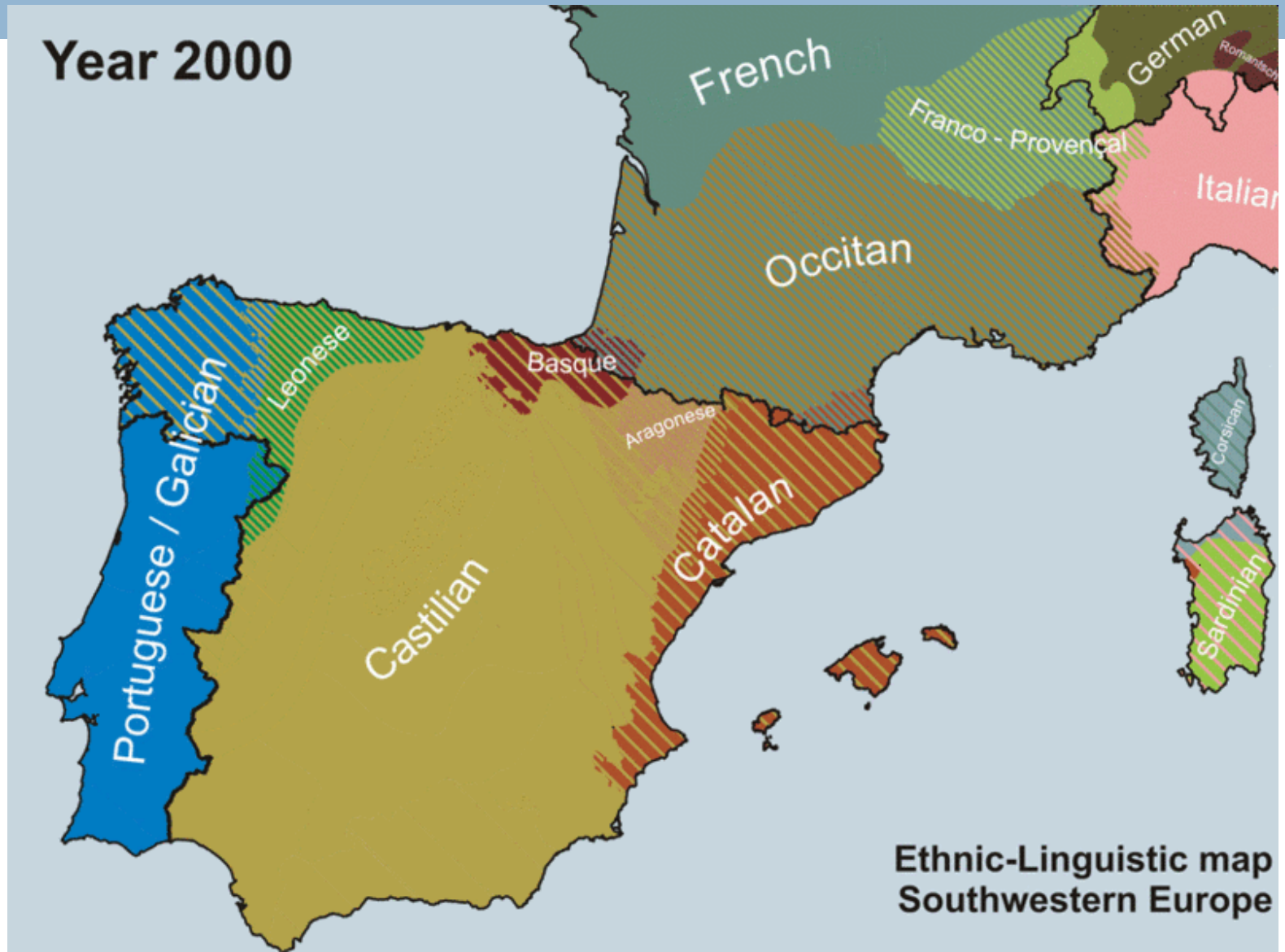


NOVAS TECNOLOXÍAS E LINGUAS MINORIZADAS: A LINGUA GALEGA NO SÉCULO XXI

Xosé Luís Regueira

Instituto da Lingua Galega (USC)

A lingua galega



Lingua minorizada

- Galego: lingua literaria e documental desde o século XIII
- Fin S. XV-inicios S. XVI: desaparece como lingua escrita.
- Imprenta de tipos móbiles: Gutenberg (1455: Biblia).
Primeiros textos en galego: 1612 (Mariño 1998).
- Primeiro libro completo: 1861 (Louis Lucien Bonaparte, *El Evangelio de S. Mateo traducido al dialecto gallego: agás a introducción*)
- 1863: Rosalía de Castro: *Cantares gallegos*.
- Máis de 400 anos de atraso.

Esfera pública

- Galego non accede á esfera pública (*Öffentlichkeit*, Habermas 1962) ata as primeiras décadas do S. XX, e só parcialmente. Cando aparece a radio en Galicia, a partir de 1932, faise en español, mais o galego está presente.
- Truncada polo golpe de estado de 1936 e a ditadura de Franco.
- Galego non aparece, por tanto, nos medios (radio, TV) ata os últimos anos da ditadura e na transición (anos 1970). Varias décadas de atraso.

A lingüística galega: inicios

- Rexurdimento literario: *Cantares gallegos* de Rosalía de Castro (1863).
- Primeiro traballo propiamente lingüístico: “Observaciones sobre la pronunciación del dialecto gallego”, do príncipe Louis Lucien Bonaparte (1861).
- 1863: *Diccionario gallego-castellano* de Francisco Javier Rodríguez.
- 1864: *Compendio de gramática gallega-castellana* de Francisco Mirás.
- 1868: *Gramática gallega* de Juan Antonio Saco Arce
- Finalidade: lingua literaria, gramática e diccionario: dignificación da lingua e da patria galega

A lingüística galega: século XX

- *Século XX: Wörter und Sachen: Fritz Krüger, Die Gegenstandskultur Sanabrias und seiner Nachbargebiete (1925)*
- Walter Ebeling: “Die landwirtschaftliche Geräte im Osten der Provinz Lugo (Spanien)” (VKR 1932)
- W. Schroeder: “Die Fischerboote von Finisterre“ (VKR 1937)
- Hans-Karl Schneider: “Studien zum Galizischen des Limiabeckens (Orense - Spanien)“, VKR11 (1938): fonética e morfoloxía.
- Esta liña remata coa fin da II GM: Krüger en Cuyo (Arxentina), Schroeder morre na guerra, Ebeling e Schneider apartados da universidade.

A lingüística galega: século XX

- Anos 40-50: etimoloxía (Harri Meier, Joseph Piel). Algúns estudos de dialectoloxía (Zamora Vicente, Dámaso Alonso).
- 1962: Dámaso Alonso: “debido al increíble atraso de los estudios de lingüística gallega, esta lengua, en su estado moderno, es casi una incógnita”.
- 1966: Carballo Calero: *Gramática elemental del gallego común*.

A lingüística galega na universidade

- Finais dos anos 60: dialectoloxía, recollida de materiais da fala
- 1971: Instituto da Lingua Galega
- 1974-76: enquisas do Atlas Lingüístico Galego (publicado a partir de 1990)
- Traballos académicos e teses de doutoramento sobre a lingua falada en diferentes lugares de Galicia.
- 1981: Estatuto de Autonomía de Galicia: galego lingua oficial (xunto co español). Non hai unha normativa ortográfica, unha gramática nin un dicionario de referencia.

Lingüística galega: anos 80

- Agora si hai lingüistas formados. O ILG constitúe un centro de referencia na documentación e na investigación lingüística.
- A RAG acode ao ILG para a elaboración das *Normas ortográficas e morfolóxicas do idioma galego* (1982).
- Foco na estandarización e na normalización (corpus planning). Gramática (1986), dicionario (1986).
- Diversificación: sociolingüística, sintaxe, lexicografía, historia da lingua, gramática histórica...

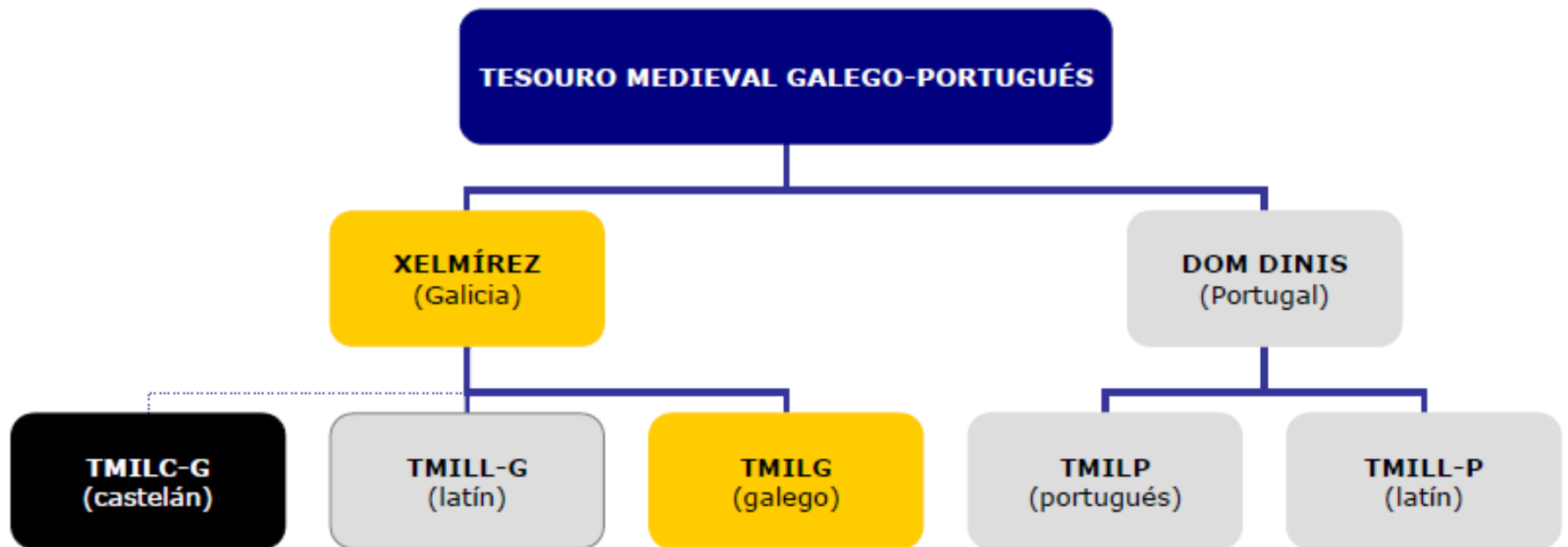
○ impacto da informática

- A partir de 1986, desenvolvemento da informática. Accesibilidade (PC): rápida informatización do traballo en lingüística.
- Programas de concordancias (OCP) e bases de datos.
- Primeiras bases de datos: corpora lexicográficos.

Tesouro Informatizado da Lingua Galega (TILG)

- Base de datos MySQL, linguaxe PHP
- Cerca de 2000 obras. Lingua escrita. Predominio de lingua literaria. Todos os clásicos. Selectiva a partir de 1980 aprox.
- Versión en liña: 26250981 rexistros.
- Obxectivos: diccionario, historia do galego escrito, historia das palabras...
- <http://ilg.usc.es/TILG/>

Tesouro Medieval Galego e Portugués



Tesouro Medieval Informatizado da Lingua Galega (TMILG)

- MySQL, PHP
- Iniciado nos anos 90. Obxectivos: gramática histórica, historia do léxico.
- 78.934.469 rexistros. 94% prosa notarial.
- : <http://ilg.usc.es/tmilg/> (require rexistro gratuíto)

Corpus de Referencia do Galego Actual (CORGA)

- Mostra representativa de xornais (533), revistas (118) e libros (473). Posteriores a 1975.
- 29115073 rexistros
- Está empezada unha versión etiquetada
- <http://corpus.cirp.es/corga/> (require rexistro gratuito)

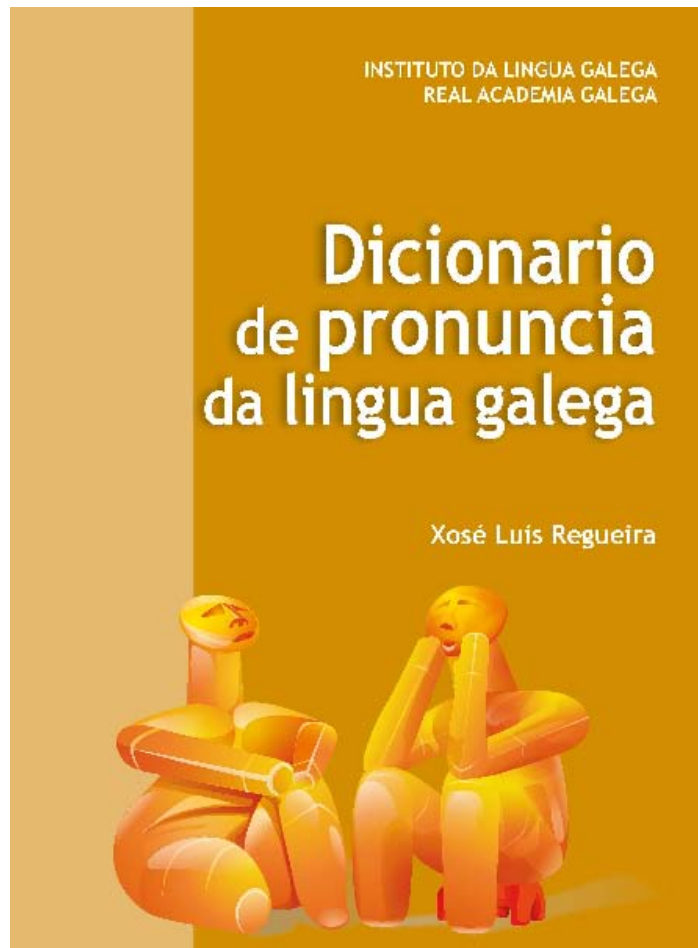
Lingua oral: AGO

- Gravacións desde 1974 ata a actualidade. A maioría dos anos 1990-2000.
- Máis de 2000 horas de gravación, parcialmente transcritas.
- Entrevista dirixida, monólogos, sobre todo. Lingua popular, rural e das xeracións maiores.
- Interese dialectolóxico, etnográfico. Cambio lingüístico.
- Limitacións: lingua urbana, xente nova, conversación. Limitacións para buscas. Non aliñada co audio.
- <http://ilg.usc.es/ago/>

Lingua oral: CORILGA

- Desenvolvido en colaboración co Grupo de Tecnoloxías Multimedia (Univ. de Vigo)
- 100 horas de gravación: transcripción ortográfica (análise da conversa). Obxectivo: 2000 horas.
- Gravacións 1965-presente: conversa, xente nova, urbana, conferencias, radio e tv...
- Obxectivo: corpus anotado multinivel e aliñado co arquivo de audio: transcripción fonética, anotacións prosódica, sintáctica, morfolóxica, léxica.
- Variación e cambio lingüístico. Análise da conversa e do discurso.
- Integrable co Corp-Oral (portugués) e co PRESEGAL (español de Galicia).
- Recurso en desenvolvemento, aínda non público:
<http://firewall.teleco.uvigo.es:56236/corilga/>

Dicionario de pronuncia da lingua galega



- ✓ Obxectivos: fornecer unha pronuncia xeralmente aceptada (aprendizaxe, ensino, medios orais, lingua pública...)
- ✓ Só léxico xeral (VOLG)
- ✓ 60000 formas fonéticas
- ✓ Transcricións :
 - pachorra** *f* pa'tʃora̝
 - pachorrento** *a* pa'tʃo'rento̝
(pa'tʃo'rento̝) | *f*:
pachorrenta pa'tʃo'renta̝
(pa'tʃo'renta̝)

Dicionario de pronuncia da lingua galega

- Léxico xeral + 2500 antropónimos + 4200 topónimos (poboacións, comarcas, parroquias, ríos, montes, illas...).
- Busca fonética
- Contacto: suxestións e preguntas.
- Redes sociais: accesos.
- @PronunciaGal
- Conversas: utilización para aplicacións de voz en programas de software libre.
- <http://ilg.usc.es/pronuncia/>
- Integrado no dicionario da RAG:
- <http://www.realacademiagalega.org/dicionario/>

Tecnoloxías da fala

- A partir de 1989: traballos de fonética acústica
- Tecnoloxías da fala: recoñecemento e síntese de voz
- Equipo de Tecnoloxías Multimedia da ETS de Telecomunicacións (UVigo).
- Transcritor automático de noticias
- Conversor texto-voz: COTOVÍA:
<http://cotovia.gts.uvigo.es/>
- Loquendo:
<http://www.nuance.es/empresas/solucion/soluciones-de-atencion-al-cliente/servicios-y-soluciones/soluciones-de-recepcion-de-llamadas/loquendo-small-business-bundle/interactive-tts-demo/index.htm>

Balance tecnoloxías da lingua

META-NET

White Paper Series Serie de Libros Blancos

THE GALICIAN LANGUAGE
IN THE DIGITAL AGE
O IDIOMA GALEGO NA
ERA DIXITAL

Carmen García Mateo
Montserrat Arza Rodríguez

 Springer

- ✓ “somos moderadamente optimistas sobre o estado actual do apoio ás tecnoloxías da lingua”
- ✓ “ferramentas dispoñibles (...) non son suficientes en calidade e cantidade para desenvolver o tipo de tecnoloxías necesarias que permiten apoiar unha sociedade do coñecemento verdadeiramente plurilingüe.”
- ✓ “o obxectivo de META-NET é introducir tecnoloxías da linguaxe de alta calidade para todas as linguas co fin de acadar unha unidade política e económica a través da diversidade cultural.”

○ impacto de internet

- Non só fai que poidamos compartir recursos, senón que afecta dunha maneira moito máis fonda a nosa maneira de estar no mundo, de entender o mundo.
- A xa vella idea de McLuhan: o medio é a mensaxe.
- Non importa o que transmitamos, os contidos, a información que fagamos circular a través dun medio. O relevante é que o medio mesmo configura a nosa maneira de percibir o mundo.

Información e coñecemento

- Hoxe temos acceso a máis información do que tivemos nunca na historia, a biblioteca mellor fornecida que puidésemos imaxinar ao noso alcance sen erguernos da cadeira.
- Mais o propio medio fainos manexar esa información dunha maneira diferente de como faciamos no papel (Carr 2010). Tamén na investigación universitaria.
- Os buscadores determinan os resultados que imos atopar e consultar: lemos o que Google nos di que é relevante.

Información e coñecemento

- Martin Heidegger advertía dos efectos da tecnoloxización sobre o pensamento: “Es gilt insofern, als die im Atomzeitalter anrollende Revolution der Technik den Menschen auf eine Weise fesseln, behexen, blenden und verblenden könnte, daß eines Tages das rechnende Denken *als der einzige* in Geltung und Übung gelte” (Heidegger 1959: 27).
- Moitas voces (p.e. Byung-Chul Han, aquí en Alemaña) advirten que estamos a substituír coñecemento por información.

Internet como ameaza

- Xavier Gómez Guinovart (2003: 3): “espazo propicio para o neocolonialismo contemporáneo globalista, homoxeneizador e lingüicida”.
- Dentro desta marea, sometida ás forzas máis poderosas ás que enfrontarse nunca, que pode facer a lingua galega neste mundo globalizado?
- Retos inéditos para unha nación con pouco poder, e para unha lingua minorizada, coma o galego.

Internet como oportunidade

- Mais a globalización tamén ofrece oportunidades inéditas para as forzas de liberación (cfr. Hardt & Negri 2001: xv).
- Mellor acceso á información: a exixencia dos dereitos lingüísticos está a ter unha difusión cada vez máis ampla.
- Posibilidades de comunicación: hoxe é moito máis doado e factible conseguir unha “independencia intelectual e cultural” respecto dos centros de poder hexemónicos do que o foi nunca no pasado.
- Os contactos, encontros, e as posibilidades de coñecemento directo doutras situacións lingüísticas e culturais: consciencia do poder e das posibilidades dos falantes.

Internet e linguas

- Primeiros anos: predominio asoballador do inglés. Mais a súa porcentaxe vaise reducindo progresivamente (1995 = 80% > 2009: 45%, Pimienta et al. 2009) e cada vez as linguas da rede son máis diversas.
- Moitas linguas ameazadas encontraron na publicación e na comunicación en liña unha ferramenta moito máis doada, barata e accesible do que podía representar a imprenta ou a radio.
- Internet presenta a contradición de poder acrecentar as desigualdades previas e, asemade, de facilitar os esforzos por eliminar esas desigualdades (Warshawer 2000).

Internet e linguas

- Cada persoa ten a posibilidade de facer a súa mensaxe visible, potencialmente, diante do mundo enteiro.
- Cada persoa pode ser non só lectora, senón tamén autora e editora: unha máquina inmensamente máis poderosa cá imprenta de Gutenberg nas mans de cada cidadán.
- Experiencia do *Dicionario de pronuncia da lingua galega*, consultado en lugares moi afastados de Galicia.

O galego na rede

- Non estamos a aproveitar as posibilidades que esta tecnoloxía nos ofrece tanto como debésemos.
- Mais Xavier Gómez Guinovart (2003) encontrou que en 2002 o galego era a lingua 37^a en número de páxinas da web, daquela cunhas cen mil páxinas.
- O galego está entre as vinte e unha linguas máis usadas en Twitter
(<http://www.webecologyproject.org/2009/09/code-release-google-language-tool/>)

O galego na rede

- A Wikipedia en galego conta xa con 113.000 artigos. A máis ampla e mellor enciclopedia que tivemos nunca.
- Traballo colaborativo, en rede: tradución de Twitter ao galego, entre outros recursos.
- Hoxe programas (Windows, Linux), aplicacións (Facebook, Whatsapp, Galegram)...
- Integración con este tipo de proxectos: Dicionario de pronuncia > aplicacións de voz de software libre.

Punto Gal

- Dominio de primeiro nivel para a lingua e a cultura galegas (o segundo, despois de .cat).
- <http://www.puntogal.org/>
- Primeiras páxinas: 25 de xullo.
- <http://www.dominio.gal/>
- Importancia para a visibilidade da lingua e a cultura galegas no mundo.

Lingua e identidade no S. XXI

- O galego tardou 400 anos en incorporarse á Galaxia Gutenberg, máis de 200 en acceder á esfera pública, e uns 50 anos en entrar na Galaxia Marconi. Mais chegou de maneira bastante normal á nova revolución da comunicación, á internet e á “nube”.
- Momento novo e crucial para demostrar que a lingua galega está ligada ao futuro.
- Globalización: no mundo de hoxe a cuestión non é “que estamos a facer?”, senón “quen somos?” (cfr. Castells 2003).
- Manuel Castells (2003: 84): “en un mundo sometido a la homogeneización cultural por la ideología de la modernización y del poder de los medios de comunicación globales, la lengua, como la expresión directa de la cultura, se convierte en la trinchera de la resistencia cultural, en el último bastión del autocontrol, en el refugio del significado identificable”.

Futuro

Hier ist
Baustelle!





Referencias

- Bonaparte, Louis Lucien (1861), "Observaciones sobre la pronunciación del dia-lecto gallego", in *El Evangelio según San Mateo* (Trad. ó galego de J. Sánchez Santamarina). Londres: L. L. Bonaparte, V-VIII.
- Carr, Nicholas (2010): *The shallows. How the internet is changing the way we think, read and remember*. London: Atlantic Books.
- Castells, Manuel (2003): *La era de la información. Vol. 2: El poder de la identidad*. 2ª ed. Madrid: Alianza.
- Gómez Guinovart, Xavier (2003): "A lingua galega en Internet", en Ana Bringas & Belén Martín (eds.): *Nacionalismo e gobalización: lingua, cultura e identidade*. Vigo: Universidade, 71-88.
- Habermas, Jürgen (1962): *Strukturwandel der Öffentlichkeit. Untersuchungen zu einer Kategorie der bürgerlichen Gesellschaft*. Darmstadt / neuwied: Luchterhand, 1983¹⁴.

Referencias

- Hardt, Michael & Antonio Negri (2001): *Empire*. Cambridge, Ma.: Harvard University Press.
- Heidegger, Martin (1959): *Gelassenheit*. Tübingen: Neske.
- Mariño Paz, Ramón (1998): *Historia da lingua galega*. Santiago de Compostela: Sotelo Blanco.
- Pimienta, Daniel, Daniel Prado & Álvaro Blanco (2009). “Twelve years of measuring linguistic diversity in the Internet: balance and perspectives”. In: *UNESCO publications for the World Summit on the Information Society*. <http://unesdoc.unesco.org/images/0018/001870/187016e.pdf>
- Warshauer, Mark (2000): “Language, identity, and the Internet”, en B. Kolko, L. Nakamura & G. Rodman (eds.): *Race in cyberspace*. New York: Routledge, 151-170 [www.gse.uci.edu/person/warschauer_m/lang.html].



Vielen Dank!