

## TECNOLOGIES LINGÜÍSTIQUES DE LA LLENGUA GALLEGA

Xavier Gómez Guinovart

Universitat de Vigo

### 1. INTRODUCCIÓ

Les tecnologies de la llengua (TL), juntament amb l'enginyeria lingüística, el processament del llenguatge natural i bona part de les indústries de la llengua, formen part de l'orientació més aplicada de la lingüística computacional. Tenen com a objectiu central desenvolupar aplicacions informàtiques d'ús general per facilitar la utilització, la traducció, l'estudi i l'aprenentatge de la llengua. Tanmateix, el desenvolupament d'aplicacions de les TL suposa una base sòlida prèvia de recursos i eines de TL. Entenem per *eines de TL* els sistemes informàtics orientats al desenvolupament d'aplicacions de TL, mentre que els *recursos de TL* constitueixen les dades lingüístiques sobre les quals es construeixen les aplicacions. Els recursos i la investigació lingüística constitueixen els pilars de les eines i de les aplicacions i, alhora, les eines i les aplicacions fan un gran servei quan es tracta d'ampliar o millorar els recursos lingüístics adquirits. Les eines, els recursos i les aplicacions són els components que integren els sistemes de TL, tot i que els recursos i les aplicacions de TL poden ser concebuts sovint com un conjunt indestruïble.

La direcció actual de la investigació i el desenvolupament en TL assenyalen dos objectius prioritaris: d'una banda, superar les barreres lingüístiques que, especialment en l'àmbit d'Internet, impedeixen un accés igualitari als continguts i —el que és realment important per a la indústria— dificulten la implantació efectiva d'un comerç electrònic realment global; d'una

altra banda, superar les «barreres arquitectòniques» d'accés a les noves tecnologies mitjançant la implementació d'interfícies d'interacció persona-ordinador en llenguatge natural oral. Les tecnologies clau que ens hauran de permetre l'ús de la llengua per accedir a totes les possibilitats d'informació, de comunicació i de consum posades a disposició dels habitants del primer món per les noves tecnologies són la identificació automàtica de la llengua, la traducció automàtica, el reconeixement i la síntesi de la parla, la recuperació interlingüística de la informació, els sistemes de diàleg i la web semàntica. En aquest sentit, qualsevol comunitat lingüística que vulgui accedir en igualtat de condicions a aquesta infraestructura social i econòmica en la qual es configura la societat de la informació, haurà de disposar d'aquestes tecnologies i estar representada en aquesta infraestructura. Per la gran incidència social, la presència d'una llengua en l'àmbit de les TL és determinant per aconseguir o conservar l'estat de llengua normalitzada.

La presència del gallec en les TL, i per mitjà de les TL a Internet, resulta també determinant per a la normalització lingüística: per un cantó, Internet és el mitjà de comunicació actual de més importància estratègica, i la presència del gallec en els mitjans de comunicació és imprescindible per normalitzar-lo; per un altre cantó, l'increment de la percepció social del prestigi i de la utilitat del gallec és fonamental per a la recuperació de la transmissió lingüística familiar (ara per ara, un dels esculls principals per a la supervivència de la llengua), i la presència del gallec a Internet

garanteix l'increment necessari d'aquesta percepció de prestigi i utilitat, i, finalment, el gallec precisa d'una identitat i una presència internacional reconeguda i respectada, i la presència del gallec a Internet i en les TL assegura la difusió idònia perquè es conegui i reconegui globalment. En darrer terme, la presència del gallec en les TL, i per mitjà de les TL en informàtica, és important per normalitzar-lo perquè pot contribuir a augmentar-ne l'ús, el prestigi i la utilitat, a més de constituir un factor simbòlic important en la percepció pròpia i aliena de la llengua.

## 2. PANORAMA DE LES TECNOLOGIES LINGÜÍSTIQUES DEL GALLEC

En el cas de la llengua gallega, els agents principals implicats en la producció de TL són els centres d'investigació pública, fonamentalment universitaris, i un conjunt reduït d'empreses de capital privat. Entre els primers, els grups més actius i amb més experiència en aquesta disciplina són el Grup de Tractament del Senyal de la Universitat de Vigo,<sup>1</sup> l'Institut de la Llengua Gallega de la Universitat de Santiago de Compostel·la,<sup>2</sup> el Laboratori de Bases de Dades de la Universitat de La Corunya<sup>3</sup> i el Seminari de Lingüística Informàtica de la Universitat de Vigo<sup>4</sup>; i, ja fora de l'àmbit universitari, el Centre d'Investigació Ramón Piñeiro,<sup>5</sup> dependent de la Xunta de Galícia. En el sector privat, les empreses que desenvolupen activitats més relacionades amb les TL són Imaxin Software,<sup>6</sup> Dimensiona<sup>7</sup> i Tagen Ata,<sup>8</sup> totes tres ubicades a Santiago de

Compostel·la. Fora de Galícia, altres empreses que han desenvolupat productes de TL per al gallec són la italiana Loquendo,<sup>9</sup> les catalanes Translendum,<sup>10</sup> AutomaticTrans<sup>11</sup> i Verbio<sup>12</sup> i, més recentment, l'estatunidenca Google<sup>13</sup>.

Presento a continuació un recull de les aplicacions i dels recursos de TL més rellevants per a la llengua gallega disponibles per mitjà d'Internet ara com ara (mitjan 2009), seguint una classificació genèrica, amb l'objectiu d'oferir, encara que sigui d'una manera molt succinta, una visió panoràmica de l'estat de la qüestió en aquest àmbit de treball de la lingüística aplicada del gallec.

### 2.1. *Corpus textuals*

#### 2.1.1. *Tesouro Informatizado da Lingua Galega (TILG)*

És el corpus textual de referència més important del gallec modern i contemporani, de lliure consulta en l'adreça web <http://www.ti.usc.es/TILG>.

Desenvolupat en l'Institut de la Llengua Gallega sota la direcció del professor Antón Santamarina, el TILG inclou gairebé la totalitat de les obres en gallec entre el 1612 i el 1980, i una àmplia representació de les obres publicades fins al 2009. Conté més de vint milions de paraules, de les quals més de dotze milions (totes les paraules lèxiques i part de les gramaticals) han estat lematitzades i etiquetades des del punt de vista morfosintàctic. A hores d'ara, s'està treballant en la lematització i l'anotació gramatical de prop de vuit milions de paraules del corpus pendents d'anàlisi en el

<sup>1</sup> <http://www.gts.tsc.uvigo.es/>

<sup>2</sup> <http://www.usc.es/ilgas/>

<sup>3</sup> <http://lbd.udc.es/>

<sup>4</sup> <http://sli.uvigo.es/>

<sup>5</sup> <http://www.cirp.es/>

<sup>6</sup> <http://www.imaxin.com/>

<sup>7</sup> <http://www.dimensiona.com/>

<sup>8</sup> <http://www.tagenata.com/>

<sup>9</sup> <http://www.loquendo.com/>

<sup>10</sup> <http://www.translendum.com/>

<sup>11</sup> <http://www.automatictrans.es/>

<sup>12</sup> <http://www.verbio.com/>

<sup>13</sup> <http://www.google.com/>

marc del projecte RILG, en curs d'integració de recursos del gallec.<sup>14</sup>

### 2.1.2. *Corpus de Referencia do Galego Actual* (CORGA)

Corpus de texts del gallec contemporani (des de 1975 fins a l'actualitat) d'uns 23 milions de paraules, desenvolupat en el Centre d'Investigació Ramón Piñeiro, sota la direcció del professor Guillermo Rojo, i accessible via Internet, prèvia alta en el sistema, en l'adreça <http://corpus.cirp.es/corga/> (López Martínez 2005).

### 2.1.3. *Corpus Lingüístico da Universidade de Vigo* (CLUVI)

Corpus de traduccions al gallec o del gallec, dels àmbits jurídic, informàtic, econòmic, literari, social i científic. Desenvolupat pel Seminari de Lingüística Informàtica de la Universitat de Vigo (Gómez Guinovart 2008) i de lliure consulta a <http://sli.uvigo.es/CLUVI/>. Aquest corpus paral·lel està alineat en l'àmbit de l'oració i supera, ara per ara, els 22 milions de paraules.

### 2.1.4. *Corpus Técnico do Galego* (CTG)

Corpus de registres especialitzats del gallec contemporani, amb texts publicats en els camps del dret, de la informàtica, de l'economia, de les ciències ambientals, de les ciències socials i de la medicina, que totalitzen més de tretze milions de paraules (Gómez Guinovart 2008). Desenvolupat pel Seminari de Lingüística Informàtica i l'Observatori de Neologia de la Universitat de Vigo, es pot consultar lliurement a <http://sli.uvigo.es/CTG/>.

### 2.1.5. *Corpus Técnico Anotado do Galego* (CTAG)

Versió etiquetada gramaticalment i lematitzada del CTG (Gómez Guinovart i López Fernández 2009). Tot i que l'anotació d'aquest corpus es troba encara en curs, ja es pot consultar en la web una secció del CTAG de més de dos milions de paraules, corresponents a texts de les ciències ambientals, en l'adreça web <http://sli.uvigo.es/CTAG/>.

### 2.1.6. *Tesouro Medieval Informatizado da Lingua Galega* (TMILG)

Corpus diacrònic del gallec, de més de nou milions de paraules, desenvolupat en l'Institut de la Llengua Gallega sota la direcció del professor Xavier Varela (Varela Barreiro 2004). Conté la totalitat de les obres no notariales publicades de la Galícia medieval (literàries, històriques, religioses, jurídiques i tècniques) i el 80 % de les obres notariales publicades. De lliure consulta, prèvia alta en el sistema, en la direcció <http://ilg.usc.es/tmilg/>.

### 2.1.7. *A Nosa Fala*

Corpus oral representatiu de totes les varietats diatòpiques del gallec, desenvolupat en l'Institut de la Llengua Gallega, sota la direcció del professor Francisco Fernández Rei i de la professora Carme Hermida, i disponible lliurement en l'adreça web <http://www.consellodacultura.org/arquivos/asg/anosafala.php/> (Fernández Rei i Hermida 2003).

## 2.2. *Lèxics i diccionaris*

### 2.2.1. *Dicionario de galego Ir Indo*

Versió electrònica del diccionari de gallec de l'editorial Ir Indo de Vigo (Ledo 2004). Conté més de 65.000 entrades amb definició,

<sup>14</sup> <http://sli.uvigo.es/RILG/>

etimologia, equivalències en portuguès, castellà, anglès i alemany, sinònims, antònims i variants, i conjugacions verbals, per la qual cosa pot ser considerat el diccionari normatiu en línia més complet del gallec. De lliure consulta a: <http://www.digalego.com/>.

### 2.2.2. *Vocabulario ortográfico da lingua galega (VOLG)*

Versió electrònica del vocabulari ortogràfic normatiu de referència per a la llengua gallega (Santamarina 2003b). De lliure consulta en l'adreça web <http://www.realacademiagallega.org/volga/>.

### 2.2.3. *Dicionario de Dicionarios*

Diccionari electrònic del gallec modern que recull les entrades de 25 obres lexicogràfiques dels segles XIX i XX. Conté 345.742 entrades de diccionari equivalents a 136.164 lemes diferents. Distribuït originàriament només en versió CD-ROM (Santamarina 2003a), des de fa poc temps també pot ser consultat lliurement a <http://sli.uvigo.es/DdD/>.

### 2.2.4. *Dicionário Estraviz*

Versió electrònica actualitzada del diccionari d'Estraviz (Estraviz 1995), que consta d'unes 90.000 entrades redactades en normativa reintegracionista. De lliure consulta a <http://www.estraviz.org/>.

### 2.2.5. *Dicionario da Real Academia Galega (DRAG)*

Versió electrònica del diccionari de referència de la RAG (García Pérez i González González 1997), que disposa d'unes 20.000 entrades no adaptades als canvis ortogràfics establerts en la reforma normativa de 2003 (Real Academia Galega i

Instituto da Lingua Galega, 2003). De lliure consulta en l'adreça: <http://www.edu.xunta.es/diccionarios/>.

### 2.2.6. *Termoteca*

Banc de coneixements terminològics desenvolupat pel Seminari de Lingüística Informàtica i l'Observatori de Neologia de la Universitat de Vigo a partir de les dades dels corpus CLUVI i CTG, amb informació sobre més de 10.000 termes en els àmbits de l'economia, de l'ecologia, de la sociologia, de la medicina i del dret (Gómez Guinovart 2008). De lliure consulta a: <http://sli.uvigo.es/termoteca/>.

### 2.2.7. *Neoteca*

Banc de neologismes del gallec desenvolupat pel Seminari de Lingüística Informàtica i l'Observatori de Neologia de la Universitat de Vigo, en constant actualització, amb més de 10.000 registres documentats en el Corpus Neo de premsa gallega (Gómez Clemente *et al.* 2005). De lliure consulta en l'adreça web <http://sli.uvigo.es/NEO/>.

### 2.2.8. *Dicionario de Dicionarios do Galego Medieval*

Diccionari electrònic del gallec medieval que recull les entrades de tretze obres lexicogràfiques del període medieval, amb un total de 53.564 lemes. Distribuït originàriament només en CD-ROM (González Seoane, Álvarez de la Granja i Boullón Agrelo 2006), es pot ara consultar lliurement també en l'adreça web <http://sli.uvigo.es/DDGM/>.

### 2.2.9. *Dicionario CLUVI inglés-galego*

Diccionari elaborat pel Seminari de Lingüística Informàtica de la Universitat de Vigo a partir de les dades del Corpus

CLUVI, amb 20.000 entrades, 30.000 traduccions i 60.000 exemples documentats en el corpus (Gómez Guinovart, Díaz Rodríguez i Álvarez Lugrís 2008). De lliure consulta en l'adreça: <http://sli.uvigo.es/diccionario/>.

#### 2.2.10. *Diccionario Aquén de Toponimia Galega*

Banc de dades de la toponímia oficial dels 315 ajuntaments, 3.794 parròquies i 37.297 llocs de Galícia. De lliure consulta en l'adreça <http://sli.uvigo.es/toponimia>.

### 2.3. *Eines de correcció lingüística*

#### 2.3.1. *Corrector Ortogal*

Corrector ortogràfic en línia del gallec desenvolupat pel Seminari de Lingüística Informàtica de la Universitat de Vigo. Permet corregir interactivament l'ortografia dels textos des de la pàgina web del corrector. Està basat en un diccionari normatiu del gallec de més de 25 milions de formes ortogràfiques, corresponents a uns 60.000 lemes, i ofereix solucions específiques per a la solució dels errors més comuns del gallec escrit, com els castellanismes, les ultracorreccions o els estrangerismes. D'accés lliure en l'adreça web <http://sli.uvigo.es/corrector/>.

#### 2.3.2. *Corrector 2.Mi3*

Última versió del corrector ortogràfic de gallec d'Imaxin Software per a Microsoft Office, basat en el recull lèxic del VOLG (Santamarina 2003b). La descàrrega del corrector és gratuïta des de la pàgina web de l'empresa, en l'adreça <http://www.imaxin.com>.

#### 2.3.3. *Correctors per a OpenOffice*

Correctors ortogràfic i gramatical per a OpenOffice.org desenvolupats per Imaxin

Software. Poden descarregar-se en l'adreça <http://www.mancomun.org>.

#### 2.3.4. *Exeria*

Corrector de llenguatge sexista en texts administratius per a OpenOffice.org, desenvolupat per Tagen Ata i Imaxin Software. Disponible en l'adreça web <http://www.exeria.net>.

### 2.4. *Traducció automàtica*

#### 2.4.1. *Traductor de la Universitat de Vigo*

Traductor automàtic espanyol-gallec i gallec-espanyol de codi obert basat en la tecnologia Apertium, desenvolupat pel Seminari de Lingüística Informàtica de la Universitat de Vigo (Alegría *et al.* 2006). D'ús lliure a <http://sli.uvigo.es/tradutor/>.

#### 2.4.2. *Traductor d'Imaxin Software*

L'empresa Imaxin Software de Santiago de Compostel·la dona accés als traductors bidireccionals *OpenTrad* espanyol-gallec i portuguès-gallec, basats en Apertium, en l'adreça web <http://opentrad.imaxin.com>.

#### 2.4.3. *Traductor de la Xunta*

La Xunta de Galícia ofereix un servei de traducció automàtica bidireccional d'espanyol-gallec i d'anglès-gallec desenvolupat per l'empresa catalana Translendum, que es pot trobar en l'adreça <http://www.xunta.es/tradutor/>.

#### 2.4.4. *Traductor de l'Institut Cervantes*

L'Institut Cervantes ofereix també el servei de traducció automàtica bidireccional d'espanyol-gallec, desenvolupat per l'empresa catalana AutomaticTrans, disponible en l'adreça <http://oesi.cervantes.es/traduccionAutomatica.html>.



#### 2.4.5. Traductor de Google

L'empresa californiana Google va incorporar el gallec al traductor automàtic estadístic el febrer de 2009, assolint la xifra de 41 idiomes. Es pot trobar de lliure accés a <http://translate.google.com> i en les diferents aplicacions d'aquest cercador que fan ús del traductor.

### 2.5. TECNOLOGIES DE LA PARLA

#### 2.5.1. Cotovia

Sistema de conversió text-veu desenvolupat pel Grup de Tecnologies del Senyal de la Universitat de Vigo (García Mateo 2003). Es pot accedir a una versió de demostració d'aquest convertidor en l'adreça web <http://www.gts.tsc.uvigo.es/cotovia/>.

#### 2.5.2. Loquendo TTS

El convertidor text-veu de l'empresa torinesa Loquendo inclou el gallec entre els més de vint idiomes del producte des de desembre de 2007. Se'n pot provar una versió de demostració en l'adreça <http://tts.loquendo.com/ttsdemo/>.

#### 2.5.3. Verbio

L'empresa catalana Verbio inclou el gallec en els productes de conversió text-veu i de reconeixement de la parla. Se'n pot provar una demostració en l'adreça [http://www.verbio.com/webverbio3/html/demos\\_ttsonline.php/](http://www.verbio.com/webverbio3/html/demos_ttsonline.php/).

### 2.6. ALTRES APLICACIONS

#### 2.6.1. FreeLing Galego

Demostració de les eines FreeLing (Atserias *et al.* 2006) per a l'anàlisi morfològica, desambiguació, lematització i anàlisi sintàctica de la llengua gallega desenvolupades pel Seminari de Lingüística

Informàtica de la Universitat de Vigo. D'accés lliure en l'adreça web <http://sli.uvigo.es/lingua>.

#### 2.6.2. é-galego

Curs multimèdia de gallec de nivell intermedi dirigit a parlants de castellà, elaborat per l'Institut de la Llengua Gallega de la Universitat de Santiago de Compostel·la (Rodríguez Malmierca i Gromaz Campos 2002). De seguiment lliure en l'adreça <http://e-galego.cesga.es>.

#### 2.6.3. Cartografía dos apelidos de Galicia

Inventari onomàstic georeferenciat dels cognoms de Galícia, desenvolupat a partir de les dades de l'Institut Nacional d'Estadística de 2001, sota la direcció de la professora Ana Boullón Agrelo i del professor Xulio Sousa Fernández, de l'Institut de la Llengua Gallega de la Universitat de Santiago de Compostel·la. De lliure consulta en l'adreça web <http://servergis.cesga.es/website/apelidos/>.

## 3. CONCLUSIONS

Tot i que l'estat actual de les tecnologies de la llengua gallega posa de manifest el treball immens que ja s'ha fet en aquest àmbit, hi ha algunes mancances importants que caldria resoldre a curt termini. Una de les mancances més importants es dona en el camp de la semàntica computacional, en el qual es constata la inexistència de components funcionals en tots els àmbits de desenvolupament. En l'àmbit bàsic dels recursos lingüístics no disposem encara ni d'una base de coneixements lexicosemàntics del gallec, ni de corpus anotats desambiguats semànticament i, paral·lelament, en l'àmbit de les eines, tampoc no disposem encara de programes informàtics per a la desambiguació semàntica de les paraules d'un text. En l'àmbit de les aplicacions, tant

la inexistència de sistemes complexos de diàleg basats en la comprensió automàtica dels enunciats com la inexistència d'aplicacions d'extracció d'informació en gallec que aprofitin el potencial del processament semàntic (com els sistemes de resposta a preguntes) estan motivades en gran mesura per aquesta absència de recursos i d'eines sobre les quals fonamentar l'anàlisi semàntica imprescindible perquè aquestes tecnologies importants funcionin.

Considero que l'objectiu de la normalització lingüística ha d'incloure dotar la societat del conjunt necessari de tecnologies de la llengua. Aquestes tecnologies constitueixen el requisit imprescindible perquè la comunitat lingüística gallega pugui accedir en peu d'igualtat i en la llengua pròpia a tots els béns i serveis de la societat de la informació. La responsabilitat principal en la promoció i gestió del desenvolupament d'aquestes tecnologies correspon a les institucions públiques gallegues, que han de fer un esforç important per situar les tecnologies de la llengua gallega al nivell de les llengües normalitzades del nostre contorn, disposant del compromís necessari de les entitats públiques i privades amb implicació social i procurant l'optimització de les inversions públiques. En aquest sentit, caldria garantir que l'obtenció dels productes estigués sempre vinculada a la formació de personal investigador propi i a la creació de teixit econòmic galleg, i caldria alhora assegurar la divulgació social dels resultats, fomentant que els productes obtinguts gràcies a les inversions públiques en investigació, desenvolupament i innovació fossin realment públics, reutilitzables, lliures, de codi obert i compartits sense restriccions innecessàries.

#### 4. REFERÈNCIES BIBLIOGRÀFIQUES

- ALEGRÍA, I.; ARANTZABAL, I.; FORCADA, M.; GÓMEZ GUINOVART, X.; PADRÓ, LL.; PICHEL, J. R.; WALIÑO, J. (2006). «Opentrad: Traducción automática de código abierto para las lenguas del estado español». *Procesamiento del Lenguaje Natural*, 37, 357-358.
- ATSERIAS, J.; CASAS, B.; COMELLES, E.; GONZÁLEZ, M.; PADRÓ, LL.; PADRÓ, M. (2006). «Freeling 1.3: Syntactic and semantic services in an open-source NLP library». *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC 2006)*, 48-55.
- ESTRAVIZ, I. A. (1995). *Dicionário da língua galega*. Santiago de Compostela: Sotelo Blanco.
- FERNÁNDEZ REI, F.; HERMIDA, C. (2003). *A nosa fala: bloques e áreas lingüísticas do galego*. Santiago de Compostela: Consello da Cultura Galega.
- GARCÍA MATEO, C. (2003). «Tecnologías del habla y lenguas minoritarias». *Procesamiento del lenguaje natural*, 31, 381-384.
- GARCÍA PÉREZ, C.; GONZÁLEZ GONZÁLEZ, M. (dir.). (1997). *Diccionario da Real Academia Galega*. A Coruña: Real Academia Galega.
- GÓMEZ CLEMENTE, X. M. et al. (2005). *Novas palabras galegas. Repertorio de creacions léxicas rexistradas na prensa e en Internet*. Vigo: Universidade de Vigo (Servizo de Publicacións).
- GÓMEZ GUINOVART, X. (2008). «A investigación en lexicografía e terminoloxía no Corpus Lingüístico da Universidade de Vigo (CLUVI) e no Corpus Técnico do Galego (CTG)». GONZÁLEZ SEOANE, E.; SANTAMARINA, A.; VARELA BARREIRO, X. (eds.). *A lexicografía galega moderna. Recursos e perspectivas*. Santiago de Compostela: Consello da Cultura Galega/Instituto da Lingua Galega, 209-228.
- GÓMEZ GUINOVART, X.; DÍAZ RODRÍGUEZ, E.; ÁLVAREZ LUGRÍS, A. (2008). «Aplicacións da lexicografía bilingüe baseada

en c rpora na elaboraci n do Dicionario CLUVI ingl s-galego». *Viceversa: Revista Galega de Traducci n*, 14, 71-87.

G MEZ GUINOVART, X.; L PEZ FERN NDEZ, S. (2009). «Anotaci n morfosint ctica do Corpus T cnico do Galego». *Linguam tica*, 1. <<http://www.linguamatica.com>>.

GONZ LEZ SEOANE, E.;  LVAREZ DE LA GRANJA, M.; BOULL N AGRELO, A. I. (eds.). (2006). *Dicionario de dicionarios do galego medieval*. Santiago de Compostela: Universidade de Santiago de Compostela (Verba, Anexo 57).

LEDO CABIDO, B. (ed.). (2004). *Dicionario de galego*. Vigo: Ir Indo.

L PEZ MART NEZ, M. S. (2005). «El Corpus de Referencia del Gallego Actual (Corga). Problemas de configuraci n y anotaci n». PUSCH, C. D.; KABATEK, J.; RAIBLE, W. (eds.). *Romanistische Korpuslinguistik II: Korpora und diachrone Sprachwissenschaft*. T bingen: Gunter Narr, 281-292.

REAL ACADEMIA GALEGA I INSTITUTO DA LINGUA GALEGA (2003). *Normas ortogr ficas e morfol xicas do idioma galego*. Santiago de Compostela: RAG/ILG.

RODR GUEZ MALMIERCA, M. J.; GROMAZ CAMPOS, M. (2002). « -galego: aprendiendo gallego en la red». *Agora Digital*, 3, <<http://www.uhu.es/agora/version01/digital/>>.

SANTAMARINA, A. (ed.). (2003a). *Diccionario de dicionarios*, 3a ed. La Coru a: Fundaci n Barri  de la Maza.

— (2003b). «O Vocabulario ortogr fico da lingua galega (VOLG): criterios de elaboraci n».  LVAREZ DE LA GRANJA, M.; GONZ LEZ SEOANE, E. (eds.). *A estandarizaci n do l xico*. Santiago de Compostela: Consello da Cultura Galega, 53-92.

VARELA BARREIRO, X. (2004). «Un proxecto do ILG no abalo da gram tica hist rica da lingua galega».  LVAREZ, R.; FERN NDEZ REI, F.; SANTAMARINA, A. (eds.). *A lingua*

*galega: historia e actualidade*. Santiago de Compostela: Instituto da Lingua Galega/Consello da Cultura Galega, 2, 649-684.