocados na mesa...	15	17
07 A cociña e os seus utensilios	12	14
08 A escola: mobles e materiais	4	10
09 A calefacción e a iluminación	17	11
10 A cidade	6	4
11 O campo	8	1
12 Os medios de transporte	11	9
13 Os traballos do campo e do xardín	18	3
14 Os animais	1	8
15 Os xogos e as distraccións	13	2
16 As profesións e os oficios	9	6
17 As cores	10	16
18 O mar	5	5

Os CI 13 e 15 destacan sobre todos os demais por seren os que acadan os rangos máis dispare, que son, como acabamos de indicar, altos na produción de lemas e baixos na de palabras (a diferenza é de 15 puntos no caso do CI 13 e de 11 puntos no caso do CI 15), o que se traduce nunha presenza significativa nestes CI de elementos léxicos proporcionados por moi poucos informantes, a miúdo mesmo por un só. Este contraste no número de lemas e palabras suxire que as áreas temáticas correspondentes son amplas pero non resultan especialmente evocadoras para os nosos informantes, que se mostraron moito máis produtivos noutros dominios léxicos, como o dos animais, os alimentos ou o corpo. Neste último caso, ademais, a situación é a contraria á dos CI 13 e 15, xa que o número de lemas non é especialmente alto (858), polo que o CI acada o rango 12 (hai unha diferenza de 9 puntos na produción de palabras e de lemas). Isto significa que estamos ante unha área temática máis ben compacta, na que se contabiliza un elevado número de palabras frecuentes, e suficientemente coñecida polos nosos informantes como para que as súas listaxes estean ben nutridas de elementos léxicos (na táboa 17 vemos que o promedio de palabras por informante é de 24,69). Non hai que esquecer, ao respecto, que recollemos o léxico dispoñible dos estudantes de 2º curso de bacharelato, e que o facemos no contexto do centro de ensino, todo o cal propicia a inclusión nas listaxes de palabras especializadas, que en moitas ocasións acadan índices de dispoñibilidade significativos: *cerebelo, cuadriceps, cúbito, esófago, esternocleidomastoideo, falanxe, farinx, larinx, neurona, rótula, tríceps, úmero* etc.

No extremo contrario ao dos CI 13 e 15 atópase o dedicado ao mar, xa que este ámbito do léxico resulta propicio tanto para a produción de palabras como de lemas, e de feito o CI 18 acada o rango 5 en ambos os dous casos. Xunto a el sitúase o CI 10, ámbito tamén caracterizado pola alta produción de palabras e de lemas. No que se refire a estes dous CI, semella que estamos ante amplas zonas do léxico que os nosos informantes dominan; ao respecto non hai que esquecer que a maior parte dos centros de ensino se sitúa en zonas urbanas ou periurbanas (o 66%) e costeiras (perto do 57%), e parece indubidable que o hábitat no que nos desenvolvemos os falantes incide de xeito importante no tipo de léxico que coñecemos e empregamos habitualmente.

Por outra banda, na táboa 20 vese con claridade que os CI con menores diferenzas na emisión de palabras e lemas (aqueles cuxos rangos non se separan máis de dous puntos) son, salvo os citados CI 18 e 10 —e talvez o CI 12 ‘Os medios de transporte’, que se situaría nunha zona media—, menos produtivos: os CI 04 ‘Os mobles da casa’, 06 ‘Os obxectos colocados na mesa para a comida’ e, en menor medida, 07 ‘A cociña e os seus utensilios’ son ámbitos do léxico moi compactos e limitados, ámbitos que por tanto evocan menos unidades léxicas ca outros máis abertos e/ou amplos.

Finalmente, cabe destacar a situación dos CI 11 ‘O campo’ e 14 ‘Os animais’, que presentan rangos bastante distanciados entre si (sete puntos) pero en sentido contrario. O CI 14 é o primeiro en produción de palabras (21.064) e o oitavo en número de lemas (1.010), o cal suxire que se trata dunha área temática ampla pero compacta e sobre a que os nosos informantes demostran ter dominio; neste caso, ademais, por riba do coñecemento que implica a localización en determinado hábitat: así, por exemplo, entre os vocábulos que exhiben altos índices de dispoñibilidade detectamos tanto nomes de animais silvestres da fauna galega (*aguia, gaivota, lobo, mosca, pomba, ra, raposo, rato, sapo, troita, xabaril* etc.) como os de animais alleos ao noso medio físico (*balea, cebra, crocodilo, elefante-a, león-oa, leopardo, mono, rinoceronte, tigre, xirafa* etc.), do que cabe inferir que neste CI inflúe máis ben a idade e a formación dos nosos informantes. Por outra parte, o CI 11 é o primeiro no número de lemas (1.626), e o oitavo en produción de palabras (16.214), e estes datos apuntan a un ámbito do léxico de gran amplitude (*vid. § 2.3.4.11*) pero menos evocador ca outros para os nosos informantes.

Estas diferenzas, esbozadas en trazos moi grosos, poden afinarse un pouco máis se incluímos na análise outros parámetros, relacionados co grao de dispoñibilidade dos vocábulos. Pero antes referirémonos ao índice de cohesión, valor normalmente manexado nos estudos sobre dispoñibilidade léxica para distinguir entre CI compactos e difusos.

3.1.2 Índice de cohesión

No que se refire ás análises de tipo cuantitativo, na metodoloxía básica da dispoñibilidade léxica destacan o índice de cohesión e a densidade, parámetros ambos que miden o grao de coincidencia entre as respostas dos informantes en cada CI (*vid. Hernández Muñoz 2004*). A densidade calcúlase dividindo o total de palabras entre o número de lemas ou palabras diferentes, mentres que o índice de cohesión obtense dividindo o promedio de respostas por suxeito entre o número de lemas. Resulta evidente que, dado que as cifras dos dous dividendos se relacionan directamente (lembramos que o promedio é, á súa vez, o resultado de dividir o número de palabras entre o número de informantes) e que o divisor é o mesmo en ambos os casos, os resultados serán equivalentes, polo que índice de cohesión e densidade ofrecen, en realidade, a mesma información. Con todo, na bibliografía adoita dárselle maior relevancia ao índice de cohesión, xa que se considera un xeito de calcular o grao de coincidencia máis elaborado ca a densidade; esta, dise, presenta unha maior limitación no que atinxe á comparación dos datos, xa que carece de teitos ou

topes referenciais: crece canto maior é o número de palabras totais, e pode facelo ata un nivel indeterminado (Hernández Muñoz 2004: 57); fronte a ela, o índice de cohesión ten como tope o 1, cifra que se acadaría no hipotético caso de que nun CI o número de lemas coincidise co número de palabras.

Aínda que en principio este tope (1) semella anular as diferenzas que poden darse entre mostras de tamaños moi dispares, o certo é que o índice de cohesión tamén se ve afectado, como a densidade, polo tamaño da mostra, diminuíndo canto maior sexa esta. Podemos comprobar este extremo se comparamos os datos de distintos grupos dunha mesma mostra establecidos segundo os parámetros de estratificación social, sempre e cando, claro, tales grupos teñan tamaños ben diferenciados; por exemplo, o grupo de estudantes dos centros de ensino público e privado da nosa mostra (672 e 128, respectivamente): tomando só os datos dos cinco primeiros CI, comprobamos que os índices de cohesión son 0,3 - 0,3 - 0,2 - 0,3 - 0,2 no primeiro grupo e 0,7 - 0,6 - 0,6 - 0,7 - 0,6 no segundo. Resulta evidente que a conclusión non pode ser que a poboación constituída polos informantes dos centros privados xera listaxes moito máis cohesionadas ca a dos informantes dos centros públicos; en realidade, a comparación debería facerse considerando as cifras dentro de cada grupo, polo que o único dato destacable neste caso sería a cifra acadada polo CI 02, algo máis cohesionado relativamente no primeiro grupo (cun valor de 0,3) ca no segundo (cun valor de 0,6)⁷⁰.

Sexa como fose, non faltan os autores que clasifican os CI en virtude deste índice cuantitativo, cualificándoos de máis compactos ou máis difusos semanticamente segundo o seu índice de cohesión se achegue máis ou menos a 1. Por exemplo, Gómez Molina / Gómez Devís (2004: 83), despois de calcular o índice de cohesión dos CI, divídenos en tres grupos:

- i. Campos semanticamente máis compactos (presentan maior concreción semántica, maior asociación conceptual).
- ii. Áreas temáticas cun grao medio de coincidencia nas respostas.
- iii. Campos semanticamente máis difusos.

Para comprobar se tal clasificación é factible e, sobre todo, se resulta útil para describir os CI da nosa mostra, partiremos dos datos expostos na táboa 21, que ordena os CI de maior a menor cohesión e densidade:

⁷⁰ Así o explica Hernández Muñoz (2004: 58): “Este índice no es ajeno a la influencia de la fluctuación en el número de informantes ya que el número de palabras diferentes, que funciona como cociente, también aumenta conforme crece el número de sujetos encuestados, por lo que cuanto más ambiciosa sea una muestra los valores del índice de cohesión tenderán a ser menores”.

Táboa 21. Índices de cohesión e de densidade

Centros de interese	Cohesión	Densidade
04 Os mobles da casa	0,031	24,63
01 O corpo humano	0,029	23,02
17 As cores	0,028	22,78
06 Os obxectos colocados na mesa...	0,028	22,27
02 A roupa	0,026	21,15
14 Os animais	0,026	20,86
05 Os alimentos e as bebidas	0,024	18,87
03 As partes da casa (sen mobles)	0,023	18,29
07 A cociña e os seus utensilios	0,023	18,24
08 A escola: mobles e materiais	0,023	18,22
12 Os medios de transporte	0,019	15,00
18 O mar	0,016	12,76
10 A cidade	0,016	12,63
16 As profesións e os oficios	0,016	12,44
09 A calefacción e a iluminación	0,015	11,26
11 O campo	0,012	9,97
15 Os xogos e as distraccións	0,010	8,18
13 Os traballos do campo e do xardín	0,009	6,93

Se aplicásemos a clasificación proposta, diríamos que os CI 04, 01, 17, 06, 02 e 14 son as áreas temáticas máis compactas semanticamente, que os CI 13, 15, 11, 09, 16, 10 e 18 son as máis difusas e que o resto ocupa a posición intermedia. Non parece, porén, que esta división sexa a máis adecuada para a nosa mostra, basicamente porque os CI intermedios están máis preto do primeiro grupo ca do segundo: en tres destes cinco CI o índice de cohesión é 0,023, cifra bastante afastada da correspondente ao CI que presenta unha maior cohesión dentro do grupo dos difusos (0,016) pero próxima á correspondente ao CI que presenta unha menor cohesión dentro do grupo dos compactos (0,026). Na nosa opinión, de facerse unha división deste tipo, na mostra do léxico dispoñible do galego só se diferenciarían dous grupos, o dos CI compactos, con índices de cohesión situados por riba da media (0,021), entre 0,031 e 0,023, e o dos CI difusos, con índices de cohesión entre 0,019 e 0,009.

De feito, os CI do primeiro grupo, en xeral, teñen denominacións que evitan a elaborar listaxes de voces baseadas nas relacións de hiponimia (os mobles, as cores, os animais, a roupa etc.) ou de meronimia (o corpo humano, as partes da casa) e moitas delas mesmo constitúen agrupacións de unidades léxicas analizadas como campos léxicos (as partes da casa, a roupa, os

alimentos) ou son taxonomías coñecidas (as cores, os animais); por contra, as denominacións dos CI do segundo grupo ou ben son moito máis xerais e abertas (a cidade, o campo, o mar) ou ben, sendo en principio tan específicas como as do primeiro (os medios de transporte, as profesións e os oficios, a calefacción e a iluminación) provocaron nos nosos informantes asociacións secundarias de moi diverso tipo, que xeraron listaxes máis ou menos heteroxéneas de voces (*vid.* § 2.3.4).

Precisamente esta distinción que observamos nos CI do segundo grupo fainos pensar que establecer unha correlación directa entre o índice de cohesión/densidade e o carácter semanticamente máis compacto ou máis difuso do CI, sen máis consideracións, non pode facerse sen riscos. Ao respecto, reproducimos a continuación unha reflexión recollida en López Meirama (2008: 107):

Debemos tener en cuenta que una misma cifra (salvo que sea 1) puede reflejar situaciones diferentes: imaginemos, por ejemplo, dos CI que presentasen la misma producción de palabras diferentes (1000) y el mismo promedio de repuestas por sujeto (20); en ambos casos el índice de cohesión sería 0.02, pero este dato cuantitativo podría no reflejar una absoluta coincidencia en la composición de los CI, ya que en uno de ellos podría combinarse la presencia de un buen número de palabras de aparición única con la existencia de un grupo también significativo de palabras emitidas por la mayoría de los informantes, lo cual sería índice de una consistencia léxica diferente de la que presentaría el otro CI si en él todas las palabras tuviesen unos porcentajes de aparición similares.

Noutros termos, para caracterizar semanticamente un CI como compacto ou difuso non é suficiente manexar o parámetro da cohesión; faise necesario, cando menos, coñecer que pasa nas zonas altas e baixas das listaxes de dispoñibilidade, é dicir, necesitamos saber se hai máis ou menos palabras emitidas por un único informante e, no outro extremo, se hai un grupo máis ou menos representativo de palabras evocadas por unha porcentaxe significativa de informantes. Á análise destes datos dedicamos o seguinte apartado.

3.1.3 Índice de dispoñibilidade

Para intentar ofrecer unha caracterización algo máis afinada das distintas áreas temáticas que compoñen a mostra necesitamos coñecer a composición das listaxes no que atinxe ao parámetro da dispoñibilidade; en particular, convén coñecer como se distribúen as súas zonas extremas e dilucidar se teñen máis ou menos peso no conxunto do CI; en principio, cabe supoñer que será máis difuso o CI no que se contabilicen moitas palabras emitidas por un único informante (obviamente, as que ocupan a parte baixa das listaxes) e,

pola contra, será máis compacto aquel que teña un número significativo de lemas con índices de dispoñibilidade altos.

3.1.3.1. Comezaremos polo primeiro dos dous datos, amosando na táboa 22 (que repite na segunda columna información xa ofrecida nas táboas 17 e 19 sobre produción de lemas) os relativos ao número de lemas que en cada CI foron proporcionados por un único informante: na terceira columna recolle-mos as cantidades absolutas e, na cuarta, a porcentaxe que estas cantidades supoñen sobre o total de lemas.

Táboa 22. Lemas producidos por un único informante

Centros de interese	Lemas	Lemas de apar. única	% Lemas apar. única
01 O corpo humano	858	360	41,96
02 A roupa	790	319	40,38
03 As partes da casa (sen mobles)	677	317	46,82
04 Os mobles da casa	451	173	38,36
05 Os alimentos e as bebidas	1078	422	39,15
06 Os obxectos colocados na mesa...	505	225	44,55
07 A cociña e os seus utensilios	720	296	41,11
08 A escola: mobles e materiais	959	456	47,55
09 A calefacción e a iluminación	945	479	50,69
10 A cidade	1330	588	44,21
11 O campo	1626	757	46,56
12 Os medios de transporte	963	454	47,14
13 Os traballos do campo e do xardín	1503	843	56,09
14 Os animais	1010	405	40,10
15 Os xogos e as distraccións	1594	825	51,76
16 As profesións e os oficios	1284	610	47,51
17 As cores	667	276	41,38
18 O mar	1319	577	43,75
Promedios	1015,5	465,67	44,95

O primeiro dato que cabe comentar dos incluídos na táboa atopámolo na derradeira fila, onde podemos comprobar que o promedio de palabras emitidas por un único informante achégase a 500, e que a porcentaxe media destas palabras é de practicamente o 45%; é dicir, que preto da metade dos lemas da mostra presenta índices de dispoñibilidade mínimos, o que significa que todos os CI conteñen un certo grao de dispersión. Con todo, esta varía en case 18 puntos entre a porcentaxe máis baixa, acadada no CI 04 ‘Os mobles da casa’

(38,36%), e a máis alta, a do CI 13 ‘Os traballos do campo e do xardín’ (56,09%), de xeito que as diferenzas entre uns CI e outros chegan a ser relevantes.

Dado que, como sinalamos anteriormente, a produción de lemas dos distintos CI é moi dispar, resulta máis adecuado que nos cinxamos ás cifras relativas, ás porcentaxes. Estas de novo sinalan o CI 13 como o máis difuso da mostra, mesmo cunha diferenza bastante avultada, no que á produción de lemas únicos se refire, respecto aos dous que lle seguen, os CI 15 ‘Os xogos e as distraccións’ e 09 ‘A calefacción e a iluminación’, ambos con máis da metade dos lemas proporcionados por un só informante. Respecto ao CI 15, xa vimos que é, xunto co 13, o que máis diferenza presenta na produción de palabras e de lemas, do que deducimos que resulta unha área temática pouco definida; no que se refire ao 09, constatamos así mesmo que se atopa entre os CI que presentan menor índice de cohesión, trazo este no que só é superado polos anteditos CI 13 e 15, e polo CI 11 (véxase a táboa 21). Deste último, pode resultar curioso que a porcentaxe de palabras únicas sexa menor que noutros catro CI (os CI 08 ‘A escola’, 16 ‘As profesións e os oficios’, 12 ‘Os medios de transporte’ e 03 ‘As partes da casa’), habida conta do seu baixo índice de cohesión, ao que acabamos de referirnos. No entanto, as diferenzas porcentuais en palabras únicas entre estes CI e o 11 son mínimas, de xeito que o dato que resulta máis interesante non é a porcentaxe do CI 11, senón precisamente a acadada polos outros catro, polo que cómpre un mínimo comentario: do CI 16 cabe sinalar que o seu baixo índice de cohesión (0,016), unido aos datos de produción de palabras e lemas (elevada en ambos os casos), nos fai pensar nunha área temática ampla na que non se detectan diferenzas salientables de dispoñibilidade; os CI 03 e 12 ocupan rangos intermedios na escala de cohesión, pero sitúanse ambos entre os que suman poucas palabras e poucos lemas, así que será necesario saber como se conforman estas dúas listaxes na súa parte superior para poder caracterizalas mellor; finalmente, o CI que máis pode chamar a atención é o 08, xa que se atopa no grupo dos que suman moitas palabras en comparación co número de lemas, polo que en principio cabería incluílo entre os máis compactos; no entanto, observamos que case a metade dos seus lemas —o 47,55%— é de aparición única (de feito, ocupa o décimo lugar na escala de cohesión), así que, de novo, a información sobre as palabras máis dispoñibles deste CI se revela necesaria para interpretar os datos (*vid.* 3.1.3.2).

No outro extremo sitúanse os CI con porcentaxes máis baixas de palabras proporcionadas por un único informante: ademais do citado CI 04 ‘Os mobles da casa’, os CI 05 ‘Os alimentos e as bebidas’, 14 ‘Os animais’, 02 ‘A roupa’, 07 ‘A cociña e os seus utensilios’, 17 ‘As cores’ e 01 ‘O corpo humano’ sitúanse preto do 40% de lemas únicos. De feito, o CI 04 é o que presenta o índice de cohesión máis elevado e, igual ca o 07, unha cantidade baixa de palabras e de

lemas: tanto os mobles como a cociña representan ámbitos do léxico bastante acoutados e limitados, que dificilmente propiciarán asociacións afastadas dos núcleos; os CI 14, 01, 02 e 17 resultan tan compactos e cohesionados coma os anteriores, pero son moito máis amplos, polo que non é de estrañar que sexan os que presentan as cantidades máis elevadas de palabras, en comparación cos lemas. Canto ao 05, que en principio simplemente incluímos entre aqueles que presentan moitas palabras e lemas, vemos que está moi preto do CI 14: ocupa o segundo lugar en produción de palabras e o sétimo na de lemas (o 14, o primeiro e o oitavo, respectivamente), ademais de ser o sétimo no que ao índice de cohesión se refire (o 14 é o sexto); ambos son tamén, por tanto, campos léxicos amplos pero con bastante concreción semántica.

3.1.3.2. Para completar esta caracterización dos CI, amosamos agora os datos sobre o número de lemas que foron proporcionados por un maior número de informantes: en primeiro lugar, na táboa 23 ofrecemos o número de lemas con índice de dispoñibilidade (*vid.* § 2.4) superior a 0,1 en cada un dos CI, así como o rango que estes ocupan en virtude deste parámetro⁷¹; en segundo lugar, na táboa 24 brindaremos o número de lemas segundo a porcentaxe mínima de informantes que os anotaron nas súas enquisas.

Táboa 23. Número de lemas con ID superior a 0,1

Centros de interese	ID > 0,1	Rango
01 O corpo humano	32	2
02 A roupa	32	2
03 As partes da casa (sen mobles)	28	4
04 Os mobles da casa	18	10
05 Os alimentos e as bebidas	37	1
06 Os obxectos colocados na mesa...	17	11
07 A cociña e os seus utensilios	22	7
08 A escola: mobles e materiais	32	2
09 A calefacción e a iluminación	20	8
10 A cidade	31	3
11 O campo	20	8
12 Os medios de transporte	23	6
13 Os traballos do campo e do xardín	18	10
14 Os animais	32	2
15 Os xogos e as distraccións	19	9
16 As profesións e os oficios	31	3
17 As cores	23	6
18 O mar	24	5

⁷¹ A coincidencia no número de lemas nalgúns CI reduce a extensión do rango, que vai do 1 ao 11.

Os datos que observamos na táboa permiten consolidar o retrato que vimos facendo dos CI 01 ‘O corpo humano’, 02 ‘A roupa’ e 14 ‘Os animais’ como ámbitos léxicos amplos e compactos, xa que entre os seus abundantes lemas contan cunha cantidade considerable dos que acadan alta dispoñibilidade, o cal significa que moitos dos informantes incluíron estes vocábulos nas súas listaxes, probablemente nos primeiros lugares; tales datos inciden tamén no carácter máis impreciso dos CI 13 e 15, áreas temáticas que, a pesar de contarse entre as que acumulan maior número de lemas (máis de 1.500 en ambos os dous casos), apenas suman un par de decenas entre as que ostentan índices de dispoñibilidade altos.

Hai, no entanto, dous CI dos que cabe facer un mínimo comentario: en primeiro lugar referímonos ao CI 04 ‘Os mobles da casa’, que a pesar de atoparse entre os que ostentan maior cohesión e menos palabras emitidas por un único informante, sitúase tamén entre os que conteñen un menor número de lemas con índice de dispoñibilidade superior a 0,1 (dezaioito); de feito, ocupa o rango 10, xunto ao CI 13. Esta aparente discordancia explícase tanto polo tamaño como polo carácter pechado deste campo léxico: os mobles da casa suman un número de lexías (relativamente) baixo, e é un estímulo que suxire apenas un pequeno feixe de palabras adecuadas ao seu núcleo significativo: *mesa, cama, armario, sofá, silla, estantería, cadeira, escritorio, sillón...* Os informantes vense obrigados a establecer asociacións secundarias para esgotar os dous minutos de que dispoñen para cubrir a enquisa, de xeito que, como xa indicamos en § 2.3.4.4, enseguida engaden electrodomésticos (*neveira, tele(visión), lavadora...*), saneamentos (*bañeira, ducha, lavabo, váter...*), obxectos auxiliares (*lámpada, espello, alfombra...*) etc. Aínda así, lembremos que é unha das áreas temáticas menos produtivas da nosa mostra, cun promedio de 13,89 palabras por informante e 451 lemas.

En segundo lugar, intentaremos aclarar a situación do CI 08 ‘A escola’, que, como sinalamos anteriormente, exhibe unhas cifras que se poderían considerar discordantes: dunha parte, presenta un elevado número de palabras en comparación co de lemas e, doutra, sitúase entre os de menor índice de cohesión e os que contabilizan máis palabras únicas. Pero se a estes datos engadimos que se atopa entre os que suman máis lemas con dispoñibilidade superior a 0,1, non parece difícil explicar a natureza deste CI: trátase dun ámbito do léxico que os nosos informantes non só dominan, senón que teñen tamén incluído na súa actividade diaria (non esquezamos que a poboación que compón a mostra é precisamente poboación estudiantil, e que ademais as enquisas foron recollidas nas aulas dos seus centros de estudo); a escola é, para un estudante, un estímulo moi importante, que probablemente xera asociacións de moi diverso tipo, e de aí a variabilidade léxica que caracteriza este CI; pero, por outra

banda, hai que ter en conta que a denominación do mesmo entraña moita concreción: o CI chámase, exactamente, ‘A escola: mobles e materiais’, así que, de feito, son mobles (*mesa, silla, cadeira, pupitre...*) e materiais (*lapis, bolígrafo, libro, goma...*) as primeiras lexías que cobren as enquisas. Non son poucas, de maneira que a parte alta das mesmas, moi homoxénea, é bastante ampla. No entanto, chega un momento no que as asociacións empezan a dispersarse, nunha área temática que pasa a ser a escola, en xeral (desprovista de especificacións), polo que en zonas máis baixas da listaxe de dispoñibilidade aparecen denominacións de espazos (*aula, biblioteca, cafetería...*), persoas (*profesora, alumno-a, mestre...*), elementos da construción (*porta, ventá, radiador, luz...*) etc., que imprimen bastante heteroxeneidade.

Para completar a imaxe da parte alta das listaxes, ofrecemos agora datos sobre os lemas máis frecuentes (que, en xeral, son tamén os máis dispoñibles): na táboa 24 proporcionamos o número de lemas que en cada CI foi achegado por un mínimo de 200 informantes (o 25% da poboación da nosa mostra), de 400 (o 50%) e de 600 (o 75%). Na columna da dereita incluímos o rango que ocupan os CI, establecido en primeiro lugar sobre o 50% e en segundo lugar sobre o 75%⁷².

Táboa 24. Número de lemas con distintas porcentaxes de frecuencia

Centros de interese	25%	50%	75%	Rango
01 O corpo humano	25	13	2	1
02 A roupa	23	9	3	4
03 As partes da casa (sen mobles)	19	6	3	8
04 Os mobles da casa	11	6	3	8
05 Os alimentos e as bebidas	24	4	1	10
06 Os obxectos colocados na mesa...	13	7	5	6
07 A cociña e os seus utensilios	18	5	-	9
08 A escola: mobles e materiais	27	9	4	3
09 A calefacción e a iluminación	10	3	2	12
10 A cidade	16	4	-	11
11 O campo	13	3	-	13
12 Os medios de transporte	19	8	4	5
13 Os traballos do campo e do xardín	8	2	-	15
14 Os animais	23	7	3	7
15 Os xogos e as distraccións	9	1	-	16
16 As profesións e os oficios	12	2	-	15
17 As cores	19	11	10	2
18 O mar	12	2	1	14

⁷² Tamén neste caso se reduce, pola coincidencia no número de lemas nos CI 03 e 04, e 13 e 16: abrangue do 1 ao 16.

Se nos fixamos en primeiro lugar nos CI que conteñen maior cantidade de lemas, vemos que a 01, 02, 08 e 14, CI xa mencionados, deberemos sumar o 17 ‘As cores’. Este CI destaca sobre todos os demais polo elevado número de lemas que foron proporcionados por un mínimo de 600 informantes, os cales comezaron a cubrir as enquisas con cores elementais (*amarelo, azul, vermello, verde, negro, branco...*) e continuaron con outras produto de combinacións e matices (*laranja, rosa, gris, marrón, violeta, beixe, lila...*). O ámbito semántico é, por tanto, moi sólido na súa zona nuclear. No entanto, fronte a outros CI, como o comentado 04, este ten moita maior amplitude, primeiro, porque desde o punto de vista referencial o espazo que cobre é un continuo, o que propicia a multiplicación de termos que establecen cortes a partir de diferentes propiedades, como o brillo, a intensidade ou o matiz; ademais, porque resulta moi suxestivo e favorece que se multipliquen as asociacións: cara ás propiedades das cores (*fosforito, brillante, (cor) cálida, apagado...*), as sensacións coas que estas se asocian (*alegría, esperanza, tristeza, paz...*), as actividades artísticas (*pintar, colorear, cadro, pincel...*) etc. É, por tanto, o CI máis compacto de todos no seu núcleo significativo, pero tamén un estímulo suficientemente amplo como para orixinar bastante variedade entre os lemas menos dispoñibles.

A revisión, en segundo lugar, dos CI que presentan un menor número de lemas con frecuencias altas, volve distinguir os CI 15 e 13, que se afianzan como os máis difusos do conxunto. De feito, ningún deles contén un só lema emitido por 600 ou máis informantes, e entre os dous só se contabilizan tres lemas cunha frecuencia superior ao 50%: *fútbol* no CI 15, con 410 aparicións e, no CI 13, *plantar* (con 420) e *regar* (con 417). Ademais destas áreas temáticas, destacan os CI 16 ‘As profesións e os oficios’ e 18 ‘O mar’, dos que xa sinalamos o elevado volume de palabras e lemas así como o baixo índice de cohesión (0,016 nos dous casos); anteriormente xa cualificamos o CI 16 como un ámbito léxico amplo sen diferenzas salientables na dispoñibilidade dos seus compoñentes; podemos caracterizar o CI 18 do mesmo xeito, mesmo podemos sinalar que, como estímulo, o mar ten menos concreción semántica ca as profesións e os oficios, razón que fai evidente a posición que ocupa.

Non quixeramos finalizar este apartado sen brindar información concreta acerca das formas léxicas de maior ocorrencia nas enquisas. Para iso, elaboramos a táboa 25, que recolle os cinco lemas máis frecuentes de cada CI, ordenados de maior a menor frecuencia absoluta. Tamén indicamos a frecuencia relativa, é dicir, o peso porcentual que o lema ten sobre o total de palabras do CI. Podemos observar que, aínda que en xeral as dúas cifras se relacionan proporcionalmente, non é infrecuente atopar frecuencias relativas elevadas

asociadas a frecuencias absolutas máis baixas. Por exemplo, *gato-a*, con 717 ocorrencias, ten unha frecuencia relativa de 3,404%, mentres que *coitelo*, con só 450, supéraa lixeiramente (3,427%). Este desequilibrio explícase pola diferenza que hai no número de palabras entre os dous CI: 717 ocorrencias constitúen case o 3,5% das 21.064 palabras que conforman o CI 14, practicamente a mesma proporción que supoñen 450 no conxunto das 13.132 palabras do CI 07. Con todo, como estamos situados na parte máis alta das listaxes, cremos que neste caso as cifras absolutas resultan máis elocuentes ca as relativas.

Táboa 25. Os lemas máis frecuentes da mostra

Lema	Centro de interese	Frecuencia absoluta	Frecuencia relativa
<i>coitelo</i>	06 Os obxectos colocados na mesa...	774	6,881%
<i>negro</i>	17 As cores	772	5,082%
<i>amarelo</i>	17 As cores	770	5,068%
<i>cociña</i>	03 As partes da casa (sen mobles)	763	6,163%
<i>culler</i>	06 Os obxectos colocados na mesa...	757	6,730%
<i>vaso</i>	06 Os obxectos colocados na mesa...	753	6,695%
<i>branco</i>	17 As cores	752	4,950%
<i>coche</i>	12 Os medios de transporte	740	5,121%
<i>avión</i>	12 Os medios de transporte	733	5,073%
<i>azul</i>	17 As cores	728	4,792%
<i>verde</i>	17 As cores	727	4,785%
<i>(cuarto de) baño</i>	03 As partes da casa (sen mobles)	718	5,800%
<i>gato-a</i>	14 Os animais	717	3,404%
<i>pantalón</i>	02 A roupa	707	4,232%
<i>mesa</i>	04 Os mobles da casa	703	6,328%
<i>bici(cleta)</i>	12 Os medios de transporte	693	4,796%
<i>auga</i>	05 Os alimentos e as bebidas	685	3,367%
<i>bolí(grafo)</i>	08 A escola: mobles e materiais	680	3,891%
<i>lapis</i>	08 A escola: mobles e materiais	677	3,874%
<i>calcetín</i>	02 A roupa	676	4,046%
<i>can</i>	14 Os animais	674	3,200%
<i>armario</i>	04 Os mobles da casa	671	6,040%
<i>ollo</i>	01 O corpo humano	670	3,392%
<i>cama</i>	04 Os mobles da casa	669	6,022%
<i>bombilla</i>	09 A calefacción e a iluminación	666	6,259%
<i>mesa</i>	08 A escola: mobles e materiais	661	3,782%
<i>camiseta</i>	02 A roupa	654	3,915%

Lema	Centro de interese	Frecuencia absoluta	Frecuencia relativa
<i>prato</i>	06 Os obxectos colocados na mesa...	643	5,717%
<i>barco</i>	18 O mar	631	3,749%
<i>pé</i>	01 O corpo humano	627	3,174%
<i>libro</i>	08 A escola: mobles e materiais	626	3,582%
<i>servilleta</i>	06 Os obxectos colocados na mesa...	619	5,503%
<i>tren</i>	12 Os medios de transporte	618	4,277%
<i>lámpada</i>	09 A calefacción e a iluminación	612	5,752%
<i>vaca</i>	14 Os animais	607	2,882%
<i>salón</i>	03 As partes da casa (sen mobles)	603	4,871%
<i>sofá</i>	04 Os mobles da casa	594	5,347%
<i>man</i>	01 O corpo humano	590	2,987%
<i>coche</i>	10 A cidade	588	3,499%
<i>perna</i>	01 O corpo humano	586	2,967%
<i>radiador</i>	09 A calefacción e a iluminación	578	5,432%
<i>dedo</i>	01 O corpo humano	577	2,921%
<i>barco</i>	12 Os medios de transporte	577	3,993%
<i>xersei</i>	02 A roupa	566	3,388%
<i>peixe</i>	18 O mar	564	3,351%
<i>edifício</i>	10 A cidade	540	3,214%
<i>habitación</i>	03 As partes da casa (sen mobles)	539	4,354%
<i>zapato</i>	02 A roupa	535	3,202%
<i>encerado</i>	08 A escola: mobles e materiais	522	2,987%
<i>pataca</i>	05 Os alimentos e as bebidas	490	2,409%
<i>herba</i>	11 O campo	489	3,016%
<i>forno</i>	07 A cociña e os seus utensilios	484	3,686%
<i>médico-a</i>	16 As profesións e os oficios	477	2,987%
<i>cabalo</i>	14 Os animais	474	2,250%
<i>vaca</i>	11 O campo	467	2,880%
<i>pota</i>	07 A cociña e os seus utensilios	466	3,549%
<i>silla</i>	04 Os mobles da casa	462	4,159%
<i>árbore</i>	11 O campo	461	2,843%
<i>(forno) microondas</i>	07 A cociña e os seus utensilios	456	3,472%
<i>tenda</i>	10 A cidade	453	2,696%
<i>coitelo</i>	07 A cociña e os seus utensilios	450	3,427%
<i>profe(sor-a)</i>	16 As profesións e os oficios	450	2,818%
<i>porco-a</i>	14 Os animais	448	2,127%
<i>garaxe</i>	03 As partes da casa (sen mobles)	444	3,586%
<i>parque</i>	10 A cidade	433	2,577%

Lema	Centro de interese	Frecuencia absoluta	Frecuencia relativa
<i>carne</i>	05 Os alimentos e as bebidas	423	2,079%
<i>plantar</i>	13 Os traballos do campo e do xardín	420	4,035%
<i>neveira</i>	07 A cociña e os seus utensilios	417	3,175%
<i>regar</i>	13 Os traballos do campo e do xardín	417	4,006%
<i>fútbol</i>	15 Os xogos e as distraccións	410	3,144%
<i>viño</i>	05 Os alimentos e as bebidas	405	1,991%
<i>auga</i>	18 O mar	397	2,359%
<i>podar</i>	13 Os traballos do campo e do xardín	392	3,766%
<i>rúa</i>	10 A cidade	381	2,267%
<i>avogado-a</i>	16 As profesións e os oficios	372	2,330%
<i>area</i>	18 O mar	369	2,193%
<i>estufa</i>	09 A calefacción e a iluminación	368	3,459%
<i>luz</i>	09 A calefacción e a iluminación	367	3,449%
<i>praia</i>	18 O mar	366	2,175%
<i>parchís</i>	15 Os xogos e as distraccións	365	2,799%
<i>zume</i>	05 Os alimentos e as bebidas	353	1,735%
<i>sementar</i>	13 Os traballos do campo e do xardín	347	3,333%
<i>tractor</i>	11 O campo	329	2,029%
<i>escondite</i>	15 Os xogos e as distraccións	323	2,477%
<i>sachar</i>	13 Os traballos do campo e do xardín	317	3,045%
<i>enfermeiro-a</i>	16 As profesións e os oficios	293	1,835%
<i>mestre-a</i>	16 As profesións e os oficios	291	1,823%
<i>(xogo de) cartas</i>	15 Os xogos e as distraccións	284	2,178%
<i>ovella</i>	11 O campo	279	1,721%
<i>baloncesto</i>	15 Os xogos e as distraccións	276	2,116%

A lectura das zonas alta e baixa desta relación achega uns datos que corroboran a sinxela caracterización que vimos dando dos CI ao longo deste apartado: dun lado observamos que entre os quince lemas máis frecuentes (os proporcionados por cando menos 700 estudantes) contabilizamos os cinco da área temática dedicada ás cores, que singularizamos coma a máis compacta da mostra (*negro, amarelo, branco, azul e verde*); en segundo lugar destaca o CI 06, que reúne os obxectos colocados na mesa para a comida, zona do léxico moi limitada no seu núcleo significativo, unha das menos produtivas da mostra, tanto en palabras como en lemas, e das máis cohesionadas (*coitelo, culler e vaso*); similar a este CI é o que agrupa as partes da casa, do que atopamos dous lemas moi frecuentes, *cociña e (cuarto de) baño*; así mesmo, do campo dos medios de trans-

porte contabilizamos dous lemas, *coche* e *avión*, aínda que este CI é máis produtivo e menos cohesionado ca os anteriores.

Do outro lado, se nos fixamos nos CI cuxos lemas máis frecuentes foron emitidos por menos da metade da poboación da mostra (400 informantes), de novo destacan o CI 15, con catro lemas (*baloncesto*, *(xogo de) cartas*, *escondite* e *parchís*) e os CI 13 e 18, con tres (*sachar*, *sementar* e *podar*; *praia*, *area* e *auga*), todos eles caracterizados como zonas do léxico moi pouco homoxéneas, aínda que con distinta produtividade. Tamén comentamos que as profesións e os oficios constitúen un campo bastante amplo no que non se detectan grandes diferenzas na dispoñibilidade dos vocábulos que o constitúen, o cal se corrobora pola presenza de tres deles no grupo dos menos frecuentes (*mestre-a*, *enfermeiro-a* e *avogado-a*), e o mesmo cabe dicir do CI 11, con dous lemas neste grupo (*ovella* e *tractor*). Canto á calefacción e a iluminación, tamén con dous lemas (*luz* e *estufa*), polos nosos datos revélase como un estímulo menor, unha zona do léxico pouco evocadora para os nosos informantes, con baixa produción de palabras e pouca cohesión.

3.1.4 Léxico galego e léxico castelán⁷³

Unha das análises máis recorrentes no ámbito da dispoñibilidade léxica é a establecida entre as diferentes sintopías que constitúen o proxecto panhispánico, e normalmente o resultado adoita confirmar a gran homoxeneidade que existe na produción do léxico —entendida en términos cuantitativos— dos mozos preuniversitarios; así mesmo, salvo excepcións, os centros de interese teñen pesos similares nas diferentes mostras. No caso do léxico dispoñible en Galicia os parámetros comparativos permútanse, xa que as poboacións están estratificadas exactamente do mesmo xeito (e só as separan tres anos), e o que varía é a lingua. Cabe preguntarse, entón, se o emprego dunha ou doutra lingua inflúe, e de que xeito, na elaboración das listaxes de léxico dispoñible: dun lado, se intervéñ na produción de palabras e, doutro, se os distintos CI, ou algúns deles, teñen peculiaridades e/ou pesos diferentes no vocabulario dispoñible dos informantes.

Empezando polo segundo aspecto, debemos dicir que as dezaioito áreas temáticas presentan as mesmas características nas dúas linguas; por exemplo, tanto en galego coma en castelán os CI 09, 13, 15 e 18 son dominios concep-

⁷³ Os datos do castelán que manexamos neste apartado proveñen de *Dispogal*, e non son absolutamente coincidentes cos ofrecidos en López Meirama (2008), xa que, para introducilos neste programa, en funcionamento desde 2010, sometémolos a un novo proceso de revisión. De todos os xeitos, as diferenzas son irrelevantes.

tuais pouco definidos; os CI 01, 02, 08 e 17 son amplos e cohesionados; os CI 03, 04 e 06 son moi compactos e limitados etc. Canto á súa dimensión, a seguinte táboa, que recupera os datos da nº 20, recolle os rangos que acadan os CI nos dous subproxectos considerando o total de palabras e de lemas. Nela podemos comprobar como as diferenzas entre ambos son mínimas, xa que non van máis alá dun punto (CI 01, 03, 04, 05, 06, 07, 09,13, 15 e 16) ou dous (02, 08, 10, 11, 17 e 18), e mesmo nalgún caso os rangos chegan a ser absolutamente coincidentes (CI 12 e 14):

Táboa 26. Rango dos CI segundo a produción de palabras e lemas nos dous subproxectos

Centros de interese	Palabras		Lemas	
	galego	castelán	galego	castelán
01 O corpo humano	3	2	12	11
02 A roupa	7	5	13	13
03 As partes da casa (sen mobles)	14	13	15	16
04 Os mobles da casa	16	16	18	17
05 Os alimentos e as bebidas	2	3	7	7
06 Os obxectos colocados na mesa...	15	15	17	18
07 A cociña e os seus utensilios	12	12	14	15
08 A escola: mobles e materiais	4	4	10	12
09 A calefacción e a iluminación	17	17	11	10
10 A cidade	6	6	4	2
11 O campo	8	10	1	1
12 Os medios de transporte	11	11	9	9
13 Os traballos do campo e do xardín	18	18	3	4
14 Os animais	1	1	8	8
15 Os xogos e as distraccións	13	14	2	3
16 As profesións e os oficios	9	8	6	6
17 As cores	10	9	16	14
18 O mar	5	7	5	5

No que se refire á produción de palabras e lemas, a táboa 27 permite constatar que en todos os CI as cifras relativas á lingua castelá son superiores ás da lingua galega, salvo no 13 e, no que atinxe só aos lemas, tamén no 15, aínda que nestes dous CI a diferenza é inapreciable:

Táboa 27. Datos da produción de palabras e lemas nos dous subproxectos

Centros de interese	Palabras		Lemas		Promedios	
	gal.	cas.	gal.	cas.	gal.	cas.
01 O corpo humano	19753	23430	858	1082	24,69	29,29
02 A roupa	16707	19751	790	901	20,88	24,69
03 As partes da casa (sen mobles)	12380	14508	677	757	15,48	18,14
04 Os mobles da casa	11109	13172	451	713	13,89	16,47
05 Os alimentos e as bebidas	20342	22798	1078	1250	25,43	28,50
06 Os obxectos colocados na mesa...	11248	13860	505	519	14,06	17,33
07 A cociña e os seus utensilios	13132	15069	720	782	16,42	18,84
08 A escola: mobles e materiais	17477	19775	959	1065	21,85	24,72
09 A calefacción e a iluminación	10640	11770	945	1092	13,30	14,71
10 A cidade	16803	18854	1330	1590	21,00	23,63
11 O campo	16214	16958	1626	1737	20,27	21,20
12 Os medios de transporte	14449	15677	963	1101	18,06	19,60
13 Os traballos do campo e do xardín	10410	10293	1503	1501	13,01	12,87
14 Os animais	21064	23815	1010	1135	26,33	29,77
15 Os xogos e as distraccións	13041	14384	1594	1512	16,30	18,00
16 As profesións e os oficios	15967	18025	1284	1431	19,96	22,53
17 As cores	15192	17497	667	885	18,99	21,90
18 O mar	16830	18395	1319	1451	21,04	22,99
Promedio	5153,22	17112,83	1015,50	1139,11	18,94	21,40

Respecto ao total de palabras producidas destacan, por un lado, os CI 01 ‘O corpo humano’, 02 ‘A roupa’, 14 ‘Os animais’ e 06 ‘Os obxectos colocados na mesa para a comida’, con máis de 2.500 palabras de diferenza e, doutro, 11 ‘O campo’ e 13 ‘Os traballos do campo e do xardín’, con menos de 1.000, ademais de que este último, como acabamos de comentar, resultou ser máis produtivo no subproxecto galego. No que se refire aos lemas, destaca de novo o CI 01, ao que hai que engadir 04 ‘Os mobles da casa’, 10 ‘A cidade’ e 17 ‘As cores’, áreas temáticas todas elas nas que contabilizamos máis de douscentos lemas de diferenza entre os dous subproxectos; do outro lado sitúanse os CI 06 e 13, con apenas catorce e dous lemas de distancia, respectivamente.

A pesar de que, como indicamos, a estratificación da mostra é exactamente a mesma nos dous subproxectos, detectamos unha diferenza concreta na recollida dos datos que puido influír nalgunha medida nos resultados finais dalgúns dos CI. Consiste no tratamento diferenciado que tiveron as marcas comerciais, consecuencia do alcance excesivo que observamos nelas nalgunhas áreas temáticas do LDEG, que nos levou a eliminar moitas for-

mas das listaxes e, sobre todo, a tomar a decisión de advertir aos informantes do subproxecto galego de que non as inclúisen nas súas enquisas. Aínda así, algúns informantes introduciron nomes de marcas nas listaxes, feito que dá conta da súa relevancia no léxico dos alumnos de segundo de bacharelato. Por tal motivo, no subproxecto galego mantivemos todas as palabras deste tipo proporcionadas polos informantes. Por exemplo, no caso do CI 15 ‘Os xogos e as distraccións’ en galego recolleamos 110 nomes de marcas, mentres que en castelán, trala eliminación, estes quedarán reducidos a 31. A diferenza, de 79 lemas, acada case os 82 que separan ambos os subproxectos. No LDEG elimináronse máis de 100 achegas, xa que só se conservaron os nomes con dez ou máis ocorrencias, aos que se sumaron *Frisbee* e *Cubo de Rubik*. Se fixésemos o mesmo en galego, en lugar de 110 nomes de marcas teríamos 17 máis algunhas voces que cabería engadir, polo seu emprego como palabras xenéricas ou por identificarse con xogos ou xoguetes únicos e universalmente coñecidos: ademais dos anteditos *Frisbee* e *Cubo de Rubik*, *Barbie*, *Twister*, *Pinball*, *Scrabble*, *Scalextric*, *iPod* e *Rummikub*, é dicir, 26, máis ou menos o mesmo número que conservamos en castelán. Así mesmo, nos CI 02 ‘A roupa’ e CI 12 ‘Os medios de transporte’ o criterio no LDEG foi o de deixar as formas das que houberse máis de cinco ocorrencias; esta decisión trouxo como consecuencia que no primeiro deles se contabilizasen 11 marcas na parte castelá, fronte ás 42 que hai na galega, as cales se reducirían a 2 de aplicarse o criterio, e que no CI 12 houberse 16 marcas en castelán e 62 en galego, lingua na que, se aplicásemos a redución, quedaríamos con 5 (*Vespa* 15, *Zodiac* 14, *Vespino* 13, *Vitrassa* 11 e *Ferrari* 5; cfr. castelán: *Zodiac* 27, *Ferrari* 19, *Mercedes* 17, *Vespino* 14, *Peugeot* 13, *Audi* 11 etc.). Outro CI que cabería destacar é o dedicado aos alimentos e as bebidas, no que contabilizamos 39 lemas que son nomes comerciais no subproxecto do castelán, trala supresión dun centenar deles, algúns de aparición frecuente, fronte aos 89 no galego.

Cremos que a razón destas disimilitudes estriba en que, como acabamos de sinalar, cando se recolleron as enquisas en lingua galega os informantes foron advertidos de que non inclúisen marcas —cousa que non se fixo no caso das enquisas do castelán—, e como consecuencia desa decisión resolvemos non eliminalas das listaxes, a pesar da nutrida presenza deste tipo de formas nalgúns CI. Con todo, é indubidable que a indicación cohibiu aos informantes á hora de encher as súas enquisas, polo que a frecuencia das marcas resultou moito menor no subproxecto galego ca no castelán. Como exemplo, constatamos que as máis frecuentes do CI 05, en ambos os dous —*Coca-cola*, *Fanta*, *Kas*, *Nestea* e *Pepsi-(Cola)*—, todas elas conservadas tamén no LDEG, suman 943 palabras no subproxecto castelán e 363 no ga-

lego, é dicir, só estes cinco lemas supoñen 580 palabras de diferenza entre unha listaxe e a outra.

Sexa como fose, a disparidade no tratamento das marcas comerciais non xustifica, en xeral, a distancia detectada entre os dous subproxectos, que, en promedios por CI, se cifra en preto de 2.000 palabras distintas, en 124 lemas e en case que 2,5 palabras por informante. Poderíamos pensar que a explicación desta diverxencia estriba en que a maior parte dos informantes emprega o castelán, sexa exclusiva ou maioritariamente, como a súa lingua de uso habitual, polo que as palabras desta lingua acoden con máis rapidez á súa memoria e por tanto escriben máis palabras nos dous minutos de que dispoñen para cada CI no LDEG ca no LDG. No entanto, os datos contradín esta explicación: se comparamos os promedios de palabras por informante e CI dos rapaces que afirman falar só galego do proxecto galego (110 informantes) e os dos que se declaran monolingües en castelán do proxecto castelán (195 informantes), as diferenzas son aínda maiores, como pode verse na táboa 28⁷⁴:

Táboa 28. Promedios de produción de palabras dos galegofalantes e castelanfalantes nos dous subproxectos

Centros de interese	Galegofalantes do sub. galego	Castelanfalantes do sub. castelán
01 O corpo humano	25,44	30,51
02 A roupa	21,00	25,37
03 As partes da casa (sen mobles)	15,41	18,65
04 Os mobles da casa	13,44	16,79
05 Os alimentos e as bebidas	25,61	29,32
06 Os obxectos colocados na mesa...	14,62	17,60
07 A cociña e os seus utensilios	16,56	19,76
08 A escola: mobles e materiais	22,77	25,73
09 A calefacción e a iluminación	13,70	14,70
10 A cidade	20,34	25,07
11 O campo	21,32	22,93
12 Os medios de transporte	18,55	20,35
13 Os traballos do campo e do xardín	14,25	13,07
14 Os animais	27,09	32,15
15 Os xogos e as distraccións	16,54	18,84
16 As profesións e os oficios	19,74	24,37
17 As cores	18,98	22,82
18 O mar	20,74	24,24
Promedio	19,23	22,35

⁷⁴ Dada a disparidade das poboacións, non manexamos as cifras absolutas, nin de palabras nin de lemas.

A distancia chega a superar os cinco puntos nos CI 01 ‘O corpo humano’ e 14 ‘Os animais’ e é maior de catro, ademais de nestes, nos CI 10 ‘A cidade’, 16 ‘As profesións e os oficios’ e 02 ‘A roupa’. Talvez a explicación estribe, entón, no tipo social de falantes dunha e doutra lingua segundo as variables que conforman as mostras: no capítulo 1 viamos que o emprego do galego é maioritario no ámbito rural e en familias con NSC baixo, de xeito que talvez sexan estas variables, e non a da lingua, as que realmente incidan na menor produción de palabras. No entanto, máis adiante poderemos comprobar que esta explicación podería resultar insuficiente; na nosa opinión, tamén pode ser relevante o contexto no que se realizan as enquisas, en relación co cal convén lembrar que os rapaces proporcionaron as súas palabras no centro educativo, máis concretamente nas clases de Lingua e Literatura, e baixo a supervisión do profesor. Deste xeito, aínda que non o fose, a actividade da realización das enquisas probablemente se asociou, quizá inconscientemente, a unha actividade lectiva, o que puido levar a moitos informantes a instalarse na lingua estándar, malia as indicacións previas que ao respecto se lles deron. E en relación con isto, consideramos que poden ser decisivas para explicar os datos atopados as deficiencias existentes no dominio do léxico galego estándar e a propia consciencia desa deficiencia, feito que puido conducir en certos casos a que algúns informantes renunciásen a introducir determinados conceptos nas listaxes (por descoñecemento da forma galega normativa⁷⁵), e noutros a que tivesen que tomar un pouco máis de tempo que cando se lles propuxo cubrir a enquisa en castelán, lingua cuxa variedade estándar, mesmo os galegofalantes, dominan en maior medida ca a do galego.

Como apoio a esta explicación, achegamos os datos dos promedios na produción de palabras en galego e en castelán nos centros de ensino públicos e privados: consideramos que a presenza moito máis ostensible do castelán nestes últimos explica que a distancia entre os promedios se eleve a case cinco puntos, que os supere en seis dos CI (01, 05, 02, 08, 03 e 14) e que mesmo chegue a 7,94 no CI 05 e a 8,65 no CI 01, así como que no CI 13 o promedio tamén sexa maior en castelán, neste caso cunha diferenza de 1,52 puntos. Fronte a estas cifras, as relativas ao ensino público están moito máis próximas entre si: a distancia media apenas supera os dous puntos, e naqueles CI nos que é maior só supera os catro puntos no CI 01 e os tres puntos nos CI 02, 06 e 14. Certamente, o número de galegofalantes nos centros privados é irrisorio (9 din falar só ou máis galego no subproxecto galego, e 10 no caste-

⁷⁵ Pode verse ao respecto Álvarez de la Granja (no prelo).

lán), pero acabamos de comprobar que a lingua non é un factor determinante (véxase a táboa 28); moito máis importante é, desde o noso punto de vista, a presenza abrumadora do castelán neste tipo de ensino, que conduce a un maior descoñecemento do galego.

Táboa 29. Datos da produción de palabras nos dous subproxectos segundo o tipo de centro

Centros de interese	Público		Privado	
	galego	castelán	galego	castelán
01 O corpo humano	24,80	28,81	24,11	32,76
02 A roupa	21,04	24,42	20,09	26,67
03 As partes da casa (sen mobles)	15,72	17,81	14,18	20,55
04 Os mobles da casa	14,00	16,14	13,30	18,86
05 Os alimentos e as bebidas	25,52	27,90	24,94	32,88
06 Os obxectos colocados na mesa...	14,06	17,23	14,05	18,00
07 A cociña e os seus utensilios	16,64	18,62	15,25	20,41
08 A escola: mobles e materiais	22,01	24,34	20,96	27,53
09 A calefacción e a iluminación	13,38	14,53	12,87	16,03
10 A cidade	20,85	23,13	21,84	27,27
11 O campo	20,47	20,80	19,19	24,09
12 Os medios de transporte	18,05	19,39	18,13	21,09
13 Os traballos do campo e do xardín	13,22	12,79	11,93	13,45
14 Os animais	26,38	29,44	26,09	32,18
15 Os xogos e as distraccións	16,36	17,88	15,99	18,90
16 As profesións e os oficios	19,75	22,28	21,07	24,41
17 As cores	18,82	21,81	19,86	22,53
18 O mar	21,06	22,91	20,91	23,59
Promedio	19,01	21,12	18,60	23,40

Finalmente, no que se refire ás palabras máis frecuentes, só cabe comentar a ausencia na parte máis alta das listaxes do galego dalgúns vocábulos cuxos equivalentes resultaron moito máis produtivos nas listaxes do castelán, ausencia que se explica ben pola menor competencia dos nosos informantes, maioritariamente castelanfalantes, e en relación con isto, por interferencias co castelán ou ben por diferenzas na estruturación do léxico entre as dúas linguas; por exemplo, *rojo* (768) no listado do castelán fronte a *vermello* (676) e *roxo* (166) no do galego; *perro* (756) fronte a *can* (674) e *perro* (77); *tenedor* (737) fronte a *tenedor* (429) e *garfo* (366); *sartén* (575) fronte a *sartén* (344) e *tixola* (275); *ola* (446) fronte a *onda* (239) e *ola* (219); *pescado* (400) fronte a *peixe* (214) e *pescado* (184); etc.

3.1.5 Recapitulación

Os datos manexados nos apartados precedentes permítenos concluír que os dezaioito centros de interese que compoñen a mostra son estímulos de moi diverso tipo: hainos moi produtivos, tanto en palabras como en formas léxicas distintas (05, 08 ou 10), e pouco produtivos (03, 04 ou 06); hainos máis cohesionados (01, 04 ou 17) e menos (09, 13 ou 15); hainos bastante limitados na produción de palabras con índices de cohesión relativamente baixos (03 ou 09); os compactos e cohesionados poden ser amplos (01, 02, 14 ou 17); etc.

Nesta heteroxeneidade inflúe, por suposto, a propia denominación do CI, que como indicamos pode propiciar a confección de listas máis ou menos pechadas de hipónimos e/ou merónimos (01, 03, 04 ou 17) ou máis ben pode sumar palabras de significados relativamente afastados entre si, agrupadas a base de amplas relacións asociativas de moi diversa natureza (10, 11 ou 18). Pero tamén inflúe o tipo de poboación que compón a nosa mostra: os rapaces preuniversitarios teñen unha idade e unha formación que os fai falantes máis competentes nunhas zonas do léxico (01, 08 ou 14) ca noutras (09 ou 13). No seguinte capítulo incidiremos precisamente nesta cuestión, pois comprobaremos se as variables de estratificación social inflúen ou non na produción de palabras e na compatibilidade na elección das unidades léxicas.

En último lugar, tamén constatamos que os datos cuantitativos da mostra galega difiren en relación cos da mostra castelá, que en xeral son algo máis elevados, e consideramos plausible xustificar a diferenza pola presenza máis acusada desta lingua nos centros de ensino de Galicia, pola falta de espazo e secundarización do galego na sociedade e polas consecuentes diferenzas no dominio de ambas as dúas linguas que se detecta nos estudantes.

3.2 As variables de estratificación social

Comentamos anteriormente que o tipo de poboación que compón a nosa mostra incide de maneira significativa na conformación dalgúns dos centros de interese; a idade e a formación dos nosos informantes explican, por exemplo, que unha palabra como *esternocleidomastoideo* alcance un elevado índice de dispoñibilidade no CI 01 ‘O corpo humano’ (ten o rango 87 dun total de 858 formas), ou que no CI 18 ‘O mar’ atopemos, entre as máis dispoñibles, formas como *plancto*, *tsunami*, *iceberg*, *fitoplancto*, *afloramento* ou *tómbolo* e tamén outras como *surf*, *bucear*, *verán*, *bañador*, *moto de auga*, *diversión*, *bikini* ou *surfista*. Parece lóxico pensar, entón, que as diferentes variables de estratificación social poden incidir na produción dos elementos léxicos, tanto cuantitativa como cualitativamente.

Dedicaremos as páxinas que seguen a esta cuestión, analizando en primeiro lugar os datos cuantitativos e, en segundo lugar, os cualitativos, para o que seleccionaremos os grupos máis disímiles entre si dentro de cada estrato.

3.2.1 As variables sociais e a produción de palabras

Dos estudos de dispoñibilidade léxica e, en particular, da análise realizada ao léxico do español de Galicia conclúese que, en xeral, as variables de estratificación social non inflúen de maneira relevante na produción de palabras, de xeito que pertencer a un ou outro sexo, falar exclusiva ou maioritariamente unha ou outra lingua, vivir nun tipo de poboación ou noutro etc., non perturban no esencial a homoxeneidade dunha poboación unificada pola idade e polo seu nivel formativo. Non o fan, cando menos, desde un punto de vista estatístico, como se indica en López Meirama (2008: 125). Neste traballo emprégase a técnica estatística da *Análise de varianza* ou ANOVA, e a significación establécese co *valor P*, fixado en 0,05: como se explica no texto (*ibíd*: 111), só se considera que existe desigualdade entre as poboacións cunha marxe de erro igual ou inferior a 0,05; de feito, canto máis preto estea o *valor P* de 1, máis arriscado resulta asumir a *Hipótese alternativa* (H_1), segundo a cal as poboacións son diferentes ou, dito doutro modo, segundo a cal o factor social en cuestión incide na produción de palabras.

Con todo, en López Meirama (2008) detéctanse algunhas diferenzas entre unhas variables e outras e, particularmente, destaca a relativa ao nivel socio-cultural dos pais como a única que semella ter algunha influencia (desbotando a estatística), no sentido de que os informantes cuxos pais teñen formación universitaria son os que proporcionan máis palabras nas súas listaxes, e aqueles cuxos pais carecen de formación ou apenas acadaron os estudos primarios son os que achegan menos palabras.

Nas páxinas que seguen comprobaremos que a situación é similar no galego, neste caso mesmo examinando con máis detalle a información relativa á lingua dos informantes, sexa como lingua inicial ou habitual, sexa como un factor tomado individualmente ou en combinación con outros.

Para levar a cabo a análise, manexaremos como valor numérico o promedio de palabras por informante en cada CI, que é o máis adecuado cando se traballa con poboacións de tamaños moi diferentes (lembramos, ao respecto, os datos de estratificación da mostra, recollidos en § 2.2).

3.2.1.1 O sexo

Na táboa 30 amosamos os promedios de produción de palabras das 427 rapazas e os 373 rapaces que realizaron as enquisas, promedios que revelan

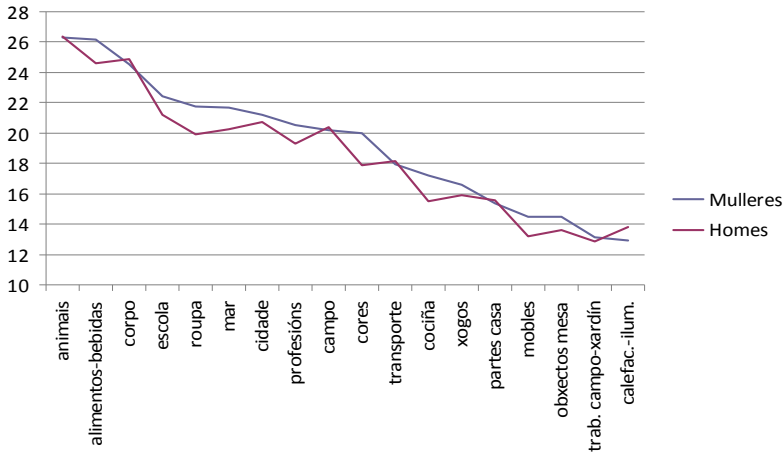
que entre unhas e outros apenas hai diferenzas; de feito, o *valor P* é 0,6126, cifra moito máis próxima a 1 que a 0,05. Con todo, as rapazas amosáronse algo máis produtivas ca os rapaces, como se desprende nos promedios globais, recollidos na derradeira fila da táboa; en doce dos dezaioito CI (preto do 70%) elas anotaron máis palabras e, ademais, en catro dos seis CI nos que os mozos as superan (11 ‘O campo’, 12 ‘Os medios de transporte’, 03 ‘As partes da casa’ e 14 ‘Os animais’), a distancia entre promedios é inapreciable. Por outra banda, aqueles CI nos que a distancia é maior (17 ‘As cores’, 02 ‘A roupa’, 07 ‘A cociña’ e 05 ‘Os alimentos e as bebidas’) sitúanse en ámbitos do léxico tradicionalmente asociados ao mundo feminino: a moda e o fogar.

De todos os xeitos, debemos insistir en que as diferenzas non son significativas, polo que concluímos que o sexo non é un factor que incida na produción de palabras entre os mozos galegos preuniversitarios. O gráfico 23 permite visualizar os datos de xeito máis directo, para o cal ordenamos os CI desde os que exhiben promedios máis altos ata os que os teñen máis baixos.

Táboa 30. Promedios de respostas segundo o sexo

Centros de interese	Mulleres	Homes
01 O corpo humano	24,53	24,87
02 A roupa	21,73	19,91
03 As partes da casa (sen mobles)	15,39	15,58
04 Os mobles da casa	14,48	13,20
05 Os alimentos e as bebidas	26,18	24,57
06 Os obxectos colocados na mesa...	14,46	13,60
07 A cociña e os seus utensilios	17,20	15,52
08 A escola: mobles e materiais	22,41	21,20
09 A calefacción e a iluminación	12,89	13,77
10 A cidade	21,23	20,74
11 O campo	20,17	20,38
12 Os medios de transporte	17,96	18,17
13 Os traballos do campo e do xardín	13,15	12,86
14 Os animais	26,32	26,34
15 Os xogos e as distraccións	16,62	15,94
16 As profesións e os oficios	20,54	19,29
17 As cores	19,97	17,87
18 O mar	21,70	20,27
Promedio global	19,27	18,56

Gráfico 23. Promedios de respostas segundo o sexo



3.2.1.2 A lingua

Como se aprecia na táboa 31 e no gráfico que a acompaña⁷⁶, a lingua inicial non é un factor relevante na produción de palabras, que é moi uniforme; no que se refire a esta variable, o *valor P* está aínda máis afastado de 0,05 ca o anterior, xa que chega a 0,7416. Ao respecto, talvez conveña lembrar que o léxico rexistrado nas listaxes non é sempre léxico galego en sentido estrito: con frecuencia atopamos palabras que parecen extraídas directamente do castelán, ben sexa por causa dun lapsus bilingüe, isto é, dunha escapada inconsciente cara a esta lingua (*vid.* Álvarez de la Granja 2011), ben pola necesidade de encher, de maneira consciente, unha lagoa léxica do galego, unha deficiencia no dominio do vocabulario desta lingua (*hueso, oreja, pajarita, perro...*); outras veces, o que rexistramos son voces do castelán adaptadas, probablemente *ad hoc*, ao galego (*vaia* —por *valla*—, *desaiuno, morcélogo, orexa...*). Por tanto, dos datos que extraemos da nosa mostra non se pode inferir que a lingua inicial non incida na capacidade dos falantes de produciren máis ou menos palabras nesa lingua ou, máis concretamente, non se pode afirmar que o feito de que a lingua inicial sexa ou non o galego non incida na súa competencia na produción de palabras galegas, posto que é moi probable que a maior parte das palabras do castelán (do tipo das arriba mencionadas) sexan especialmente frecuentes en informantes que tiveron o castelán como lingua inicial. Se cadra, se suprimísemos esas palabras do cómputo as dife-

⁷⁶ Eliminamos desta análise as dúas informantes que declararon aprender a falar nunha lingua diferente do galego e do castelán, xa que constitúen unha poboación tan pequena que as cifras dos promedios non poden ser representativas.

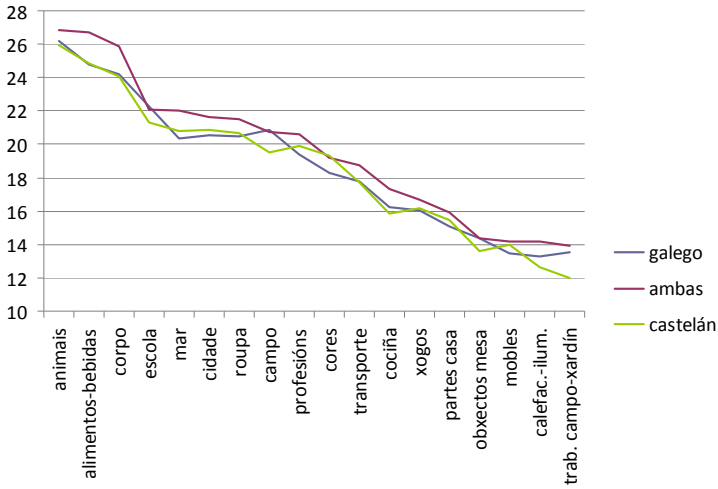
renzas entre os informantes que teñen o castelán como lingua inicial e aqueles que teñen o galego serían maiores. O mesmo comentario cabe facer dos datos relativos á lingua habitual, que achegaremos máis adiante.

A táboa 31 revela que os rapaces que declararon aprender a falar en ambas as dúas linguas presentan uns promedios lixeiramente superiores, non só porque o global, 19,59, supere case nun punto aos que aprenderon en galego ou en castelán, senón tamén porque en catorce dos dezaoto CI (preto do 80%) os promedios son máis altos. Só nos CI 06 ‘Os obxectos colocados na mesa para a comida’, 08 ‘A escola’ e 11 ‘O campo’ son superiores no grupo dos que aprenderon a falar en galego, e no CI 17 ‘As cores’ éo no grupo dos que aprenderon en castelán. De todos os xeitos, a través do gráfico 24 vese claramente que as diferenzas son inapreciables.

Táboa 31. Promedios de respostas segundo a lingua inicial

Centros de interese	Galego	Ambas	Castelán
01 O corpo humano	24,18	25,91	24,10
02 A roupa	20,47	21,49	20,69
03 As partes da casa (sen mobles)	15,09	15,94	15,44
04 Os mobles da casa	13,48	14,17	13,98
05 Os alimentos e as bebidas	24,80	26,73	24,86
06 Os obxectos colocados na mesa...	14,38	14,36	13,62
07 A cociña e os seus utensilios	16,21	17,35	15,82
08 A escola: mobles e materiais	22,29	22,08	21,34
09 A calefacción e a iluminación	13,31	14,15	12,65
10 A cidade	20,54	21,64	20,88
11 O campo	20,84	20,76	19,50
12 Os medios de transporte	17,79	18,74	17,74
13 Os traballos do campo e do xardín	13,52	13,94	11,97
14 Os animais	26,22	26,86	25,93
15 Os xogos e distraccións	16,03	16,70	16,17
16 As profesións e os oficios	19,36	20,58	19,90
17 As cores	18,31	19,21	19,32
18 O mar	20,32	22,02	20,81
Promedio global	18,73	19,59	18,60

Gráfico 24. Promedios de respostas segundo a lingua inicial

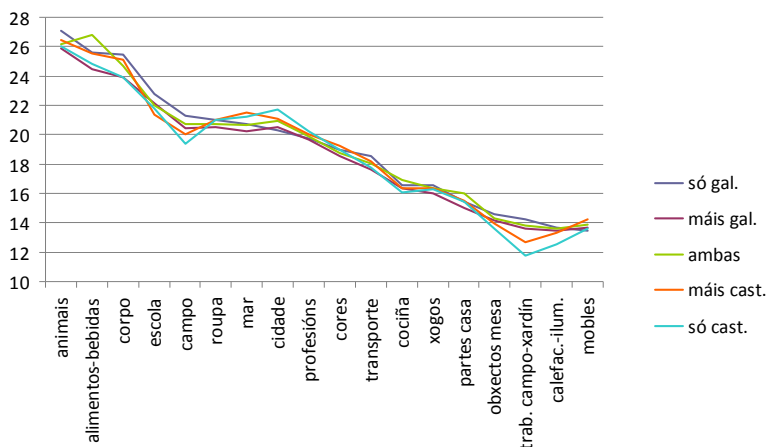


Canto á lingua habitual, os resultados son similares aos da lingua inicial, tal e como se pode ver na táboa 32: neste caso, é o grupo dos rapaces que din falar exclusivamente galego o que presenta o promedio global máis alto, pero só se eleva unha décima por riba do que corresponde ao grupo dos que declaran empregar ambas as linguas indistintamente; por outro lado, podería chamar a atención o feito de que o promedio máis reducido sexa o dos estudantes que din falar maioritariamente galego, por baixo dos que din falar maioritariamente castelán e mesmo dos que din falar exclusivamente castelán; no entanto, as diferenzas son tan pequenas que na nosa opinión este dato carece de relevancia. De feito, observamos que no gráfico 25 as liñas que representan cada un dos cinco grupos están practicamente superpostas.

Táboa 32. Promedios de respostas segundo a lingua habitual

Centros de interese	Só gal.	Máis galego	Ambas	Máis castelán	Só cast.
01 O corpo humano	25,44	23,88	24,71	25,09	23,90
02 A roupa	21,00	20,49	20,73	21,00	20,99
03 As partes da casa (sen mobles)	15,41	15,01	15,98	15,51	15,44
04 Os mobles da casa	13,44	13,67	13,86	14,25	13,63
05 Os alimentos e as bebidas	25,61	24,47	26,78	25,52	24,81
06 Os obxectos colocados na mesa...	14,62	14,14	14,34	13,94	13,60
07 A cociña e os seus utensilios	16,56	16,35	16,95	16,34	16,09
08 A escola: mobles e materiais	22,77	22,12	21,98	21,39	21,82
09 A calefacción e a iluminación	13,70	13,45	13,62	13,32	12,56
10 A cidade	20,34	20,53	20,97	21,11	21,70
11 O campo	21,32	20,44	20,74	20,05	19,42
12 Os medios de transporte	18,55	17,64	18,08	18,18	17,75
13 Os traballos do campo e do xardín	14,25	13,57	13,84	12,66	11,73
14 Os animais	27,09	25,91	26,14	26,43	26,03
15 Os xogos e as distraccións	16,54	16,02	16,36	16,32	16,27
16 As profesións e os oficios	19,74	19,66	19,90	20,05	20,23
17 As cores	18,98	18,54	18,73	19,25	19,00
18 O mar	20,74	20,25	20,67	21,49	21,21
Promedio global	19,23	18,67	19,13	18,99	18,68

Gráfico 25. Promedios de respostas segundo a lingua habitual



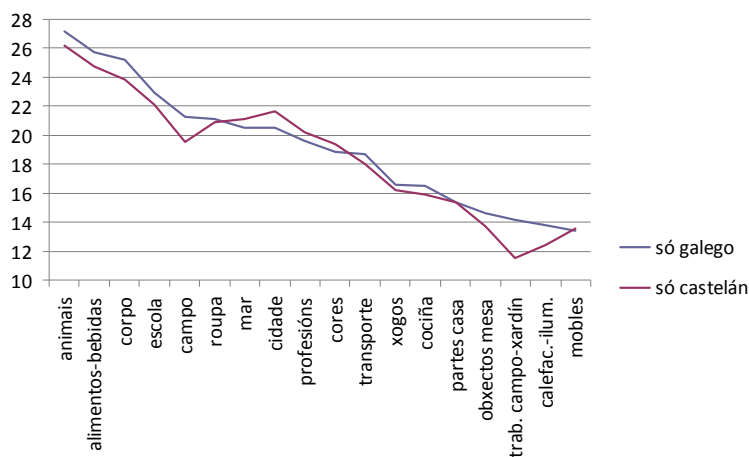
Se comparamos as dúas poboacións máis afastadas aplicando unha combinación de ambas as variables, que son as constituídas polos informantes que declararon aprender a falar en galego e empregar esta lingua exclusivamente (98 rapaces) e os que afirmaron aprender en castelán e falar só cas-

telán (121), de novo as diferenzas son inapreciables; tanto, que nin siquiera necesitamos o referendo da estatística (o valor P chega a 0,7072) para comprobalo: basta con observar as cifras que acadan os promedios de produción de palabras que se expoñen na táboa 33 e no gráfico 26.

Táboa 33. Promedios de respostas segundo a lingua inicial e habitual

Centros de interese		Só gal.	Só cast.
01	O corpo humano	25,19	23,83
02	A roupa	21,09	20,88
03	As partes da casa (sen mobles)	15,37	15,40
04	Os mobles da casa	13,40	13,57
05	Os alimentos e as bebidas	25,73	24,73
06	Os obxectos colocados na mesa...	14,65	13,70
07	A cociña e os seus utensilios	16,48	15,93
08	A escola: mobles e materiais	22,96	22,11
09	A calefacción e a iluminación	13,76	12,40
10	A cidade	20,51	21,64
11	O campo	21,26	19,50
12	Os medios de transporte	18,67	18,00
13	Os traballos do campo e do xardín	14,18	11,53
14	Os animais	27,17	26,20
15	Os xogos e as distraccións	16,58	16,17
16	As profesións e os oficios	19,63	20,20
17	As cores	18,82	19,36
18	O mar	20,55	21,09
Promedio global		19,22	18,68

Gráfico 26. Promedios de respostas segundo a lingua inicial e habitual



Talvez si caiba comentar que os CI nos que se aprecia maior diferenza entre os dous grupos son o 13 'Os traballos do campo e do xardín', no que a distancia acada os 2,65 puntos, e o 11 'O campo', no que chega a 1,76, en ambos os casos a favor dos estudantes galegofalantes, precisamente dúas áreas temáticas (ou máis ben unha, pois en realidade 13 está contido en 11) que se relacionan dun xeito máis directo co ambiente no que o galego se desenvolve con máis vigor, o ambiente rural. Pola contra, a cidade e o mar, asociados en Galicia a un uso maioritario do castelán, son CI nos que os castelanfalantes proporcionan máis elementos léxicos, aínda que neste caso as diferenzas son francamente pequenas. Vexamos a continuación se os distintos tipos de poboación inflúen ou non na produción de palabras.

3.2.1.3 O tipo de poboación

Como xa indicamos no capítulo 2, en Tipo de poboación diferenciamos aquel no que está situado o centro de ensino do que corresponde ao domicilio familiar. En ambos os dous casos estamos ante variables sociais sen influencia na produción de palabras, segundo se manifesta nas táboas 34 e 35, a través das cales comprobamos tamén que apenas hai diferenzas nos resultados dunha variable e outra. Por outra banda, que sexan os estudantes dos centros rurais os que nos dous casos presentan os promedios máis baixos non nos parece significativo, dada a proximidade que hai entre todos os promedios; desde logo, non o é estatisticamente, xa que o valor P chega a 0,8618 na variable Tipo de poboación do centro de ensino e a 0,8939 na variable Tipo de poboación do domicilio familiar, cifras que se afastan máis de 0,05 das que levamos vistas ata o de agora e que mesmo se sitúan moi preto de 1.

Táboa 34. Promedios de respostas segundo o tipo de poboación do centro de ensino

Centros de interese	Urbana	Periurb.	Vila	Rural
01 O corpo humano	24,89	25,42	25,41	22,66
02 A roupa	21,00	22,48	20,71	19,48
03 As partes da casa (sen mobles)	15,24	17,15	15,69	14,78
04 Os mobles da casa	14,02	14,25	13,78	13,27
05 Os alimentos e as bebidas	25,20	27,31	25,41	24,80
06 Os obxectos colocados na mesa...	14,08	15,23	13,83	13,39
07 A cociña e os seus utensilios	16,32	17,19	16,21	16,39
08 A escola: mobles e materiais	21,95	20,40	22,38	21,98
09 A calefacción e a iluminación	13,56	13,36	13,01	12,71
10 A cidade	21,67	19,94	20,59	20,01
11 O campo	20,40	20,43	20,22	19,77
12 Os medios de transporte	18,01	17,84	18,22	18,21
13 Os traballos do campo e do xardín	12,50	13,93	13,74	13,25
14 Os animais	26,39	27,05	25,99	25,98
15 Os xogos e as distraccións	16,51	17,48	15,78	15,30
16 As profesións e os oficios	20,31	19,78	19,38	19,55
17 As cores	19,16	19,09	19,11	18,19
18 O mar	21,12	22,25	20,43	20,53
Promedio global	19,02	19,48	18,88	18,35

Táboa 35. Promedios de respostas segundo o tipo de poboación da residencia familiar

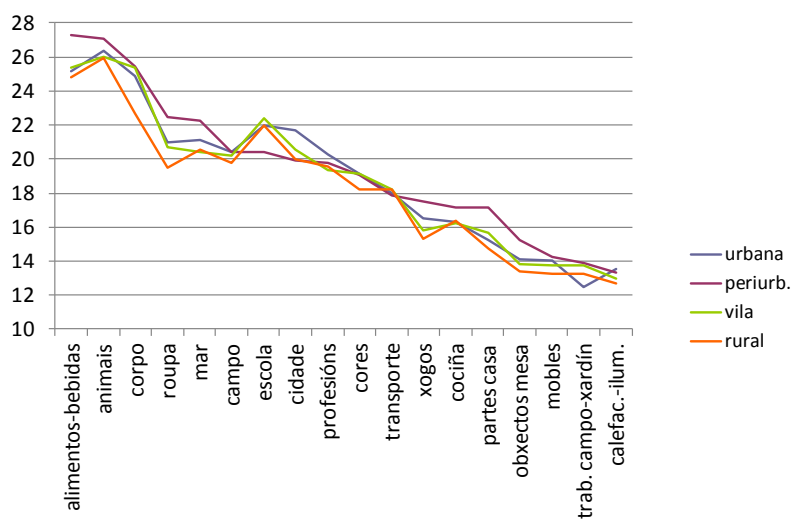
Centros de interese	Urbana	Periurb.	Vila	Rural
01 O corpo humano	24,85	25,74	24,78	24,04
02 A roupa	20,77	22,16	20,82	20,43
03 As partes da casa (sen mobles)	15,16	16,46	15,80	15,38
04 Os mobles da casa	14,13	13,46	13,83	13,90
05 Os alimentos e as bebidas	24,89	26,75	25,52	25,63
06 Os obxectos colocados na mesa...	14,03	14,98	13,92	13,74
07 A cociña e os seus utensilios	16,28	16,56	16,53	16,63
08 A escola: mobles e materiais	21,86	20,31	22,50	22,41
09 A calefacción e a iluminación	13,52	12,93	13,07	13,30
10 A cidade	21,76	20,26	20,90	20,35
11 O campo	20,30	20,33	19,30	20,75
12 Os medios de transporte	18,14	17,55	18,51	18,11
13 Os traballos do campo e do xardín	12,37	13,54	13,59	13,68
14 Os animais	26,22	27,41	25,72	26,17
15 Os xogos e as distraccións	16,53	16,91	15,94	15,87
16 As profesións e os oficios	20,35	20,01	19,59	19,61
17 As cores	19,20	19,01	19,68	18,32
18 O mar	20,95	22,38	20,72	20,74
Promedio global	18,96	19,26	18,93	18,84

Na táboa 34 obsérvase que os estudantes que se forman en centros rurais producen promedios inferiores en doce dos dezoito CI e os que asisten a centros en poboacións periurbanas ostentan os promedios máis altos noutras doce, e este dato dalgún xeito singulariza os dous tipos de poboación, e constitúe o único trazo diferencial que podemos detectar no que a esta variable se refire; no entanto, tal trazo é inexistente cando o tipo de poboación se refire á residencia familiar, xa que neste caso non podemos destacar ningún dos catro tipos. De feito, a poboación periurbana baixa de doce a oito CI con promedios máis elevados, e ademais en catro CI revélase como a menos produtiva (no caso da anterior variable érao en dous); así mesmo, a poboación rural baixa de doce a cinco CI con promedios máis baixos, e ademais en tres deles, ‘A cociña’, ‘O campo’ e ‘Os traballos do campo e do xardín’, resulta ser a máis produtiva (no caso da anterior variable non o era en ningún).

Examinamos tamén os datos combinando as dúas variables, de xeito que seleccionamos os informantes que viven e estudan no mesmo tipo de poboación (véxase a táboa 6), e puidemos comprobar que os resultados son moi similares aos anteriores: de novo destacan —pero tamén minimamente, iso si— as poboacións periurbana e rural como as que presentan as porcentaxes máis altas e máis baixas, respectivamente.

Dado que as diferenzas, insistimos, son mínimas, a visualización das mesmas faise bastante difícil a través do gráfico, no que as liñas aparecen cruzándose e superpoñéndose, produto da similitude dos promedios. Por iso só ofrecemos o relativo á variante Tipo de poboación do centro, no que se evidencia a súa irrelevancia.

Gráfico 27. Promedios de respostas segundo o tipo de poboación do centro de ensino



3.2.1.4 O tipo de centro

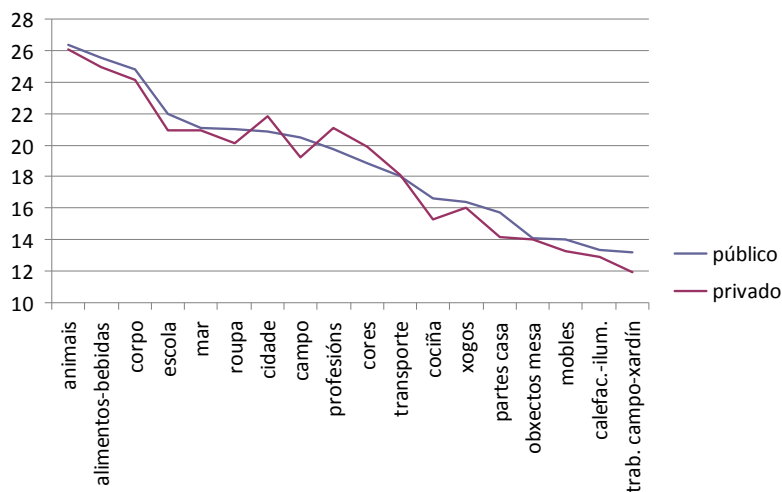
Na mostra tamén diferenciamos os dous tipos de ensino, público e privado, e xa sinalamos no capítulo 2 que con esta variable se establecen dous grupos de tamaños moi desiguais (véxase o gráfico 4), o cal é debido á presenza pouco relevante que ten o ensino privado no nivel do bacharelato en Galicia. Este é, por tanto, un dos casos nos que se revela a conveniencia de traballar con promedios, e non con cifras absolutas.

De novo os datos evidencian a pouca ou nula influencia que esta variable ten na produción de palabras, que apenas se distancia, nos promedios globais, en 0,40 puntos. De feito, en once dos dezoito CI (máis do 60%) non chega a un punto, e nos CI 18 ‘O mar’, 12 ‘Os medios de transporte’ e 06 ‘Os obxectos colocados na mesa para a comida’ a desigualdade é case inexistente. A maior separación, que advertimos no CI 03 ‘As partes da casa’, só chega a 1,54 puntos. Por iso, ao aplicar ANOVA a esta variable, constatamos que o *valor P* é 0,7745.

Táboa 36. Promedios de respostas segundo o tipo de centro

Centros de interese	Público	Privado
01 O corpo humano	24,80	24,11
02 A roupa	21,04	20,09
03 As partes da casa (sen mobles)	15,72	14,18
04 Os mobles da casa	14,00	13,30
05 Os alimentos e as bebidas	25,52	24,94
06 Os obxectos colocados na mesa...	14,06	14,05
07 A cociña e os seus utensilios	16,64	15,25
08 A escola: mobles e materiais	22,01	20,96
09 A calefacción e a iluminación	13,38	12,87
10 A cidade	20,85	21,84
11 O campo	20,47	19,19
12 Os medios de transporte	18,05	18,13
13 Os traballos do campo e do xardín	13,22	11,93
14 Os animais	26,38	26,09
15 Os xogos e as distraccións	16,36	15,99
16 As profesións e os oficios	19,75	21,07
17 As cores	18,82	19,86
18 O mar	21,06	20,91
Promedio global	19,01	18,60

Gráfico 28. Promedios de respostas segundo o tipo de centro



Se en todas as variables anteriores os datos coinciden, no esencial, cos relativos á lingua castelá (*vid.* López Meirama 2008: 111 e ss.), os que proporcióna o tipo de centro apuntan noutra dirección, xa que no LDEG os mozos formados nos centros de ensino privados produciron máis palabras ca os dos centros públicos (con promedios globais de 23,40 e 21,12 palabras, respectivamente, recollidos na táboa 29), mentres que, como estamos vendo, no léxico galego dispoñible son estes últimos os que presentan promedios máis altos, aínda que agora as diferenzas sexan menores⁷⁷. Habida conta dos desequilibrios no peso que os tres niveis socioculturais teñen nos dous tipos de centro (causados, no esencial, pola preferencia dos pais universitarios por educar aos seus fillos en centros privados), en López Meirama (2008) cruzáronse os datos de ambas as variables, e a conclusión foi que “la única variable social que parece tener una cierta relevancia en la producción de palabras, en términos globales, es la relativa al nivel sociocultural” (*ibíd*: 125). Comprobaremos a continuación se a situación é a mesma no subproxecto galego.

3.2.1.5 O nivel sociocultural

Lembremos que as tres poboacións que integran este valor establécense tomando como base a formación académica dos pais —só secundariamente se ten en conta a súa actividade laboral—, de xeito que o NSC alto corresponde a pais universitarios, o NSC baixo aos que carecen de formación ou teñen estudos primarios e o NSC medio aos que cursaron formación profesional ou

⁷⁷ De todos os xeitos, os resultados do LDEG tampouco son relevantes estatisticamente, cun *valor P* establecido en 0,2029.

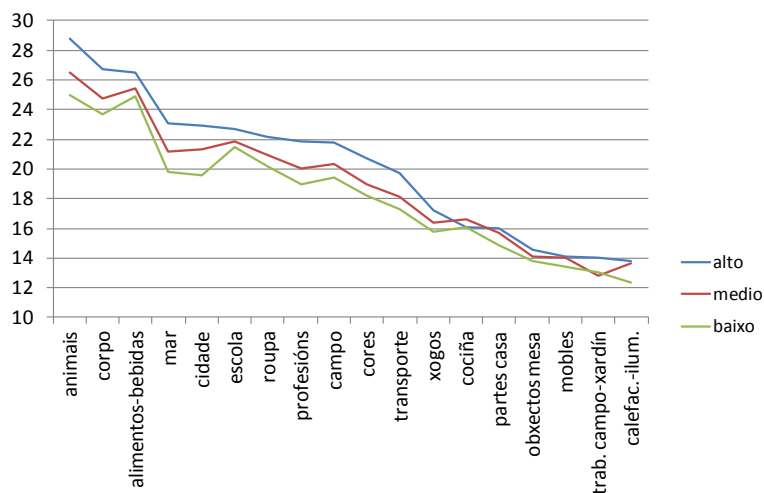
bacharelato. Aínda que existe, como é obvio, unha relación directa entre o nivel de estudos e o ámbito profesional, e daquela o nivel de ingresos, interesa resaltar que nos proxectos de léxico dispoñible de Galicia está establecido, como dicimos, basicamente en termos de formación académica, de maneira que, por exemplo, un informante cun proxenitor mestre e outro enfermeiro corresponde ao NSC alto, mentres que outro cuxa familia rexenta un establecemento de hostelería que proporciona ingresos elevados pode corresponder ao baixo.

Na derradeira fila da táboa 37 observamos que o promedio global de palabras por CI diminúe desde o NSC alto ao baixo en case dous puntos, e se nos fixamos en todos e cada un dos CI, comprobamos que o NSC baixo presenta os promedios máis baixos en todos eles e o NSC alto os máis elevados, coa excepción do CI 07 ‘A cociña e os seus utensilios’, no que son os informantes do NSC medio os que proporcionan máis palabras. Esta excepción explícase polas características do propio CI, moi compacto e limitado, que estorban a variabilidade; de feito, son precisamente este tipo de CI os que producen, en xeral, as diferenzas menos acusadas: ademais do 07, os CI 03 ‘As partes da casa’, 04 ‘Os mobles da casa’ e 06 ‘Os obxectos colocados na mesa para a comida’.

Táboa 37. Promedios de respostas segundo o NSC

Centros de interese	Alto	Medio	Baixo
01 O corpo humano	26,75	24,77	23,67
02 A roupa	22,18	20,96	20,18
03 As partes da casa (sen mobles)	15,99	15,67	14,85
04 Os mobles da casa	14,12	14,06	13,42
05 Os alimentos e as bebidas	26,53	25,46	24,89
06 Os obxectos colocados na mesa...	14,58	14,09	13,79
07 A cociña e os seus utensilios	16,12	16,63	16,09
08 A escola: mobles e materiais	22,73	21,86	21,45
09 A calefacción e a iluminación	13,82	13,62	12,39
10 A cidade	22,95	21,30	19,55
11 O campo	21,79	20,34	19,46
12 Os medios de transporte	19,74	18,10	17,28
13 Os traballos do campo e do xardín	14,03	12,79	13,05
14 Os animais	28,80	26,47	24,99
15 Os xogos e as distraccións	17,25	16,36	15,78
16 As profesións e os oficios	21,85	20,04	19,00
17 As cores	20,72	19,00	18,23
18 O mar	23,09	21,21	19,81
Promedio global	20,17	19,04	18,22

Gráfico 29. Promedios de respostas segundo o NSC



A pesar de que, segundo se visualiza no gráfico 29, semella que hai unha clara relación entre a produtividade léxica dos informantes e o nivel de estudos dos seus proxenitores, esta non chega, nin con moito, a ser estatisticamente significativa, xa que o *valor P* se establece en 0,3888, moi lonxe de 0,05, valor que aseguraría un 95% de probabilidades de acerto da hipótese segundo a cal o factor do nivel formativo dos pais inflúe na produción de palabras. Tampouco o contraste entre o NSC alto e o baixo ofrece resultados relevantes, xa que, aínda que neste caso o *valor P* baixa, só o fai ata 0,1878. Destacan, iso si, catro CI con distancias superiores aos 3 puntos: 01 ‘O corpo humano’, 18 ‘O mar’, 10 ‘A cidade’ e 14 ‘Os animais’, centro este último que case chega aos catro.

A situación é similar á que se observa no subproxecto do castelán, pero neste as diferenzas entre as medias son máis acusadas, ademais de que o NSC alto se distingue sobre os outros dous, pois existe maior distancia entre este nivel e o medio que entre o medio e o baixo, circunstancia que non se dá no subproxecto galego. Por iso a comparación entre os dous NSC extremos revela máis desigualdade no LDEG: a media de distancia entre ambos establécese en 3 puntos, cifra que se supera en oito dos CI; a media mesmo é maior de catro nalgúns deles: 18 ‘O mar’ (4,14 puntos), 14 ‘Os animais’ (4,54), 01 ‘O corpo humano’ (4,63) e 10 ‘A cidade’ (4,96). Estas cantidades supoñen entre o 15% e o 25% de diferenza na produción de palabras entre ambos os grupos, polo que non é de estrañar que o cálculo do *valor P* dea como resulta-

do 0,0824. Por tanto, como indicamos liñas atrás, o NSC destaca como a única variable que, ata certo punto —e sen o referendo da estatística— inflúe na produción de palabras no subproxecto do castelán. No entanto, non estamos seguras de poder dicir o mesmo respecto á lingua galega, xa que neste caso a desemeianza é menor. Unha posible explicación a esta disimilitude entre ambos os dous subproxectos podería atoparse no feito de que boa parte dos informantes do NSC baixo son galegofalantes, circunstancia que atenuaría as diferenzas no subproxecto do galego e mesmo podería incrementalas no subproxecto do castelán.

Debemos concluír, pois, que ningún dos factores sociais manexados na mostra intervén en que os rapaces teñan máis ou menos palabras dispoñibles na súa memoria cando pensan en galego, o cal, como xa indicamos, non pode traducirse en que todos os estudantes de segundo curso de bacharelato, vivan onde vivan, falen o que falen, sexan do sexo que sexan ou medren nunha familia cun nivel formativo ou outro, van acadar o mesmo nivel de competencia nesta lingua; lembremos que o castelán está presente de maneira relativamente importante nas enquisas. A análise das distintas variantes que asociamos a cada lema das dezaioito listaxes en relación coas variables de estratificación social pode achegar datos valiosos a este respecto, e por tanto haberá que encaralo no futuro (*vid.* Álvarez de la Granja (no prelo)). O que si podemos asegurar agora é que ningunha destas variables funciona como un factor que condicione aos nosos informantes á hora de encher as enquisas; en xeral, os promedios de respostas por suxeito son moi homoxéneos en cada CI.

3.2.2. As variables sociais e a compatibilidade léxica

Os programas empregados na informatización dos datos, incluído *Dispogal*, permiten realizar análises comparativas respecto á selección léxica dos informantes, xa que ofrecen información sobre a cantidade de palabras comúns e diverxentes en cada CI. Esta información manexa sempre as cifras relativas a dous grupos; así, se tomamos, por exemplo, o CI 01 ‘O corpo humano’ e comparamos o número de lemas do grupo dos falantes exclusivos de galego co do grupo dos falantes exclusivos de castelán, o programa achega a información que recollemos na táboa 38. Nela, debe entenderse por proxecto 1 o constituído polos informantes que declararon falar só galego e por proxecto 2 o dos informantes que afirmaron falar só castelán:

Táboa 38. Compatibilidade no CI 01 dos falantes de galego e castelán

Cardinalidade do proxecto 1:	342	Cardinalidade do proxecto 2:	401
Cardinalidade da intersección:		229	
Cardinalidade da unión:		514	
Complemento do proxecto 1:	113	Complemento do proxecto 2:	172
Suma disxuntiva:		285	
Compatibilidade:		44,55%	
Contribución na intersección:	66,96%	Contribución na intersección:	57,11%

A táboa interprétase da seguinte maneira: os informantes que falan só galego produciron 342 lemas distintos, e os que falan só castelán, 401; dos 342 lemas do primeiro grupo, 229 coinciden cos dos segundo, mentres que 113 son exclusivos deles; e dos 401 lemas do segundo grupo, 229 coinciden cos dos primeiro, mentres que 172 son exclusivos deles. Se unimos os dous grupos, a cantidade de lemas elévase a 514, exactamente o resultado da suma dos lemas comúns (229) máis os exclusivos de cada grupo (113+172=285). Os lemas comúns constitúen o 44,55% de 514, polo que dicimos que a compatibilidade acada o 44,55%. Esta é, precisamente, a cifra que empregaremos na análise que a continuación expoñemos.

Como dicimos, o programa só aplica a comparación a dúas poboacións, pero dáse o caso de que a maior parte das variables de estratificación social dividen a mostra en máis de dous grupos; isto significa que teremos que deixar fóra boa parte dos grupos, limitándonos a contrastar as poboacións máis afastadas entre si de cada variable. A selección amosámola na táboa 39, na que indicamos tamén o número de informantes que corresponde a cada grupo:

Táboa 39. Grupos sociais escollidos e número de informantes

Sexo	feminino	427	masculino	373
Lingua inicial	galego	228	castelán	247
Lingua habitual	só galego	110	só castelán	142
Poboación do centro de ensino	urbana	432	rural	128
Poboación da residencia familiar	urbana	368	rural	207
Tipo de ensino	público	672	privado	128
Nivel sociocultural	alto	95	baixo	226

Na táboa 40 recolleemos os datos relativos aos dezaioito CI da mostra: as cifras das columnas representan a compatibilidade entre os dous grupos de cada variable de estratificación social, é dicir, reflicten a porcentaxe de lemas comúns sobre o total de lemas de ambos os dous grupos.

Táboa 40. Compatibilidade segundo as variables sociais

Centros de interese	Sexo	Ling. inic.	Ling. habit.	Pob. centro	Pob. resid.	Tipo ensino	NSC
01 O corpo humano	46,27	46,21	44,55	38,49	45,59	35,43	44,69
02 A roupa	46,08	46,32	43,46	38,90	42,33	33,92	40,77
03 As partes da casa (sen mobles)	41,80	39,22	34,36	32,48	40,39	27,92	39,49
04 Os mobles da casa	51,00	48,85	41,25	42,01	47,38	37,47	48,24
05 Os alimentos e as bebidas	50,83	49,72	45,83	43,13	48,68	36,64	46,08
06 Os obxectos colocados na mesa...	44,36	41,63	43,33	36,26	43,84	30,50	43,17
07 A cociña e os seus utensilios	49,03	46,29	46,44	42,68	45,88	34,31	43,82
08 A escola: mobles e materiais	43,38	40,43	39,32	32,15	38,02	26,49	39,97
09 A calefacción e a iluminación	38,62	35,60	29,73	30,99	35,71	21,27	32,98
10 A cidade	41,28	40,84	36,47	34,15	39,36	30,23	38,57
11 O campo	43,17	38,53	32,70	34,34	38,67	30,14	35,61
12 Os medios de transporte	39,98	39,40	36,35	36,44	39,31	27,41	37,58
13 Os traballos do campo e do xardín	32,29	26,69	21,38	23,76	28,20	18,38	26,45
14 Os animais	49,21	50,66	47,38	44,11	49,25	37,03	46,40
15 Os xogos e as distraccións	33,81	33,70	31,76	29,43	34,23	24,72	31,49
16 As profesións e os oficios	41,67	40,19	37,55	35,60	41,39	31,46	38,83
17 As cores	45,58	42,73	37,28	35,87	42,43	33,73	42,89
18 O mar	46,59	43,26	40,92	36,90	41,67	33,00	40,60
Medias	43,61	41,68	38,34	35,98	41,24	30,56	39,87

O primeiro que cabe destacar, observando na derradeira fila as medias de compatibilidade de cada variable, é que non se aprecian diferenzas considerables entre unhas variables e outras, xa que tales medias oscilan entre o 30,56% de compatibilidade media na variable Tipo de ensino e o 43,61% na variable Sexo. Máis amplas son as disimilitudes detectadas nas porcentaxes de compatibilidade dos distintos CI en cada variable: van desde os 26 puntos de distancia na variable Lingua habitual (na que coinciden o 47,38% dos lemas do CI 14 ‘Os animais’ e só o 21,38% do CI 13 ‘Os traballos do campo e do xardín’) ata os 18,71 puntos na variable Sexo (desde o 32,29% do CI 13 ao 51% do CI 04 ‘Os mobles da casa’). Semella, entón, que o tipo de ámbito temático inflúe dun xeito máis determinante ca as diferenzas sociais, e de feito algúns CI tenden a exhibir ou moita coincidencia ou, pola contra, moi pouca, independentemente da variable social considerada. Por exemplo, o CI 13 é o que presenta menores porcentaxes de coincidencia en todas e cada unha das variables sociais, seguido do CI 09 ‘A calefacción e a iluminación’ (nas variables Lingua habitual e Tipo de ensino) e do CI 15 ‘Os xogos e as distraccións’ (nas outras cinco variables). Todos eles, cómpre lembrar, ámbitos léxicos caracterizados no apartado

anterior coma os menos cohesionados da mostra. No outro extremo destaca o CI 14 ‘Os animais’ como o que presenta maior coincidencia de lemas nas variables relativas á lingua e ao tipo de poboación, así coma o CI 04 ‘Os mobles da casa’, nas variables Sexo, Tipo de ensino e NSC. E estes dous CI cóntanse, como xa vimos, entre os que exhiben maior consistencia léxica.

Na táboa 41 presentamos os rangos que acadan os CI en cada variable segundo a compatibilidade. Nela vese claramente a situación dos xa comentados CI 13, 09 e 15, canda os que poderíamos situar os CI 10 ‘A cidade’, 11 ‘O campo’ e 12 ‘Os medios de transporte’, todos eles caracterizados tamén como zonas do léxico amplas e difusas. Así mesmo, percíbese sen dúbidas que, ademais dos CI 14 e 04, o 05 ‘Os alimentos e as bebidas’ e o 07 ‘A cociña e os seus utensilios’ se manifestan de novo como áreas temáticas ben compactas e cohesionadas.

Táboa 41. Rangos das porcentaxes de compatibilidade en cada variable

Centros de interese	Sexo	Lingua inic.	Ling. habit.	Pob. centro	Pob. resid.	Tipo ensino	NSC
01 O corpo humano	6	6	4	6	5	4	4
02 A roupa	7	4	5	5	8	6	8
03 As partes da casa (sen mobles)	12	14	14	14	11	13	11
04 Os mobles da casa	1	3	7	4	3	1	1
05 Os alimentos e as bebidas	2	2	3	2	2	3	3
06 Os obxectos colocados na mesa...	9	9	6	9	6	10	6
07 A cociña e os seus utensilios	4	5	2	3	4	5	5
08 A escola: mobles e materiais	10	11	9	15	15	15	10
09 A calefacción e a iluminación	16	16	17	16	16	17	16
10 A cidade	14	10	12	13	12	11	13
11 O campo	11	15	15	12	14	12	15
12 Os medios de transporte	15	13	13	8	13	14	14
13 Os traballos do campo e do xardín	18	18	18	18	18	18	18
14 Os animais	3	1	1	1	1	2	2
15 Os xogos e as distraccións	17	17	16	17	17	16	17
16 As profesións e os oficios	13	12	10	11	10	9	12
17 As cores	8	8	11	10	7	7	7
18 O mar	5	7	8	7	9	8	9

Con todo, os rangos que corresponden a algúns centros de interese resultan, cando menos, un tanto peculiares: en particular, chama a atención que o CI que caracterizamos como o máis compacto da mostra, 17 ‘As cores’, oscile entre os rangos 7 e 11, e que un dos máis cohesionados, 03 ‘As partes da casa’, o faga entre os rangos 11 e 14. Claro que estes CI son moi homoxéneos na par-

te alta, a parte na que se sitúan as palabras máis dispoñibles, pero menos na baixa, onde se agrupan as palabras recollidas ao final das enquisas e emitidas por moi poucos informantes.

Ademais disto, existe un dato moi importante que debemos ter en conta: se relacionamos as medias de compatibilidade correspondentes ás distintas variables de estratificación social co número de informantes que acada cada un dos dous grupos comparados nesas variables, recollido na táboa 39, podemos constatar que, deixando a un lado a excepción da lingua, a maior ou menor compatibilidade está relacionada directamente coa diferenza de tamaño entre os dous grupos analizados en cada variable, de xeito que canto maior é a distancia entre tales grupos, menor é a porcentaxe de palabras coincidentes. Isto pode explicar que atopemos a maior compatibilidade media na variable Sexo, xa que o número de mulleres e de homes é moi próximo, ou que a compatibilidade media máis baixa corresponda á variable Tipo de ensino, sendo que, como indicamos, a cantidade de informantes dos centros privados é moi baixa en comparación coa dos informantes dos centros públicos (a proporción aproximada é de 1/6).

Na táboa 42 recollemos, a modo de exemplo, o número de lemas contabilizados en cada un dos grupos no primeiro dos ámbitos léxicos da mostra, 'O corpo humano'. Nela podemos constatar que, como sería de esperar, os grupos máis nutridos son os que producen máis lemas (p.e., os estudantes formados nos centros de ensino público ou nos centros de ensino das cidades) e, pola contra, os máis reducidos son os que producen menos lemas (p.e., os dos centros privados, os do NSC alto ou os que só falan galego). Sendo isto así, resulta evidente que a compatibilidade sempre será moito menor cando se contrasten as cifras dos grupos máis disímiles entre si, xa que como mínimo os lemas que haxa de diferenza entre un grupo e outro serán lemas diverxentes. Por exemplo, no caso da variable Tipo de ensino, hai un mínimo de 450 lemas diverxentes, que son os que separan os emitidos polos rapaces que empregan o ensino público dos proporcionados polos que asisten a centros privados (806-356).

Táboa 42. Número de lemas no CI 01 segundo as variables sociais

Sexo	feminino	635	masculino	620
Lingua inicial	galego	509	castelán	551
Lingua habitual	só galego	342	só castelán	401
Poboación do centro de ensino	urbana	670	rural	359
Poboación da residencia familiar	urbana	594	rural	495
Tipo de ensino	público	806	privado	356
Nivel sociocultural	alto	371	baixo	474

Por outra banda, convén recordar que nas listaxes se contabiliza un bo número de lemas que son palabras proporcionadas por un único informante, lemas que, a fin de contas, poderíamos considerar residuais. Obviamente, todas estas palabras vanse sumar ao grupo de lemas diverxentes; de aí que as porcentaxes medias de compatibilidade que vemos na táboa 40 apenas se sitúen, aproximadamente, entre o 30% e o 40%.

Todo isto lévanos a realizar un segundo cálculo da compatibilidade, esta vez limitándoo ás palabras máis dispoñibles das listaxes; concretamente, ás que acadan un índice de dispoñibilidade superior a 0,02. Así, o número de lemas por grupo está moito máis próximo, independentemente da distancia que haxa no tamaño dos grupos, como podemos comprobar na táboa 43, na que amosamos, como exemplo, o número de lemas con dispoñibilidade superior a 0,02 no CI 01:

Táboa 43. Número de lemas con ID > 0,02 no CI 01 segundo as variables sociais

Sexo	mulleres	93	homes	103
Lingua inicial	galego	107	castelán	94
Lingua habitual	só galego	102	só castelán	98
Poboación do centro de ensino	urbana	102	rural	98
Poboación da residencia familiar	urbana	103	rural	102
Tipo de ensino	público	103	privado	96
Nivel sociocultural	alto	102	baixo	102

As cifras todas sitúanse ao redor dos cen lemas: a súa proximidade asegura, cremos, unha análise máis efectiva con resultados máis convincentes. Comprobamos, na táboa 44, que as porcentaxes de coincidencia aumentan considerablemente, como non podía ser doutro xeito, habida conta de que nos limitamos ás formas léxicas proporcionadas por un maior número de informantes. Se observamos a fila que recolle as medias, verificaremos que as variables Sexo, Lingua inicial e Poboación da residencia familiar seguen

a ser as que exhiben os promedios de compatibilidade máis altos; mentres que as variables Nivel sociocultural e Lingua habitual pasan de ocupar unha posición media a ser as que menor compatibilidade presentan. Pero entre todas as variables destaca a do Tipo de ensino, que do 30,56% de compatibilidade de todos os lemas —a máis baixa— pasa ao 67,84% nas palabras máis dispoñibles.

Táboa 44. Compatibilidade das palabras máis dispoñibles segundo as variables sociais

Centros de interese	Sexo	Ling. inic.	Ling. habit.	Pob. centro	Pob. resid.	Tipo ensino	NSC
01 O corpo humano	76,58	77,88	75,44	72,41	75,21	73,04	70,00
02 A roupa	71,00	70,33	60,40	72,09	72,09	79,35	75,00
03 As partes da casa (sen mobles)	73,24	71,23	55,70	76,47	72,46	74,63	72,22
04 Os mobles da casa	84,00	68,83	59,49	68,00	76,00	77,63	65,82
05 Os alimentos e as bebidas	75,00	69,57	62,15	66,27	70,91	67,66	59,67
06 Os obxectos colocados na mesa...	75,81	68,75	55,38	61,82	69,64	65,63	61,19
07 A cociña e os seus utensilios	78,43	62,38	59,22	58,93	63,30	70,00	54,72
08 A escola: mobles e materiais	83,16	78,72	66,67	67,35	69,70	70,30	72,45
09 A calefacción e a iluminación	64,10	64,71	60,81	63,89	68,83	61,64	58,90
10 A cidade	72,44	70,73	66,13	65,38	64,39	62,32	57,34
11 O campo	62,84	47,88	40,34	58,17	59,48	55,26	52,87
12 Os medios de transporte	77,22	70,13	59,76	65,43	61,63	68,75	60,00
13 Os traballos do campo e do xardín	71,76	62,96	53,19	58,89	62,22	67,47	57,45
14 Os animais	67,53	65,00	62,28	64,14	63,16	61,64	57,58
15 Os xogos e as distraccións	61,16	66,12	57,26	67,59	62,93	64,04	62,60
16 As profesións e os oficios	70,80	73,48	60,81	65,44	74,62	67,14	61,90
17 As cores	82,76	83,93	71,67	80,70	81,36	72,58	73,85
18 O mar	70,97	64,78	55,84	54,94	62,91	62,11	60,82
Medias	73,27	68,75	60,14	66,00	68,38	67,84	63,02

Por outra banda, podemos contrastar a diferenza que existe nos rangos que ocupan os CI en cada variable: se examinamos a táboa 45, comprobaremos que hai poucas coincidencias respecto aos establecidos coas listaxes completas de lemas (expostos na táboa 41): podemos dicir que 01 ‘O corpo humano’ é un dos poucos CI que en ambos os casos se sitúa entre os máis compatibles, ou que tanto 11 ‘O campo’ coma 13 ‘Os traballos do campo e do xardín’ ocupan os postos máis baixos, aínda que neste caso haxa un intercambio de rangos. Con todo, destacan moito máis as diverxencias: por exemplo, os CI 05 ‘Os alimentos e as bebidas’, 06 ‘Os obxectos colocados na mesa para a comida’, 14 ‘Os animais’ ou 18 ‘O mar’ ocupan rangos inferiores de compatibilidade cando

se examinan as palabras máis dispoñibles, mentres que os CI 08 ‘A escola’ ou 16 ‘As profesións e os oficios’ ocupan rangos máis elevados⁷⁸.

Táboa 45. Rangos das porcentaxes de compatibilidade das palabras máis dispoñibles en cada variable

Centros de interese	Sexo	Ling. inic.	Ling. habit.	Pob. centro	Pob. resid.	Tipo ensino	NSC
01 O corpo humano	6	3	1	3	3	4	5
02 A roupa	12	7	9	4	6	1	1
03 As partes da casa (sen mobles)	9	5	15	2	5	3	4
04 Os mobles da casa	1	10	11	5	2	2	6
05 Os alimentos e as bebidas	8	9	6	8	7	9	12
06 Os obxectos colocados na mesa...	7	11	16	14	9	12	9
07 A cociña e os seus utensilios	4	17	12	15	12	7	17
08 A escola: mobles e materiais	2	2	3	7	8	6	3
09 A calefacción e a iluminación	16	15	8	13	10	16	13
10 A cidade	10	6	4	11	11	14	16
11 O campo	17	18	18	17	18	18	18
12 Os medios de transporte	5	8	10	10	17	8	11
13 Os traballos do campo e do xardín	11	16	17	16	16	10	15
14 Os animais	15	13	5	12	13	17	14
15 Os xogos e as distraccións	18	12	13	6	14	13	7
16 As profesións e os oficios	14	4	7	9	4	11	8
17 As cores	3	1	2	1	1	5	2
18 O mar	13	14	14	18	15	15	10

No entanto, o dato máis interesante que podemos extraer da visualización desta táboa, desde o noso punto de vista, é a enorme distancia que ás veces separa os rangos que acadan as poboacións dunhas e outras variables no mesmo CI. Por exemplo, o NSC e o Tipo de ensino semellan influír pouco na selección léxica dos informantes no que atinxe á área temática da roupa; cando menos, é o CI no que estes grupos de informantes chegan ás maiores porcentaxes de compatibilidade (por iso ten nos dous casos o rango 1); pola contra, a variable Sexo si parece influír, ou polo menos faino máis ca noutros CI, xa que nesta variable ocupa o rango 12 de compatibilidade. Outro exemplo que poderíamos destacar é o do CI 04 ‘Os mobles da casa’, que ten rangos altos de compatibilidade en practicamente todas as variables salvo nas dúas relativas á lingua, onde ocupa rangos medio-baixos; pola contra, a lingua semella influír menos na selección das voces asociadas á cidade, mentres que

⁷⁸ Consideramos que o rango 1 é o máis elevado porque é o que se relaciona coa maior compatibilidade e, á inversa, o 18 é o máis baixo porque se asocia á menor compatibilidade.

outras variables, nomeadamente o Tipo de ensino e o NSC si poderían ter máis influencia; cando menos, iso parecen dicir os rangos que o CI ocupa nelas (14 e 16, respectivamente). Estas diferenzas suxiren que, no que se refire ás palabras máis dispoñibles, as variables de estratificación social poden contribuir á converxencia léxica nas diferentes zonas do léxico, ou, cando menos, poden facelo nalgunhas desas zonas e en certa medida.

Se analizamos, por exemplo, a variable Sexo, comprobamos que o CI 15 ‘Os xogos e as distraccións’ é o que presenta menos compatibilidade, e se observamos as partes máis altas das dúas listaxes (concretamente, os primeiros corenta elementos), na relativa ás rapazas atopamos palabras como *bailar*, *boneca*, *cantar* ou *mariola*, palabras todas que ocupan postos máis baixos na listaxe dos rapaces, mentres que nesta última vemos outras como *billar*, *deporte*, *futbolín* ou *natación*, que, inversamente, acadan índices menores de dispoñibilidade na listaxe das rapazas. Así mesmo, no CI 11 ‘O campo’, que ocupa o rango 17 de compatibilidade, palabras como *boi*, *esterco* ou *fouciño* son máis dispoñibles para os homes ca para as mulleres, mentres que *naturaleza*, *planta* ou *toxos* sono para elas; no CI 09 ‘A calefacción e a iluminación’, que ocupa o rango 16, palabras como *fusible*, *gas natural*, *lareira* ou *neon* son máis dispoñibles para eles, mentres que a elas lles chegaron antes á mente outras como *claridade*, *enerxía*, *fogueira* ou *iluminar*. Todas estas voces evocan diferenzas —probablemente moi leves— entre os dous grupos no que respecta ás súas preferencias ou mesmo ao seu modo de interpretar diversos aspectos da realidade: semella que os rapaces están máis interesados polos deportes, que interpretan máis facilmente o campo como un medio de vida e que os aspectos da tecnoloxía lles interesan máis ca a elas. Naturalmente, hai ámbitos nos que apenas se detectan diferenzas entre un grupo e outro; é o caso, por exemplo, dos mobles da casa, a escola ou as cores, tres CI nos que a compatibilidade supera o 80%.

Canto ás dúas variables relativas á lingua, constatamos que, en xeral, coinciden bastante nos rangos que ocupan os CI: así, as áreas temáticas que acadan maior compatibilidade en ambas son os CI 17 ‘As cores’, 08 ‘A escola’ e 01 ‘O corpo humano’, mentres que outros CI, como o 11 ‘O campo’, o 13 ‘Os traballos do campo e do xardín’ ou o 18 ‘O mar’ se sitúan entre os que manifestan menos compatibilidade. Non obstante, hai algúns CI que presentan rangos afastados entre si: por exemplo, o 07 ‘A cociña e os seus utensilios’ só chega ao rango 17 na variable Lingua inicial e ocupa o 12 na variable Lingua habitual; e o CI 09 ‘A calefacción e a iluminación’ ten o rango 15 na Lingua inicial e o 8 na Lingua habitual. Moito maiores son as diferenzas de rango nos CI 14 ‘Os animais’ e 03 ‘As partes da casa’, que acadan, respectivamente,

os oito e os dez puntos. Tales diferenzas poden explicarse, cando menos o caso do CI 14, pola proximidade das porcentaxes de compatibilidade en cada variable: os rapaces que aprenderon a falar en galego e os que o fixeron en castelán presentan o 65% de compatibilidade, e os que falan só galego e só castelán chegan ao 62,28%, porcentaxes ambas moi próximas entre si, pero que, situadas en cada variable, ocupan rangos distantes. No entanto, no CI 03 as porcentaxes de compatibilidade están bastante afastadas entre si: do 71,23% na variable *Lingua inicial* báixase ao 55,70% na *Lingua habitual*. En principio non parece haber razón ningunha que explique esta discordancia; desde logo, non se atopa no propio centro de interese, que presenta porcentaxes de compatibilidade similares (entre o 71% e o 76%, aproximadamente) en todas as variables excepto na da *Lingua habitual*. A disimilitude talvez poida xustificarse pola diferente densidade que presentan as dúas listaxes na súa parte máis alta: os informantes cuxa lingua habitual é exclusivamente o galego proporcionan 63 lemas cun índice de dispoñibilidade superior a 0,02, e estes lemas suman 1.395 palabras (a densidade é 22,14), mentres que os que teñen como lingua habitual só o castelán proporcionan 60 lemas que suman 1.823 palabras (a densidade sobe a 30,43). Isto significa que hai menos coincidencia nas respostas das palabras con maior dispoñibilidade no caso dos falantes de galego ou, dito doutro xeito, significa que os lemas máis dispoñibles da súa listaxe presentan frecuencias absolutas máis baixas ca os lemas máis dispoñibles da listaxe correspondente aos castelanfalantes. Sendo isto así, é obvio que os lemas diverxentes de ambas as listaxes teñen que sumar un número pequeno de palabras, e esta situación facilita que aumenten as diferenzas. De feito, na parte alta da listaxe dos 228 informantes que declaran expresarse exclusivamente en galego contabilizamos 14 lemas exclusivos, que suman 140 palabras: *ladrillo* (13 palabras), *tella* (13), *columna* (12), *horta* (12), *cemento* (11), *corte* (11), *azulexo* (10), *eira* (10), *bañeira* (9), *bufarda* (9), *azotea* (8), *oficina* (8), *tabique* (7) e *viga* (7); e na dos 247 informantes que afirman falar só castelán atopamos unicamente 7 lemas exclusivos, que suman 116 palabras: *lavabo* (19), *fachada* (18), *ximnasio* (18), *piso* (16), *sala de estudo* (16), *vestíbulo* (15) e *habitación de invitados* (14).

Esta xustificación, con todo, non é unha explicación cumprida, porque deixa sen aclarar por que é precisamente neste CI no que a parte alta da listaxe dos galegofalantes resulta ser máis difusa ca a dos castelanfalantes; ao respecto, só se nos ocorre aducir o tamaño das poboacións, que talvez non sexa o suficientemente amplo como para que non se dean algúns casos que rompan a tendencia xeral. De feito, detectamos que esta situación tamén se dá nas variables que se refiren ao Tipo de poboación, entre as que de novo atopamos

moita similitude: son os mesmos CI en ambas as variables os que exhiben as porcentaxes de compatibilidade máis elevadas (17 ‘As cores’, 03 ‘As partes da casa’, 01 ‘O corpo humano’, 02 ‘A roupa’ e 04 ‘Os mobles da casa’), as máis baixas (07 ‘A cociña e os seus utensilios’, 13 ‘Os traballos do campo e do xardín’, 11 ‘O campo’ e 18 ‘O mar’) e as medias, exceptuando os CI 15 ‘Os xogos e as distraccións’ e 12 ‘Os medios de transporte’, que presentan rangos inferiores cando a variable considerada é a da Poboación da residencia familiar. Centrándonos nos CI con menos compatibilidade, en concreto nas primeiras corenta palabras de cada listaxe, constatamos que os lemas disímiles revelan, aínda que debilmente, que o léxico dispoñible ata certo punto reflicte o medio no que se desenvolve o informante; así, os rapaces que estudan e viven no rural teñen máis dispoñibles no CI 07 voces como *cociña de leña* (os que asisten a clase no medio rural, tamén *cociña de ferro*), mentres que os que estudan e viven nas cidades pensaron antes en palabras como *lavadora*, *lavaplatos*, *tesoiras* ou *tostadora* (os que asisten a clase no medio urbano, tamén *lavalouza*); no CI 13 detectamos as voces *arado*, *fouciño*, *herba*, *muxir*, *pataca*, *rastrillo* e *terra* tanto nos que estudan como nos que habitan no rural, fronte a *agricultura*, *gandería*, *labrador*, *labrego*, *limpar*, *pastora*, *recolleita*, *talar* e *xardinería* nos que se educan e habitan nas cidades: observamos que as palabras do primeiro grupo denotan actividades, utensilios ou produtos concretos, mentres que as voces dos rapaces urbanos presentan un carácter máis xeral, o cal evidencia que estes teñen menor contacto cos labores do campo. Similar é o caso do CI 11, que nas listaxes urbanas conta con formas como *aldea*, *bosque* ou *planta*, fronte ás rurais, onde atopamos voces como *casa rural* ou *esterco*, así como *gandeiro* (só na listaxe da variable Centro de ensino), *burro*, *toxo* e *xabaryl* (só na listaxe da variable Domicilio familiar). Finalmente, no CI 18 as palabras que son máis dispoñibles para os rapaces que viven e estudan nas cidades suxiren unha concepción máis aberta do mar, que, entre outras cousas, tamén se ve como fonte de pracer: *centolo*, *golfinho*, *iate*, *océano*; mentres que as de maior índice de dispoñibilidade dos estudantes que se desenvolven no ámbito rural evocan un medio de vida; máis concretamente, a pesca de baixura: *cana*, *sardiña*, *xurelo*.

Polo que respecta ao Tipo de ensino e ao NSC, comprobamos que os rangos de compatibilidade que ocupan os CI se explican polas propias características destes, de xeito que os máis compactos, densos e homoxéneos son os que presentan os rangos máis altos en ambas as dúas variables: 01 ‘O corpo humano’, 02 ‘A roupa’, 03 ‘As partes da casa’, 04 ‘Os mobles da casa’, 08 ‘A escola’, 17 ‘As cores’; mentres que, pola contra, os máis difusos, heteroxéneos e amplos son os que ocupan os rangos máis baixos: 09 ‘A calefacción e a ilu-

minación', 10 'A cidade', 11 'O campo', 14 'Os animais'. Nalgúns CI atopamos certa distancia entre os rangos: os CI 07 'A cociña e os seus utensilios' e 13 'Os traballos do campo e do xardín' acadan rangos máis altos de compatibilidade na variable Tipo de ensino (10 e 7, respectivamente, fronte a 15 e 17 no NSC), e os CI 15 'Os xogos e as distraccións' e 18 'O mar' preséntanos máis altos na variable NSC (7 e 10, respectivamente, fronte a 13 e 15 no Tipo de ensino). No entanto, tras unha revisión dos lemas máis frecuentes de todas estas áreas temáticas nos distintos grupos, non atopamos ningunha tendencia en ningunha das listaxes que poida xustificar sequera minimamente estas diferenzas, así que supoñemos que, de novo, estamos ante pequenas anomalías ocasionadas polo tamaño das poboacións, moi reducido no caso do ensino privado e do nivel sociocultural alto. Só podemos concluír, entón, que ningunha destas dúas variables inflúe de xeito significativo na selección léxica dos informantes; cando menos, non o fan no que respecta ás voces máis dispoñibles.

Despois de todo o anterior, concluimos que a variabilidade léxica que podemos atopar entre as palabras con índices altos de dispoñibilidade depende en moita maior medida da área temática de que se trate, e non tanto da variable de estratificación social considerada. Semella que a idade e o nivel educativo dos nosos informantes son factores de homoxeneidade enormemente fortes, polo menos no que se refire ás palabras que lles veñen antes á memoria; sono moito máis, desde logo, ca o sexo, o tipo de ensino, o nivel sociocultural e mesmo que o lugar no que viven ou estudan e a lingua que declaran falar. Naturalmente, o que puidemos indicar nestas páxinas non pasa da simple suxestión, porque para chegar a conclusións máis firmes necesitaríamos desenvolver análises pormenorizadas e concienciadas, fóra do noso alcance aquí, e que por tanto reservamos para o futuro.

4 O DICCIONARIO DE LÉXICO DISPOÑIBLE: *DISPOGAL*

En § 2.4.2 explicamos sucintamente a información fornecida polas listaxes de dispoñibilidade a que se ten acceso a través de *Dispogal* ([www.http://gramatica.usc.es/disvogal/](http://gramatica.usc.es/disvogal/)), que constitúen o dicionario de léxico dispoñible, o núcleo e o obxectivo principal deste libro. Tales listaxes ofrécese, en forma de táboa, ordenadas polo índice de dispoñibilidade, de xeito que os lemas que ocupan os primeiros lugares corresponden ás formas léxicas máis dispoñibles, pero tamén poden ordenarse pola frecuencia (absoluta ou relativa) dos lemas, ou por estes mesmos, seguindo a orde alfabética. Para cambiar o criterio de ordenación, basta con pulsar sobre o título correspondente que aparece na cabeceira da táboa.

Pero, con ser a máis importante, esta tan só é unha pequena parte da información que se pode obter da aplicación. Na seguinte captura de pantalla aparecen, na segunda liña, as catro opcións de busca que se poden levar a cabo: Dispoñibilidade xeral, Índices cuantitativos, Conxuntos e Conxuntos avanzados. Da primeira das opcións, que é a que aparece por defecto, extráense as listaxes dos lemas de cada un dos CI, en principio ordenados por dispoñibilidade. No momento en que se elixe un CI en particular, despréganse unha serie de lapelas que serven para seleccionar grupos sociais específicos da mostra. Por exemplo, as rapazas de NSC medio que estudan no ensino público, ou os rapaces que só falan galego da provincia de Lugo.

Castellana - Galego

[Sair](#)

Inicio Datos Resultados

Disponibilidade xeral Índices cuantitativos Conxuntos Conxuntos avanzados

Disponibilidade xeral

PROXECTO: --- Selecciona un proxecto ---

Centro de estudio: --- Centro de estudio ---

CENTRO DE INTERESE: --- Centro de interese ---

VARIABLES:

Calcular

Universidade de Santiago de Compostela

A continuación ofrecemos un fragmento de pantalla cunha selección feita, a modo de exemplo. Nela pode observarse que, no caso das variables sociais que presentan máis de dous valores (a lingua habitual e a inicial, o NSC, a provincia e o tipo de poboación do centro educativo e do domicilio familiar), ofrecemos a posibilidade de seleccionar un ou dous destes valores, de xeito que as opcións das buscas se multiplican. No exemplo escollemos os informantes que falan só galego ou maioritariamente galego e que viven e estudan nas cidades ou en poboacións periurbanas.

Dispoñibilidade xeral						
PROXECTO:		Léxico dispoñible do galego ▼				
Centro de estudio:		--- Centro de estudio --- ▼				
CENTRO DE INTERESE:		A cidade ▼				
VARIABLES:		Sexo --- Seleccionar --- ▼ Lingua habitual Só galego ▼ Lingua habitual (2) Máis galego ca castelán ▼ Lingua inicial --- Seleccionar --- ▼ Lingua inicial (2) --- Seleccionar --- ▼ Nivel sociocultural --- Seleccionar --- ▼ Nivel sociocultural (2) --- Seleccionar --- ▼ Tipo de centro --- Seleccionar --- ▼ Provincia do centro --- Seleccionar --- ▼ Provincia do centro (2) --- Seleccionar --- ▼ Situación xeográfica do centro --- Seleccionar --- ▼ Tipo de poboación do centro Urbana ▼ Tipo de poboación do centro (2) Periurbana ▼ Tipo de poboación do domicilio familiar Urbana ▼ Tipo de poboación do domicilio familiar (2) Periurbana ▼				
<input type="button" value="Calcular"/>						
Media de respostas: 22,73 Total de informantes: 55 Total de respostas: 1.250		Gardar resultados: <input type="text"/> <input type="button" value="Gardar"/>			Exportar resultados	
Nº	Palabra	Dispoñibilidade	Aparición	F. Absoluta	F. Relativa	F. Acumulada
1	edificio	0,58633	76,364%	42	3,360%	3,360%

A segunda opción de busca, denominada Índices cuantitativos, ofrece os datos sobre produción de palabras dos grupos sociais que se seleccionen. Empregamos estes datos na elaboración do apartado 3.2.1 deste libro. No exemplo que ofrecemos a continuación, escollemos un grupo constituído polas rapazas que aprenderon a falar en castelán e galego, que habitualmente falan tanto galego como castelán e que viven na provincia da Coruña.

Nº	Centro Interese	Palabras Totais	P. Diferentes	P. por Informante	Índice de Cohesión
1	O corpo humano	303	111	21,64	0,19
2	A roupa	262	113	18,71	0,17
3	As partes da casa (sen mobles)	228	79	16,29	0,21
4	Os mobles da casa	195	75	13,93	0,19
5	Os alimentos e as bebidas	367	176	26,21	0,15
6	Os obxectos colocados na mesa para a comida	197	60	14,07	0,23
7	A cociña e os seus utensilios	258	115	18,43	0,16
8	A escola: mobles e materiais	284	115	20,29	0,18
9	A calefacción e a iluminación	179	97	12,79	0,13
10	A cidade	264	134	18,86	0,14
11	O campo	256	161	18,29	0,11
12	Os medios de transporte	226	85	16,14	0,19
13	Os traballos do campo e xardín	182	94	13,00	0,14
14	Os animais	315	146	22,50	0,15
15	Os xogos e as distraccións	201	133	14,36	0,11
16	As profesións e os oficios	247	108	17,64	0,16
17	As cores	266	86	19,00	0,22
18	O mar	276	160	19,71	0,12

As opcións de busca terceira e cuarta ofrecen, esencialmente, o mesmo tipo de información (a cantidade de palabras comúns e diverxentes nun CI) pero no caso da cuarta, ‘Conxuntos avanzados’, os datos sométense a tres tipos posibles de filtrado: ou ben se escollen as palabras con maior frecuencia que sumen unha frecuencia acumulada do 80%, ou ben as cen palabras máis dispoñibles, ou ben as que acadan un índice de dispoñibilidade maior de 0,02. Esta información é contrastiva, e mesmo se poden comparar os datos do proxecto galego cos do proxecto do español de Galicia.

Castellano - Galego

Inicio Datos Resultados [Salir](#)

[Dispoñibilidade xeral](#) [Índices cuantitativos](#) [Conxuntos](#) [Conxuntos avanzados](#)

Conxuntos avanzados

PROXECTO: --- Selecciona un proxecto --- --- Selecciona un proxecto ---

Centro de estudio: --- Centro de estudio --- --- Centro de estudio ---

CENTRO DE INTERESE: --- Centro de interese --- --- Centro de interese ---

VARIABLES:

TIPO DE FILTRADO: Palabras con maior frecuencia que sumen unha frecuencia acumulada do 80%

Calcular

Universidade de Santiago de Compostela

Os datos que así se obtéñen utilizámoslos en § 3.2.2, onde explicamos como se extrae a información sobre a compatibilidade léxica empregando distintos factores comparativos, que nos permiten contrastar dous “proxectos”. Por “proxecto” pode entenderse un proxecto de léxico dispoñible propiamente dito (como acabamos de indicar, o do galego ou o do español de Galicia), un grupo social establecido a través das variables de estratificación social e mesmo un centro de interese.

As buscas realizadas empregando como factor comparativo o proxecto propiamente dito revelan que a compatibilidade léxica entre ambas as dúas linguas oscila entre o 10,65% no CI 13 ‘Os traballos do campo e do xardín’ e o 22,63% no CI 02 ‘A roupa’, e que na maior parte dos CI a compatibilidade supera o 15%.

As buscas realizadas empregando como factor comparativo o propio centro de interese poñen de manifesto que o punto de intersección léxica entre dúas áreas temáticas pode ser bastante alto, como no caso dos CI 06 ‘Os obxectos colocados na mesa para a comida’ e 07 ‘A cociña e os seus utensilios’, nos que as palabras comúns chegan ao 20,10% do total, ou pode ser mínimo, como no caso dos CI 01 ‘O corpo humano’ e 08 ‘A escola: mobles e materiais’, onde estas apenas acadan o 1,45% do total de palabras.

Con todo, as máis interesantes son as buscas que utilizan como factor comparativo as variables de estratificación social. Estas foron, precisamente, as empregadas en § 3.2.2. A táboa 38 deste apartado mostra a información

que se obtén da aplicación cando se contrastan dúas poboacións, e vai seguido dun parágrafo que proporciona unha interpretación dos datos. A continuación presentamos outro exemplo, que ofrece os datos relativos ao CI 17 ‘As cores’, no que o “proxecto 1” identifica os informantes que aprenderon a falar en galego e só falan galego e o “proxecto 2” remite aos que aprenderon a falar en castelán e só falan castelán.

Cardinalidade do proxecto 1:	251	Cardinalidade do proxecto 2:	259
Cardinalidade da intersección:		135	
Cardinalidade da unión:		375	
Complemento do proxecto 1:	116	Complemento do proxecto 2:	124
Suma disxuntiva:		240	
Compatibilidade		36,00%	
Contribución na intersección:	53,78%	Contribución na intersección:	52,12%

Universidade de Santiago de Compostela

5 PRESENTACIÓN DE APÉNDICES

5.1 Listaxes de lemas e variantes por CI

Como xa indicamos, un dos aspectos máis destacados do LDEG e do LDG é que, fronte a outros léxicos dispoñibles, permite coñecer cales foron as respostas concretas ofrecidas polos informantes. Este feito é, se cabe, máis relevante no LDG, dada a gran variabilidade dos rexistros, de tal xeito que baixo un mesmo lema podemos atopar formas estándares galegas, variantes dialectais, palabras do castelán ou adaptacións destas ao galego, coas correspondentes desviacións gráficas en todos os casos. Consideramos que esta información pode ser de gran utilidade para os investigadores interesados no estudo do léxico galego, desde diferentes perspectivas (coñecemento do léxico estándar, adaptación de estranxeirismos, presenza do castelán...).

No Apéndice I ([Apéndice 1.pdf](#)) presentamos o conxunto de lemas de cada CI (en letra grosa), con indicación do número de variantes asociadas (en azul). Baixo o lema, e en letra normal, figuran as variantes concretas con cadanseu rexistro numérico, que indica a cantidade total de respostas para cada unha delas nas listaxes dos informantes. Presentamos un breve extracto do Apéndice, que recolle sete lemas do CI 02 ‘A roupa’, coas súas correspondentes variantes:

pantalóns de pinzas	1
pantalóns de pinzas.....	1
(pantalóns) piratas	67
pantalon pirata.....	2
pantalón pirata.....	12
pantalones piratas	3
pantalóns pirata.....	1
pantalons piratas.....	3
pantalóns piratas.....	4

pantalos pirata	1
pirata.....	12
piratas.....	28
pirátas.....	1
(pantalóns) vaqueiros.....	168
baqueiro	4
baqueiros	9
baqueros	1
baqueros	2
p vaqueiro	1
pantalón baqueiro.....	1
pantalon vaqueiro.....	2
pantalón vaqueiro.....	4
pantalon vaquero.....	1
pantalón vaquero.....	7
pantalón vaqueros	1
pantalóns vaqueiro.....	1
pantalóns vaqueiros.....	7
pantalóns vaqueros	1
vaqueiro	29
vaqueiros	64
vaquero.....	11
vaqueros.....	22
pantufila	5
pantufas	1
pantufilas.....	4
panty	51
panti.....	2
panties.....	1
pantis.....	41
panty.....	1
pantys.....	6
pañal.....	2
pañais..	1
pañal....	1
pañoleta.....	17
pañola..	2
pañoleta...	9
pañoleta	1
pañoleta	5

Ademais de mostrar información relevante per se (formas de adaptación do estranxeirismo *panty*, rexistro da variante *pañoleta*, convivencia de *vaqueiro* e *vaquero*, problemas ortográficos...), a presentación das variantes permite amosar como se realizou a lematización en todos e cada un dos casos e explica algunha das decisións tomadas: por exemplo, o emprego de parénteses e plural en *(pantalóns) vaqueiros* e *(pantalóns) piratas*, dada a presenza

nos dous exemplos de rexistros sen o núcleo *pantalón* e dada a maior cantidade de respostas en plural ca en singular. Remitimos a § 2.3.2 e § 2.3.3 para unha explicación detallada dos criterios empregados na selección dos lemas, que xustifican a variación que se pode atopar de CI a CI e mesmo dentro dun mesmo centro de interese.

5.2 Listaxe xeral de lemas ordenados por frecuencia

No Apéndice II ([Apéndice 2.pdf](#)) ofrecemos o total de lemas do LDG, ordenados por frecuencia, con indicación do CI ao que pertencen e do número total de rexistros asociados a cada un deles. Debe terse en conta que, como xa sinalamos en varias ocasións, a lematización é independente para cada CI e este tratamento, por suposto, mantense nos apéndices. Isto implica que un mesmo lema pode aparecer en varios lugares da listaxe porque se rexistra en máis dun CI. Podemos exemplificar este feito co lema *luz*, que, como se observa nos tres fragmentos reproducidos a seguir, figura, entre outros, no CI 09 (con 367 respostas, isto é, case o 46% dos informantes ofrecen esta palabra no dito CI), no 08 (con 73 respostas, 9% dos informantes) e no 10 (46 respostas, 5'75%).

09. A calefacción e a iluminación.....	estufa	368
09. A calefacción e a iluminación.....	luz	367
06. Os obxectos colocados na mesa.....	garfo	366
(...)		
14. Os animais.....	lebre.....	73
08. A escola: mobles e materiais.....	luz	73
15. Os xogos e as distraccións.....	natación.....	73
(...)		
11. O campo.....	labrego.....	46
10. A cidade.....	luz	46
05. Os alimentos e as bebidas	manteiga.....	46

Como pode observarse nos dous últimos fragmentos, o segundo criterio de ordenación, empregado cando o número de rexistros coincide, é o alfabético.

5.3 Listaxe xeral de lemas ordenados alfabeticamente

No Apéndice III ([Apéndice 3.pdf](#)) presentamos o total de lemas do LDG, ordenados alfabeticamente, con indicación do CI ao que pertencen así como da frecuencia absoluta de cada un deles. O que acabamos de comentar en relación coa independencia de cada CI é, por suposto, aplicable tamén aquí. Ademais, neste apéndice a presenza de lemas repetidos faise máis evidente,

habida conta de que o criterio alfabético determina que os lemas coincidentes presentes en varios CI aparezan en contigüidade. Véxase o seguinte fragmento, no que se inclúen os diferentes lemas *luz* que figuran en nove CI:

07. A cociña e os seus utensilios.....	luva de cociña.....	1
04. Os mobles da casa	luxoso	1
09. A calefacción e a iluminación.....	luz	367
08. A escola: mobles e materiais	luz	73
10. A cidade.....	luz	46
17. As cores.....	luz	12
12. Os medios de transporte	luz	10
03. As partes da casa (sen mobles)	luz	8
07. A cociña e os seus utensilios.....	luz	7
18. O mar.....	luz	5
11. O campo.....	luz	5
09. A calefacción e a iluminación.....	luz aforradora.....	1
09. A calefacción e a iluminación.....	luz amarela	2
09. A calefacción e a iluminación.....	luz artificial.....	5
09. A calefacción e a iluminación.....	luz azul	1
09. A calefacción e a iluminación.....	luz branca	6

5.4 Mostra da enquisa

No Apéndice IV ([Apéndice 4.pdf](#)) ofrecemos, a modo de exemplo, unha das enquisas realizadas a un dos oitocentos informantes que participaron na elaboración do LDG. Nas dúas primeiras páxinas recóllese a información sobre as diferentes variables de estratificación social consideradas (*vid.* § 2.2). Esta información foi proporcionada polos estudantes no momento da realización das enquisas. Na terceira páxina figuran as respostas fornecidas para o CI empregado de mostra, ‘Os deportes’, que non foi integrado no LDG. A seguir poden atoparse as respostas proporcionadas polo informante nos dezaoto CI analizados neste libro, ordenados de acordo coa numeración que lles outorgamos ao longo do traballo. Para máis información sobre a enquisa, consúltese § 2.1.

6 REFERENCIAS BIBLIOGRÁFICAS

ÁLVAREZ DE LA GRANJA, MARÍA (2011): “La presencia del gallego en el léxico disponible del español de Galicia. Análisis formal y funcional”, en BELÉN LÓPEZ MEIRAMA (ed.): *Estudios sobre disponibilidad léxica en el español de Galicia*. Santiago de Compostela: Universidade de Santiago de Compostela, pp. 17-102.

ÁLVAREZ DE LA GRANJA, MARÍA (no prelo): “O coñecemento do vocabulario normativo galego. Unha aproximación a través do léxico dispoñible”, en *Actas do Simposio “Un galego común fiel a si mesmo. Trinta anos de estandarización: 1982-2012”* (Santiago de Compostela: 11-20 de decembro de 2012). Santiago de Compostela: Instituto da Lingua Galega / Consello da Cultura Galega.

CARCEDO GONZÁLEZ, ALBERTO (1998): “Tradición y novedad en las aportaciones hispánicas a los estudios de disponibilidad léxica”, *ALFAL, Lingüística* 10, pp. 5-68.

DCG = REAL ACADEMIA GALEGA (2004): *Diccionario castelán-galego da Real Academia Galega*. [A Coruña]: Fundación Pedro Barrié de la Maza / Real Academia Galega.

DIGALEGO = *Diccionario de galego Digalego*. Vigo: Ir Indo, 2008.

<<http://www.digalego.com/>>

DRAE = Real Academia Española (2001²²): *Diccionario de la lengua española*. Madrid: Real Academia Española.

<<http://lema.rae.es/drae/>>

DRAG = REAL ACADEMIA GALEGA (1997): *Diccionario da Real Academia Galega*. A Coruña: Real Academia Galega.

DRAG² = REAL ACADEMIA GALEGA (2012²): *Diccionario da Real Academia Galega*. A Coruña: Real Academia Galega.

<<http://www.realacademiagalega.org/diccionario#inicio.do>>

- GALLOSO CAMACHO, M^a VICTORIA (2002): *El léxico de los estudiantes preuniversitarios en el distrito universitario de Salamanca*. Salamanca: Universidad de Salamanca.
- GALLOSO CAMACHO, M^a VICTORIA (2003): *El léxico disponible de Ávila, Salamanca y Zamora*. Burgos: Instituto Castellano y Leonés de la Lengua.
- GDXL = CARBALLEIRA ANLLO, XOSÉ M^a (coord.) (2009²): *Gran dicionario Xerais da lingua*. Vigo: Xerais.
- GÓMEZ MOLINA, JOSÉ RAMÓN / M^a BEGOÑA GÓMEZ DEVÍS (2004): *La disponibilidad léxica de los estudiantes preuniversitarios valencianos. Estudio de estratificación sociolingüística*. València: Universitat de València.
- GONZÁLEZ GONZÁLEZ, MANUEL / ANTÓN SANTAMARINA FERNÁNDEZ (2004): *Vocabulario ortográfico da lingua galega (Volga)*. Santiago de Compostela: Real Academia Galega / Instituto da Lingua Galega.
<<http://www.realacademiagalega.org/recursos-volga/>>
- GONZÁLEZ MARTÍNEZ, ADOLFO (2002): *La disponibilidad léxica de los alumnos preuniversitarios en la provincia de Cádiz*. Cádiz: Universidad de Cádiz.
- GOUGENHEIM, GEORGES / RENÉ MICHEÁ / PAUL RIVENC / AURÉLIEN SAUVAGEOT (1964): *L'élaboration du Français Fondamental (1^{er} degré). Étude sur l'établissement d'un vocabulaire et d'une grammaire de base*, ed. revisada e aumentada (1956). Paris: Didier.
- HERNÁNDEZ MUÑOZ, NATIVIDAD (2004): *El léxico disponible de los estudiantes conguenses*. Salamanca: Universidad de Salamanca.
- LABOV, WILLIAM (1966): *The Social Stratification of English in New York City*. Washington DC: Center of Applied Linguistics.
- LDEG = LÓPEZ MEIRAMA, BELÉN (2008).
- LÓPEZ CHÁVEZ, JUAN / CARLOS STRASSBURGER FRÍAS (1991): "Un modelo para el cálculo del índice de disponibilidad léxica individual", en HUMBERTO LÓPEZ MORALES (ed.): *La enseñanza del español como lengua materna*. Río Piedras: Ed. de la Universidad de Puerto Rico, pp. 91-112.
- LÓPEZ MEIRAMA, BELÉN (2008): *Léxico disponible en el español de Galicia*. Santiago de Compostela: Universidade de Santiago de Compostela.
- LÓPEZ MORALES, HUMBERTO (1973): *Disponibilidad léxica de los escolares de San Juan*. MS.
- LÓPEZ MORALES, HUMBERTO (1995): "Los estudios de disponibilidad léxica: pasado y presente", *Boletín de Filología de la Universidad de Chile (Homenaje a Rodolfo Oroz)* 35, pp. 245-259.
- LÓPEZ MORALES, HUMBERTO (1999): *Léxico disponible de Puerto Rico*. Madrid: Arco/Libros.

- MACKAY, WILLIAM F. (1971): *Le vocabulaire disponible du français*, 2 vols. Paris-Bruzelles-Montréal: Didier.
- MATEO GARCÍA, M.^a VICTORIA (1998): *Disponibilidad léxica en el COU almeriense. Estudio de estratificación social*. Almería: Universidad de Almería.
- MICHÉA, RENÉ (1953): “Mots fréquents et mots disponibles. Un aspect nouveau de la statistique du langage”, *Les langues modernes* 47, pp. 338-344.
- PAZÓ FERNÁNDEZ, EDITH (2005): *Diccionario galego de televisión*. Santiago de Compostela: Centro Ramón Piñeiro para a investigación en Humanidades.
- SAMPER HERNÁNDEZ, MARTA (2009): *Evolución de la disponibilidad léxica en estudiantes grancanarios de enseñanza primaria y secundaria*. Las Palmas de Gran Canaria: Cabildo de Gran Canaria.
- SAMPER PADILLA, JOSÉ ANTONIO (1998): “Criterios de edición del léxico disponible: sugerencias”, *Lingüística* 10, pp. 311-333.
- SAMPER PADILLA, JOSÉ ANTONIO / JUAN JOSÉ BELLÓN FERNÁNDEZ / MARTA SAMPER HERNÁNDEZ (2003): “El proyecto de estudio de la disponibilidad léxica en español”, en RAÚL ÁVILA *et al.*: *Pautas y pistas en el análisis del léxico hispano(americano)*. Frankfurt-Madrid: Vervuert-Iberoamericana, pp. 27-140.
- SAMPER PADILLA, JOSÉ ANTONIO / MARTA SAMPER HERNÁNDEZ (2006): “Aportaciones recientes de los estudios de disponibilidad léxica”, *LynX. Panorama de Estudios Lingüísticos* 5, pp. 5-95.
- SANTAMARINA FERNÁNDEZ, ANTÓN (ed.) (2003): *Diccionario de diccionarios* (versión 3). A Coruña: Fundación Pedro Barrié de la Maza.
<<http://sli.uvigo.es/ddd/index.html>>
- SARALEGUI, CARMEN / CRISTINA TABERNEIRO (2008): “Aportación al Proyecto pan-hispánico de léxico disponible: Navarra”, en Inés Olza Moreno / Manuel Casado Velarde / Ramón González Ruiz (eds.): *Actas del XXXVII Simposio Internacional de la Sociedad Española de Lingüística (SEL)*. Pamplona: Universidad de Navarra, 745-761.
- SEMINARIO DE LINGÜÍSTICA INFORMÁTICA (2006-2008): *Neoteca. Banco de Datos de Neologismos da Universidade de Vigo*. Vigo: Universidade de Vigo.
<<http://sli.uvigo.es/neoteca/>>

7 ACCESO A APÉNDICES

- Apéndice I Listaxes de lemas e variantes por CI
<[www.usc.es/libros/descargas/Léxico dispoñible. Apéndice 1.pdf](http://www.usc.es/libros/descargas/Léxico_dispoñible_Apéndice_1.pdf)>
- Apéndice II Listaxe xeral de lemas ordenados por frecuencia
<[www.usc.es/libros/descargas/Léxico dispoñible. Apéndice 2.pdf](http://www.usc.es/libros/descargas/Léxico_dispoñible_Apéndice_2.pdf)>
- Apéndice III Listaxe xeral de lemas ordenados alfabeticamente
<[www.usc.es/libros/descargas/Léxico dispoñible. Apéndice 3.pdf](http://www.usc.es/libros/descargas/Léxico_dispoñible_Apéndice_3.pdf)>
- Apéndice IV Mostra da enquisa
<[www.usc.es/libros/descargas/Léxico dispoñible. Apéndice 4.pdf](http://www.usc.es/libros/descargas/Léxico_dispoñible_Apéndice_4.pdf)>

