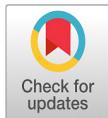


Activismo lingüístico digital para lenguas minoritarias: el caso del catalán



Sergi Álvarez-Vidal

sergi.alvarez@uab.cat

<https://orcid.org/0000-0002-6355-4559>

Universitat Autònoma de Barcelona, España

Resumen

El auge de Internet y la evolución de las tecnologías del lenguaje, en particular los sistemas de traducción automática y los modelos masivos de lenguaje, han transformado de manera radical las condiciones de acceso al conocimiento, la producción textual y la circulación global de información. Estas innovaciones han facilitado la comunicación multilingüe, pero también han reforzado dinámicas de concentración lingüística que pueden afectar negativamente la visibilidad, el prestigio y la funcionalidad de las lenguas minoritarias, generando lo que se ha denominado “brecha lingüística digital”. Este fenómeno pone en riesgo la sostenibilidad de dichas lenguas en el ecosistema digital, al limitar su presencia en plataformas tecnológicas y restringir su capacidad de adaptación a los nuevos entornos comunicativos. El artículo analiza críticamente el impacto de esta brecha sobre la diversidad lingüística y examina el papel del activismo lingüístico digital como forma de acción colectiva orientada a contrarrestar estas asimetrías. En este marco, se estudia el caso del catalán, a través de dos iniciativas emblemáticas: la Viquipèdia y Softcatalà, una organización sin ánimo de lucro pionera en el desarrollo de tecnologías de la lengua en catalán. Ambas propuestas ilustran cómo el compromiso sostenido de comunidades lingüísticas puede traducirse en la creación de infraestructuras digitales, recursos lingüísticos abiertos y espacios de uso significativo en línea, más allá de las políticas gubernamentales.

Palabras clave: activismo lingüístico, brecha lingüística digital, catalán, lenguas minoritarias, Softcatalà, Viquipèdia

Digital Linguistic Activism for Minority Languages: The Case of Catalan

Abstract

The boom of the Internet and the developments in language technologies, particularly machine translation systems and large language models, have dramatically changed the conditions of access to knowledge, textual production, and the global exchange of information. These developments have made multilingual communication easier but only by reinforcing dynamics of linguistic concentration that can negatively affect the visibility, prestige, and functionality of minority languages, generating what has been called the “digital language divide.” This phenomenon jeopardizes minority language sustainability in the digital ecosystem by limiting their presence on technological platforms and restricting their ability to adapt to new communicative environments. The article critically analyzes the impact of this gap on language diversity and examines the role of digital language activism as a form of collective action aimed at counteracting these asymmetries. Within this framework, the



case of Catalan is studied through two representative efforts: Viquipèdia and Softcatalà, the latter being a non-profit organization pioneer in the development of language technologies in Catalan. Both efforts illustrate how the sustained engagement of linguistic communities may give room to the creation of digital infrastructures, open linguistic resources and spaces for meaningful use online, beyond government policies.

Keywords: language activism, digital linguistic divide, Catalan, minority languages, Softcatalà, Viquipèdia

L'activisme linguistique numérique pour les langues minoritaires : le cas du catalan

Résumé

L'essor de l'Internet et l'évolution des technologies langagières, en particulier les systèmes de traduction automatisée et les grands modèles de langage, ont radicalement transformé les conditions d'accès à la connaissance, la production textuelle et la circulation mondiale de l'information. Ces innovations ont facilité la communication multilingue, mais ont également renforcé les dynamiques de concentration linguistique qui peuvent affecter négativement la visibilité, le prestige et la fonctionnalité des langues minoritaires, générant ce que l'on a appelé la « fracture numérique linguistique ». Ce phénomène met en péril la durabilité de ces langues dans l'écosystème numérique en limitant leur présence sur les plateformes technologiques et en restreignant leur capacité à s'adapter aux nouveaux environnements de communication. Cet article analyse de manière critique l'impact de cette fracture sur la diversité linguistique et examine le rôle de l'activisme linguistique numérique en tant que forme d'action collective visant à contrecarrer ces asymétries. Dans ce cadre, le cas du catalan est étudié à travers deux initiatives emblématiques : Viquipèdia et Softcatalà, une organisation à but non lucratif pionnière dans le développement des technologies linguistiques en catalan. Les deux propositions illustrent comment l'engagement soutenu des communautés linguistiques peut se traduire par la création d'infrastructures numériques, de ressources linguistiques ouvertes et d'espaces pour une utilisation significative en ligne, au-delà des politiques gouvernementales.

Mots clés : activisme linguistique, fracture linguistique numérique, catalan, langues minoritaires, Softcatalà, Viquipèdia

Ativismo lingüístico digital para idiomas minoritários: o caso do catalão

Resumo

O boom da Internet e a evolução das tecnologias de linguagem, em especial os softwares de tradução automática e os grandes modelos de linguagem, transformaram radicalmente as condições de acesso ao conhecimento, a produção textual e a circulação global de informações. Essas inovações facilitaram a comunicação multilíngue, mas também reforçaram a dinâmica da concentração linguística, o que pode afetar negativamente a visibilidade, o prestígio e a funcionalidade dos idiomas minoritários, gerando o que foi chamado de “exclusão linguística digital”. Esse fenômeno põe em risco a sustentabilidade desses idiomas no ecossistema digital, limitando sua presença em plataformas tecnológicas e restringindo sua capacidade de adaptação a novos ambientes comunicativos. Este artigo analisa criticamente o impacto dessa lacuna na diversidade linguística e examina o papel do ativismo linguístico digital como uma forma de ação coletiva destinada a neutralizar essas assimetrias. Nessa estrutura, o caso do catalão é estudado por meio de duas iniciativas emblemáticas: Viquipèdia e Softcatalà, uma organização sem fins lucrativos pioneira no desenvolvimento de tecnologias linguísticas em catalão. Ambas as propostas ilustram como o engajamento sustentado das comunidades linguísticas pode se

traduzir na criação de infraestruturas digitais, recursos linguísticos abertos e espaços para uso significativo on-line, além das políticas governamentais.

Palavras-chave: ativismo linguístico, exclusão linguística digital, catalão, idiomas minoritários, Softcatalà, Viquipédia

Introducción

El auge de Internet y el desarrollo de las tecnologías del lenguaje —especialmente desde la aparición de los traductores automáticos y, más recientemente, de los modelos masivos de lenguaje (*large language models*)— han tenido un impacto considerable en la sociedad. Estas tecnologías no solo modifican las barreras espaciotemporales de acceso a la información, sino que también afectan a las relaciones sociales y a las prácticas comunicativas, aumentando la velocidad de las interacciones y la sensación de cercanía entre hablantes (Hohenstein *et al.*, 2023; Wang *et al.*, 2022).

Si bien estas herramientas prometen facilitar la comunicación multilingüe, la realidad es que se desarrollan fundamentalmente para un número reducido de lenguas dominantes, mientras que muchas otras apenas cuentan con presencia digital. Esta situación plantea riesgos para la diversidad y la justicia sociolíngüísticas, ya que refuerza desigualdades existentes entre lenguas y comunidades (Piller, 2016; Stroud, 2018).

El catalán, como lengua minoritaria y minorizada en el contexto del Estado español, constituye un ejemplo paradigmático de esta tensión: dispone de una amplia comunidad de hablantes y de cierto reconocimiento institucional, pero se enfrenta a procesos estructurales de subordinación que limitan su desarrollo digital. Aunque se han implementado políticas de fomento y protección desde las instituciones europeas y regionales (Arzoz, 2008), su efectividad está condicionada por factores políticos, económicos y sociales.

Ante este escenario, han surgido formas de activismo lingüístico digital que buscan revertir estas desigualdades y promover el uso pleno de las lenguas minorizadas en el entorno digital, movilizando diferentes actores, imaginarios y estrategias (De Korne, 2021). En este artículo analizamos cómo dos iniciativas impulsadas

por la sociedad civil —la Viquipèdia y Softcatalà— ilustran estas formas de activismo en el caso del catalán. A través de estos ejemplos, ofrecemos una reflexión crítica sobre el activismo lingüístico digital, así como sobre su potencial transformador en favor de la justicia social y lingüística.

1. Lenguas minoritarias y activismo lingüístico

Aunque el catalán es hablado por más de nueve millones de personas y tiene estatus cooficial en varias comunidades autónomas, sigue teniendo una situación de subordinación respecto al castellano en muchos ámbitos institucionales, mediáticos y tecnológicos. Esta condición se inscribe en un proceso histórico de minorización lingüística, entendido como la restricción del uso, visibilidad y prestigio de una lengua en favor de otra dominante (Strubell, 1999). Como ocurre con muchas lenguas minoritarias, esta subordinación se acentúa en el entorno digital, dominado por unas pocas lenguas globales.

Más allá de las políticas institucionales, la implicación activa de las comunidades resulta esencial para la revitalización efectiva de las lenguas minoritarias (Wiltshire *et al.*, 2024). En esa medida, el activismo lingüístico, entendido como una forma de acción colectiva, adopta múltiples formas y moviliza una diversidad de actores: desde hablantes y educadores hasta desarrolladores y agentes culturales (De Korne, 2021). Algunos se identifican explícitamente como activistas, mientras que otros contribuyen sin asumir esa etiqueta (Stroud, 2018). En todos los casos, su labor trasciende lo estrictamente lingüístico, al confrontar desigualdades sociales vinculadas a la discriminación lingüística (Piller, 2016).

Nettle y Romaine (2000) diferencian entre los enfoques ascendentes y descendentes. En el primer caso, el activismo se centraría en llevar a cabo actividades de revitalización por las

comunidades de hablantes, mientras que, en el segundo, el foco principal de la actividad sería elaborar micro-políticas lingüísticas. Este último activismo se podría vincular directamente con otros movimientos en defensa del medio ambiente o los derechos humanos, representados por grupos como Greenpeace o Amnistía Internacional.

Spolsky (2009), en cambio, asume que el activismo lingüístico hace parte del entramado de la gestión lingüística. Su ideología apoya el mantenimiento o la revitalización de una lengua amenazada y su labor busca influir en los hablantes de la lengua para que continúen con su uso, además de persuadir a los Gobiernos para que apoyen sus planteamientos. En este contexto, se considera que el activismo lingüístico es un eslabón que permite enlazar las comunidades de hablantes y las políticas gubernamentales, aunque en muchas ocasiones estos dos niveles estén contrapuestos o en tensión.

Por su parte, Florey *et al.* (2009) describen el activismo lingüístico como un elemento necesario en la documentación y la revitalización de una lengua. Los autores defienden que forma parte de la responsabilidad de todos los actores implicados y trazan una clasificación del activismo en función de la relación que establece con el resto de los hablantes de la lengua minorizada y su posicionamiento respecto a la comunidad lingüística dominante.

A su vez, Della Porta y Diani (2020) plantean que el activismo lingüístico puede asumirse como un movimiento social cuando los actores buscan transformar el orden social y político, promoviendo la visibilidad y el uso de lenguas minoritarias en esferas públicas y privadas. La importancia de este tipo de activismo radica en su capacidad para influir en la sociedad, no solo en los aspectos de elección lingüística, sino también al desafiar y modificar las estructuras de poder que perpetúan la marginalización de las lenguas minoritarias.

En Cataluña, el activismo lingüístico ha sido fundamental en la lucha por la autonomía cultural y política, así como por la preservación y la promoción del catalán frente a las políticas centralistas del Estado español. El activismo lingüístico en Cataluña no solo aboga por el uso del catalán en la administración pública y la educación, sino que también se extiende a otros ámbitos de la vida cotidiana, como los medios de comunicación, la cultura popular y las tradiciones locales. Iniciativas como *Voluntariat per la llengua*, que organiza encuentros de personas que quieren aprender y practicar catalán, y les *colles de diables*, grupos culturales que promueven el uso del catalán en contextos informales y festivos, ejemplifican cómo el activismo lingüístico se puede expresar tanto en espacios institucionales como en contextos más informales y comunitarios.

A esto se suma la aparición de nuevas generaciones de creadores de contenido digital en catalán —como *youtubers*, *tiktokers* o *streamers*— que promueven el uso de la lengua desde una lógica de autoafirmación y entretenimiento, y que han logrado conectar con audiencias jóvenes en formatos alejados de los canales tradicionales (Direcció General de Política Lingüística, 2021). Estos espacios funcionan como “zonas seguras”, donde los hablantes nativos y los nuevos hablantes pueden subvertir las normas lingüísticas dominantes, creando nuevos modos de interacción social en catalán (Puigdevall Serralvo *et al.*, 2022), y que suponen a menudo un trampolín hacia nuevas socializaciones y generan nuevos discursos y prácticas lingüísticas.

Estos ejemplos muestran la lucha del activismo lingüístico por revitalizar la lengua, al cuestionar las jerarquías lingüísticas y al abogar por nuevas formas de participación comunitaria (O'Rourke *et al.*, 2015). Estos espacios de socialización incluyen cada vez más los espacios digitales, en los que se reproducen parte de los patrones de interacción con las lenguas, pero se plantean también patrones

sociolingüísticos diferenciados (Gorri Oria, 2016).

2. Brecha lingüística digital

La vitalidad de una lengua no depende únicamente del número de hablantes, sino del uso de esa lengua en el ámbito social, cultural y político, además del ámbito digital (Ferré-Pavia *et al.*, 2018). Como mostró Kornai (2013) hace algo más de diez años, menos del 5 % de las entre 7000 y 8000 lenguas del mundo tienen una representación significativa en Internet, y a pesar de los avances de la última década, la brecha apenas se ha reducido (Joshi *et al.*, 2020). Kornai (2013) también estableció tres signos que resultan buenos indicadores de la muerte digital de una lengua: la pérdida de la función (que una lengua se use para todos los propósitos y en todos los contextos), la pérdida de prestigio (directamente vinculado con la actitud que tienen en especial los jóvenes y la presencia de este lengua en entornos web y redes sociales) y la pérdida de competencia (disminución del dominio lingüístico completo entre sus hablantes).

En el contexto actual de globalización y transformación digital, la presencia de una lengua minoritaria en el entorno digital se ha convertido en un factor clave para su vitalidad, percepción social y supervivencia a mediano y largo plazo. La expansión de las tecnologías de la información y la comunicación ha modificado profundamente el modo en que nos comunicamos, al reducir las barreras espacio-temporales de la interacción humana y favorecer la circulación global de contenidos que nos permite el acceso a millones de documentos desde cualquier lugar en todo momento (Sendra, 2025). Este fenómeno ha impactado de manera negativa en la diversidad lingüística, al reforzar el dominio de lenguas internacionales como el inglés, que se ha consolidado como lengua de prestigio y de acceso privilegiado a los medios de comunicación, la educación, la economía o la investigación (Crystal, 2003;

Khan, 2023). En consecuencia, las lenguas minoritarias se encuentran a menudo en una posición asimétrica dentro del ecosistema mediático digital, lo que las hace especialmente vulnerables a los cambios desfavorables en el paisaje comunicativo (Moring, 2014).

Es lo que se ha denominado “brecha lingüística digital” (Zaugg *et al.*, 2022), distinta de la brecha digital general, ya que no se refiere al acceso a la tecnología, sino a la falta de recursos y servicios digitales, así como a la desigual disponibilidad de contenidos, tecnologías del lenguaje y servicios accesibles en distintas lenguas, lo que ha generado una nueva forma de exclusión comunicativa que incide directamente en el valor percibido de una lengua por parte de sus hablantes, especialmente de las generaciones jóvenes (Acharya, 2017). Esta brecha se expresa en varios niveles: desde la desigualdad en el acceso a la información y los servicios digitales, hasta la exclusión del desarrollo tecnológico o la restricción en el uso cotidiano de las lenguas en entornos digitales (Fung *et al.*, 2025). La consecuencia de esta situación no es solo la invisibilidad digital de dichas lenguas, sino también la generación de dinámicas de desplazamiento lingüístico, reforzadas por el atractivo funcional de las lenguas hegemónicas.

A ello se suma el hecho de que el desarrollo de las tecnologías lingüísticas requiere grandes volúmenes de datos de entrenamiento —corpus paralelos, grabaciones transcritas, recursos terminológicos estructurados—, los cuales suelen estar disponibles solo para lenguas con fuerte presencia institucional o mercado asociado (Moshagen *et al.*, 2024). Esta situación se ve exacerbada con las tecnologías lingüísticas basadas en inteligencia artificial, especialmente los modelos masivos de lenguaje. Estos modelos han revolucionado la generación automática de texto, pero su funcionamiento depende de enormes cantidades de datos para entrenar redes neuronales profundas (Melero Nogués, 2018). En este contexto, la competencia entre modelos no se basa tanto en los algoritmos,

que son en gran parte públicos o conocidos, sino en el acceso a grandes volúmenes de datos de entrenamiento, su configuración y la potencia de los recursos computacionales (Bender y Hanna, 2024; Bender *et al.*, 2021). Así, se perpetúa un círculo vicioso: las lenguas con menos representación digital generan menos datos, lo que impide el desarrollo de tecnologías adaptadas a ellas, y esto, a su vez, limita aún más su uso en la red.

Así, el problema para las lenguas minoritarias es estructural: no es que las grandes empresas tecnológicas no quieran usar sus datos, sino que muchas veces estos datos simplemente no existen en cantidades suficientes. La escasez de corpus escritos, recursos digitales y materiales lingüísticos impide entrenar modelos robustos (Moshagen *et al.*, 2024). Y a diferencia de las lenguas mayoritarias, esta escasez no se resolverá de manera espontánea, ya que la mayoría de estas lenguas no tiene suficientes hablantes y carecen de infraestructuras digitales de producción de contenido.

El caso de Internet resulta paradigmático para entender el impacto de las lenguas mayoritarias en el ámbito digital. En sus inicios, hacia el año 1998, el 80 % de los sitios web estaba en inglés. Aunque actualmente otras lenguas han ganado presencia, el inglés sigue siendo la lengua dominante y representa el 49,2 % (W³Techs, s. f.) del contenido. Hay un conjunto de diez lenguas —formadas por el chino, el español, el árabe, el portugués, el malayo, el japonés, el ruso, el francés y el alemán— que, sumadas, concentran el 80 % de todos los contenidos en la red (Meleiro Nogués, 2018). Así, aunque Internet parezca cada vez más multilingüe y diversa, sigue habiendo una fuerte concentración lingüística. La lengua utilizada condiciona la experiencia del usuario, el acceso a la información, la calidad de los servicios disponibles y las comunidades con las que puede interactuar. A pesar de la apariencia de creciente diversidad, la infraestructura digital sigue reproduciendo jerarquías lingüísticas y excluyendo a las lenguas

minoritarias, que apenas logran hacerse visibles en el entorno digital (Soria, 2017).

Si nos atenemos al argumento del tamaño de las lenguas minoritarias y cómo afecta a la brecha lingüística digital, se hace evidente la importancia de la dimensión política y social. Por ejemplo, el número de páginas en Wikipedia en kiswahili, una de las principales lenguas africanas, con unos 80 millones de hablantes, es similar al del bretón, una lengua celta en peligro de desaparición hablada en el oeste de Francia por unas 200 000 personas (Helm *et al.*, 2024). La primera, aunque ampliamente hablada, recibe poco apoyo, mientras que la segunda se beneficia de programas de preservación cultural.

Otro ejemplo de impulso de una comunidad minoritaria en el ámbito digital es Common Voice (<https://commonvoice.mozilla.org/ca>),

[...] una base de datos libre de voces en diferentes idiomas impulsada por la Fundación Mozilla para entrenar programas de reconocimiento de voz con una amplia pluralidad de lenguas y voces. De los 124 idiomas registrados en la plataforma, el catalán es el que más horas de contenido tiene, con más de 3700 horas acumuladas, unas 300 más que el inglés. Todo esto ha sido posible gracias a la participación de personas voluntarias y al impulso de entidades públicas y privadas. (Redacció, 2024; traducción propia)

Además, el enfoque basado exclusivamente en datos plantea serios dilemas éticos. Las grandes compañías suelen recolectar información textual de forma masiva y sin consentimiento, y no existe un equilibrio justo entre quienes producen esos datos y quienes los utilizan con fines comerciales (Hummel *et al.*, 2021). Esta lógica se inscribe en la ideología del *dataism*, que reduce el lenguaje a un conjunto de datos descontextualizados y cuantificables, ignorando su dimensión social y política (Erdocia *et al.*, 2024). En el caso de las lenguas minoritarias, aumenta el riesgo de que una gran parte

del contenido digital no sea de buena calidad o contenga errores. Esto no solo compromete la fiabilidad del entorno digital en esa lengua, sino que además obstaculiza procesos fundamentales, como la normalización lingüística, la educación y la producción cultural en lengua propia (Zaugg *et al.*, 2022).

Sin embargo, a pesar de estas dificultades estructurales, numerosos estudios subrayan también el potencial de las tecnologías digitales como herramientas para la revitalización lingüística. El uso de medios digitales —en particular, Internet y las redes sociales— puede reforzar la visibilidad, la modernidad y la funcionalidad de las lenguas minoritarias, con incidencia directa en la percepción que los propios hablantes tienen de ellas y favorecer procesos de apoderamiento individual y colectivo (Belmar, 2020; Crystal, 2003; Cunliffe, 2009; Lema-Blanco y Meda González, 2016).

Esto, a su vez, tiene un impacto claro en la forma de interactuar y usar las lenguas. Por ejemplo, Internet permite superar las limitaciones geográficas tradicionales, lo que facilita que hablantes de lenguas con diferentes estatus puedan conectarse, compartir contenidos y formar comunidades lingüísticas activas en línea, sin necesidad de estar físicamente en el mismo lugar (Moriarty, 2015). Así, la presencia de una lengua minoritaria en el ámbito digital no solo influirá de manera directa en su supervivencia a mediano o largo plazo, sino que se relaciona estrechamente con la visibilidad y la modernidad de esa lengua, y afecta a la percepción que tienen de ella los hablantes, por lo que también se convierte en un factor de apoderamiento (Belmar, 2020).

3. El caso del catalán

Según un informe reciente del Gobierno catalán (Departament de Política Lingüística, 2024), esta lengua cuenta con más de 10 millones de hablantes y es la segunda lengua más conocida en España. Aunque se trata de una

lengua minoritaria, tiene más hablantes que algunas lenguas oficiales de Estados europeos, como el luxemburgoés o el danés. Es una lengua estandarizada, con abundantes recursos gramaticales y léxicos, y se utiliza en la mayoría de los ámbitos.

Sin embargo, su situación actual varía mucho en función del territorio. Se habla en Andorra (donde es la única lengua oficial), en Cataluña, las Islas Baleares, la Comunidad Valenciana, algunas localidades del sur de Francia, la ciudad sarda de Algúer y ciertos pueblos fronterizos de Aragón y Murcia. Aun así, su uso cotidiano depende en gran medida de cada comunidad e incluso de cada localidad. Se trata de una lengua minorizada porque, incluso en los territorios donde es oficial, se encuentra en una posición de desventaja respecto a la lengua mayoritaria (Pla, 2000). En otras regiones, como Algúer y algunas ciudades valencianas, también ha sufrido una fuerte sustitución lingüística a lo largo del siglo xx (Blas-Arroyo, 2007; Cases-noves Ferrer, 2010).

Tras la muerte del presidente Francisco Franco y la consiguiente transición a la democracia en 1975, se produjo una restauración del catalán como lengua oficial en varios territorios, lo que permitió su reintroducción en el sistema educativo, la administración pública y los medios de comunicación. No obstante, aún persisten problemas, como la reducción de hablantes jóvenes (Institut d'Estadística de Catalunya, 2023) y la sustitución lingüística en ciertos ámbitos, especialmente la justicia y la administración, donde el catalán sigue siendo marginal en comparación con el castellano.

En el ámbito digital, el catalán hace parte de las lenguas de trabajo de productos de traducción automática como Google Translate y Bing, y han proliferado las iniciativas que integran la estrategia de inteligencia artificial (por ejemplo, Catalonia.ai), liderada por el Departamento de Políticas Digitales de la Generalitat de Catalunya, que ha aprobado el proyecto

AINA (<https://projecteaina.cat/>), cuyo objetivo es desarrollar aplicaciones tecnológicas en catalán con el apoyo del Barcelona Supercomputing Centre.

En este contexto, el activismo lingüístico digital ha desempeñado un papel crucial en la defensa del catalán, buscando contrarrestar las amenazas que enfrenta la lengua en un mundo globalizado y dominado por lenguas más fuertes, como el inglés y el español. Iniciativas como la edición catalana de la Wikipedia y los proyectos desarrollados por organizaciones como Softcatalà se erigen como ejemplos paradigmáticos de cómo el activismo digital y comunitario ha contribuido a la revitalización y la promoción del catalán, en una lógica comparable a la que analiza Gorri Oria (2016) en su estudio sobre la comunidad catalana de Ubuntu. Estos proyectos no solo han facilitado el acceso a contenidos en catalán, sino que también han contribuido a consolidar una identidad lingüística que se proyecta más allá de las fronteras físicas de los territorios donde se habla.

3.1. Wikipedia: visibilidad digital y activismo distribuido

Wikipedia se ha consolidado como uno de los proyectos colaborativos más influyentes en la historia reciente de internet (Miquel-Ribé y Laniado, 2021). Desde su lanzamiento en 2001, se ha convertido en un espacio digital global, donde lectores y editores voluntarios colaboran en la construcción de un corpus enciclopédico común, accesible y multilingüe. El proyecto se rige por principios como la neutralidad del punto de vista, la verificabilidad y la edición colaborativa, y se apoya en la idea de que el conocimiento es un bien común, producido colectivamente y accesible para toda la humanidad. Su estructura multilingüe, que incluye ediciones separadas para cada lengua, ha creado un entorno perfecto para la promoción de lenguas minoritarias como el catalán. Aunque el inglés es la lengua mayoritaria en la Wikipedia, la existencia de la edición

catalana, la Viquipèdia (<https://ca.wikipedia.org/>), asegura la representación de esta lengua y cultura, así como su preservación, en el espacio digital (Mamadouh, 2020).

Más allá de su formato enciclopédico, la Wikipedia representa una ecología informativa distribuida que permite conectar disciplinas y saberes gracias al hipertexto, superando las clasificaciones tradicionales propias de los manuales impresos. A través de proyectos hermanos como Wikidata, la información se estructura de manera interoperable y escalable, lo que fomenta la actualización sincronizada de contenidos y la reutilización masiva de datos estructurados. Sin embargo, la naturaleza multilingüe del proyecto presenta todavía importantes asimetrías. Aunque existen más de 300 ediciones lingüísticas activas, menos del 30 % de los artículos de la Wikipedia en inglés tienen un equivalente en otras lenguas (Dittus y Graham, 2022).

Por lo tanto, la promesa universal del conocimiento libre no está exenta de tensiones. Aunque Wikipedia aspira a la neutralidad, sus contenidos están inevitablemente marcados por la subjetividad de sus editores y por los desequilibrios estructurales del entorno digital. Las lenguas con menos hablantes, menor presencia institucional o menor implantación tecnológica tienden a estar infrarepresentadas, lo que perpetúa una jerarquía lingüística global también en la esfera digital. Estudios empíricos como el de Pellejero *et al.* (2011) demuestran que la dimensión de cada edición lingüística de Wikipedia guarda una alta correlación con variables sociolingüísticas, económicas y tecnológicas, como el número de hablantes, el producto interno bruto o el acceso a internet. A pesar de ello, algunas lenguas logran desafiar esta lógica mediante un activismo digital sostenido. Este es el caso del catalán.

La Wikipedia en catalán (Viquipèdia) fue la segunda edición lingüística en crearse, apenas un día después de la inglesa, el 16 de marzo de 2001. A pesar de representar una comunidad

lingüística de unos 10 millones de hablantes, esta edición se encuentra entre las más activas del proyecto, tanto en número de artículos como en calidad editorial. Según Hinojo (2020), este fenómeno no puede explicarse únicamente por datos demográficos o económicos, sino que además debe entenderse como el resultado de una comunidad que hace del activismo lingüístico una forma de presencia digital estratégica. En lugar de resignarse a la diglosia digital, los editores catalanohablantes han convertido la Viquipèdia en un instrumento de normalización lingüística y de visibilidad internacional.

La especificidad del caso catalán se explica por el concepto de *comunidad lingüística digital*, entendido como la parte de una comunidad lingüística con acceso efectivo a Internet y capacidad de producción de contenidos en línea en dicha lengua (Pellejero *et al.*, 2011). Pellejero *et al.* (2011) demuestran que si bien el número total de hablantes de una lengua explica parte de la dimensión de la Wikipedia de dicha lengua, la variable con mayor poder predictivo es el número de hablantes con acceso efectivo a la red. Así, aunque el chino supera al inglés en hablantes nativos, la comunidad digital angloparante es mucho mayor, debido a una mayor penetración de Internet. En este marco, el catalán representa una anomalía positiva: con una comunidad digital relativamente pequeña, ha conseguido un volumen de artículos y una calidad editorial que superan con creces lo predicho por los modelos estadísticos.

La Viquipèdia no solo es un repositorio de conocimiento en catalán, sino también un espacio de representación cultural. En ella se documentan aspectos de la cultura, la historia, la ciencia y las instituciones catalanas que difícilmente encontrarían lugar en otras ediciones lingüísticas. Esta estrategia de autosuficiencia enciclopédica refuerza la identidad cultural digital de la comunidad y genera un archivo abierto de saberes situados.

Para reforzar esta dinámica comunitaria, en 2008 se fundó Amical Wikimedia, una organización de voluntariado dedicada a promover el conocimiento libre en catalán y la cultura catalana en todas las lenguas. Esta asociación fue reconocida en 2013 por la Fundación Wikimedia como organización temática, lo que supuso un hito dentro del movimiento wiki, tradicionalmente organizado en torno a Estados. La existencia de Amical demuestra que es posible articular modelos de gobernanza wiki basados en criterios lingüísticos y culturales, en lugar de geopolíticos. Además, su trabajo ha sido reconocido con numerosos premios y ha servido de modelo para otras comunidades minorizadas.

El caso catalán también destaca por la coherencia discursiva y estratégica de su comunidad editora. A diferencia de otras ediciones, donde las contribuciones se producen de manera más dispersa, la Viquipèdia ha desarrollado proyectos temáticos, colaboraciones con instituciones culturales y actividades formativas con una fuerte cohesión interna. Esta estructura, basada en el principio de buena fe y en la autoorganización distribuida, ha permitido que la comunidad se mantenga activa y resiliente ante los cambios en el ecosistema digital.

Desde una perspectiva global, el éxito de la Wikipedia en catalán desafía la narrativa dominante que asocia visibilidad digital con hegemonía lingüística. Lejos de ser una excepción anecdótica, la Viquipèdia demuestra que la agencia comunitaria puede contrarrestar las lógicas del mercado lingüístico y construir espacios de representación simbólica en condiciones de minoría. Tal como subraya Hinojo (2020, p. 145), editar Wikipedia no es solo un acto cognitivo, sino también político y ético. En este sentido, la Viquipèdia encarna una forma de resistencia activa frente a la concentración del poder digital, un modelo de cómo las lenguas minorizadas pueden ocupar un lugar central en la red a través de la acción colectiva y la inteligencia distribuida.

Ante los discursos que predicen la desaparición de las lenguas minorizadas en la era digital, la experiencia catalana ofrece una alternativa: una comunidad cohesionada, tecnológicamente alfabetizada y éticamente comprometida puede multiplicar su impacto más allá de lo que predicen las métricas tradicionales. No se trata solo de traducir contenidos, sino también de construir un conocimiento situado, enraizado en códigos culturales propios, pero accesible desde una lógica de apertura y colaboración. Este modelo no solo fortalece la lengua catalana, sino que además amplía los márgenes de lo posible para otras lenguas en situación de vulnerabilidad.

3.2. Softcatalà: desarrollo tecnológico comunitario y empoderamiento lingüístico

Otro ejemplo de cómo trabajar por normalizar y ampliar el uso del catalán en el ámbito tecnológico es Softcatalà. Como se indica en el portal web de la asociación <https://www.softcatala.org/que-es-softcatala/>), esta asociación sin ánimo de lucro, fundada en 1998, trabaja para normalizar el uso del catalán en el ámbito de la informática, Internet y las nuevas tecnologías. Lo hace principalmente a través de la traducción de *software* libre, el desarrollo de herramientas lingüísticas y la creación de recursos abiertos y accesibles para toda la comunidad. Uno de sus pilares es la participación en proyectos de traducción de *software* libre al catalán.

La asociación se creó porque sus miembros detectaban la ausencia del catalán en el ámbito de las tecnologías de la información y la comunicación, que carecía de *software* y de recursos tecnológicos disponibles en esta lengua. Un grupo de voluntarios de Softcatalà comenzó traduciendo el navegador Netscape Navigator para que estuviera localizado al catalán, y se creó un sitio web para dar a conocer este trabajo (Ivars-Ribes y Sánchez-Cartagena, 2011). Con el paso del tiempo, el portal de Softcatalà (www.softcatala.org) ha ido

creciendo e incorporando cientos de aplicaciones y herramientas para facilitar el uso del ordenador a los hablantes de catalán.

Para llevar a cabo su actividad, Softcatalà se guía por una serie de valores fundamentales que sustentan su actividad diaria. La transparencia es un principio esencial en una comunidad basada en el voluntariado y resulta clave para facilitar una comunicación fluida entre sus diferentes miembros. También se promueve la excelencia, poniendo el máximo cuidado en la calidad y corrección de todos los desarrollos y las actividades que se llevan a cabo.

Softcatalà se define como una comunidad abierta a cualquier persona, institución o entidad del ámbito empresarial que desee colaborar, siempre que respete los valores y las normas básicas de convivencia. Además, mantiene una independencia total respecto a empresas, Gobiernos e instituciones, lo que le permite actuar de forma coherente con sus principios. Su labor se dirige a todo el ámbito lingüístico catalán y se basa en la utilización de la variante estándar general de la lengua.

Además de las actividades mencionadas, se ha desarrollado y se sigue actualizando un conjunto de herramientas pensadas para ayudar a los desarrolladores de *software* a traducir sus aplicaciones al catalán. Entre ellas se encuentran un glosario terminológico, una guía de estilo, una memoria de traducción y un corrector ortográfico. Estas herramientas han sido ampliamente utilizadas por traductores autónomos, empresas y otras organizaciones, consolidándose como un referente en el campo de la traducción tecnológica en catalán (Irigoyen *et al.*, 2020).

En su página de memorias se recogen proyectos que suman alrededor de 10 millones de palabras, muchas de ellas incorporadas por recomendación de colaboradores. De este total, 2,6 millones de palabras corresponden a proyectos gestionados directamente por

Softcatalà o realizados por sus miembros. Entre estos destacan GNOME (730 000 palabras), LibreOffice (650 000 palabras), GIMP (320 000 palabras) y Mozilla (215 000 palabras).

Durante el 2024, la actividad de Softcatalà fue intensa. Según los datos disponibles en su portal web (Softcatalà, 2025), las herramientas más destacadas en número de visitas fueron las siguientes: el corrector ortográfico (52 millones), el traductor automático (24 millones) y el diccionario de sinónimos (18 millones). Además, la herramienta de transcripción automática de vídeos y audios, lanzada en 2023, ha crecido hasta superar las 85 000 visitas, gracias a importantes mejoras en su rapidez y funcionalidad.

En términos generales, el total de visitas al portal ha duplicado sus cifras, pasando de 69,2 millones en 2020 a 116 millones en 2024, lo que refleja un incremento notable en la utilización de las herramientas ofrecidas. Este aumento muestra cómo las visitas al sitio web se distribuyen entre las distintas herramientas, subrayando el interés y la adopción por parte de la comunidad.

También se ha mejorado la calidad de la traducción automática neuronal en varios pares lingüísticos (catalán-español, francés, inglés y alemán) y se ha incorporado un nuevo par catalán-euskera. De hecho, algunos de los modelos desarrollados han mostrado mejores resultados de calidad que los modelos entrenados por grandes corporaciones, como Google Translate (Brieva Iglesias, 2022). El corrector, por su parte, ha estrenado una interfaz más cómoda e intuitiva.

En reconocimiento a su labor, la Generalitat de Catalunya otorgó a Softcatalà la Creu de Sant Jordi, uno de los máximos galardones del país, por su contribución a la presencia del catalán en el mundo digital.

Así, vemos cómo la actividad de Softcatalà no es solo un trabajo por dar visibilidad y

presencia digital al catalán, sino que con los años que ha convertido en una forma de articular una comunidad digital comprometida con la lengua y que lucha porque esta lengua minoritaria sea una lengua con presencia propia en las tecnologías lingüísticas.

5. Conclusiones

El análisis de los casos de la Viquipèdia y Softcatalà permite identificar una serie de elementos clave que ayudan a conceptualizar el activismo lingüístico digital como una forma de intervención social organizada, distribuida y tecnológicamente mediada. Estas iniciativas ilustran cómo una comunidad lingüística puede apropiarse críticamente de las tecnologías digitales para construir espacios de uso significativos y simbólicos para una lengua minorizada como el catalán.

Ambos proyectos se sustentan en la producción colaborativa de conocimiento y tecnología al margen de las grandes corporaciones o instituciones estatales. Viquipèdia se organiza en torno a una comunidad horizontal de editores que garantiza la presencia del catalán en un repositorio global de conocimiento, mientras que Softcatalà promueve el acceso igualitario a herramientas lingüísticas y digitales mediante el desarrollo y la localización de software libre.

Los agentes implicados son hablantes organizados, programadores voluntarios, activistas culturales y educadores que, sin jerarquías rígidas, articulan discursos de legitimación y resistencia. Sus ámbitos de actuación van desde la edición de contenidos y el desarrollo de tecnologías de la lengua hasta la alfabetización digital o la innovación educativa. Así, estas iniciativas combinan funciones instrumentales, simbólicas y políticas, generando un impacto que no puede reducirse a métricas cuantitativas, sino que abarca también la consolidación de identidades lingüísticas y la creación de comunidades cohesionadas.

Ahora bien, estas propuestas presentan también limitaciones importantes. Su impacto sobre la revitalización efectiva del catalán —especialmente en lo que se refiere al uso cotidiano o la transmisión intergeneracional— es desigual y depende en gran medida del contexto territorial. Además, su desarrollo se enmarca en una realidad europea con acceso tecnológico generalizado y cierto apoyo institucional, lo que dificulta su extrapolación directa a contextos de mayor precariedad sociotecnológica.

Pese a ello, el caso catalán sí puede ofrecer claves útiles para otras lenguas en situación de vulnerabilidad. La coherencia organizativa, la apropiación crítica de las tecnologías abiertas y la capacidad de movilización social son elementos replicables, al menos parcialmente, en otros escenarios de minorización.

En un contexto global de creciente concentración lingüística y tecnológica, el activismo lingüístico digital aparece como una herramienta valiosa para democratizar el acceso a los recursos comunicativos, garantizar el derecho a existir y expresarse en la propia lengua, y construir formas de soberanía digital desde las lenguas no hegemónicas.

Referencias

- Ahora bien, estas propuestas presentan también limitaciones importantes. Su impacto sobre la revitalización efectiva del catalán —especialmente en lo que se refiere al uso cotidiano o la transmisión intergeneracional— es desigual y depende en gran medida del contexto territorial. Además, su desarrollo se enmarca en una realidad europea con acceso tecnológico generalizado y cierto apoyo institucional, lo que dificulta su extrapolación directa a contextos de mayor precariedad sociotecnológica.

Pese a ello, el caso catalán sí puede ofrecer claves útiles para otras lenguas en situación de vulnerabilidad. La coherencia organizativa, la apropiación crítica de las tecnologías abiertas y la capacidad de movilización social son elementos replicables, al menos parcialmente, en otros escenarios de minorización.

En un contexto global de creciente concentración lingüística y tecnológica, el activismo lingüístico digital aparece como una herramienta valiosa para democratizar el acceso a los recursos comunicativos, garantizar el derecho a existir y expresarse en la propia lengua, y construir formas de soberanía digital desde las lenguas no hegemónicas.

Referencias

Acharya, B. B. (2017). Conceptual evolution of the digital divide: A systematic review of the literature. *World of Media Journal. Journal of Russian Media and Journalism Studies*, 6, 41-74. <https://doi.org/10.20381/ruor-21028>

Arzoz, X. (Ed.). (2008). *Respecting linguistic diversity in the European Union*. John Benjamins. <https://doi.org/10.1075/wlp.2>

Belmar, G. (2020). Digital presence and language revitalization: Attitudes towards and use of minority languages on social media. En L. Scyriha (Ed.), *Comparative studies in bilingualism and bilingual education* (pp. 201-218). Cambridge Scholars Publishers.

Bender, E. M., y Hanna, A. (2024). *The AI con: A fight for our lives, our jobs, and the meaning of life in the age of AI*. Harper Business.

Bender, E., Gebru, T., McMillan-Major, A. y Mitchell, M. (2021). On the dangers of stochastic parrots: Can language models be too big? *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. <https://dl.acm.org/doi/10.1145/3442188.3445922>

Blas-Arroyo, J. L. (2007). Spanish and Catalan in the Balearic Islands. *International Journal of the Sociology of Language*, (184), 79-93. <https://doi.org/10.1515/IJSL.2007.015>

Bofill Caralt, C., e Hinojo Sánchez, À. (2018). El treball col·laboratiu en terminologia: un projecte de l'àmbit musical entre el TERM-CAT i Amical Wikimedia. En M. González González, M. D. Sánchez-Palomino, e I. Veiga Mateos, (Eds.), *Terminoloxía: A necesidade da colaboración* (pp. 229-242). Verlagsgesellschaft. <https://doi.org/10.31819/9783964568090-011>

Briva Iglesias, V. (2022). English-Catalan neural machine translation: State-of-the-art technology, quality, and productivity. *Revista Tradumática*, (20), 149-176. [https://ddd.uab.cat/pub/tradumatica/tradumatica_a2022n20p149.pdf](https://ddd.uab.cat/pub/tradumatica/tradumatica_a2022n20/tradumatica_a2022n20p149.pdf)

Casesnoves Ferrer, R. (2010). Changing linguistic attitudes in Valencia: The effects of language planning measures. *Journal of Sociolinguistics*, 14(4), 477-500. <https://doi.org/10.1111/j.1467-9841.2010.00450.x>

Crystal, D. (2003). *English as a global language* (2.ª ed.). Cambridge University Press. <https://doi.org/10.1017/CBO9780511486999>

Cunliffe, D. (2009). The Welsh language on the internet: Linguistic resistance in the age of the network society. En G. Goggin y M. McLelland (Eds.), *Internationalizing internet studies. Beyond anglophone paradigms* (pp. 96-111). Routledge.

De Korne, H. (2021). *Language activism: Imaginaries and strategies of minority language equality*. De Gruyter Mouton. <https://doi.org/10.15439/9783110700000>

- www.degruyterbrill.com/document/doi/10.1515/9781501511561/html
- Della Porta, D. y Diani, M. (2020). *Social movements: An introduction* (3.^a ed.). Wiley-Blackwell. <https://www.wiley.com/en-ie/Social+Movements%3A+An+Introduction%2C+3rd+Edition-p-9781119167655>
- Departament de Política Lingüística, Generalitat de Catalunya. (2024). *Situació del català, l'occità aranès i la llengua de signes catalana*. https://llengua.gencat.cat/web/.content/temes/politica_linguistica_a_catalunya/2024-informe-situacio-i-legislatura.pdf
- Direcció General de Política Lingüística. (2021). *El català i els youtubers. Anàlisi i reptes de futur*. Generalitat de Catalunya.
- Dittus, M. y Graham, M. (2022). *The language geography of Wikipedia. Internet's Languages report*. <https://internetlanguages.org/en/numbers/wikipedia-language-geography/>
- Erdocia, I., Migge, B., & Schneider, B. (2024). Language is not a data set—Why overcoming ideologies of dataism is more important than ever in the age of AI. *Journal of Sociolinguistics*, 28(5), 20-25. <https://doi.org/10.1111/josl.12680>
- Ferré-Pavia, C., Zabaleta, I., Gutiérrez, A., Fernández-Astobiza, I. y Zamardo, N. (2018). Internet and social media in European minority languages: Analysis of the digitalization process. *International Journal of Communication*, 12, 1065-1086. <https://ijoc.org/index.php/ijoc/article/view/7464/2285>
- Florey, M., Penfield, S. y Tucker, B. (2009). Towards a framework for language activism. *1st International Conference on Language Documentation and Conservation (ICLDC)*. [diapositivas]. <https://scholarspace.manoa.hawaii.edu/server/api/core/bitstreams/f982e082-a390-4998-8fb6-69ae86206655/content>
- Fung, K. K., Lai, C. Y., Hung, S. L., Yu, Y. y He, L. (2025). A systematic review of the digital divide experienced by migrant women. *Journal of International Migration and Integration*, 26, 1177-1204. <https://doi.org/10.1007/s12134-024-01222-0>
- Gorri Oria, K. (2016). La identitat catalana en el món digital. El cas de la comunitat d'Ubuntu en català. *Treballs de Sociolinguística Catalana*, (26), 55-69. <https://raco.cat/index.php/TSC/article/view/316430>
- Heidemann, K. A. (2015). Theorizing school-based language activism in the Basque Country and beyond. En R. DePalma, D. B. Napier, y W. Dze-Ngwa (Eds.), *Revitalizing minority voices: Language issues in the New Millennium* (pp. 71-90). SensePublishers. https://doi.org/10.1007/978-94-6300-187-8_5
- Helm, P., Bella, G., Koch, G. y Giunchiglia, F. (2024). Diversity and language technology: How language modeling bias causes epistemic injustice. *Ethics and Information Technology*, 26, 8. <https://doi.org/10.1007/s10676-023-09742-6>
- Hinojo, À. (2020). Somien els viquipedistes en encyclopédies elèctriques? Present i futur de la Viquipèdia i el rol de la comunitat catalanoparlant. *Revista de Llengua i Dret*, (73), 133-145. <https://doi.org/10.2436/rld.i73.2020.3424>
- Hohenstein, J., Kizilcec, R. F., DiFranzo, D., Aghajari, Z., Mieczkowski, H., Levy, K., Naaman, M., Hancock, J. y Jung, M. F. (2023). Artificial intelligence in communication impacts language and social relationships. *Scientific Reports*, 13(1), e5487. <https://doi.org/10.1038/s41598-023-30938-9>
- Hummel, P., Braun, M., Tretter, M. y Dabrock, P. (2021). Data sovereignty: A review. *Big Data & Society*, 8(1), 2053951720982012. <https://doi.org/10.1177/2053951720982012>
- Institut d'Estadística de Catalunya. (2023). *Enquesta d'usos lingüístics de la població*. <https://www.idescat.cat/pub/?id=eulp&n=7195>
- Irigoyen, M. R., Ribes, X. I., Esteve, P. O., Camacho, J. M., Hernández, J. M., & Cremades, A. V. (2020). Softcatalà: nous reptes per garantir la vitalitat del català a les tecnologies. *Revista de Llengua i Dret*, (73), 146-153. <https://doi.org/10.2436/rld.i73.2020.3396>
- Ivars-Ribes, X. y Sánchez-Cartagena, V. M. (2011). A widely used machine translation service and its migration to a free/open-source

- solution: The case of Softcatalà. En F. Sánchez-Martínez y J. A. Pérez-Ortiz (Eds.), *Proceedings of the Second International Workshop on Free/Open-Source Rule-Based Machine Translation* (pp. 61-68). <https://aclanthology.org/2011.freeopmt-1.10/>
- Joshi, P., Santy, S., Budhiraja, A., Bali, K. y Choudhury, M. (2020). The State and fate of linguistic diversity and inclusion in the NLP world. En D. Jurafsky, J. Chai, N. Schluter y J. Tetreault (Eds.), *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 6282-6293). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2020.acl-main.560>
- Khan, A. (2023). Language and globalization: A critical study on language, culture, and power dynamics. *International Journal of English Learning & Teaching Skills*, 5(3), 3412-3423. <https://www.ijeltsjournal.org/wp-content/uploads/2023/04/LANGUAGE-AND-GLOBALIZATION-A-CRITICAL-STUDY-ON-LANGUAGE-CULTURE-AND-POWER-DYNAMICS-1.pdf>
- Kornai, A. (2013). Digital language death. *PLoS One*, 8(10), e77056. <https://doi.org/10.1371/journal.pone.0077056>
- Lema-Blanco, I. y Meda González, M. (2016). Linguistic diversity and communication rights: The role of community media in the promotion of regional or minority languages in Europe. *Radio, Sound & Society Journal*, 1(1), 26-41. <http://hdl.handle.net/2183/18410>
- Mamadouh, V. (2020). Wikipedia: Mirror, microcosm, and motor of global linguistic diversity. En S. D. Brunn y R. Kehrein (Eds.), *Handbook of the changing world language map* (pp. 3773-3799). Springer International Publishing. https://doi.org/10.1007/978-3-030-02438-3_200
- Melero Nogués, M. (2018). El futur de les llengües en l'era digital: Oportunitats i bretxa lingüística. *Revista de Llengua i Dret*, (70), 152-165. <https://doi.org/10.2436/rld.i70.2018.3201>
- Miquel-Ribé, M. y Laniado, D. (2021). The Wikipedia diversity observatory: Helping communities to bridge content gaps through interactive interfaces. *Journal of Internet Services and Applications*, 12(1), 10. <https://doi.org/10.1186/s13174-021-00141-y>
- Moriarty, M. (2015). Globalizing language policy and planning: An Irish language perspective. Palgrave Macmillan, 9-23. <https://doi.org/10.1057/9781137005618>
- Moring, T. (2014). Media markets and minority languages in the digital age. *Journal on Ethnopolitics and Minority Issues in Europe*, 12(4), 34-53. https://www.ecmi.de/fileadmin/redakteure/publications/JEMIE_Datens%C3%A4tze/jemie-datens%C3%A4tze_2013/Moring.pdf
- Moshagen, S. N., Antonsen, L., Wiechetek, L., y Trosterud, T. (2024). Indigenous language technology in the age of machine learning. *Acta Borealia*, 41(2), 102-116. <https://doi.org/10.1080/08003831.2024.2410124>
- Nettle, D. y Romaine, S. (2000). *Vanishing voices: The extinction of the world's languages*. Oxford University Press. <https://doi.org/10.1093/oso/9780195136241.001.0001>
- O'Rourke, B., Pujolar, J. y Ramallo, F. (2015). New speakers of minority languages: The challenging opportunity – Foreword. *International Journal of the Sociology of Language*, 2015(231), 1-20. <https://doi.org/10.1515/ijsl-2014-0029>
- Pellejero, B., Sorolla, N. y Pich, M. N. (2011). La dimensió de les llengües a la Wikipedia i la seua relació amb els elements socials. *Digital hum*, (13), 37-49. <https://raco.cat/index.php/Digitalhum/article/view/245133/328336>
- Piller, I. (2016). *Linguistic diversity and social justice: An introduction to applied sociolinguistics*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199937240.001.0001>
- Pla, J. T. i. (2000). La Llengua catalana a Andorra: problemes i perspectives. *Treballs de Sociolinguística Catalana*, 14, 165-167. <https://raco.cat/index.php/TSC/article/view/224115>
- Puigdevall Serralvo, M., Pujolar Cos, J. y Colombo, A. (2022). Linguistic safe spaces and stepping stones: Rethinking mudes to Catalan through the lens of space. *Journal of Multilingual and Multicultural Development*, 43(1), 21-31. <https://doi.org/10.1080/01434632.2021.1974463>

- Redacció. (2024, mayo 16). El català, la llengua líder de Common Voice. Metadata. <https://www.metadata.cat/noticia/4516/catala-llengua-lider-common-voice>
- Sendra, M. (2025). Contra la diglòssia digital. Polítiques públiques orientades a promoure la presència de les llengües minoritzades i minoritàries al món digital. Fundació Josep Irla. <https://irla.cat/wp-content/uploads/2025/01/estudi-contra-diglossia-digital-fundacioirla-coppietersfoundation.pdf>
- Softcatalà. (2025, febrero 3). Resum de l'any 2024 a Softcatalà. <https://www.softcatala.org/noticies/resum-de-lany-2024-a-softcatala/>
- Soria, C. (2017). What is digital language diversity and why should we care? En *Linguapax review 2016: Digital media and language revitalisation* (pp. 13-28). Linguapax International.
- Spolsky, B. (2009). *Language management*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511626470>
- Stroud, C. (2018). 1. Linguistic citizenship. En *The Multilingual Citizen* (pp. 17-39). Multilingual Matters. <https://doi.org/10.21832/9781783099665-004>
- Wang, H., Wu, H., He, Z., Huang, L. y Church, K. W. (2022). Progress in machine translation. *Engineering*, 18, 143-153. <https://doi.org/10.1016/j.eng.2021.03.023>
- Wiltshire, B., Bird, S. y Hardwick, R. (2024). Understanding how language revitalisation works: A realist synthesis. *Journal of Multilingual and Multicultural Development*, 45(9), 3946-3962. <https://doi.org/10.1080/01434632.2022.2134877>
- W3Techs. (s. f.). *Usage statistics of content languages for websites*. https://w3techs.com/technologies/overview/content_language
- Zaugg, I., Hossain, A. y Molloy, B. (2022). Digitally-disadvantaged languages. *Internet Policy Review*, 11(2), 1-11. <https://doi.org/10.14763/2022.2.1654>

Cómo citar este artículo: Álvarez-Vidal, S. (2025). Activismo lingüístico digital para lenguas minoritarias: el caso del catalán. *Mutatis Mutandis. Revista Latinoamericana de Traducción*, 18(2), 558-573. <https://doi.org/10.17533/udea.mut.v18n2a14>