

XESTIÓN DA BASE DE DATOS DO PROXECTO TOPONIMIA DE GALICIA

Rut Vázquez González, Xosé M. Otero, Fernando García Pazos
S.A. para o Desenvolvemento Comarcal de Galicia. SITGA

A cantidade inxente de referencias toponímicas existentes no territorio recollidas mediante a realización de enquisas formarán un gran banco de datos toponímico. Toda esta información é sometida a unha serie de procesos na procura de conseguir un «corpus» toponímico único onde se recollen aspectos xeográficos e lingüísticos do topónimo como identificador do territorio. Podemos diferenciar dúas fases no traballo da xestión das bases de datos:

1ª Fase. Procesos de control na formación e integración de información na Base de Datos de Toponimia. Realízase un seguimento da aplicación da metodoloxía de traballo deseñada e un control dos datos introducidos nas bases polos técnicos de campo para garantir a homoxeneidade da información.

2ª Fase. Procesos de normalización lingüística. Isto supón sistematizar e axilizar mediante distintos pasos os traballos propios para normalizar unha cantidade enorme de nomes. É necesario fixar criterios sobre problemas ou formas toponímicas recorrentes.

1. Proceso de supervisión e control da calidade da información

A información toponímica que se recolle no campo, mediante a enquisa realizada a unha serie de individuos coñecedores do territorio, necesariamente ten que pasar unha serie de procesos que garantan a homoxeneidade da información e a coherencia cunhas liñas metodolóxicas previamente establecidas. Son distintas as variables que interveñen na recollida de datos, desde a propia persoa que realiza a entrevista ata a dispoñibilidade de bos informantes, pasando polas propias particularidades socioeconómicas do territorio modificadoras das vinculacións do home co seu medio.

Por todo isto é imprescindible realizar unha serie de procesos coa información

Rut Vázquez, Xosé M. Otero e Fernando García

procedente da recolleita de campo introducida nunha base datos. Podemos diferenciar os seguintes:

- Supervisión e control da integridade da base de datos en todos os seus campos.
- Análise do volume da información en función da densidade de topónimos recolleitos no territorio.
- Estudo da información segundo os parámetros que identifican a súa tipoloxía como elemento identificador dunha realidade territorial.
- Comprobación da localización xeográfica a partir da fotografía aérea e verificación da identificación dos elementos toponímicos tanto puntuais (pendedos, mullíños, pontes, igrexas) coma espaciais, nos que se delimita o seu ámbito de identificación.

Con este control da base de datos quérese evitar posibles erros de forma na introdución de datos, topónimos repetidos, campos de datos non introducidos, erros de grafías, etc. Tamén se ten en conta a comprobación da documentación cartográfica e bibliográfica existente que poida servir para verificar se a recolleita realizada é a máis axeitada tanto pola densidade de información coma pola súa localización.

Para a realización deste control da información introducida na base de datos foi elaborado un formulario de revisión que a continuación explicamos. Tamén facilita a elaboración de informes sobre a documentación cartográfica revisada.

Neste formulario podemos ver toda a información introducida na base de datos en diferentes formatos segundo a súa natureza:

Xestión da base de datos do Proxecto Toponimia de Galicia

a) **Datos estatísticos xerais:**

Situados no recadro superior e con distinta cor, proporcionánnos a seguinte información:

- O número total de fotos que temos introducidas
- O número total de informantes
- O número total de topónimos
- Observacións: amósanos todas as fotos introducidas e, entre corchetes, indícanos o primeiro e o último número de topónimo introducido, sinalando tamén os números que faltan por asignar, entre o primeiro e o último, na numeración correlativa de topónimos (no suposto de que polo medio faltara algún número).

b) **Informes temáticos:**

Situados na parte esquerda do formulario e con distinta cor. Cada botón mostra un informe sobre o que nos indica:

- Listado de fotos. Indícanos o número das fotos aéreas utilizadas no traballo co número de topónimos que conteñen.
- Listado de informantes. Permítenos visualizar cada unha das fichas correspondentes ós datos dos informantes.
- Listado de topónimos. Facíltanos as características de cada topónimo
- Listado de xenéricos: Pequeno glosario de termos léxicos utilizados nunha zona de traballo para describir determinadas realidades territoriais: agra, fraga, corga, etc.
- Listado de dúbidas lingüísticas.
- Listados de documentación. Resultado da verificación da información contida no IGN, catastro, parcelaria, cartografía municipal e outra documentación.

c) **Consultas:**

Permítenos facer determinadas consultas á base de datos para facilitarnos a manipulación ou modificación da información xa introducida. Podemos realizar buscas través dos seguintes campos:

- Número de Foto
- Numero Topónimo
- Cuadrícula Central
- Tipoloxía
- Topónimo

O resultado da consulta será a ficha do topónimo que se corresponda co criterio da nosa selección. Estas consultas pódense facer:

- Sobre unha parte do texto mediante a utilización de asteriscos.

Exemplo: Queremos coñecer tódolos nomes coa denominación *Breas* [*Brea*];

Rut Vázquez, Xosé M. Otero e Fernando García

Sabemos como comeza pero non como remata o topónimo
 [Barr*]
 Podemos preguntar polo final dunha palabra
 [*ido]

A partir desta análise da información realízanse as correccións necesarias para a súa posterior preparación para os procesos de normalización lingüística.

2. Proceso de normalización e aspectos lingüísticos da toponimia

A información toponímica recollida segundo unha metodoloxía homoxénea pasa á fase de normalización na que se fixa a forma escrita do topónimo. Correspóndelle á Comisión de Toponimia revisar e ditaminar cal debe ser a graffa dos topónimos de acordo coa forma da oralidade e seguindo as regras ortográficas e gramaticais do estándar.

Nesta fase do traballo concrétnanse os seguintes procesos:

- Sistematizar, axilizar o traballo e fixar as propostas normalizadas derivadas do ditame da Comisión de Toponimia.
- Supervisar incoherencias na normalización de topónimos repetidos.
- Establecer unha serie de orientacións e criterios de normalización de formas problemáticas ou formas toponímicas recorrentes.
- Introducir filtros que sirvan para depurar *a grosso modo* algunha información lingüística que posibilite a explotación para investigacións filolóxicas e futuras codificacións (semánticas, morfolóxicas).

a) A codificación lingüística

O elevado número de topónimos que se rexistra no noso territorio fai que a tarefa de propoñer unha ortografía estándar para cada topónimo -afnda que o índice de repetición das formas é elevado- sexa inabordable sen ningún tipo de control e sistematización. Aproximadamente, o 35% das formas recolleitas na enquisa oral presenta dúbidas ou problemas no momento de normalizalas. Para este inxente labor cómpre establecer unha serie de criterios de estandarización e aplicalos segundo unha sistemática controlada, ademais supón un profundo coñecemento da toponimia galega, a lingua e a súa historia.

A codificación lingüística xorde para cubrir esta necesidade e ten como obxectivos:

- Servir de apoio para axilizar a revisión e o ditame da Comisión de Toponimia
- Detectar os problemas recorrentes na toponimia segundo a análise da Comisión
- Fixar unha serie de criterios para normalizar a toponimia, segundo as propostas da Comisión
- Evitar incoherencias na normalización de formas parellas.

Xestión da base de datos do Proxecto Toponimia de Galicia

A cada un dos topónimos asígnaselle un código numérico que é indicativo de:

- estado de revisión do topónimo,
- tipo de problema rexistrado no momento de normalizar o topónimo.

Deste xeito, establécese para os topónimos unha dobre marca: un código co estado de revisión que indica o grao de normalización do topónimo despois da revisión da Comisión de Toponimia (Táboa 1) e un código co que se identifican os problemas detectados na revisión do topónimo no momento de normalizalo (Táboa 2).

Táboa 1

| ESTADO DE REVISIÓN | CODIFICACIÓN | EXEMPLO |
|---|--------------|------------|
| Topónimo cunha proposta normalizada | 1 | A Casilla |
| Topónimo \pm opaco pendente de discusión e de forma normalizada: casos de limbo | 2 | O Concixil |
| Topónimo pendente de verificar e confirmar co técnico de campo ou en fontes documentais | 3 | Bordecedo |

Esta tipoloxía empregada para a codificación dos problemas na toponimia (Táboa 2) non se elaborou cunha finalidade de clasificación científica senón práctica, para axilizar e sistematizar a súa revisión, na medida do posible. A análise da Comisión sobre as dúbidas dos técnicos de campo e outros problemas atopados na toponimia revisada utilizáronse como base para establecela.

Táboa 2

| TIPO | PROBLEMAS LINGÜÍSTICOS | | |
|-----------|--|--------|--|
| | DESCRIZIÓN | CÓDIGO | EXEMPLO |
| GRÁFICOS | Problemas de graffas no topónimo | 20 | O Baltrigo de Abaixo (O Val Trigo de Abaixo) |
| | Topónimos con B/V | 21 | As Avesadas; O Bordecedo |
| | Topónimos con H | 22 | As Horticas |
| FONÉTICOS | Problemas fonéticos | 30 | Nameló (variante de Lamelo) |
| | <i>Vogais:</i> a) vacilacións nas tónicas | 31 | O Curro (variante de Corro) |
| | b) vacilacións nas átonas | 32 | A Devesa (fonética: A Dovesa) |

Rut Vázquez, Xosé M. Otero e Fernando García

| PROBLEMAS LINGÜÍSTICOS | | | |
|------------------------|---|--------|---|
| TIPO | FENÓMENO | CODIGO | EXEMPLO |
| FONÉTICOS | <i>Ditongos e hiatos:</i> a) variantes nos ditongos b) vacilacións de hiatos ↔ ditongos | 33 | Meancifios (fonética: Miansifios) |
| | Metátese | 34 | A Bertonia (variante de A Bretoña) |
| | Consoantes: | 35 | O Salleiro (variante de O Xalleiro) |
| | <i>Gheada</i> | 36 | Guelura (fonética: Ghelura) |
| | <i>Seseo</i> | 37 | As Calzadiñas (fonética: As Calsadiñas) |
| | <i>Rotacismo</i> | 38 | Os Dices (fonética: ORdices) |
| MORFOSINTÁCTICOS | <i>Problemas morfosintácticos:</i> | 50 | |
| | Artigo | 51 | (A) Canella do Porto ² As Gorbias (fonética: A Esgorbía) |
| | Variante morfolóxicas | 52 | Os Nabais (forma oral: Os Nabales) |
| LÉXICOS | <i>Problemas léxicos</i> | 60 | |
| | Topónimos compostos xuntos/separados | 61 | Entrasaugas (/Entre as augas) Lama Pereira |
| | Castelanismos | 62 | Igrexa de Santa Eufemia (forma oral: Iglesia de Santa Eufemia) |

b) Procesos para establecer a forma normalizada dos topónimos

Os topónimos, antes de ter unha forma gráfica definitiva sofren varios procesos de revisión. Por isto resulta necesario establecer unha dinámica de traballo en fases ben definidas que garantan o control e o rigor no momento de fixar a forma normalizada do topónimo. As fases son as seguintes:

- 1ª fase: Comparación dos topónimos pendentes de proposta normalizada cun corpus de toponimia normalizada³. Deste xeito, as formas que se repiten quedan automaticamente codificadas e evítase volver sobre o revisado.
- 2ª fase: Revisión individual de cada membro da Comisión de Toponimia.
- 3ª fase: Informe global do gabinete que selecciona a toponimia codificada e comentada persoalmente polos membros da Comisión e integra tódolos problemas analizados coa súa correspondente codificación.

Xestión da base de datos do Proxecto Toponimia de Galicia

- 4ª fase: Reunión da Comisión para debater e fixar propostas normalizadas dos topónimos comentados. Finalmente, os topónimos terán unha forma máis ou menos normalizada, segundo os casos (véxase Táboa 1).
- 5ª fase: Nova codificación da toponimia segundo o comentado na reunión. O resto dos topónimos non comentados pola Comisión actualízanse como normalizados.
- 6ª fase: A toponimia codificada de novo segundo o ditame da Comisión pasa a engrosar o corpus de toponimia normalizada.
- 7ª fase: Informe final elaborado sobre cada bloque de toponimia normalizada pola Comisión, que recolle as incidencias e conclusións derivadas de todo o proceso, sobre todo, o referente ó debate e á fixación das formas normalizadas.

3. Aplicación informática para revisión lingüística

As fases do proceso de normalización e a codificación da toponimia realízanse sobre unha base de datos específica que recolle a toponimia sobre a que xa se realizou a revisión xeral e a homoxeneización antes descritas. Nesta base de datos imos atopar os campos básicos na identificación e caracterización dos topónimos (campos comúns á base xeral) e ademais, unha serie de campos específicos onde van: a codificación, os comentarios técnicos sobre o topónimo e a súa forma gráfica normalizada (Táboa 3).

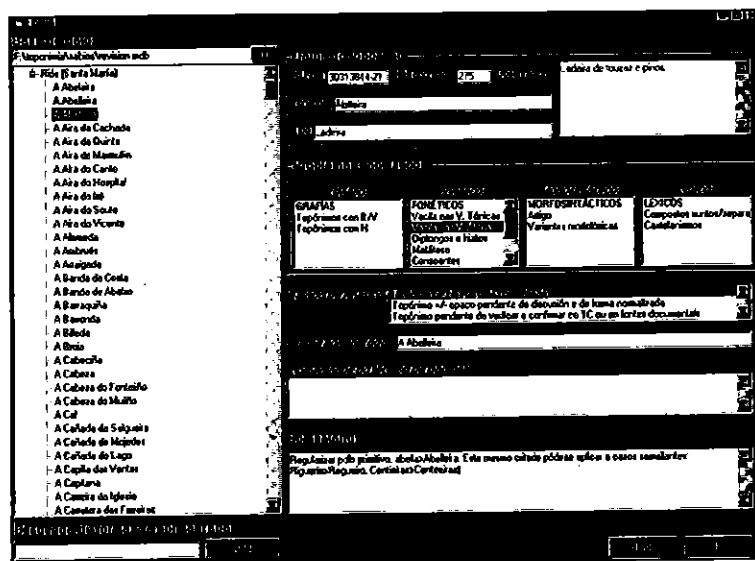
Táboa 3

| B.D. CODIFICACIÓN LINGÜÍSTICA | |
|--------------------------------------|-------------------------------------|
| <i>CAMPOS B.D.</i> | |
| <i>COMÚNS</i> | <i>ESPECÍFICOS</i> |
| Foto_nº | G20 ⁴ |
| Nº topónimo | F30 ⁵ |
| Topónimo | M50 ⁶ |
| Fonética | L60 ⁷ |
| Observacións xerais | Estado de Revisión (ER) |
| Tipo xeográfico | Forma Normalizada (FN) |
| Concello | Comentarios_Comisión Toponimia (CT) |
| Parroquia | Comentarios_Técnicos Gabinete (TG) |
| Aldea | |

Rut Vázquez, Xosé M. Otero e Fernando García

Ademais, á Comisión de Toponimia facilitóuselle un formulario específico para a súa revisión na aplicación de bases de datos. Deseguido, ofrécese unha mostra do formulario, cos campos comúns á base de datos xeral e os campos específicos. Tamén dispón do corpus de toponimia normalizada (*).

Formulario para a revisión lingüística da Comisión de Toponimia



Deste xeito trátase de facilitar o traballo da Comisión de Toponimia e tamén se procuran establecer e seguir uns criterios específicos para os traballos de normalización lingüística. Son milleiros de topónimos os que terán que ser analizados e estudados polo que é necesario realizar un enorme esforzo no asentamento dunha codificación lingüística para a formación dunha base de datos de toponimia con rigor científico.

Xestión da base de datos do Proxecto Toponimia de Galicia

NOTAS

1. As formas que aparecen en primeiro lugar son as normalizadas pola Comisión de Toponimia.
2. Presenza /ausencia do artigo.
3. Formas normalizadas e xa revisadas pola Comisión de Toponimia coa súa correspondente codificación, que se van ampliando a medida que avanza os traballos de revisión.
4. Problemas gráficos, levan códigos correspondentes á decena 20.
5. Problemas fonéticos, levan códigos correspondentes á decena 30.
6. Problemas morfolóxicos, levan códigos correspondentes á decena 50.
7. Problemas léxicos, levan códigos correspondentes á decena 60.
8. Nos dous casos, os códigos numéricos correspondentes ós tipos de problemáticas e ó estado de revisión, respectivamente, aparecen desenvoltos en letra.