

# A avaliación da neoloxidade en terminoloxía

Susana López Fernández, Xavier Gómez Guinovart,  
Xosé María Gómez Clemente, Ana Belén Crespo Bastos  
(Grupo TALG, Universidade de Vigo)

## 1. Introducción

Antes de describir a nosa investigación<sup>1</sup>, cómpre analizarmos unha cuestión central no tratamento das novas unidades léxicas que rexistramos tanto na lingua xeral coma na especializada. Esta cuestión é a de establecer un concepto de neoloxismo que permita determinar un índice cuantitativo de neoloxidade das unidades léxicas consideradas.

As definicións de neoloxismo poden ser variadas pero todas están centradas no feito de unha unidade ser considerada como “nova” nun sistema lingüístico determinado. Cabré (1993: 444) afirma que “o neoloxismo se pode definir como unha unidade léxica de formación recente, unha acepción nova dun termo xa existente ou, finalmente, un termo emprestado hai pouco dun sistema lingüístico estranxeiro”); Rey (1976: 4) defíneo como “une unité nouvelle, de nature lexicale, dans un code linguistique défini” e Boulanger (1989: 202) entende que “[o neoloxismo é o] resultado tanxible da operación de produción lingüística inédita [...], é dicir, a unidade nova capaz de encher cada baleiro detectado introducíndose no uso corrente ou socioprofesional”. A novidade á que nos referiamos establécese basicamente con respecto a un criterio temporal (unha unidade é nova se apareceu recentemente nunha lingua), aínda que tamén pode referirse á ausencia desa palabra nun grupo de dicionarios ou á opinión sobre esa novidade que teñen os falantes. Sablayrolles (2002) afirma que, xa que a neoloxía non é un concepto discreto e que implica distintos graos, o concepto de neoloxismo está suxeito a unha variabilidade de xuízos e tamén explica a presenza de unidades que non serían espontaneamente e unanimemente consideradas como neoloxismos. Esta consideración vai ter reflexo directo nos resultados da investigación que presentamos, xa que, conscientes da variabilidade e da gradualidade do concepto, pretendemos nos achegar á súa sistematización mediante a ponderación dos parámetros cuantificables que interveñen na condición neolóxica.

Existen tres criterios tradicionalmente empregados na identificación de neoloxismos que son abondo coñecidos por seren os utilizados nas principais investigacións que se están a desenvolver:

(a) Criterio lexicográfico. Está baseado na presenza ou ausencia dunha unidade léxica nunha serie de dicionarios: as unidades que non se rexistran neles son consideradas como “neoloxismos lexicográficos”. Este criterio, que á vez define o carácter do concepto de “neoloxismo”, presenta certos problemas. Xa Corbeil (1971: 135-136) afirmaba que un dicionario nunca pode recoller todas as palabras dunha lingua e polo tanto o carácter neolóxico dunha unidade establecido con este criterio pode ser posto en dúbida. Ademais, os lexicógrafos, cando elaboran as súas obras, réxense por unha serie de normas que son máis ou menos permeables ás novas creacións léxicas, dependendo do tipo de dicionario que se elabore, da ideoloxía, da tradición, etc. Malia constataremos estes feitos, o criterio de exclusión lexicográfica ten unha vantaxe que é a súa facilidade

---

<sup>1</sup> Este traballo foi financiado polo Ministerio de Educación y Ciencia e o Fondo Europeo de Desenvolvemento Rexional (FEDER), dentro do proxecto “Deseño e implementación dun servidor de recursos integrados para o desenvolvemento de tecnoloxías da lingua galega (RILG)” do Plan Nacional de I+D+I, 2006-2009 (ref. HUM2006-11125-C02-01/FILO), proxecto coordinado da Universidade de Vigo (Grupo TALG) coa Universidade de Santiago de Compostela (Instituto da Lingua Galega).

de aplicación. Porén, a fiabilidade non é abonda e, deste feito, xorde o interese en conxugar este criterio con outros para poñer unhas estremas, o máis claras posibles, entre o que se pode considerar neoloxismo e o que non.

(b) Criterio temporal ou diacrónico. Baséase na constatación de que unha unidade léxica leva pouco tempo instalada nunha determinada lingua (velaquí a cerna do concepto de “neoloxismo”). Compróbase coas datas de aparición dunha palabra nun dicionario ou nun corpus textual para a atribución dunha teórica primeira ocorrencia.

(c) Criterio psicolóxico-pragmático. Rey (1976), cando fala do neoloxismo, di que é imposible consideralo *in abstracto* como un elemento novo dentro dun sistema, independente do funcionamento concreto da lingua. Os neoloxismos defínense en relación cun modelo de comunicación determinado e, polo tanto, unha forma funcional limitada na súa función a un subsistema (dialecto, sociolecto) pode pasar a outro subsistema, no que será percibido como neoloxismo. É dicir, un neoloxismo tamén pode ser aquela unidade que ben por moda, por unha intervención institucional ou por outras causas pasa dun grupo de falantes a outro, ou dun rexistro de lingua a outro, independentemente da súa novidade obxectiva. O criterio pragmático é interesante sobre todo para a neoloxía terminolóxica en contextos de linguas minorizadas onde hai políticas planificadas para establecer o corpus. Nestas situacións (e a lingua galega está inserida nunha delas), a recuperación de arcaísmos ou dialectalismos, coma tal, constitúen verdadeiros actos de creación neolóxica, que teñen que ver con recuperacións denominativas ou coa neoloxía semántica. No corpus que nós manexamos non detectamos ningún procedemento de anovamento léxico que poida ser considerado á maneira de Rey, feito que non deixa de ser desacougante no sentido de que indica que non se utiliza o fondo léxico da lingua. O lexicógrafo francés opina, ademais, que basear a neoloxía en criterios cronolóxicos ou lexicográficos presupón unha visión parcial dos fenómenos neolóxicos; daquela, fala do concepto de “impresión neolóxica”, ou percepción da neoloxía por parte dos falantes, que debe estar baseada nun xuízo colectivo que supere os diferentes niveis de competencia que teñen os falantes. Este concepto de impresión neolóxica é difícil de cuantificar, e as investigacións feitas nesta liña son poucas e insatisfactorias (Sablayrolles, 2002).

Velaquí os criterios tradicionais para a definición e para a identificación de neoloxismos. Dos tres, o último é case que imposible utilizalo; o primeiro adquiriu unha relevancia extraordinaria nos métodos de traballo dos Observatorios de Neoloxía, por razóns obvias, e o segundo só se pode aplicar, con xusteza, se se posúen corpus textuais o suficiente grandes para incluíren textos dun amplo leque temporal.

Daquela, cómpre utilizarmos outros parámetros que encaixen mellor cunha nova interpretación do concepto de neoloxismo. Faura (2007), cando analiza o criterio lexicográfico para a identificación dos neoloxismos, afirma que o concepto de neoloxismo ten que ir alén da utilización dun criterio lexicográfico. Propón, como Cabré et al. (2004), a fixación de criterios obxectivos e mesurables para medir a estabilidade dunha palabra nova. Entre eses criterios están o número de ocorrencias, a regularidade das ocorrencias, a previsibilidade da nova forma, a formación segundo as regras da gramática e a estabilidade do concepto que designan. É interesante a súa proposta de definir un novo status para as palabras que, sen estar rexistradas nos dicionarios, teñen unha presenza abonda no uso. A seguir expoñemos os criterios utilizados no noso traballo para a avaliación da neoloxidade, exemplificando a súa aplicación á caracterización dos termos extraídos dun corpus de especialidade.

## 2. Avaliación da neoloxidade

Na elaboración dun método cuantitativo para a avaliación do índice de neoloxidade (v) dunha unidade léxica, tivemos en conta unicamente factores cuantitativos, co fin de que o seu valor se puidese derivar automaticamente dos datos empíricos obtidos dos corpóra consultados para a determinación da primeira documentación, da frecuencia de uso e da dispersión textual das unidades léxicas consideradas. Os factores de neoloxidade utilizados na avaliación da súa neoloxidade foron a exclusión lexicográfica da unidade léxica, a súa primeira documentación ( $\alpha$ ), a súa frecuencia relativa ( $\varphi$ ), a súa dispersión relativa ( $\delta$ ) e a exclusión lexicográfica dos seus compoñentes ( $\kappa$ ). As unidades léxicas consideradas foron os 1.444 termos de ecoloxía e ciencias ambientais repertoriados na *Termoteca* (<http://sli.uvigo.es/termoteca>), extraídos dun corpus de especialidade de 2.349.362 palabras (953 textos) que forma parte do *Corpus Técnico do Galego* (<http://sli.uvigo.es/CTG>) (Crespo et al., 2008; Gómez Guinovart, 2007).

Mediante o criterio da exclusión lexicográfica da unidade léxica, identificamos un total de 777 neónimos lexicográficos (o 53,8% dos 1.444 termos) non repertoriados en ningunha destas tres obras de referencia: o *Léxico do medio* (2002) do galego, publicado na rede polo proxecto da Unión Latina e da Unión Europea "Linmiter" (<http://www.linmiter.net>), o *Gran Diccionario Século XXI da Lingua Galega* (2005) e o *Gran Diccionario Xerais da Lingua* (2000). A primeira das obras citadas, o *Léxico do medio*, é a única obra en galego que conta cun léxico específico dedicado ás ciencias ambientais. O feito de conter un reducido número de entradas (235) motivou a necesidade de completar a nosa peneira coa consulta de dous dicionarios de lingua xeral. Escollemos o *Gran Diccionario Século XXI*, por ser unha obra de recente publicación e fonte de referencia na incorporación de novas palabras do galego. Alén diso, utilizamos tamén o *Gran Diccionario Xerais da Lingua* porque é a obra lexicográfica cun maior número de entradas e subentradas na nosa lingua (95.000) e porque ofrece un tratamento da terminoloxía máis completo. A utilización do criterio lexicográfico como primeira peneira posúe a vantaxe da súa facilidade de aplicación e presenta o problema de deixar fóra da análise un conxunto de unidades que se considerarían neolóxicas a teor da súa datación, frecuencia, dispersión e composición, mais que, por razóns indeterminadas, forman parte do corpus de exclusión.

Dos 1.444 termos analizados, 667 aparecían no corpus de exclusión como entradas ou subentradas e foron desbotados deste estudo (malia poderen ser analizados no que atinxe á súa neoloxidade). Algúns exemplos deste grupo son *biotecnoloxía*, *xestión ambiental*, *tecnoloxía limpa*, *risco ambiental*, *residuo doméstico* ou *reciclaXe*. Os 777 neónimos lexicográficos restantes son (a) termos pluriléxicos que non aparecen como sintagmas no corpus de exclusión, aínda que si aparecen todas as palabras formantes, como *aforro enerxético*, *punto verde*, *programa de educación medioambiental* ou *recursos naturais renovables*; (b) termos pluriléxicos que non aparecen como sintagmas no corpus de exclusión, aínda que si aparecen algúns dos seus formantes, como *sistema de ecoetiquetado*, *combustible de recombustión*, *planta de coxeneración*, *sostibilidade do sistema* ou *central minihidráulica*; ou (c) termos monoléxicos que non aparecen no corpus de exclusión, como *bioacumulación*, *agroecosistema*, *bioseguridade*, *ecotaxa*, *microhábitat*, *hidrobiólogo*, *eucaliptización*, *ambientalización* ou *agroecoloxía*. Para cada neónimo lexicográfico así identificado, tomamos nota da súa primeira documentación ( $\alpha$ ), da súa frecuencia relativa ( $\varphi$ ), da súa dispersión relativa ( $\delta$ ) e da exclusión lexicográfica relativa dos seus compoñentes ( $\kappa$ ), combinando todos estes factores nun índice de neoloxidade (v), obtido para cada neónimo lexicográfico mediante a aplicación da seguinte fórmula:

$$v = \frac{(((\alpha - Li(\alpha)) / (Ls(\alpha) - Li(\alpha))) \times 50) + (((Ls(\varphi) - \varphi) / Ls(\varphi)) \times 20) + (((Ls(\delta) - \delta) / Ls(\delta)) \times 30) - 50}{50} + \kappa$$

Tomamos como primeira documentación ( $\alpha$ ) dun neónimo lexicográfico o ano no que se documenta por primeira vez o neoloxismo nos corpóra de referencia consultados. De modo xeral, canto máis recente sexa un neoloxismo, maior debe ser o seu índice de neoloxicidade. A fórmula elaborada para o cálculo do índice  $v$  da un valor entre 0 e 1 para os termos considerados neolóxicos, e un valor 0 ou inferior para os termos considerados non neolóxicos. Establecemos un límite inferior para a variable  $\alpha$  –isto é,  $Li(\alpha)$ – no ano 1975, e un límite superior  $-Ls(\alpha)$ – no ano 2007. Os valores de  $\alpha$  inferiores a 1975 darán como resultado un índice de neoloxicidade negativo. Os corpóra de referencia consultados son o *Corpus Técnico do Galego (CTG)*, o *Tesouro Informatizado da Lingua Galega (TILG)* e o *Corpus de Referencia do Galego Actual (CORGA)*. O *CTG* é un corpus técnico do galego contemporáneo composto de textos monolingües especializados nos eidos do dereito, da informática, da economía, das ciencias ambientais, da socioloxía e da medicina, cunha extensión actual de máis de 12 millóns de palabras. O *TILG* e o *CORGA* son corpóra xerais de lingua galega duns 20 millóns de palabras. Establecer a data da primeira ocorrencia supón un problema porque os corpóra textuais, as únicas ferramentas que temos para determinala, non recollen exhaustivamente todos os posibles textos onde o neoloxismo puido aparecer, e aínda menos tratándose de neónimos. O *CTG* recolle textos desde 1990, o *TILG* desde 1612 e o *CORGA* desde 1975 pero, nestes últimos, a presenza de textos de especialidade é limitada por seren corpóra de referencia. Ora ben, se consignamos esta primeira data, malia ser teoricamente aproximada, temos un dato obxectivo ao que nós, nesta investigación, lle damos un alto valor. Este criterio temporal, leva directamente a outras cuestións que teñen relevancia: unha palabra rexistrada nun corpus textual, non nos dicionarios, en época temperá, durante canto tempo mantén a súa condición neolóxica? Guilbert cría que a presenza na lingua dunha unidade léxica durante cinco anos evitaba a súa consideración de neoloxismos. Quizais esta consideración poida ser tida en conta para linguas de grande difusión, ou se cadra non, pero non é aplicable a unha lingua como o galego. Hai que establecer unha fronteira, sempre subxectiva, e nós fixámola en 1975. Lémbrese que estamos a falar de neoloxismos lexicográficos que levan rexistrados na lingua máis dun cuarto de século. A reflexión lévanos directamente ás prácticas lexicográficas e, de certo, á ben coñecida lentitude con que se aceptan as novas palabras nos dicionarios normativos e mesmo de uso. Pero este feito é fariña doutra muiñada. Se a isto engadimos a falta de dicionarios especializados, sistemáticos, en determinados dominios, metémonos nun sarillo. Porén, o que é primordial é o problema que isto presenta para a utilización única do criterio de exclusión lexicográfica.

Canto á frecuencia de uso do neónimo, utilizamos para o cálculo o dato da súa frecuencia relativa ( $\varphi$ ), determinada polo número de veces en que se documenta o neoloxismo no corpus de referencia consultado, en relación co tamaño do corpus en número de palabras. Cantas máis veces se documente o neoloxismo no corpus, menor debe ser o seu índice de neoloxicidade. Neste caso, o corpus de referencia consultado é o *Corpus Técnico do Galego*. Calculamos a frecuencia relativa ( $\varphi$ ) a partir da frecuencia absoluta ( $\Phi$ ) ou número de veces en que se documenta o neoloxismo no *CTG*, multiplicando por 1.000.000 o resultado de dividir  $\Phi$  polo número de palabras ( $\pi$ ) do *CTG*, isto é,  $\varphi = (\Phi/\pi) * 1000000$ . Xa en Cabré et al. (2004), reflexionando sobre os filtros que se deben aplicar aos neoloxismos lexicográficos, fálase de que a frecuencia é un criterio que se debe empregar para a consideración da calidade neolóxica dunha

unidade léxica. A baixa frecuencia relativa dunha unidade contribúe en certa maneira a aumentar o seu carácter neoloxico porque a pouca difusión pode ser un indicador de que non tivo unha aceptación entre os falantes ou que non foi adecuadamente difundida. Con todo, esta afirmación debe ser matizada para os neónimos extraídos dun corpus especializado, xa que a baixa frecuencia pode estar motivada pola estrutura interna do corpus e ligada a unha escasa presenza dos textos dunha determinada póla da árbore conceptual asociada ao dominio.

O terceiro factor considerado no cálculo do índice de neoloxicidade ( $v$ ) é a dispersión relativa ( $\delta$ ) do neónimo, isto é, o número de textos distintos nos que se documenta o neoloxismo no corpus de referencia consultado, en relación co tamaño do corpus calculado en número de textos. Cantos maior sexa o número de textos que conteñan a unidade léxica, menor debe ser o seu índice  $v$ . De novo, o corpus de referencia consultado é o *Corpus Técnico do Galego*. Calculamos a dispersión relativa ( $\delta$ ) a partir da dispersión absoluta ( $\Delta$ ) ou número de textos distintos nos que se documenta o neoloxismo no corpus de referencia, multiplicando por 1.000 o resultado de dividir  $\Delta$  polo número de textos ( $T$ ) do corpus de referencia, isto é,  $\delta = (\Delta/T)*1000$ . Evidentemente, entendemos que un neónimo utilizado por varios especialistas en distintos textos ten unha calidade neoloxica menor, xa que o uso implica difusión, e o coñecemento dunha unidade léxica por determinados usuarios, especializados ou non, reduce a percepción da súa neoloxicidade. Ao contrario, un neónimo utilizado por poucos autores en poucos textos terá un maior carácter neoloxico, e aínda máis no caso dos *hápx*, de grande interese en terminoloxía por implicaren unha necesidade denominativa específica que responde a un concepto. Os métodos lexicométricos utilizados por Quirion (2003) para a análise da terminoloxía poden ser de grande axuda para determinar o carácter neoloxico dunha palabra. Entendemos que o criterio de dispersión ten máis importancia que o criterio da frecuencia porque as ocorrencias dunha unidade léxica poden corresponder a un número moi limitado de textos, ou mesmo a un único texto.

En relación con estes dous factores de dispersión e frecuencia, e para axustar o índice de neoloxicidade a uns parámetros de calidade neoloxica axeitados, introducimos no cálculo de  $v$  dúas constantes discriminatorias dos usos neoloxicos:  $Ls(\varphi)$  ou o valor límite de  $\varphi$  a partir do cal o índice  $v$  toma un valor negativo, e  $Ls(\delta)$  ou o valor límite de  $\delta$  a partir do cal o índice  $v$  toma un valor negativo. Para o noso estudo, asumimos que  $Ls(\varphi)$  e  $Ls(\delta)$  coinciden cos límites máis altos de  $\varphi$  e de  $\delta$  documentados na *Termoteca*. Neste caso, os límites máis altos de  $\varphi$  e de  $\delta$  documentados danse no termo "calidade de vida", que aparece 652 veces en 208 textos diferentes, cun valor de  $\varphi$  de 53,17 e un valor de  $\delta$  de 45,80. Así mesmo, aos efectos de reflectir no índice de neoloxicidade a importancia outorgada á primeira documentación ( $\alpha$ ), á frecuencia relativa ( $\varphi$ ) e á dispersión relativa ( $\delta$ ), os valores de  $\alpha$ ,  $\varphi$  e  $\delta$  son ponderados nunha proporción de 50, 20 e 30 sobre 100, respectivamente.

Por último, consideramos o criterio da exclusión lexicográfica dos compoñentes do neónimo como un factor de corrección do índice de neoloxicidade obtido mediante a avaliación dos criterios de antigüidade, frecuencia e dispersión, polo que introducimos no cálculo de  $v$  a exclusión lexicográfica relativa dos seus compoñentes ( $\kappa$ ), entendida como a cantidade de compoñentes do neoloxismo que non están repertoriados no corpus lexicográfico de exclusión, en relación coa cantidade de compoñentes identificables no neoloxismo. Canto maior sexa a proporción de compoñentes do neoloxismo non repertoriados, maior debe ser o seu índice  $v$ . Para o cálculo de  $v$ , asignámoslle a  $\kappa$  un valor proporcional entre 0 e 0,02 en función da porcentaxe de elementos do neónimo

recollidos no corpus de exclusión (por exemplo,  $\kappa=0$  se están todos recollidos,  $\kappa=0,01$  se faltan a metade dos compoñentes, e  $\kappa=0,02$  se ningún compoñente está recollido).

Na fórmula matemática deseñada para o cálculo do índice  $v$ , substituíndo  $Li(\alpha)$ ,  $Ls(\alpha)$ ,  $Ls(\varphi)$  e  $Ls(\delta)$  polos valores empregados neste estudo, a fórmula operativa que aplicamos no cálculo do índice  $v$  na *Termoteca* é a seguinte (tendo en conta que 32 son os anos entre 2007 e 1975):

$$v = \frac{(((\alpha - 1975) / 32) \times 50) + (((53,17 - \varphi) / 53,17) \times 20) + (((45,80 - \delta) / 45,80) \times 30)) - 50}{50} + \kappa$$

Despois de presentar os parámetros utilizados que nos permiten falar da neoloxidade dunha unidade léxica, cómpre establecemos unha aproximación á definición de neoloxismo que fundamenta este traballo. Entendemos, con Sablayrolles (2002), que o concepto de neoloxismo é un concepto gradual que nos permite falar de unidades non rexistradas en dicionarios, cunha data recente de aparición, de pouco uso en poucos textos, e que posúen elementos constituíntes que, á súa vez, poden ser considerados novos. A nosa proposta ten unha base denominativa e exclúe consideracións cognitivas e psicolóxicas, malia seren estas importantes, especialmente a primeira, para unha consideración da neoloxía terminolóxica.

### 3. Análise dos resultados

Como dixemos anteriormente, partindo dos datos tirados da *Termoteca* da Universidade de Vigo, referidos ao subcorpus de ecoloxía e ciencias ambientais do Corpus Técnico do Galego, recollemos 777 unidades que non cumpriron o criterio inicial de estaren incluídos no corpus de exclusión. Isto significa que, dos 1444 termos censados na *Termoteca* no campo das ciencias ambientais, un 53,8 % é considerado neoloxismo. Unha cifra tan alta está condicionada, evidentemente, polo corpus de exclusión. Para cada un dos neónimos así identificados, analizamos a súa antigüidade, a súa frecuencia, a súa dispersión e a exclusión lexicográfica dos seus compoñentes, e destes datos derivamos o seu índice de neoloxidade ( $v$ ). Presentaremos a seguir algúns dos resultados xerais que se desprenden desta análise.

(a) *Datación*. Nos 777 neónimos lexicográficos identificados, a datación máis serodia é de 2006; a máis temperá, de 1859. Hai 185 neoloxismos lexicográficos (23,8 %) que se recollen entre 2001 e 2006; 269 (34,6 %) entre 2000 e 1995; 155 (19,9 %) entre 1994 e 1989; 43 (5,5 %) entre 1983 e 1988; 54 (6,9 %) entre 1977 e 1982 e 71 (9,1 %) entre 1859 e 1976. Sen poder afirmar que existe unha data a partir da cal un neoloxismo deixa de selo, podemos observar que o 58,4 % dos neoloxismos lexicográficos identificados están rexistrados nos últimos 11 anos, feito que lles outorga, en principio, un trazo de novidade.

(b) *Frecuencia*. No *Corpus Técnico do Galego* (dunha extensión total de 12.326.694 de palabras, no momento de facer a consulta), as frecuencias absolutas dos neónimos identificados son as seguintes: hai 40 neoloxismos que teñen máis de 100 ocorrencias (de 101 a 652) (5,1 %), 54 entre 51 e 100 (6,9 %), 238 entre 11 e 50 (30,6 %) e 445 entre 1 e 10 (57,3 %). Destes últimos, 64 posúen unha única ocorrencia (8,2 %). Con estes datos enxérgase que os exemplos máis frecuentes son minoría (o 5,1 %), o que, de consideramos a frecuencia un esteo fundamental para establecer o índice de neoloxidade, podería levar á conclusión de que polo menos un 57,3 % dos exemplos analizados terían un alta calidade neolóxica.

(c) *Dispersión*. Dos 777 neónimos identificados, a partir dun total de 4.547 textos que compoñen o CTG (no momento da consulta), 8 neoloxismos rexístranse en máis de 100 textos (1 %), o máis “disperso” dos cales está en 208 textos e todos menos 1 (*avaliación do impacto ambiental*) obteñen un índice negativo de neoloxicidade; 22 aparecen en 50-99 textos (2,8 %), 6 dos cales con índice de neoloxicidade negativo; 219, en 10-49 textos (28,2 %), 39 con índice negativo, e 528 en menos de 10 textos (68 %), dos cales 19 posúen un índice negativo de neoloxicidade e 94 só aparecen nun texto (12,1 %). Isto tamén, en principio, podería indicar que a meirande parte dos neoloxismos identificados teñen un alto nivel de neoloxicidade por apareceren en moi poucos textos (68 % en menos de 10).

(d) *Compoñentes*. Un dos parámetros que utilizamos para completar os datos da fórmula foi o da exclusión lexicográfica dos compoñentes. Só 45 exemplos dos 777 (o 5,8 %) tiñan algún compoñente que non estaba rexistrado no dicionario (nos compostos cultos e nos sintagmáticos). O resto (o 94,2 %) tiñan todos os seus compoñentes repertoriados. Estes resultados están moi relacionados co corpus de exclusión empregado, xa que non recolle a meirande parte dos compostos sintagmáticos, pero si os compoñentes individuais.

(e) *Índice de neoloxicidade (v)*. Para analizar o índice de neoloxicidade (v), establecemos arbitrariamente tres segmentos, obtidos simplemente pola división en terzos do índice máis alto de neoloxicidade (0,9652519), para facelos coincidir cun concepto difícil de definir, doado de entender e aínda máis doado de criticar, pero con utilidade metodolóxica: a existencia dun nivel alto (N3), medio (N2) e baixo (N1) de neoloxicidade. No listado completo de unidades analizadas, dispoñible para consulta pública vía web na *Termoteca*, pódese comprobar que as extremas entre o nivel máis baixo do grupo superior e máis alto do inferior están pouco definidas porque as diferenzas son insignificantes. Porén, entre os niveis máis altos de neoloxicidade de N3 e N2 e N1, e entre o máis alto de N3 e o máis baixo de N1, evidentemente, si que hai diferenzas significativas que máis abaixo comentaremos. As características destes tres grupos, e mais do grupo N0 de neoloxismos lexicográficos con índice de neoloxicidade v negativo, son as seguintes:

(N3) *Alto nivel de neoloxicidade*. Agrupa 351 neónimos lexicográficos (o 45 % dos 777 identificados) cun valor de v entre 0,9652519 e 0,6420747. Neste grupo a data máis recuada é a de 1996 (*ambiente acuático*), a frecuencia máis alta é 85 (*factor de emisión*) e a maior dispersión 24 (*avaliación de efectos ambientais*). Velaquí os neónimos dentro deste grupo que obtiveron un maior índice de neoloxicidade neste estudo, seguidos pola data da súa primeira documentación, a súa frecuencia absoluta, a súa dispersión absoluta e o seu índice v.

metanoxénico	2006	1	1	0,9652519
xustiza ambiental	2006	1	1	0,9652519
ecosistema hídrico	2006	2	1	0,9646384
benzopireno	2006	11	2	0,9562323
xestión ecolóxica do lixo	2006	7	3	0,9558017
composto xenobiótico	2005	3	3	0,9370057
sistema de ecoetiquetado	2005	6	3	0,9351652
bioacumulativo	2005	1	1	0,9340019
elevada alcalinidade	2005	1	1	0,9340019
praga masiva	2005	1	1	0,9340019

Como se pode observar nestes dez exemplos, a datación é o criterio determinante, de acordo co peso superior que se lle deu a esta variable na fórmula: *metanoxénico*, *xustiza ambiental*, *ecosistema hídrico*, *benzopireno* e *xestión ecolóxica do lixo* pertencen ao ano 2006 e nin sequera, *benzopireno* ou *xestión ecolóxica do lixo*, cun número relativamente alto de ocorrencias en 2 e 3 textos respectivamente, se ven afectados por este feito. Comparémolos con *bioacumulativo* un neoloxismo de 2005, con baixa recorrencia e pouca dispersión, que ten menor índice de neoloxicidade. Repárase, con todo, que estamos a falar de variacións mínimas nos valores do índice.

(N2) *Nivel medio de neoloxicidade*. Corresponde a 251 neónimos lexicográficos cun índice entre 0.6412282 e 0.3193399, que representan o 32 % dos neónimos analizados. Neste grupo a data máis recuada é 1987 (*actividade extractiva*), a frecuencia máis alta é 336 (*autorización ambiental integrada*) e a maior dispersión 66 (*autoridade ambiental*). Pódense contemplar a continuación os neónimos que, dentro deste grupo, obtiveron un maior índice de neoloxicidade no noso estudo, no formato de táboa xa explicado.

sanidade ambiental	1997	19	12	0,6412282
arrastre peláxico	1996	6	4	0,6409655
residuo gandeiro	1997	29	10	0,6408624
monumento natural	1997	23	12	0,6387742
recursos de auga	1996	5	5	0,6387594
substrato xeolóxico	1996	5	5	0,6387594
planta termoeléctrica	1996	7	5	0,6375324
sistema de aforro	1996	7	5	0,6375324
turismo verde	1996	7	5	0,6375324

Nestes exemplos observamos que todos os exemplos son de 1997 ou 1996. As ocorrencias condicionan o nivel de neoloxicidade: *sanidade ambiental* e *monumento natural*, ambos de 1997 e cunha dispersión diferente (12 e 10) están condicionados polo número de rexistros, 19 no primeiro caso (maior neoloxicidade) e 23 no segundo (menor neoloxicidade). No caso de *arrastre peláxico*, xa do ano 96, a dispersión en textos é relevante, xa que en comparanza con *turismo verde*, aquel ten un maior índice de neoloxicidade polo feito de estar recollido nun texto menos (a diferenza de ocorrencias é tamén dun só exemplo). Este caso sería paradoxal se aplicásemos o criterio da “impresión neolóxica”, xa que o concepto de *arrastre peláxico* pode ser considerado desde este punto de vista moito menos neolóxico có de *turismo verde*.

(N1) *Nivel baixo de neoloxicidade*. Recolle 92 neónimos (o 12 % do total analizado) cun valor para o índice  $v$  entre 0,3169862 e 0,0000631. Neste grupo a data máis recuada é 1976 (*animal silvestre*), a frecuencia máis alta é 531 (*parque eólico*) e a maior dispersión 107 (*avaliación do impacto ambiental*). Dentro deste grupo, os neónimos con maior índice de neoloxicidade son os seguintes:

residuo tóxico	1990	64	39	0,3169862
microclimático	1986	12	7	0,3161957
valor ecolóxico	1991	68	51	0,3111668
substancia contaminante	1992	108	54	0,3092231
sistema de transporte	1990	69	41	0,3081495
gas de efecto invernadoiro	1995	305	45	0,3080757
arrastre de profundidade	1985	4	2	0,3042398
factores naturais	1985	4	3	0,3013922
ave rapaz	1985	7	5	0,2937034



recursos hídricos                      1991            118    47    0,2920304

Neste feixe de exemplos, *arrastre de profundidade*, *factores naturais* e *ave rapaz* de 1985 teñen un menor índice fronte a outros como *gas de efecto invernadoiro* de 1995 e *valor ecolóxico* de 1991, malia estaren estes últimos rexistrados nun amplo número de textos con múltiples ocorrencias.

(N0) *Neoloxidade negativa*. 83 neónimos lexicográficos (o 11 % do total analizado) obteñen un índice de neoloxidade v negativo. Neste grupo, a data máis recuada é 1859 (*fango*) –ou 1878 (*auga do mar*), se excluímos os castelanismos–; a frecuencia máis alta é de 652 (*calidade de vida*), exemplo que tamén ten a maior dispersión (208); e o neoloxismo lexicográfico máis moderno é de 1992 (*sistema de xestión*). Como neónimos con maior índice de neoloxidade dentro deste grupo, aparecen os seguintes:

aproveitamento da enerxía	1977	23	19	-0,0067024
residuo sólido	1990	292	104	-0,0103911
balsa de decantación	1975	7	3	-0,0129483
bioloxía mariña	1975	4	4	-0,0139924
balsa de homoxeneización	1975	10	4	-0,0176734
política enerxética	1972	47	29	-0,2062382
ordenación do territorio	1977	149	66	-0,2192957
actividades humanas	1974	82	56	-0,2439964
recursos pesqueiros	1972	132	36	-0,2785777
residuo urbano	1975	346	64	-0,3968853
sector industrial	1976	207	109	-0,4101669

Este grupo de neoloxismos lexicográficos non posúen un índice positivo de neoloxidade polo feito de seren rexistrados en anos afastados (1975, 4 exemplos, 1974, 1 exemplo e 1972, 2 exemplos), ou por teren un alto número de ocorrencias (*residuo urbano*, *sector industrial*) que ademais teñen unha ampla dispersión textual. Un caso curioso é o de *residuo sólido* de 1990 que, malia estar recollido nun ano serodio, ten un índice negativo pola súas ocorrencias e dispersión.

Hai outros exemplos como *avaliación ambiental* cun índice de -0,06 e datado en 1998, *contaminación ambiental* (-0,06/1975), *tratamento biolóxico* (-0,08/1975), *axente contaminante* (-0,09/1975), *aproveitamento de recursos* (-0,13/1972), *política enerxética* (-0,20/1972), *ambiente* (-0,86/1973), que poderían ser considerados neoloxismos segundo o criterio de impresión neolóxica. Isto fainos albiscar a posibilidade de que, nun futuro, a avaliación da percepción neolóxica teña que ser considerada para o establecemento dun índice de neoloxidade, dunha maneira que for obxectiva.

#### 4. Conclusións

Aínda que os resultados non son concluíntes, si se comproba que o índice v de neoloxidade é un complemento útil do criterio lexicográfico na caracterización e selección da neoloxía porque nos permite ofrecer uns resultados mesurables. Coa fórmula aplicada quixemos poñer en xogo todos os criterios que en diversos artigos teóricos se mencionaron para a consideración do concepto de neoloxidade. É a nosa intención ampliar o tamaño dos corpóra técnicos consultados para a determinación da primeira documentación, da frecuencia e da dispersión, e incluír a dispersión cronolóxica das ocorrencias do neónimo como un factor máis na formulación do índice

v. Así mesmo, pretendemos aplicar o método aquí explicado á neoloxía xeral para, a partir dos resultados, seguir por este vteiro ou establecer as correccións pertinentes.

## 5. Referencias bibliográficas

- Boulanger, J.-C. (1989): “L’*évolution du concept de néologie de la linguistique aux industries de la langue*”. En C. de Schaetzen (ed.), *Terminologie diachronique. Actes du Colloque organisé à Bruxelles les 25 et 26 mars 1988*. Paris/Bruxelles: Conseil Int. de la Langue Française / Ministère de la Communauté Française de Belgique, 193-211.
- Cabré, M.T. (1993): *La terminología. Teoría, metodología, aplicaciones*. Barcelona: Empúries.
- Cabré, M.T. et al. (2004): “La lexicografía i la identificació automatitzada de neologia lèxica”. En P. Battaner e J. DeCesaris (eds.), *De Lexicografía: actes del I Symposium Internacional de Lexicografía*. Barcelona: Universitat Pompeu Fabra, 287-294.
- Carballeira Anxo, X. M. (coord.) (2000): *Gran diccionario Xerais da Lingua*. Vigo: Xerais.
- Corbeil, J.-C. (1971): “Aspects du problème néologique”, *La Banque des Mots* 2, 123-136.
- Crespo Bastos, A., X.M. Gómez Clemente, X. Gómez Guinovart e S. López Fernández (2008): “XML-based Extraction of Terminological Information from Corpora”. En J.C. Ramalho, J. Correia e S. Abreu (eds.), *Actas da 6ª Conferência Nacional XATA2008. XML: Aplicações e Tecnologias Associadas*. Universidade de Évora, 28-39.
- Faura i Pujol, N. (2007): “El criteri lexicogràfic per a la identificació dels neologismes i els corpus d’exclusió”. En Lorente, M. et al. (eds.), *Estudis de lingüística i de lingüística aplicada en honor de M. Teresa Cabré Castellví*. Barcelona: Universitat Pompeu Fabra, 277-282.
- Gardin, B. et al. (1974): “À propos du ‘sentiment néologique’”, *Langages* 36, 45-72.
- Gómez Clemente, X.M. e X. Gómez Guinovart (dir.) (2006-2008): *Corpus Técnico do Galego*. Vigo: Universidade de Vigo. <http://sli.uvigo.es/CTG/>.
- Gómez Clemente, X.M. e X. Gómez Guinovart (dir.) (2006-2008): *Termoteca - Banco de Datos Terminolóxico da Universidade de Vigo*. Vigo: Universidade de Vigo. <http://sli.uvigo.es/termoteca/>.
- Gómez Guinovart, X. (2007): “A investigación en lexicografía e terminoloxía no Corpus Lingüístico da Universidade de Vigo (CLUVI) e no Corpus Técnico do Galego (CTG)”. En *Actas do Simposio 'A Lexicografía galega do século XXI'*. Santiago de Compostela: Universidade de Santiago de Compostela.
- López Martínez, M. e F. García Gondar (dir.): *Corpus de referencia do galego actual*. <http://balteira.cirp.es/corga/>.
- Pena, X. A. (2005): *Gran diccionario Século XXI da lingua galega*. Vigo: Editorial Galaxia / Edicións do Cumio.
- Quirion, Jean(2003): “Methodology for the Design of a Standard Research Protocol for Measuring Terminology Usage”, *Terminology* 9.1, 29-49.
- Rey, A. (1976): “Le néologisme: un pseudoconcept?”, *Cahiers de Lexicologie* 28, 3-17.
- Sablayrolles, J.F. (2002): “Fondements théoriques des difficultés pratiques du traitement des néologismes”. *Revue française de linguistique appliquée VII-1*, 97-111.
- Santamarina Fernández, A. (dir.): *Tesouro informatizado da lingua galega*. <http://www.ti.usc.es/TILG/>.
- Unión Latina (coord.) (2002): *Léxico do medio*. París/Barcelona: Unión Latina/Termcat.