



CONVENIO PLURIANUAL DE COLABORACIÓN ENTRE A CONSELLERÍA DE CULTURA, EDUCACIÓN E ORDENACIÓN UNIVERSITARIA E A UNIVERSIDADE DE SANTIAGO PARA A REALIZACIÓN DUN PROGRAMA DE ACCIÓNS EN MATERIA DE INVESTIGACIÓN A TRAVÉS DO CENTRO RAMÓN PIÑEIRO PARA A INVESTIGACIÓN EN HUMANIDADES

Santiago de Compostela , 13 de marzo de 2018.

REUNIDOS:

Don , en virtude do seu cargo de consellerio de Cultura, Educación e Ordenación Universitaria, para o que foi nomeado polo Decreto 148/2016, do 13 de novembro, no exercicio das facultades que lle confire o artigo 34 da Lei que actúa en virtude das facultades atribuídas no artigo 34 da Lei 1/1983, do 22 de febreiro, reguladora da Xunta e da súa Presidencia, modificada pola Lei 11/1988, do 20 de outubro e pola Lei 12/2007, do 27 de xullo.

Don , reitor magnífico da Universidade de Santiago de Compostela, en virtude do establecido nos seus estatutos e do nomeamento como reitor polo Decreto 62/2014, do 28 de maio, (DOG nº 105 de 4 de xuño), actuando en nome e representación da devandita entidade.

As partes actúan no exercicio dos seus respectivos cargos e coa representación competencial que ostentan, e recoñecéndose mutuamente capacidade para se obrigaren nos termos do presente convenio,

EXPOÑEN:

1. Que a recuperación da nosa lingua depende en gran medida do apoio que lle presten as institucións e entidades, contribuíndo así á súa difusión e ao proceso de normalización lingüística.
2. Que para a potenciación do emprego do galego (artigos 5º.3 do Estatuto de autonomía de Galicia e 6º.3 da Lei de normalización lingüística) cómpre dispor de estudos sobre diversos eidos da lingua galega, que contribúan a protexela, promovela e modernizala para que sexa un medio de comunicación eficiente en cada momento.
3. Que o Estatuto de autonomía de Galicia establece, no seu artigo 27.20, a competencia exclusiva da Comunidade Autónoma de Galicia en materia de promoción e ensino da lingua galega.
4. Que a Universidade de Santiago de Compostela (en diante USC) é unha entidade que leva a cabo actividades de investigación, docencia e desenvolvemento científico e tecnolóxico, e ten como un dos seus fins o de procurar o seu enraizamento na sociedade galega que a sustenta e, en xeral, contribuír ao seu desenvolvemento social, económico e cultural.



5. Que os ditos fins son apoiados e compartidos pola Consellería de Cultura, Educación e Ordenación Universitaria (en diante consellería) e que, con data do 11 de febreiro de 1993, foi creado o Centro de Investigacións lingüísticas e literarias Ramón Piñeiro, modificado o seu nome polo Decreto 330/1997, do 13 de novembro (DOG nº 227, do 24 de novembro), polo que pasa a denominarse Centro Ramón Piñeiro para a Investigación en Humanidades (en diante centro), co fin de levar a cabo tarefas de estudo filolóxico e conservación das fontes lingüísticas, literarias e antropolóxicas que permitan unha investigación profunda da nosa historia lingüística e literaria, que ademais de garantir a utilidade da lingua galega na sociedade actual, sirvan, asemade, de soporte para outros investigadores.

6. Que, posto que os seus intereses son concorrentes e perseguen a obtención dunha finalidade común, a USC e o CRPIH veñen colaborando desde o ano 1995 (sinatura do primeiro convenio) no desenvolvemento conxunto de actuacións de investigación relacionadas coa lingua e a literatura galegas, para o cal xa se formalizaron diversos convenios de colaboración de xeito ininterrompido -o cal amosa un destacado compromiso das partes en continuar cooperando en materia de investigación lingüística.

Así, a USC é quen dispón dos coñecementos necesarios, en clave de visión xeral do nivel actual do estado das investigacións, para propiciar contornas de colaboración tendentes a favorecer unha axeitada continuidade dos proxectos de investigación en curso que se desenvolven no centro.

7. Que, co propósito de que as tarefas encomendadas ao centro acaden resultados axeitados, a consellería desexa continuar contando coa colaboración dos investigadores de maior reputación con que contamos en cada un destes ámbitos e dadas as particularidades e características específicas dos proxectos de investigación de que se trata, tendentes a indagar e afondar sobre a identidade colectiva de Galicia e sobre estudos lingüísticos, literarios e filolóxicos galegos -co fin de contribuír a un maior prestixio e a unha crecente normalización da lingua galega- tendo en conta que os departamentos de Filoloxía da USC contan con recursos tecnolóxicos axeitados e con destacados investigadores de recoñecida capacidade científica -en relación á natureza dos proxectos incluídos no anexo deste convenio- os obxectivos perseguidos non poderían acadarse co mesmo nivel de excelencia nun escenario de concorrencia pública.

8. Que é interese das partes asinantes, no marco da colaboración establecida, incorporar novas perspectivas, co fin de reforzar o obxectivo de facilitar un servizo de carácter público e marcado interese xeral, ampliando o alcance e a accesibilidade dos traballos e resultados da execución destes proxectos, facendo deles unhas ferramentas modernas e eficaces de referencia científica mediante a introdución e mellora, en termos de alcance, no manexo de novos usos asociados ás tecnoloxías da información e da comunicación (TIC).

9. Que a USC dispón de investigadores de valía profesional e prestixio recoñecido que se dedican sectorialmente á realización deste tipo de actividades e cos que se pode establecer un espazo común de intereses no eido da investigación, que permiten a utilización da figura do convenio de colaboración como fórmula negocial apropiada para articular as súas relacións, dado que non se manifesta nunha contraposición de intereses, senón que se trata de establecer unha colaboración institucional para levar a cabo unha actuación en resposta de obxectivos compartidos e os intereses de ambas as dúas partes teñen por causa a consecución dun fin común (a investigación lingüística, literaria e antropolóxica que potencie o prestixio crecente da lingua galega, así como a súa canalización, difusión e proxección) en beneficio da comunidade.

Así mesmo, ao abeiro do establecido no artigo 4 –“*Negocios e contratos excluídos*”-, do Real Decreto Legislativo 3/2011, do 14 de novembro, polo que se aproba o Texto Refundido da Lei



de Contratos do Sector Público, este caso considérase excluído e plenamente aplicable á figura do convenio de colaboración debido ás prestacións de investigación e desenvolvemento que, con respecto aos proxectos incluídos no anexo do presente convenio, van ser levadas a cabo por parte da Universidade e que son susceptibles de se incluír no apartado r), en relación, á súa vez, co apartado c) – “*Convenios de colaboración*”- do citado artigo.

Por todo isto, a consellería e mais a USC establecen o presente convenio de colaboración que se rexerá polas seguintes

ESTIPULACIÓNS:

Primeira. Obxecto do convenio

O fin último desta cooperación é o de desenvolver proxectos de investigación que sexan estratéxicos para a garantir a utilidade da lingua galega en todos os ámbitos da vida actual e de proxectos que sexan de investigación fundamental e substancial para outros investigadores.

A través deste convenio establécense as condicións polas que se rexerá a colaboración entre a consellería e a USC, a través da Secretaría Xeral de Política Lingüística, para proseguir co desenvolvemento das actividades do programa de accións en materia de investigación, especialmente no ámbito lingüístico, que se está levando a cabo no CRPIH que se recolle no anexo citado na estipulación segunda que forma parte do convenio.

Segunda. Obrigas da USC

As actividades que ha desenvolver a USC en colaboración co centro, recóllense no ANEXO I ao presente convenio onde se especifican os traballos que se realizarán en cada unha das accións que conforman o programa de investigación.

A USC cooperará coa consellería e coa Secretaría Xeral de Política Lingüística na realización das actividades do programa de actuacións en materia de investigación científica do centro e comprométese a realizar -a través dos seus investigadores- os distintos proxectos que nel se desenvolven:

- PROXECTO RECURSOS PARA O DESENVOLVEMENTO DAS TECNOLOXÍAS DA FALA
- PROXECTO TERMINOLOXÍA CIENTÍFICO-TÉCNICA (TERMIGAL).
- PROXECTO BASE DE DATOS DO ATLAS LINGÜÍSTICO ROMÁNICO (ALIR)
- PROXECTO BIBLIOGRAFÍA INFORMATIZADA DA LINGUA GALEGA (BILEGA)
- PROXECTO CORPUS DE REFERENCIA DO GALEGO ACTUAL (CORGA)
- PROXECTO CORPUS DOCUMENTALE LATINUM GALLAECIAE (CODOLGA)
- PROXECTO DICCIONARIO GALEGO- ITALIANO / ITALIANO-GALEGO
- PROXECTO ARQUIVO GALICIA MEDIEVAL (ARGAMED)
- PROXECTO LÍRICA GALEGO-PORTUGUESA
- PROXECTO PROSA LITERARIA
- PROXECTO CANTIGAS DE SANTA MARÍA
- BIBLIOGRAFÍA DE REFERENCIA DA LITERATURA GALEGA MEDIEVAL (BIRMED)
- PROXECTO DICCIONARIOS DE LITERATURA
- PROXECTO BIBLIOGRAFÍA DA LITERATURA GALEGA
- PROXECTO INFORME ANUAL DE LITERATURA GALEGA
- PROXECTO DE INVESTIGACIÓN EN LITERATURA INFANTIL E XUVENIL



• PROXECTO DIMENSIÓNS DA IDENTIDADE COLECTIVA DE GALICIA

A USC colaborará na realización das distintas accións que integran este programa de investigación cunha contribución en especie, consistente en horas de dedicación á investigación, que serán as estimadas para a consecución dunha axeitada execución das actuacións.

A estes efectos, a valoración económica da devandita contribución equivalerá ó resultado do cálculo –en horas– sobre a retribución dun docente investigador -catedrático de universidade (T.C.)-. En consecuencia, a Universidade asumirá, pola súa conta, a cantidade de corenta e dous mil douscentos sesenta e seis (42.266 €) derivada dos custos do seu persoal docente investigador, distribuída en dúas anualidades (2018 e 2019), por un importe de 21.133 € cada unha.

Ademais, a USC adoptará as actuacións de soporte técnico necesarias que permitan un intercambio de datos, co fin de contribuír a un mellor desenvolvemento das investigacións, en termos de alcance.

A USC deberá cumprir coas estipulacións xurdidas do vínculo xerado coa celebración do presente convenio no xeito e termos establecidos.

Terceira. Obrigas da Consellería de Cultura, Educación e Ordenación Universitaria

A Consellería ha de propiciar a participación dos profesionais especialistas do CRPIH e a súa contribución activa nos proxectos obxecto de investigacións conxuntas.

Contribuirá á difusión e divulgación dos resultados das investigacións, co fin de canalizar e proxectar os estudos resultantes, con efectos sobre o conxunto da sociedade en xeral e dos profesionais especializados, en particular, non só no ámbito autonómico, senón tamén no nacional e mesmo no internacional.

Facilitará a transferencia de coñecemento científico propio do CRPIH nas áreas temáticas obxecto das investigacións.

Velará polo cumprimento das disposicións, que han de ser executadas no xeito e termos establecidos, en favor do patrimonio público.

A consellería, a través da Secretaría Xeral de Política Lingüística colaborará na realización destas accións cunha achega económica de douscentos setenta e nove mil douscentos catorce euros (279.214 €), repartidas en dúas anualidades coa seguinte distribución:

Exercicio 2018: 139.607 € con cargo á aplicación orzamentaria 10.30.151A.640.1

Exercicio 2019: 139.607 € con cargo á aplicación orzamentaria 10.30.151A.640.1

Estas cantidades só cubrirán o custo da realización das actividades, polo que non levan asociado ningún beneficio económico para USC.

Para atender os compromisos de gasto derivados das accións establecidas existe crédito axeitado e suficiente na aplicación orzamentaria sinalada, dos presupostos xerais da comunidade autónoma para os citados exercicios orzamentarios.

O pagamento, en recibindo a certificación correspondente dos traballos de investigación realizados, efectuarase en catro entregas repartidas do seguinte xeito:



- 1ª.- 69.803,50 € (1 de xuño de 2018)
- 2ª.- 69.803,50 € (1 de novembro de 2018)
- 3ª.- 69.803,50 € (1 de xuño de 2019)
- 4ª.- 69.803,50 € (1 de novembro de 2019)

Cuarta. Custo total do convenio

Tendo en conta as obrigas da USC, relacionadas na cláusula 2ª, e as da consellería, estipuladas na cláusula 3ª, o custo total do convenio ascende a unha cantidade de trescentos vinte e un mil catrocentos oitenta euros (321.480 €), distribuídos anualmente tal e como se recolle no Anexo II

Quinta. Libramento

Para recibir as achegas da Consellería de Cultura, Educación e Ordenación Universitaria que se indican anteriormente, a USC entregará na Secretaría Xeral de Política Lingüística, antes das datas indicadas, a seguinte documentación:

- Certificación por duplicado do/da xerente da universidade que acredite a execución das actividades do convenio ao longo do exercicio correspondente, xunto coas pertinentes notas de cargo emitidas polo servizo que lle corresponda á USC para a súa tramitación, así como unha relación valorada das horas dedicadas a cada unha das accións do programa de investigación.
- Unha vez executadas as actividades do convenio, e antes do 15 de decembro de cada ano, os investigadores responsables destas emitirán por duplicado unha memoria detallada das actuacións realizadas.

No caso de que a acreditación dos custos sexa inferior ao importe do convenio, a achega da Consellería de Cultura, Educación e Ordenación Universitaria diminuiríase na mesma proporción.

A USC facilitará toda a información que lle sexa requirida pola Consellería de Cultura, Educación e Ordenación Universitaria, así como pola Intervención Xeral da Comunidade Autónoma, Consello de Contas e o Tribunal de Contas no exercicio das súas funcións de fiscalización e control.

Antes de proceder á sinatura do presente convenio e do seu pagamento, a USC deberá presentar unha declaración responsable, asinada polo órgano competente, de non ter contraído débedas, por ningún concepto, coa Administración pública da Comunidade autónoma galega, así como de estar ao día das obrigas tributarias e ao corrente das obrigas sociais.

Sexta. Propiedade ou titularidade dos traballos e publicidade

A propiedade, titularidade e uso dos traballos será común polas partes e só poderán usalos para os fins de cada una delas sempre de xeito gratuíto.

Nas actuacións informativas e divulgativas, nas publicacións que se fagan relacionadas coas actividades deste convenio ou cando a USC ou calquera dos investigadores que participe nas actividades obxecto deste convenio desexa utilizar os resultados acadados para a súa publicación, deberán facer constar, nun lugar destacado, a colaboración da Consellería de Educación e Ordenación Universitaria, a través do CRPIH, e da USC, para o que se usarán os anagramas aprobados pola normativa de imaxe corporativa de ambas as institucións.



Sexa cal for o medio de difusión, cada unha das partes comprométese a facer mención da outra parte e deste convenio.

Sétima. Regulación da participación dos investigadores

A participación dos investigadores da USC nas actividades recollidas no presente convenio, efectuarase conforme á regulamentación propia da USC recollida nos seus estatutos e no Regulamento de cursos de especialización e convenios, e de acordo coa Lei de incompatibilidades do persoal ao servizo das administracións públicas e coa Lei orgánica 6/2001, do 21 de decembro de universidades.

Oitava. Utilización da lingua galega

Todas as accións e actividades que se realicen ao abeiro deste convenio deberanse levar a cabo de acordo co contemplado nas vixentes normas ortográficas e morfolóxicas do idioma galego, aprobadas pola Real Academia Galega no ano 2003, tal e como se establece na disposición adicional da Lei 3/1983, do 15 de xuño, de normalización lingüística. Así mesmo, coidarase especialmente do respecto rigoroso da toponimia oficial nos termos previstos no artigo 10 da citada lei.

Novena. Vixencia

Este convenio entrará en vigor o mesmo día da súa sinatura e terá vixencia durante os anos 2018 e 2019, sen posibilidade de prórroga, e poderá ser denunciado por calquera das partes asinantes por incumprimento do que nel se establece, cunha antelación dun mes polo menos, e a consellería poderá, se fose o caso, esixir o reintegro das cantidades achegadas para a súa realización.

Décima. Comisión de seguimento

Crearase unha comisión de seguimento das accións obxecto do presente convenio- facultada para proporlles ás partes as resolucións de cantas dúbidas xurdan na súa execución- presidida polo secretario xeral de Política Lingüística e integrada polo catedrático de Filoloxía Románica da USC, investigador principal e coordinador deste convenio, e polo secretario do Consello científico e executivo do CRPIH, que actuará como secretario da mesma.

Décimo primeira. Responsabilidades das actuacións

A sinatura deste convenio non implica relación laboral, contractual ou de calquera outro tipo entre os profesionais que vaian desenvolver as actividades previstas e a Xunta, de tal xeito que a esta non se lle pode esixir responsabilidade ningunha, nin directa, nin indirecta, nin subsidiaria, polos actos ou feitos acaecidos no desenvolvemento do convenio.

Décimo segunda. Rexistro de convenios

A sinatura do presente convenio suporá o consentimento expreso do interesado á Administración para incluír e facer públicos os datos referidos ao convenio, de conformidade co artigo 15 da Lei 1/2016, do 18 de xaneiro, de transparencia e bo goberno de Galicia e co Decreto 132/2006, do 27 de xullo, polo que se regulan os rexistros públicos de convenios da Xunta de Galicia.



XUNTA DE GALICIA

CONSELLERÍA DE CULTURA, EDUCACIÓN
E ORDENACIÓN UNIVERSITARIA
Secretaría Xeral de Política Lingüística

R/ San Roque, Nº 2
15704 Santiago de
Compostela
Telf.: 881 996 139
Fax.: 881 996 140
E-mail: crpih@cirp.gal

Décimo terceira. Normativa aplicable e xurisdición

Este convenio ten carácter administrativo e os seus efectos rexeranse polo establecido nas súas estipulacións, e quedará suxeito ao disposto na disposición adicional segunda da Lei 1/2016, do 18 de xaneiro, de transparencia e bo goberno de Galicia.

Aínda que é un dos excluídos do ámbito de aplicación do Real Decreto legislativo 3/2011, do 14 de novembro, polo que se aproba o texto refundido da Lei de Contratos do sector público, segundo especifica o seu artigo 4, ao non previsto aplicaráselle os principios recollidos na citada lei, co fin de resolver dúbidas e omisións que puidesen presentarse. Así mesmo e supletoriamente, terase en conta o Decreto legislativo 1/1999, do 7 de outubro, polo que se aproba o texto refundido da Lei de réxime financeiro e orzamentario de Galicia e demais normativa vixente na materia.

A xurisdición contencioso administrativa será a competente para dirimir as cuestións que puidesen xurdir da aplicación deste convenio.

En proba de conformidade cos termos deste convenio, as partes comparecentes asinan por triplicado, no lugar e data mencionados.

Pola consellería de Cultura,
Educación e Ordenación Universitaria

Pola Universidade de
Santiago de Compostela



ANEXO I

As actividades de investigación que ha desenvolver a USC, ás que se fai referencia na cláusula segunda do presente convenio de colaboración, relaciónanse a continuación:

1. ÁREA DE LINGÜÍSTICA

1.1. LINGÜÍSTICA COMPUTACIONAL

ACCIÓN 1.1.1

• RECURSOS PARA O DESENVOLVEMENTO DAS TECNOLOXÍAS DA FALA

Despois da conclusión e posta en código aberto en Sourceforge do sintetizador de voz Cotovía, o Grupo de Tecnoloxías da fala, integrado por enxeñeiros de telecomunicacións da Universidade de Vigo e por lingüistas da Universidade de Santiago, centra as súas actuacións durante os anos 2018 e 2019 na elaboración e potenciación de recursos para o desenvolvemento das tecnoloxías da fala en lingua galega.

Durante este período continuarase coa recollida, análise e tratamento de rexistros de voz, así como de rexistros de audio e vídeo, coa correspondente anotación nos distintos niveis.

Dada a necesidade da polivalencia dos recursos elaborados para que sexan útiles a proxectos diversos, os distintos niveis de anotación serán discernibles e utilizables por separado, en función das aplicacións para as que sexan usados.

Estes recursos, aínda que están especialmente orientados ao desenvolvemento e mellora das tecnoloxías da fala, serán tamén útiles para investigadores doutros ámbitos, particularmente do campo da lingüística.

• PROXECTO TERMIGAL (Servicio Galego de Terminoloxía Científico-Técnica).

Este servizo continuará coa atención ás tres seccións en que está estruturado: a) documentación de textos de especialidade orientados á elaboración de terminoloxía especializada.

b) Elaboración de terminoloxía sistemática

c) Elaboración de terminoloxía puntual.

Á parte dos proxectos de terminoloxía sistemática e da atención ao servizo de consultas sobre problemas denominativos, especialmente do ámbito científico e técnico, durante os anos 2018 e 2019 continuarase alimentando a base de datos TERGAL, proseguirase co fornecemento para a Comisión de Validación da Real Academia Galega do estudo previo dos termos sobre as linguas de especialidade que se estean a tratar, e continuarase co labor de revisión e correccións de vocabularios de especialidade elaborados fóra de Termigal.



• PROXECTO BASE DE DATOS DO ATLAS LINGÜÍSTICO ROMÁNICO (ALIR)

Continuación cos labores de investigación asociados á elaboración do Atlas Lingüístico Románico e de elaboración, revisión e enriquecemento da base de datos dialectais que reúna o léxico deste proxecto.

Particularmente, nos anos 2018 e 2019:

- Incorporaranse os datos dos novos mapas que elabore o proxecto europeo neste período.
- Estructuraranse os datos, para facilitar o seu tratamento desde o punto de vista categorial, morfolóxico, etimolóxico e semántico.
- Elaboraranse ao redor dunhas 15 sínteses galegas para fornecer as sínteses románicas deste proxecto.

ACCIÓN 1.1.2

• PROXECTO BIBLIOGRAFÍA INFORMATIZADA DA LINGUA GALEGA (BILEGA)

Tendo en conta que BILEGA é un repertorio dixital que alberga os datos bibliográficos dos traballos de investigación e divulgación/opinión sobre a lingua galega publicados desde o século XVI ata a actualidade, xunto cunha sinopse do seu contido e as recensións e notas de publicación a que deron lugar e que desde o comezo da súa construción aspirou a ser un repertorio exhaustivo e a estar sempre actualizado, a tarefa fundamental durante os dous próximos anos será seguir recollendo datos tanto en dirección retrospectiva (inclusión de edicións que pasaron desapercibidas ata agora, corrección de erros e elaboración de sinopses de traballos cuxa referencia xa está incluída na base de datos), coma prospectiva (inclusión de novas edicións e dos novos traballos que se vaian publicando, así como elaboración das correspondentes sinopses), continuando así as tarefas de compilación analítica con ese dobre obxectivo da exhaustividade e a actualización.

Arestora, a base de datos alberga algo máis de 20.000 rexistros, que conteñen 48.277 asentamentos bibliográficos para os que se implementaron 163.432 valores de busca nos campos tema, enfoque, período, variedade de lingua, ámbito, tipo de traballo, lingua do documento e na sección de identificadores do campo documento. Na inclusión de novos traballos ou de novas edicións de traballos xa referenciados na BILEGA, considérase fundamental seguir a explorar a rede Internet para identificar novas edicións dixitais en liña, das que arestora xa están incluídas as referencias dunhas 18.000.

Por outra banda, co obxectivo de contribuír á divulgación da extensa e detallada información que alberga a BILEGA, tamén se completarán as sinopses dos rexistros que posteriormente entrarán a formar parte do caderno nº 3 da colección CADERNOS DA BILEGA, que aparecerá co título O idioma galego no eido das novas tecnoloxías lingüísticas: Bibliografía analítica, e levaranse a cabo os labores de preparación do texto dunha Bibliografía analítica da fraseoloxía galega, que constará dunhas 800 entradas e que está previsto publicar no próximo número 18 dos Cadernos de Fraseoloxía Galega.



• PROXECTO CORPUS DE REFERENCIA DO GALEGO ACTUAL E ETIQUETADOR /LEMATIZADOR DO GALEGO ACTUAL (CORGA/XIADA)

Estes recursos computacionais abertos e dispoñibles en liña son sumamente importantes para o desenvolvemento de estudos do galego dende distintas perspectivas: morfolóxica, sintáctica, discursiva, terminolóxica, fraseolóxica, incorporación de neoloxismos, fixación da normativa etc. Así mesmo, con este proxecto a lingua galega consolídase no mundo das novas tecnoloxías e no do procesamento da linguaxe natural.

A versión 3.0 (presentada en novembro de 2017) -dispoñible para a súa consulta en liña no enderezo <http://corpus.cirp.gal/corga->, consta de 36,8 millóns de palabras, das que algo máis de 570.000 son formas diferentes. Nela unifícanse os dous sistemas de consulta en liña dispoñibles anteriormente (o do CORGA e o do CORGAetq).

Entre as numerosas novidades que achega esta versión destaca a incorporación dunha pequena mostra representativa do rexistro oral en cuxas transcricións se aliña o texto coa voz, de xeito que non só ofrece a posibilidade de realizar pesquisas en textos orais, senón tamén de escoitar o fragmento de audio que corresponde ao contexto dos resultados obtidos.

Porén, sen dúbida, a característica máis sobresaínte desta versión do corpus é a etiquetaxe automática da totalidade dos documentos que o conforman empregando o *Etiquetador/Lematizador do Galego Actual (XIADA)*.

Así, o CORGA presenta un motor de busca excepcionalmente potente que sitúa o galego na vangarda dos corpus lingüísticos.

Outras melloras destacables da versión 3.0 son as seguintes: incremento do volume textual coa inclusión de novos documentos pertencentes aos últimos lustros; emprego dunha interfaz gráfica amigable na que salienta a visualización dos resultados mediante concordancias e gráficas, as múltiples posibilidades de ordenación dos resultados ou a opción de deshabilitar a sensibilidade a acentos e maiúsculas nas procuras; clasificación tipolóxica dos documentos (se a orixe é oral ou escrita, se pertence ao bloque da ficción ou non, xénero ao que se remite e subtipo), ou a engádegas de abundante información nas páxinas estáticas sobre o corpus: descrición, documentos que o integran, historial de versións, publicacións que o referencian, guía de uso, descarga das listaxes de palabras, elementos gramaticais, lemas ou etiquetas etc.

Porén, co fin de situar o galego no mesmo nivel das linguas peninsulares e na vangarda da lingüística computacional, cómpre continuar desenvolvendo o proxecto. Por un lado, mediante a preparación e incorporación de novos documentos ao corpus, para progresivamente aumentar o seu volume textual e conseguir que siga sendo representativo do uso do galego actual.

Neste sentido, no 2018 e 2019, continuarase coa incorporación de textos correspondentes aos últimos dous lustros dos xéneros xa definidos: xornalístico, ensaístico e ficción. E tamén se impón a necesidade de aumentar a introdución no CORGA de textos procedentes de dúas vías iniciadas en anos recentes, e das que xa temos representación na versión 3.0:

a) Internet. A presenza do galego en Internet esixe que se documente este no corpus, tanto en relación co vocabulario procedente das novas tecnoloxías como co uso da lingua nun soporte amplamente difundido e que non se recolle en ningún outro corpus de galego.

b) Oral. Procederase á transcripción de programas de radio do período 2010-2017. O obxectivo, como se pode comprobar no sistema de consultas en liña da versión 3.0 é ofrecer ao usuario a dobre posibilidade de realizar buscas en textos de procedencia oral e



de escoitar a parte do audio no que se localizan os resultados obtidos. Isto supón un proceso totalmente novidoso que converte o CORGA nun recurso aínda máis valioso e único, ademais de fortalecer a presenza do galego na vangarda das tecnoloxías asociadas ao procesamento da linguaxe natural.

Dado que o CORGA debe estar etiquetado e lematizado automaticamente para facilitar a recuperación e extracción de información tanto por forma gráfica como por lema e etiqueta, é preciso dispor dun etiquetador. Con esta finalidade desenvólvese un proxecto paralelo denominado Etiquetador/Lematizador do Galego Actual (XIADA), ferramenta coa que se etiquetou a versión 3.0 do CORGA (versión actualizada en novembro de 2017). O etiquetador precisa, para inferir a gramática do galego, un corpus de adestramento desambiguado manualmente que, xunto con outros recursos de tipo lingüístico, serve para etiquetar automaticamente todo o CORGA. Para mellorar a calidade da etiquetaxe do CORGA e dado que o etiquetador segue un modelo estatístico, a investigación para os anos 2018 e 2019 debe centrarse:

1. No incremento do lexicón que serve de base para o recoñecemento e caracterización das palabras,
2. Na creación de regras lingüísticas que corrixa etiquetaxes anómalas recorrentes para as que a estatística non atina,
3. Na posibilidade de mellorar o tratamento dos nomes propios e as unidades multipalabra variables, co obxectivo de mellorar a porcentaxe de acerto do etiquetador e obter unha alta fiabilidade na etiquetaxe automática do CORGA.
4. Para incrementar as posibilidades de recuperación de información do CORGA, revisarase a asignación de hiperlemas no lexicón e deseñarase a súa aplicación no sistema de consultas en liña. O obxectivo é empregar o campo *hiperlema* no sistema de consultas como un parámetro máis. Este parámetro afianzará aínda máis a posición do galego na vangarda da lingüística de corpus, pois é un concepto non aplicado ata o de agora en corpus etiquetados.

Naturalmente, o labor lingüístico das epígrafes anteriores require de mantemento, actualizacións e implementación computacional. Neste ámbito, ademais, considérase imprescindible nos anos 2018-19 realizar as seguintes tarefas:

5. Publicar e liberar o código do etiquetador, contribuíndo á disposición pública e libre de ferramentas de procesamento da linguaxe natural.
6. Crear unha páxina web para explicar e promocionar o etiquetador.
7. Estudar a viabilidade de mellorar as funcionalidades de busca no corpus CORGA para visualizar: coaparicións, gráficas ponderadas, listaxe de frecuencias a demanda e ordenación dos resultados por número de elemento.
8. Solucionar os problemas que xurdan no procesamento de toda a base documental do CORGA.

ACCIÓN 1.1.3

• PROXECTO CORPUS DOCUMENTALE LATINUM GALLAECIAE (CODOLGA)

O Corpus Documentale Latinum Gallaeciae é un corpus textual con documentación latina medieval relacionada con Galicia no período cronolóxico que abrangue desde o século VI ata o XV, ambos inclusive.



As súas investigacións van encamiñadas á recompilación e estudo de documentos en latín referidos a Galicia desde o século VI ao XV, e edición e publicación dos mesmos nun corpus textual nunha base de datos denominada GALADOC que se actualiza anualmente.

Despois de publicar en rede o número 13 (2016) de CODOLGA, para o que se engadiron 1.307 novos documentos, súmase xa un total de máis de 13.000 documentos do entorno da Gallaecia postos a disposición dos investigadores, estudosos e interesados pola lingua, cultura e historia de Galicia. A situación é a seguinte, a piques xa de se publicar o número 14 (decembro de 2017) de CODOLGA:

Nº documentos novos no 2016: 1316 ; no 2016, 672 e, no 2017, 1.121

Total de documentos publicados na rede: 14.920 (nº 14)

Engadiranse tamén no novo número tres textos de fontes históricas importantes: Gesta Berengarii; Crónica Najerense; Chronicon Mundi de Lucas de Túi.

É de salientar o feito da estreita colaboración estes anos de CODOLGA e os outros proxectos derivados do noso, como CODOLCAT: *Corpus Documentale Latinum Cataloniae*, que se inscribe no proxecto *Glossarium Mediae Latinitatis Catalonia* (GMLC), centrado na documentación notarial e textos literarios, xurídicos e científicos escritos en latín e producidos no ámbito catalán entre os anos 700 e 1250 (<http://gmlc.imf.csic.es/codolcat/>). Tamén é fillo de CODOLGA o *Corpus Documentale Latinum Portucalese* ou CODOLPOR, que pretende poñer á disposición do público os documentos medievais do territorio portugués escritos en latín anteriores ó ano 1200 (<http://codolpor.ul.pt/>), e o *Corpus Documentale Latinum Valencie* (CODOLVA), que ten o obxectivo de crear unha base documental de textos latinos do Reino de Valencia, así como realizar edicións críticas da documentación diplomática, e analizar as interaccións entre a lingua latina e as vernáculos coas que convive (<https://web.ua.es/va/codolva/corpus-documentale-latinum-valencie-codolva.html>).

Estes grupos de investigación (CODOLGA, CODOLCAT, CODOLPOR e CODOLVA) teñen trabaxado de forma colaborativa dende os seus inicios, e este traballo culminou na creación da rede de excelencia *Corpus Documentale Latinum Hispaniarum*, coordinada dende a Universitat de Barcelona, e enmarcada no «Programa Estatal de Fomento de la Investigación Científica y Técnica de Excelencia» (Subprograma Estatal de Generación de Conocimiento, 01/01/2017-31/12/2018).

Continuarase a traballar na preparación da nova documentación latina medieval para o seu estudo e incorporación nos próximos números do proxecto CODOLGA, seguindo un camiño sistemático baseado nos datos do CODIPHIS (García de Cortázar, José Ángel, José Antonio Munita y Luis Javier Fortún: *Codiphis. Catálogo de colecciones diplomáticas hispano-lusas de época medieval*. 2 vols. Santander: Fundación Marcelino Botín, 1999) e na bibliografía específica galega máis relevante.

Por outra banda seguirase a preparar a edición propia dos *Documentos reais anteriores ó ano mil*, que está xa moi avanzada pero aínda chea de dificultades.

Un dos obxectivos prioritarios para o novo bienio é a preparación de tódolos materiais cara á migración a unha nova plataforma e entorno informático máis áxiles, que favorezan o labor do equipo de traballo e a consulta da base de datos por parte dos usuarios.



Outro dos grandes obxectivos é a integración funcional da rede **Corpus Documentale Latinum Hispaniarum** que permita a comparación dos resultados, e dos datos galegos mesmos, fronte ós das restantes bases de datos vinculadas ó noso proxecto.

1. 2. LEXICOGRAFÍA CONTRASTIVA

ACCIÓN 1.2.1.

• PROXECTO DICCIONARIO GALEGO- ITALIANO / ITALIANO-GALEGO

O Dicionario italiano-galego/galego-italiano, é unha obra lexicográfica orientada a falantes galegos que desexen producir textos en lingua italiana e viceversa. O leuario consta de 20.000 voces.

I) Liñas de investigación previstas para os anos 2018-2019

- 1) Continuar coas actuacións necesarias para a selección de novas entradas.
- 2) Corrección das mesmas.
- 3) Redacción dos artigos que compoñen a obra lexicográfica e as correccións pertinentes das palabras redactadas no ano anterior.

II) Actividade científica de tema lexicográfico

- 1) Continuar coa redacción de artigos para revistas especializadas en Lexicografía.
- 2) Participación en reunións científicas, congresos, encontros, cursos e outros eventos.
- 3) Participación en Redes Temáticas de Lexicografía: RELEX (ROMANYCOM, GLEX, FILGA, HUMBOLDT, ITALIANÍSTICA, Léxico e Lexicografía Especializada, TALG e FRASEONET) e cinco de outras universidades (Lexicografía, Lexicoloxía e Historiografía Lingüística da Universidad de La Laguna; IULATERM, do IULA, Universitat Pompeu Fabra; da Università de Bologna, campus de Forlì; GRILS, da Università di Milano, e o grupo de Lexicografía Monolingüe e Bilingüe Italiana, da Università di Torino). <http://relex.udc.es/> e do grupo de investigación LICEI, que se ocupa de proxectos e publicacións de temas lexicográfico, desde a perspectiva contrastiva italo-españolas.
- 4) Divulgación en xornais para achegar á xente dun modo divulgativo a relación entre Galicia e Italia pensando, sobre todo, no aspecto contrastivo entre a lingua galega e a italiana e darlle a coñecer ao público algúns temas da investigación na universidade

2. ÁREA LITERATURA

2.1. LITERATURA MEDIEVAL

ACCIÓN 2.1.1

• PROXECTO ARQUIVO GALICIA MEDIEVAL (ARGAMED)

Este Arquivo constitúe unha referencia internacional para os investigadores polas súas bases de datos *MedB* (Base de datos da Lírica Profana Galego-Portuguesa e *BiRMED* (Bibliografía de Referencia do Arquivo Galicia Medieval), que seguen a ser un dos obxectivos fundamentais do proxecto.



Para os vindeiros anos 2018 e 2019 proseguirase cos labores de redacción das publicacións –15 ata a actualidade– que estudan aspectos concretos da nosa literatura medieval e preparación de participacións e colaboracións en reunións científicas relacionadas coa materia.

Á súa vez, forman parte deste macroproxecto (**Argamed**), outros catro proxectos diferentes:

• PROXECTO LÍRICA GALEGO-PORTUGUESA

Nacido co obxectivo de proporcionar instrumentos de traballo fiables e de fácil acceso para os estudosos da lírica galego-portuguesa, na súa primeira fase recompiláronse case 1.700 cantigas que integran o *corpus e, posteriormente*, Posteriormente, procedeuse a preparar o material para convertelo nunha base de datos (MedDB v.1), que pode ser consultada na Internet.

Ao mesmo tempo, deseñouse unha base de datos para incorporar a bibliografía sobre a nosa lírica trobadoresca, con posibilidade de buscas temáticas mediante o recurso a palabras-clave e coa inclusión dun índice de citas, que deu lugar á actual BIRMED.

Continúase cos labores de recompilación, selección e asesoramento para a elaboración da base de datos (*MedB*, que xa vai pola súa terceira versión) na que se van recollendo todos os manuscritos orixinais das cantigas galego-portuguesas, desde o século XII ata o XIV, ademais dos estudos e edicións críticas sobre trovadores e xograres.

Continuación dos labores de actualización, revisión e mellora da base de datos *MedDB* (que en 2017 puxo en rede a súa terceira versión e está deseñando a cuarta), na que, ademais de ofrecer múltiples e relevantes opcións de busca relativas á totalidade do corpus, se recollen todos os manuscritos orixinais das cantigas galego-portuguesas, desde o século XII ata o XIV, ademais dos estudos e edicións críticas sobre trovadores e xograres.

Elaboración dunha base de datos complementaria (e conectada á anterior), que conterá a transcripción paleográfica de todos os testemuños para facilitar buscas de carácter paleográfico, filolóxico e lingüístico.

Publicación de novas edicións críticas e de novos estudos.

• PROXECTO BIBLIOGRAFÍA DE REFERENCIA DA LITERATURA GALEGA MEDIEVAL

Ante a envergadura da elaboración, revisión e actualización das bases de datos *MedB* e *BirMED*, procedeuse a crear un novo proxecto que terá como obxectivo a revisión completa e actualización da base de datos *BiRMED* ('Bibliografía de Referencia do Arquivo Galicia Medieval'), que contén, entre outras cousas, a posibilidade de coñecer as citas recibidas polos traballos que a compoñen.

Tamén se responsabilizará da incorporación de novas opcións de consulta na pantalla de busca e de recuperación da información, máis sinxelas e intuitivas, e de novas ferramentas que suporán un salto cualitativo nas bases de datos dedicadas aos estudos literarios.

Traballos de probas de integridade da nova versión da base de datos ('database integrity testing') para poder poñela á disposición do público.



• PROXECTO PROSA LITERARIA

Este proxecto pretende conservar e difundir a través de edicións impresas e electrónicas os textos en prosa escritos en galego durante a Idade Media. Ten como obxectivo prioritario a edición, estudo, caracterización e anotación das obras medievais que integran o xénero que dá título ao proxecto.

Ademais da elaboración (ou revisión, no caso de que xa existan traballos realizados con criterios filolóxicos rigorosos) das edicións paleográfica e crítica de cada texto, o proxecto contempla tamén o estudo dos aspectos filolóxicos e literarios que caracterizan a cada obra

No futuro, proseguirase coa recompilación dos textos medievais escritos en galego, xunto co seu estudo e redacción, sobre os seus aspectos filolóxico e literario. Estes traballos teñen como obxectivo a elaboración dunha base de datos dispoñible na rede.

Co obxectivo de proporcionar unha obra fiable e de fácil acceso, preparárase unha edición de *Os Miragres de Santiago*, obra senlleira da nosa cultura, da que aínda non existe unha boa edición do texto, o que permitirá realizar futuros estudos máis rigorosos e exhaustivos.

Paralelamente, comezase a articular unha base de datos de prosa galega medieval que tamén estará dispoñible na páxina do Centro Ramón Piñeiro.

Obxectivos para o 2018 e 2019:

- i) Continuación dos traballos de recompilación do material medieval escrito en galego e estudo e redacción sobre os seus aspectos filolóxico e literario, que ten como fin unha base de datos de acceso internacional.
- ii) Ademais, levaranse a cabo as tarefas necesarias para a edición dunha das principais obras do noso acervo cultural, os *Miragres de Santiago*, xa que aínda non existe unha boa edición do texto, co obxectivo de proporcionar unha obra fiable e de fácil acceso.
- iii) Estudo de diferentes aspectos da obra *Miragres de Santiago* e publicación deses traballos en revistas e monografías.

• PROXECTO CANTIGAS DE SANTA MARÍA

O *Proxecto Cantigas de Santa María* incorpora a parte de poesía relixiosa á lírica profana, tanto no que respecta aos textos afonsinos á MedDB, como á complementación da base de datos bibliográfica. O proxecto aspira a atender tamén os outros dous aspectos fundamentais desta obra literaria: o estudo musicolóxico e o iconográfico. Ata agora non existe unha aproximación exhaustiva a este monumento literario baixo esta triple perspectiva, que se vería completada coa dixitalización íntegra dos cancioneiros.

Proseguirase coas tarefas de estudo das *Cantigas de Santa María* e a revisión das transcricións dos códices.

Seguindo a metodoloxía da Lírica profana, estanse incluíndo nas bases de datos cerca de 500 cantigas relixiosas. O obxectivo máis inmediato é a dixitalización íntegra dos Cancioneiros e a súa edición.

Como xa se viña facendo con anterioridade, proseguirá a colaboración co proxecto Informes



de literatura para a elaboración de entradas específicas de literatura medieval galega.

Continuación das tarefas de estudo do corpus das *Cantigas de Santa María*.

- Transcrición paleográfica íntegra dos catro manuscritos.
- Revisión e unificación de caracteres gráficos das transcricións dos códices.
- Preparación das transcricións para a súa posta en liña.
- Inicio do traballo para a edición crítica dos textos.
- Revisión das fichas bibliográficas que se van incluíndo na base bibliográfica BiRMED - coas referencias relativas ás *Cantigas de Santa María*.

2.2. LITERATURA CONTEMPORÁNEA

ACCIÓN 2.2.1

• PROXECTO DICIONARIOS DE LITERATURA

Consiste na elaboración dun gran dicionario xeral de termos literarios, que inclúe voces referidas a Retórica, Métrica, Semiótica, Poética, Narratoloxía, Ciberliteratura, Hermenéutica, etc, do que saíran outros máis especializados de cada unha desas parcelas ou disciplinas do coñecemento.

Fronte á intención inicial de elaborar un gran dicionario xeral de literatura galega, optouse por parcelar máis o campo e comezar a pensar en dicionarios para cada unha das etapas que ten asentadas a periodoloxía da literatura galega: Rexurdimento, Época Nós, Ditadura, Restauración democrática, etc. Despois de rematada a redacción estase traballando na revisión do *Dicionario sistémico-literario do Rexurdimento*. Ao mesmo tempo, elabórase unha base de datos que permitirá a súa consulta a través de Internet.

Estase a preparara edición do 3º volume do *Dicionario de literatura (I-M)*, que xa se pode consultar a través da base de datos DiTerLi. Simultaneamente, se están redactando novas voces que conformarán o 4º volume do dicionario. Tamén se está prestando unha atención detallada aos novos paradigmas teóricos, especialmente os que teñen que ver coa nova realidade tecnolóxica e dixital.

• PROXECTO BIBLIOGRAFÍA DA LITERATURA GALEGA

Este proxecto pretende reunir do xeito máis exhaustivo posible toda a información relacionada coa literatura galega contemporánea, dende 1800 até a actualidade, para a súa posterior publicación. A orientación actual está na dirección de confluír co proxecto *Informes de literatura*. Preténdese continuar cos labores de selección, compilación e revisión da base de datos. De momento, só é accesible para os distintos proxectos do Centro, pero estase a traballar para ampliar a súa dispoñibilidade.

• PROXECTO INFORME ANUAL DE LITERATURA GALEGA

O proxecto recolle nun informe anual, toda a produción literaria en galego (narrativa, poesía, teatro, ensaio, crítica; literatura infantil e xuvenil; literatura de transmisión oral; Literatura Medieval; revistas; premios) que se rexistre en libros, revistas ou prensa (tanto en papel como electrónica). Este informe elabórase dende o ano 1995 e pretende apoiar a investigación, difundir esta información e observar a construción do sistema literario galego dende todos os ámbitos da comunicación literaria.



No traballo do informe recompílanse todas as informacións e reflexións que, atinxindo directa ou indirectamente ao ámbito da Literatura galega, apareceron por unha parte en soporte impreso durante o ano e, dentro deste tipo de soporte, nas súas máis diversas manifestacións: monografías, libros colectivos, revistas, xornais etc.

Ofrécese un comentario descriptivo, non valorativo nin crítico, de cada unha das entradas rexistradas, é dicir, tanto dos libros relacionados coa literatura que viron a luz ao longo do ano como das recensións que sobre eles apareceron nas revistas e nos xornais durante ese mesmo período de tempo.

A partir de agora, o Informe de Literatura editarase exclusivamente como publicación electrónica en liña.

• PROXECTO DE INVESTIGACIÓN EN LITERATURA INFANTIL E XUVENIL

Este proxecto, creado a finais do ano 2012, acolle os traballos que anualmente realizan sobre Literatura Infantil e Xuvenil galega e doutros ámbitos lingüísticos a Rede Temática "As Literaturas Infantís e Xuvenís do Marco Ibérico e Iberoamericano" (LIXMI); a Asociación de Investigadores en Literatura Infantil e X/Juvenil (ELOS), ELOS-GALICIA. Asociación de investigación en Literatura Infantil e Xuvenil e o Grupo de Investigación da Universidade de Santiago de Compostela LITER2I (GI-1819) que conta entre as súas liñas de investigación "Literatura Infantil e Xuvenil. Investigacións literarias, artísticas, interculturais e educativas".

Deste proxecto, naceu no CRPIH a Rede Ibérica de Literatura Infantil e Xuvenil, da que forman parte especialistas de toda a xeografía lingüística da península e que celebran un congreso todos os anos no CRPIH.

Realización de tarefas de difusión relacionadas directamente co ámbito da literatura infantil e xuvenil do marco ibérico e iberoamericano e organización de seminarios internacionais sobre o a mesma.

Participación activa en actividades científicas e publicacións resultantes dos traballos de investigación levados a cabo.

3. ÁREA DE ANTROPOLOXÍA

ACCIÓN 3.1

• PROXECTO DIMENSIÓNS DA IDENTIDADE COLECTIVA DE GALICIA

Este proxecto emprende unha análise dos parámetros sociais da identidade galega e da súa evolución desde a autonomía, mediante técnicas de prospección estatística en documentos da historia recente e a súa comparación con estudos doutras épocas ou doutros países.

O proxecto dimensións da identidade colectiva en Galicia procederá, durante os anos 2018 e 2019, sobre a base da base de datos estatísticos elaborada en anos anteriores, á elaboración do mapa socioelectoral de Galicia que permitirá, concello a concello, a consulta automática de datos electorais das diferentes consultas canda os datos sociodemográficos de máis relevo.



ANEXO II

O custo total do convenio, coas achegas correspondentes de ambas as partes asinantes, distribuirase anualmente do seguinte xeito:

Accións	Custo total 2018	Custo total 2019	Achega Consellería 2018	Achega Consellería 2019	Contribuc. USC 2018	Contribuc. USC 2019
1.1.1.	39.540,00 €	39.540,00 €	36.800,00 €	36.800,00 €	2.740,00 €	2.740,00 €
1.1.2.	73.440,00 €	73.440,00 €	69.902,00 €	69.902,00 €	3.538,00 €	3.538,00 €
1.1.3.	14.340,00 €	14.340,00 €	12.090,00 €	12.090,00 €	2.250,00 €	2.250,00 €
1.2.1.	4.500,00 €	4.500,00 €	2.627,00 €	2.627,00 €	1.873,00 €	1.873,00 €
2.1.1.	14.220,00 €	14.220,00 €	10.508,00 €	10.508,00 €	3.712,00 €	3.712,00 €
2.2.1.	10.200,00 €	10.200,00 €	7.680,00 €	7.680,00 €	2.520,00 €	2.520,00 €
3.1.	4.500,00 €	4.500,00 €	0,00 €	0,00 €	4.500,00 €	4.500,00 €
TOTAL ANUAL	160.740,00 €	160.740,00 €	139.607,00 €	139.607,00 €	21.133,00 €	21.133,00 €
TOTAL CONVENIO BIANUAL	321.480,00 €		279.214,00 €		42.266,00 €	