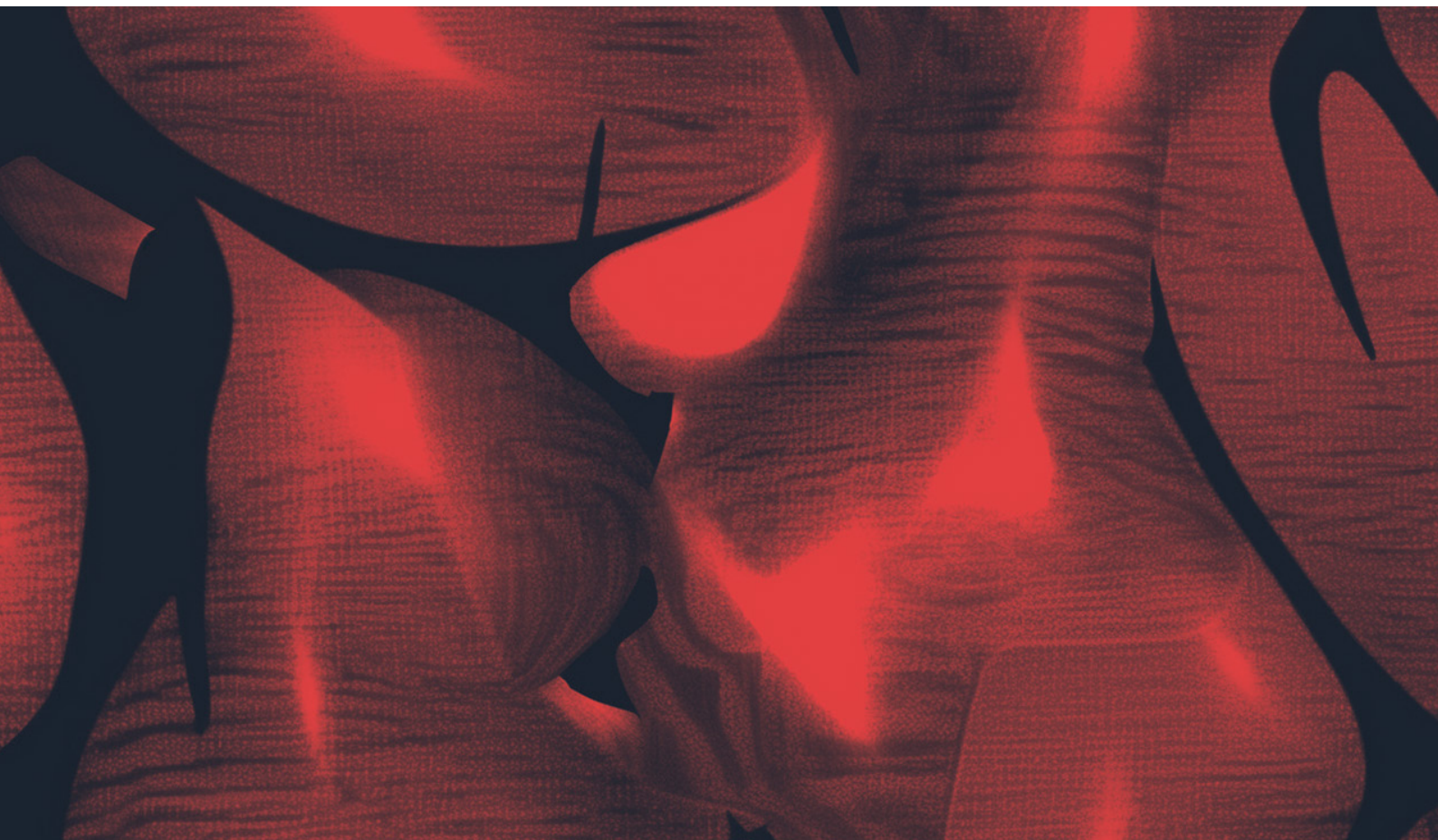


Informe 2025

EL ESPAÑOL: LENGUA PARA EL MUNDO



© 2025, INSTITUTO CERVANTES

Libreros, 23. 28801 Alcalá de Henares (Madrid)

Alcalá, 49. 28014 Madrid

E-mail: informa@cervantes.es

<https://www.cervantes.org>

CONSEJO EDITORIAL:

Luis García Montero, director del Instituto Cervantes

Carmen Noguero Galilea, secretaria general del Instituto Cervantes

Álvaro García Santa-Cecilia, director académico del Instituto Cervantes

María Jesús García González, directora del Gabinete de Dirección del Instituto Cervantes

Antonio Lázaro Gozalo, director del Gabinete Técnico de Secretaría General del Instituto Cervantes

Francisco Moreno Fernández, Director, Observatorio Global del Español

EL ESPAÑOL EN EL MUNDO. ANUARIO DEL INSTITUTO CERVANTES 2025:

DIRECCIÓN: Álvaro García Santa-Cecilia

COORDINADORA EDITORIAL: Rebeca Gutiérrez Rivilla

REVISIÓN Y CORRECCIÓN: Germán Hita Barrenechea, Miguel Ángel Carrón Sánchez,
Almudena Labrador Bravo y María Dolores Pérez Vidal

EL ESPAÑOL, LENGUA PARA EL MUNDO 2025:

AUTOR: Observatorio Global del Español

EDITOR: Francisco Moreno Fernández

ADAPTACIÓN DE MAQUETA: Departamento de Cultura Digital

INSTITUTO CERVANTES

NIPO (PDF): 110-25-051-3

Diseño de cubierta y maquetación: Departamento de Cultura Digital

Queda prohibida, salvo excepción prevista en la ley, cualquier forma de reproducción, distribución, comunicación pública y transformación de esta obra sin contar con la autorización de los titulares de la propiedad intelectual. La infracción de los derechos mencionados puede ser constitutiva de delito contra la propiedad intelectual (arts. 270 y ss. del Código Penal).

Informe 2025

EL ESPAÑOL: LENGUA PARA EL MUNDO

Sumario

EL ESPAÑOL: LENGUA PARA EL MUNDO 2025	
Observatorio Global del Español	6
I. Demografía del español en el mundo 2025	
Francisco Moreno Fernández y Héctor Álvarez Mella	9
II. Análisis complementarios	
El valor económico del español ante un nuevo escenario internacional	
Juan Carlos Jiménez Jiménez	87
Inteligencia artificial y lengua española: panorama actual y perspectivas	
Elena González-Blanco García	108
Terminología en español y multilingüismo digital	
Elea Giménez Toledo	129
La orientación del voto hispano en las elecciones presidenciales de Estados Unidos en 2024	
Francisco javier Pueyo Mena	149
III. Conclusiones finales y recomendaciones	178

EL ESPAÑOL: LENGUA PARA EL MUNDO

+635

millones de personas (635.743.644) son usuarios potenciales de español en el mundo.



En los países hispanohablantes residen cerca de

460

millones de hablantes con dominio nativo de español y 25 millones con competencia limitada.

La población hispanohablante nativa estimada alcanzará su máximo entre

2050 y 2070

El grupo de hablantes con competencia limitada ha crecido desde 2012

↑ 79 %

Esta población podría superar los **200 millones a finales del siglo XXI**, en su estimación más conservadora.



78 %

de los aprendices de español se concentran en **Estados Unidos, la Unión Europea y Brasil.**

El español es la

3.^a

lengua materna del mundo, tras el chino mandarín y el hindi y la **2.^a** entre las lenguas oficiales de la ONU.

La comunidad con dominio nativo del español supera por primera vez

500 millones

(+22 % desde 2012)

El espacio hispanohablante está conociendo un intenso **proceso de urbanización** que alcanzará al



86 %

de la población en 2050.

Uno de cada diez hablantes **nativos de español reside en países no hispanohablantes**, dato que apunta a su relevancia como lengua migratoria.



Fuera de los países hispánicos hay **más de 127 millones** de hablantes potenciales de español.

Desde 2012, el número de aprendices de lengua española **ha crecido un**

↑ 36 %

gracias a la institucionalización del ELE y a su creciente presencia en espacios educativos y sociales.

EL ESPAÑOL: LENGUA PARA EL MUNDO 2025

OBSERVATORIO GLOBAL DEL ESPAÑOL

20 claves del español en el mundo

1. Los hablantes potenciales de español sobrepasan los 630 millones en todo el mundo.
2. La comunidad con dominio nativo de la lengua española supera en 2025 por primera vez los 500 millones de hablantes y llega hasta los 520 millones.
3. La comunidad de hablantes de español con dominio nativo ha pasado a ser la tercera mayor del mundo, por detrás de las comunidades de hablantes nativos de chino mandarín y de hindi. La comunidad hablante de hindi es la segunda comunidad nativa más grande del mundo debido al intenso crecimiento demográfico de la India.
4. Los hablantes nativos de español son cerca de 460 millones solamente en los países hispanohablantes, donde los hablantes con habilidades limitadas suman alrededor de 25 millones.
5. Uno de cada diez hablantes nativos de español reside en países no hispanohablantes, dato que apunta a su relevancia como lengua migratoria.
6. La cifra de hablantes potenciales de español fuera de los países hispánicos supera los 120 millones.
7. La Unión Europea en su conjunto es lugar de residencia para más de 45 millones de hispanohablantes, al margen de los residentes en España.

8. El porcentaje total de crecimiento de hablantes con dominio nativo desde 2012 es del 22 %; su evolución demográfica natural se ha ido ralentizando, al tiempo que ha aumentado la emigración hacia y entre los países hispanohablantes.
9. El grupo de hablantes con competencia o habilidades limitadas ha crecido desde 2012 un 79 %, impulsado por la enseñanza de español en los sistemas educativos de varios países de Europa, Estados Unidos y Brasil.
10. El porcentaje total de crecimiento de los aprendices de lengua española desde 2012 es del 36 %, debido a la institucionalización del español como lengua extranjera y a su creciente presencia en espacios educativos y sociales.
11. El número de aprendices de español entre 2024 y 2025 ha crecido un 1,5 %. El ritmo de crecimiento mantenido podría conducir a una cifra de 100 millones de estudiantes antes de que termine el siglo, si el nivel de institucionalización de la enseñanza es el adecuado.
12. Los territorios que aportan mayores cantidades al conjunto de aprendices de español son Estados Unidos, Brasil y varios países de la Unión Europea.
13. El espacio hispanohablante está conociendo un intenso proceso de urbanización, que alcanzará al 86 % de la población en 2050. Este hecho va en detrimento del uso de las lenguas minoritarias, en la misma medida que beneficia la adquisición del español en grupos de hablantes a los que les es ajeno.
14. La población hispanohablante nativa estimada alcanzará su máximo en las décadas de los cincuenta y sesenta del siglo xxi.
15. La población con competencia o habilidades limitadas en español podría superar los 200 millones a finales del siglo xxi, en su estimación más conservadora: uno de cada tres podría ser hablante de español con dominio limitado (segunda lengua o extranjera) en países no hispanohablantes.

16. La combinación de las lenguas españolas e inglesa es la más prometedora con vistas a la comunicación internacional, tanto por el peso conjunto de sus comunidades nativas (alrededor de 1.000 millones de hablantes), como por el número de países en los que una u otra es oficial (más de 75), como por el prestigio de las culturas y los conocimientos a los que sirven de expresión.
17. En conjunto, los países hispanohablantes son menos desiguales entre sí en cuanto a los valores de desarrollo humano (IDH) que el conjunto de los países hablantes de otras lenguas, como el inglés o el francés.
18. La declaración del inglés como el idioma oficial de Estados Unidos no tiene por qué afectar a la cotidianeidad de la vida colectiva e individual estadounidense, pero sí puede ir en detrimento de derechos sociales y lingüísticos alcanzados, heredados o adquiridos por la población hispana.
19. El porcentaje de páginas web con contenido en español es el segundo más alto del mundo, si bien a gran distancia del número de páginas en inglés.
20. El análisis del lenguaje utilizado en los comentarios de usuarios en redes sociales revela una tendencia al acortamiento de los enunciados, a la vez que a la creación léxica.

I. DEMOGRAFÍA DEL ESPAÑOL EN EL MUNDO 2025¹

1. Introducción

Este informe analiza con detalle la configuración en 2025 de la comunidad de hablantes potenciales de español; es decir, del conjunto de personas que, en distintas condiciones y circunstancias, son capaces de mantener interacciones en español. De hecho, es posible identificar hablantes de lengua española prácticamente en todo el mundo, si bien con diferentes perfiles y en muy distinta proporción.

En el momento actual, la demografía de la lengua española revela dos datos cuantitativamente significativos: la consolidación de una comunidad de hablantes potenciales de más de 630 millones y la superación de los 500 millones de hablantes con dominio nativo de la lengua por primera vez en su historia. Estos datos se obtienen gracias a la disponibilidad de unas fuentes y a la aplicación de una metodología de demografía de las lenguas que están sujetas a continua revisión y refinamiento, pero que hoy por hoy son las que mayor fiabilidad ofrecen.

La demografía es uno de los factores de mayor peso a la hora de describir y analizar cualquier comunidad lingüística. La demografía sustenta la vitalidad de una lengua, cuando es pujante, garantizando la transmisión intergeneracional, al tiempo que revela la naturaleza de los diferentes perfiles de sus hablantes. Por esta razón, si se aspira a conocer la configuración y la evolución sociolingüísticas del español o de cualquier idioma, resulta esencial prestar atención a la demografía de sus agrupaciones de hablantes. El análisis

1. Informe elaborado desde el Observatorio Global del Español por Francisco Moreno Fernández y Héctor Álvarez Mella, con la colaboración de Mariel Elizondo Romo y Jana Weckesser.

demolingüístico de la lengua española, junto al de otros factores cuantitativos y cualitativos, permite obtener una imagen de conjunto de la situación de esta lengua en la actualidad.

Como en el informe correspondiente a 2024 —y los de años anteriores—, aquí se presentan datos de los hablantes que conforman un *grupo de dominio nativo* de español (GDN) en los países donde esos hablantes tienen una mayor presencia. Entre esos países se incluyen tanto aquellos que histórica y sociopolíticamente son considerados como hispanohablantes, más allá de la oficialidad o no de la lengua, como aquellos en los que se identifican contingentes de diferente volumen de hablantes nativos de español.

Asimismo, este informe atiende también a aquellos hablantes que, sin alcanzar un dominio nativo de la lengua, disponen de un conocimiento lo suficientemente amplio como para comunicarse en español con nativos y con no nativos. Este conjunto de hablantes se denomina aquí *grupo de competencia limitada* (GCL) por las limitaciones o restricciones de adecuación sociolingüística o estilística, así como de fluidez, que puedan mostrar en el empleo de la lengua para las interacciones comunicativas. Tales limitaciones resultan difíciles de calibrar por cuanto las habilidades lingüísticas, así como su desarrollo, se desdibujan en la complejidad de las capacidades cognitivas y en la multiplicidad de funciones sociales e interpersonales que la lengua puede cumplir en cada contexto. En este grupo es posible incluir diferentes perfiles de hablantes, desde los hablantes de herencia (descendientes de inmigrantes hispanohablantes) que no han disfrutado de condiciones adecuadas para el mantenimiento de la lengua, hasta los egresados de sistemas de enseñanza que no llegaron a alcanzar niveles muy altos de dominio lingüístico.

Finalmente, este informe atiende a la demolingüística de aquellas personas que están inmersas en procesos de aprendizaje de la lengua institucionalizados o regulados de un modo u otro. El análisis del *grupo de aprendices de lengua española* (GALE) también se enfrenta a obstáculos complicados de salvar porque la enseñanza de la lengua tiene infinidad de variantes en las geografías del español y en los espacios internacionales, sin que existan datos

completos, equiparables y actualizados para cada una de ellas. Las dificultades metodológicas señaladas se han superado mediante el empleo de recursos e instrumentos de investigación demolingüística, así como mediante la toma de decisiones epistemológicas, que quedarán oportunamente explicadas.

En un momento en que la comunidad de hablantes potenciales de español ha sobrepasado los 630 millones, adquiere mayor relevancia, si cabe, la comparación de esta lengua y sus comunidades con otras comunidades de hablantes tanto de *lenguas internacionales*, como de otras lenguas de gran vitalidad. Denominamos lenguas internacionales a aquellas que sirven como vehículo de comunicación entre hablantes de diferentes países en condición de lenguas nativas, lenguas segundas o extranjeras, generalmente en espacios geográficos transcontinentales. Esto significa que una lengua de gran vitalidad no ha de ser necesariamente considerada como internacional. Sobre tales supuestos, este informe propone un análisis de la situación del español comparada con la de otras lenguas del mundo, especialmente con las oficiales en el sistema de la Organización de Naciones Unidas (ONU) (inglés, francés, chino mandarín, ruso y árabe), pero también con otras lenguas de gran vitalidad.

La comparación que proponemos no pretende cubrir todas las dimensiones externas de una lengua: economía, cultura, educación, edición, medios, tecnología... Un objetivo así superaría con mucho los límites de la publicación en que se enmarca este informe, más allá de las dificultades que entraña conseguir una información realmente comparable para un número significativo de lenguas. Por este motivo, nuestro análisis comparativo se ceñirá a algunos aspectos relevantes de las lenguas, entre los que destaca la propia demografía. Los índices de desarrollo de las comunidades lingüísticas, su evolución más reciente y su proyección futura ocuparán asimismo un lugar destacado. A este respecto, contamos con la ventaja de disponer de los informes anuales del Instituto Cervantes, realizados con una metodología común o afín desde hace más de una década, en los que se incluyen algunas informaciones que no se han visto alteradas de forma notable recientemente, bien porque los datos actuales son similares, bien porque no se dispone de datos novedosos.

Como ya se hizo en el informe «El español: lengua para el mundo 2024» (Observatorio Global del Español, 2024), este de 2025 incluye un dossier con varios análisis monográficos dedicados a diferentes aspectos relativos a la situación, el uso y la difusión internacionales de la lengua española.

El primero de los análisis se centra en cuestiones relativas a la economía de la lengua española y en él se analizan las previsiones del español en el PIB mundial, en los flujos de comercio y en el ámbito de la economía digital. El segundo ofrece un diagnóstico sobre el momento actual de la inteligencia artificial aplicada a la lengua española, prestando atención a las estrategias públicas, los recursos, el ecosistema empresarial, los grupos de investigación y el impacto de las tecnologías sobre el español en su diversidad geográfica. El tercero analiza el peso del español en los recursos lingüísticos contenidos en infraestructuras digitales europeas y profundiza en la relevancia y utilidad de las terminologías en español como un tipo de recurso lingüístico clave para el desarrollo de distintas tecnologías del lenguaje. Finalmente, se ofrece un análisis y una interpretación de la configuración del voto hispano o latino en las últimas elecciones presidenciales de Estados Unidos.

Estos informes, elaborados por especialistas en cada una de las materias tratadas, junto al informe demográfico, fruto del trabajo de un equipo de perfil universitario, confirman la voluntad de conferir un carácter colegiado, no personalista, a la investigación, intentando obtener ventaja de la diversidad de saberes y perspectivas de las personas implicadas en la tarea. Para todo ello contamos, además, con el minucioso trabajo de revisión, verificación y edición de la Unidad de Hispanismo del Instituto Cervantes. De este modo, el Observatorio Global del Español pretende cumplir con dos de los fines que su normativa fundacional establece: actuar como órgano permanente de recogida y análisis de información en torno a la situación del español en diferentes ámbitos a escala internacional; y promover estudios e informes técnicos de diagnóstico de la situación del español para fomentar el conocimiento y prestigio de la lengua y las culturas hispánicas.

2. Demolingüística fundamental del español

2.1 El español en los países hispanohablantes

El cuadro 1 presenta las cifras poblacionales y el detalle demolingüístico de la comunidad hispanohablante en los países en que el español es lengua oficial, nacional, vehicular o, en todo caso, mayoritaria. El cuadro incluye datos, actualizados a 2025, de los hablantes de español con dominio nativo y con competencia limitada. La base fundamental de estos recuentos es la información censal sobre la población de cada uno de estos países, junto a diversas informaciones aportadas por sus respectivos institutos de estadística. Para la elaboración del cuadro 1 se han seguido las siguientes pautas:

1. Los datos de población están actualizados a 2025, de acuerdo con la información ofrecida por las oficinas del censo o institutos de estadística de cada país. Cuando no se cuenta con cifras oficiales para 2025, se adopta la proyección realizada por otros organismos oficiales o internacionales. En general, se da prioridad a las fuentes nacionales sobre las internacionales, salvo cuando las fuentes nacionales presentan carencias evidentes.
2. La cantidad de hablantes integrantes del Grupo de Dominio Nativo (GDN) se calcula a partir de la población total, a la que se restan los hablantes potencialmente monolingües en otras lenguas y los hablantes que integran el Grupo de Competencia Limitada. Se parte de la idea de que el dominio nativo de la lengua española no es óbice para que también se tenga un dominio nativo de otra u otras lenguas.
3. La cantidad de hablantes integrantes del Grupo de Competencia Limitada (GCL) se obtiene generalmente como resultado de la suma de los individuos adscritos a pueblos indígenas u originarios que han sido escolarizados en lengua española hasta un nivel primario, excluyendo los monolingües en lenguas indígenas identificados como tales. A esa cantidad se le suma la de extranjeros originarios de países no hispanohablantes censados en cada país, especialmente los no nacionalizados, de los que no existe constancia de que hayan alcanzado un dominio nativo

de la lengua, aunque pueden haberlo hecho. También cabe incluir aquí a los hablantes de otras lenguas nativas en los que se evidencia o a quienes han reconocido un uso limitado de la lengua española.

4. Las cifras, tanto en este cuadro como en las demás referencias cuantitativas, se dan con la exactitud derivada de los cálculos. Esto no significa que se trate de números exactos de hablantes, ya que es imposible conocerlos, por la propia naturaleza del objeto cuantificado. Existe la posibilidad de redondear las cifras de acuerdo con los criterios matemáticos habituales, pero consideramos que los lectores de estos informes pueden realizar tal ejercicio con facilidad, sin prescindir de los datos concretos que emanan de las correspondientes operaciones, que así podrían ser objeto de ulteriores comprobaciones.

En la elaboración del cuadro 1 se han mantenido las proporciones de hablantes del GDN y del GCL calculadas para 2024. En el anterior informe, los porcentajes se calcularon *ex novo* de acuerdo con los datos actualizados aportados por las fuentes demográficas pertinentes. Es bien cierto que tales proporciones no son estáticas, sino dinámicas, ya que están sujetas a las condiciones derivadas de muy distintos procesos sociales o migratorios. Sin embargo, no es previsible que en un año se hayan modificado radicalmente esas proporciones; además, las fuentes primarias no se han actualizado de un modo que obligue a la revisión detallada de tales proporciones. Salvo casos de drásticas modificaciones de las condiciones políticas, legislativas, educativas o migratorias de la población de un país, mantendremos durante un tiempo prudencial estas proporciones de hablantes, a la espera de una consistente actualización de los datos aportados por fuentes primarias.

En los recuentos demolingüísticos de los territorios hispanohablantes, los países para los que se ofrecen datos más inseguros son Cuba, Venezuela y Guinea Ecuatorial. En el caso de los países caribeños, existen fuertes movimientos migratorios, motivados por causas políticas y económicas, con un significativo dinamismo en cuanto al número de exiliados, migrantes y residentes, que dificultan el recuento de hablantes de español. Las fuentes oficiales de estos países, cuando son accesibles, ofrecen datos poco

actualizados o de incierta verosimilitud, como se desprende de su fuerte contraste con fuentes internacionales. En el caso de Guinea Ecuatorial, al problema de las fuentes se le une el de la interpretación dada al concepto de dominio de la lengua española que, en este informe, se dirime, como se ha explicado, manejando el concepto de «dominio de la lengua» equivalente al de un nativo y no el concepto de «lengua materna», siempre delicado y motivo de dispares interpretaciones (Moreno Fernández, 2024).

Junto a las dificultades comentadas, hallamos otras, que hemos sorteado de la forma más adecuada en cada caso: por un lado, los residentes en el extranjero pueden o no estar incluidos en los censos nacionales de los países hispanohablantes, por lo que sería posible incurrir en doble contabilidad, como también puede ocurrir con las segundas generaciones de inmigrantes o la población escolarizada; por otro lado, los hablantes de origen hispánico, al nacionalizarse en un país no hispanohablante, suelen ver desaparecer estadísticamente la información acerca de su origen, lo que impide identificarlos como hispanohablantes nativos, a menos que los censos incluyan información sobre la lengua hablada en el hogar. En cualquiera de estos casos, la metodología aplicada siempre intenta evitar las duplicidades y las malinterpretaciones, muy especialmente cuando se trata de grupos poblacionales de cierta entidad.

CUADRO 1 POBLACIÓN, HABLANTES CON DOMINIO NATIVO Y CON COMPETENCIA LIMITADA EN PAÍSES HISPANOHABLANTES

País	Estimación Población 2025 (Cálculo)	Grupo de Dominio Nativo (GDN)	Hablantes nativos (%)	Grupo de Competencia Limitada (GCL)
México	138.070.271	130.034.675	94,18 %	7.952.848
Colombia	53.110.609	52.090.983	98,08 %	1.019.724
España	49.270.000	47.250.026	95,90 %	1.675.180
Argentina	47.473.760	46.866.195	98,72 %	607.664
Perú	34.412.393	29.601.626	86,02 %	977.312

País	Estimación Población 2025 (Cálculo)	Grupo de Dominio Nativo (GDN)	Hablantes nativos (%)	Grupo de Competencia Limitada (GCL)
Venezuela	29.395.334	28.631.153	97,40 %	764.279
Chile	20.206.953	19.315.922	95,59 %	891.127
Ecuador	18.103.660	16.963.223	93,70 %	568.455
Guatemala	18.079.810	13.722.652	75,90 %	3.283.293
Bolivia	11.425.746	9.197.806	80,50 %	2.079.486
República Dominicana	10.878.267	10.323.570	94,90 %	554.792
Cuba	10.055.968	9.970.591	99,15 %	1.006
Honduras	10.038.054	9.548.292	95,12 %	489.857
Paraguay	7.758.263	5.222.154	67,31 %	2.488.851
Nicaragua	7.218.143	7.001.696	97,00 %	216.544
El Salvador	6.366.086	6.358.547	99,88 %	7.639
Costa Rica	5.567.811	5.510.005	98,96 %	57.905
Panamá	4.544.816	3.925.898	86,38 %	606.278
Uruguay	3.588.937	3.557.812	99,13 %	31.224
Puerto Rico	2.980.532	2.834.581	95,10 %	146.046
Guinea Ecuatorial	1.682.611	1.245.206	74,00 %	437.479
TOTAL	490.228.024	459.172.613	93,67 %	24.856.989

FUENTE: elaboración propia. Los datos en este cuadro 1 se ordenan por el volumen de la población total de cada país.

FUENTES E INFORMACIÓN ADICIONAL DEL CUADRO 1

Argentina

Datos del Instituto Nacional de Estadística y Censos de la República Argentina (INDEC, 2023) proyectados a 2025 (INDEC, 2024). EL GCL incluye hablantes de lenguas indígenas alfabetizados con estudios primarios incompletos, así como extranjeros de origen no hispanohablante.

Bolivia

Datos del Instituto Nacional de Estadística de Bolivia a partir del censo 2024, que corrigen las previsiones que existían para 2025 (INE Bolivia, 2024). La proyección de población para 2025 se consigue aplicando el 1 % de crecimiento poblacional anual que prevé *The World Factbook* (Central Intelligence Agency, s. f.). El GCL incluye hablantes de lenguas indígenas no monolingües y de escolarización baja, así como extranjeros de origen no hispanohablante.

Chile

A partir del Instituto Nacional de Estadísticas de Chile (INE Chile, 2017a). No se incorporan aún datos del Censo de Población y Vivienda de 2024. De acuerdo con informaciones oficiales, la publicación del primer set de indicadores del censo de 2024 se realiza en junio de 2025. La proyección de población para 2025 procede del INE Chile (2024). El GCL incluye hablantes de lenguas indígenas no monolingües y de escolarización baja, así como extranjeros de origen no hispanohablante. Acceso a datos a través de REDATAM (CEPAL, s. f.-b).

Colombia

Datos del Departamento Administrativo Nacional de Estadística de Colombia (DANE, 2025). El porcentaje de población indígena en 2018 (4,4 %) se aplica a 2025. Se estima que no hablan español los indígenas no alfabetizados.

Costa Rica

Información del Instituto Nacional de Estadística y Censos de Costa Rica a partir del censo de 2022 (INEC Costa Rica, s.f.). Proyección del INEC para 2025 (INEC Costa Rica, 2024). El porcentaje de población indígena es del 1,3 %. El analfabetismo entre la población indígena es del 4,8 %, que podría no hablar español.

Cuba

La cifra de población incluye los residentes registrados como tales (Agencia Cubana de Noticias, 2024; Figueredo Reinaldo *et al.*, 2024), si bien pueden tener residencia parcial en otros países (Agencia Cubana de Noticias, 2024). Se estima que los residentes permanentes en Cuba son 8.620.000 personas (Álvarez Quiñones, 2023). Estas cifras vienen a coincidir con las aportadas por la ONU (2024b). Sus proyecciones apuntan la posibilidad de que Cuba vea reducida su población en torno a los 6 millones de habitantes para 2100.

Ecuador

A partir de datos del Instituto Nacional de Estadística y Censos de Ecuador (INEC Ecuador, 2022) y de la proyección de población para 2025 del propio instituto (INEC Ecuador, 2024). En enero de 2025, la población según el INEC ya era superior a los 8 millones. El GCL incluye hablantes de lenguas indígenas no monolingües y de escolarización baja, así como extranjeros de origen no hispanohablante.

El Salvador

Datos del Banco Central de Reserva de El Salvador (BCR, 2022) y de la Oficina Nacional de Estadística y Censos de El Salvador (ONEC, s. f.). El GCL incluye extranjeros de origen no hispanohablante.

España

Datos del Instituto Nacional de Estadística de España (INE España, 2024a, 2024c). Para el GCL, de la población total se calcula la proporción de personas que no hablan nada de español o castellano y las que lo hablan con dificultad a partir de datos de 2021 (INE España, 2021).

Guatemala

Datos procedentes de la Encuesta Nacional de Calidad y Bienestar de los Hogares (ENCABIH) (INE Guatemala 2024). La proyección de población para 2025 es del Instituto Nacional de Estadística de Guatemala (INE Guatemala, 2020; INE Guatemala, Oficina Nacional de Estadística de la República Dominicana y Fundación para el Análisis Estratégico y Desarrollo de la Pequeña y Mediana, 2024). El GCL incluye hablantes de lenguas indígenas no monolingües y de escolarización baja, así como extranjeros de origen no hispanohablante.

Guinea Ecuatorial

Datos a partir del censo de 2015 (República de Guinea Ecuatorial, 2016; Instituto Nacional de Estadística de Guinea Ecuatorial INEGE, 2023). Hay previsión de elaboración de un nuevo censo en 2025. Se ha aplicado la tasa de crecimiento poblacional prevista por el Departamento de Asuntos Económicos y Sociales de la ONU (2022a). El Gobierno de Guinea Ecuatorial proyectaba para 2023 una población de 1.612.677 (INEGE, 2024a). El GCL incluye hablantes de lenguas nativas, criollas y europeas. Como es sabido, el nivel de competencia lingüística es un criterio que provoca discrepancias en cuanto a la estimación de la cantidad y proporción de hablantes de español en Guinea. En este informe se da continuidad al criterio manejado en análisis anteriores (Moreno Fernández y Otero Roth, 2006) con el fin de posibilitar una valoración de la evolución de la comunidad de hablantes, a la espera de una revisión a fondo de esta estimación. Consideramos miembros del Grupo de Dominio Nativo a aquellos que son nativos, que manejan la lengua con un dominio equivalente al de un nativo y que pueden ser considerados como tales en las interacciones con otros hablantes. Este criterio se fundamenta en argumentos ajenos al mero orden de adquisición. En relación con este aspecto, conviene tener en cuenta: 1) que en Guinea Ecuatorial no se maneja otra lengua africana (suajili, wolof, hausa...) como lengua franca nacional (Eberhard, Simons y Fennig, 2025), como sí ocurre en otros países; 2) que el porcentaje de población alfabetizada en 2024, según cifras oficiales, es del 90,1 % (INEGE, 2024b) y que la alfabetización se desarrolla en español.

Honduras

Datos a partir de la Encuesta Permanente de Hogares de Propósitos Múltiples (EPHPM) del Instituto Nacional de Estadística de Honduras (INE Honduras, 2017, 2024a), así como de BASEINE (INE Honduras, 2024b). La tasa de crecimiento demográfico se ha obtenido del Departamento de Asuntos Económicos y Sociales de la ONU (2022a). El GCL incluye hablantes de lenguas indígenas y población negra anglohablante, así como extranjeros de origen no hispanohablante.

México

Datos de población del Instituto Nacional de Estadística y Geografía de México (INEGI, s. f., 2020), del Consejo Nacional de Población de México (CONAPO, 2024) y del Gobierno de México. Proyección para 2025 del CONAPO y la Secretaría de Medio Ambiente y Recursos Naturales de México (CONAPO-SEMARNAT, 2020). El GCL incluye la proporción correspondiente de hablantes indígenas con dominio precario de español y extranjeros de origen no hispanohablante (INEGI, 2022).

Nicaragua

Datos del Instituto Nacional de Información de Desarrollo de Nicaragua (INIDE, s. f., 2025), dado que el último censo es de 2005 (INIDE, 2005). Se ha tenido en cuenta la ficha país de la Oficina de Información Diplomática del Ministerio de Asuntos Exteriores (OID, 2025).

Panamá

Datos procedentes del Instituto Nacional de Estadística y Censo de Panamá (INEC Panamá, s. f.). La proyección a 2025 se ha realizado a partir de informaciones de la Comisión Económica para América Latina y el Caribe (CEPAL, s. f.-a). El GCL incluye hablantes de lenguas indígenas no monolingües y de

escolarización baja, y extranjeros de origen no hispanohablante. Información adicional del Consejo del Sector Privado para la Asistencia Educacional (COSPAE, 2025).

Paraguay

Datos a partir del censo de 2022 (INE Paraguay 2022; 2024a). Proyección de la población nacional para 2025 del Instituto Nacional de Estadística de Paraguay (INE Paraguay, 2024b). El GDN incluye hablantes monolingües en español y bilingües de español y guaraní, con datos obtenidos a través del sistema REDATAM (CEPAL, s. f.-b).

Perú

El último censo peruano es de 2017, lo que ha obligado a realizar proyecciones y estimaciones no siempre coincidentes. Los datos más recientes del Instituto Nacional de Estadística e Informática de Perú (INEI Perú, 2024) reducen la población proyectada en informaciones anteriores. En cualquier caso, las informaciones del INEI y del Ministerio de Salud (MINSA, 2024) no coinciden plenamente. El GDN incluye población indígena con estudios secundarios y universitarios en español. Acceso a datos a través del sistema REDATAM (CEPAL, s. f.-b). El GCL incluye hablantes de lenguas indígenas no monolingües y de escolarización baja, así como extranjeros de origen no hispanohablante.

Puerto Rico

Datos a partir del censo realizado por el Instituto de Estadísticas de Puerto Rico (State Data Center-Puerto Rico, 2023). La proyección de población para 2025 es del mismo Instituto (State Data Center-Puerto Rico, 2025). Para el cálculo del GCL se restan del GDN los hablantes de inglés y de otras lenguas en el hogar.

República Dominicana

Datos de la Oficina Nacional de Estadística de la República Dominicana (ONE, 2023). Con estimación de la propia ONE para 2025 (ONE, 2025). Se estima que la población dominicana crece en torno al 1 % en 2025 (Domínguez, 2024). El GCL incluye extranjeros de origen no hispanohablante, principalmente haitianos.

Uruguay

Proyección del Instituto Nacional de Estadística de Uruguay, a partir del censo de 2023 (INE Uruguay, 2023). El GCL incluye extranjeros de origen no hispanohablante.

Venezuela

Ante la inseguridad de las cifras aportadas por diversas fuentes, mantenemos la misma población anotada en nuestro informe de 2024. Existen discrepancias entre la información aportada por el Instituto Nacional de Estadística de Venezuela, cuyas páginas electrónicas a menudo no funcionan, y otras fuentes, que estiman una población entre 26 y 30 millones. Unas fuentes dan una población de 26,68 millones, a partir de datos del Fondo Monetario Internacional (Statista, 2024b); otras fuentes dan una población de 30,55 millones, supuestamente a partir de datos del Departamento de Asuntos Económicos y Sociales de la ONU (Countrymeters, s. f.). La Plataforma de Coordinación Interagencial para Refugiados y Migrantes de Venezuela da la cifra de 7,89 millones de venezolanos refugiados o migrantes en el mundo (R4V, 2024). Hasta donde sabemos, no hay proyección oficial de población ni previsión de crecimiento para 2025. Las proyecciones de población del Centro Latinoamericano de Demografía (CELADE) son de 1983 y no pueden tenerse en cuenta (CEPAL, 1983). El GCL incluye hablantes de lenguas indígenas no monolingües y de escolarización baja, así como extranjeros de origen no hispanohablante.

Al margen de los casos de Cuba y Venezuela, especialmente de este último país, la comparación de los datos demolingüísticos de 2024 y 2025 nos permite observar algunos hechos que merecen comentario. El primero es que la demografía de los países hispanohablantes ha crecido a un promedio del 1,3 % anual, lo que ratifica o consolida el peso de la propia comunidad nativa en el conjunto de hablantes potenciales de lengua española. Este crecimiento del total poblacional supone un crecimiento proporcional del GDN y del grupo de hablantes con competencia limitada (GCL): el primero de ellos se sitúa al borde de los 460 millones de hablantes nativos y el segundo, en torno a los 25 millones.

La proporción de hablantes nativos, en el conjunto de los países hispanohablantes, se sitúa por encima del 90 %, como ocurría también en años anteriores. Al mantenerse las proporciones de hablantes nativos para 2025, Guatemala, Bolivia, Paraguay y Guinea Ecuatorial siguen mostrando un porcentaje inferior al 85 % dentro de este grupo, los más bajos de todo el espacio hispánico. En cuanto al GCL, continúan destacando las bajas cifras registradas en Cuba, por lo reducido de su población extranjera, y las altas proporciones de México, Guatemala, Bolivia o Paraguay, por la dimensión de sus comunidades indígenas bilingües, así como de España, por su población inmigrante de origen no hispanohablante.

Ahora bien, para comprender mejor el significado de los datos presentados en el cuadro 1, resulta interesante conocer el panorama general de los países donde el español es de uso mayoritario, así como de su entorno inmediato. La región de América Latina y el Caribe tiene en 2025 una población superior a los 600 millones de habitantes, que naturalmente incluye tanto a hispanohablantes, como a lusohablantes y hablantes de otras lenguas (Banco Mundial, s. f.-a, 2024), aunque los primeros sean mayoritarios.

A efectos demolingüísticos, es interesante apreciar cómo el porcentaje anual de crecimiento demográfico se ha ido reduciendo paulatinamente. En el conjunto de América Latina y el Caribe, se ha pasado de una tasa de crecimiento del 1,5 % en el año 2000, a una tasa del 1,1 % en 2012 y del 0,7 %

en la actualidad. En el caso de los países hispanohablantes, esa tasa se sitúa alrededor del 1 %, lo que es un claro indicador de que la comunidad hispánica poco a poco ve debilitado su crecimiento demográfico.

Asimismo, en el futuro también habrían de tenerse en cuenta las consecuencias de las deportaciones desde Estados Unidos. Esas migraciones de retorno forzado pueden ser un factor que influya en las demografías de los países hispanohablantes en los próximos tiempos, al tiempo que podrían tener un efecto potenciador de las migraciones de América Latina hacia Europa, sobre todo hacia España, como sucede cada vez que se endurecen las políticas migratorias de Estados Unidos.

Otros indicadores demográficos revelan una progresiva mejora de las condiciones de vida de esa amplia región americana. Así, la tasa de mortalidad infantil (menores de 5 años) se ha reducido desde un 34,4 % en 2000 a un 15,7 % en 2024, y el porcentaje de población que tiene acceso a internet ha pasado de un 42,2 % en 2012 a un 75,6 % en 2024. No obstante, frente a la mejora revelada por estos datos, se hallan otros indicadores que muestran un retroceso o un estancamiento cuantitativo. Así, el crecimiento del PIB per cápita, en porcentaje anual, se ha reducido de un 2,1 % en 2000 a un 1,4 % en 2024, con una mejora de apenas el 0,2 % desde 2012. Al mismo tiempo, la esperanza de vida al nacer se ha estancado desde 2012 hasta la actualidad, pasando de los 77 años en las mujeres y los 76,7 años en los hombres, a los 76,7 años en las primeras y los 70,5 años en los segundos. En materia educativa, la tasa de finalización de la escuela primaria, en porcentaje, se ha reducido desde 2012 en un 2,3 % —del 96,9 % al 94,6 %— (CEPAL, 2022).

Este sucinto panorama nos muestra una región mundial en la que, en el último cuarto de siglo, se ha ido produciendo una mejora general de sus indicadores de progreso y bienestar, pero que, en los últimos doce años, ha conocido el deterioro de varios ámbitos. Todo ello en entornos políticos muy diversos que alejan a América Latina de la imagen de región homogeneizada que ha podido proyectar en otros periodos.

Finalmente, desde una perspectiva sociolingüística, es apreciable, tanto en el Atlántico occidental como en el oriental, el desarrollo de tendencias de aproximación de las hablas más regionales o locales a estándares lingüísticos nacionales e internacionales. En ello es seguro que están influyendo factores derivados de la facilidad en la movilidad y en los contactos entre poblaciones de diverso origen a través de unos medios de comunicación interregionales y de las redes sociales: los contactos suelen producir igualamiento y búsqueda de usos comunes. Junto a ello, además, debe valorarse el aumento en el porcentaje de la población escolarizada en todo el territorio hispánico y muy especialmente del número de universitarios, que también produce una aproximación a los estándares lingüísticos abiertamente más prestigiosos (Bravo García, 2008; Masuda, 2020).

2.2 El español en los países no hispanohablantes

Resulta muy complicado conocer en cuántos países del mundo viven o residen hispanohablantes, con cualquiera de los perfiles lingüísticos que este informe considera: probablemente puedan existir en todos los países y territorios del mundo. Un simple dato servirá como síntoma de esta extensión: para la prueba de Conocimientos Constitucionales y Socioculturales de España (CCSE), con exámenes administrados por el Instituto Cervantes y destinados a acceder a la ciudadanía española por residencia, se han presentado entre 2024 y 2025 candidatos de 200 países. Sin embargo, la recogida de informaciones actualizadas y fidedignas de hablantes de español en todos los países del mundo se presenta como imposible en unos casos, dificultosa en otros y, en fin, vana o inocua en no pocos. Por este motivo, este informe presta atención de forma individualizada a una treintena de países, fuera del espacio hispánico propiamente dicho, en los que el español tiene una presencia significativa o que podrían ser importantes para la difusión del español en el futuro. Como se ha explicado en informes anteriores, en unos casos se trata de territorios en los que España u otros países hispanohablantes han mantenido contingentes de población hablante de español durante un periodo histórico, como en los actuales Estados Unidos o Filipinas; en otros países o territorios, la presencia

hispanohablante tiene su origen en movimientos migratorios más recientes, como en Canadá, Australia, el Reino Unido, Suiza o la propia Unión Europea; por último, en otros casos el conocimiento de español puede estar motivado por la contigüidad geográfica, por el dinamismo transfronterizo, por la importancia de los intercambios socioeconómicos entre países limítrofes (Belize, Brasil, Andorra...) o por razones históricas o simbólicas de diversa índole y surgidas en diferentes épocas. El resto de los países o territorios, más allá de la treintena anunciada, han sido objeto de una valoración de conjunto con base en las fuentes que mayor solidez ofrecen.

CUADRO 2 POBLACIÓN HISPANOHABLANTE EN PAÍSES Y TERRITORIOS NO HISPÁNICOS

País	Grupo Dominio Nativo (GDN)	Grupo Competencia Limitada (GCL)	Total de hablantes potenciales de español
Estados Unidos	45.481.557	20.000.000	65.481.557
UE-27	11.433.891	34.301.673	45.735.564
Reino Unido	217.974	5.564.106	5.782.080
Brasil	1.111.502	1.931.480	3.042.982
Marruecos	136.892	1.751.733	1.888.625
Canadá	321.407	1.043.892	1.365.299
Australia	185.209	386.573	571.782
Suiza	213.776	351.259	565.035
Filipinas	6.834	547.696	554.530
Israel	79.827	360.150	439.977
Belize	38.908	199.256	238.164
Rusia, Bielorrusia, Kazajistán y Kirguistán	55.540	153.776	209.316

País	Grupo Dominio Nativo (GDN)	Grupo Competencia Limitada (GCL)	Total de hablantes potenciales de español
Antillas Neerlandesas	36.651	140.693	177.344
Argelia	1.348	173.600	176.948
Trinidad y Tobago	48.800	68.417	117.217
Aruba	17.975	78.140	96.115
Japón	74.413	17.004	91.417
Andorra	35.232	38.756	73.988
Guam e Islas Marianas del Norte (Estados Unidos)	-	42.300	42.300
China	19.849	12.510	32.359
Noruega	27.367	4.090	31.457
Nueva Zelanda	27.523	92	27.615
Islandia	2.985	19.554	22.539
Moldavia	10.632	5.203	15.835
Corea del Sur	4.395	10.268	14.663
Turquía	6.858	5.115	11.973
India	7.466	3.446	10.912
Jamaica	8.514	330	8.844
Egipto	938	47	985
Palestina	657	95	752
Otros países	327.725	-	327.725
TOTAL	59.942.645	67.211.254	127.153.899

FUENTE: elaboración propia. Países ordenados según número de hablantes potenciales de español.

FUENTES E INFORMACIÓN ADICIONAL DEL CUADRO 2

Andorra

Para los hablantes de dominio nativo, sobre la población de Andorra para 2025 (88.082) se ha aplicado un porcentaje del 40 %, que corresponde a los hablantes que declaran el español como su lengua materna (Govern d'Andorra, 2022, p. 10). Para hablantes con competencia limitada, sobre la población total de Andorra (Govern d'Andorra, 2022, p. 10), se ha aplicado un porcentaje del 44 %, correspondiente a los hablantes que declaran el catalán como su lengua materna, si bien estos podrían tener igualmente un dominio nativo de español. Téngase en cuenta que la investigación publicada en 2021 por Jiménez Salcedo revelaba que, entre la población andorrana, el castellano es «la lengua que todo el mundo sabe hablar», «lengua franca», «lengua común» o «lengua por defecto» del país (Jiménez Salcedo 2021). La nueva Ley de la lengua propia y oficial en Andorra, aprobada el 25 de abril de 2024 en la sesión ordinaria del Consejo General, determina que, a partir de 2026, se exigirá una formación de 30 horas de catalán para renovar el permiso de residencia y trabajo. La población de Andorra a 31 de diciembre de 2024 ya era de 87.097 personas, de las cuales un 16,7 % era de origen extranjero, porcentaje que se ha duplicado en los últimos años (*La Ciutat*, 2025).

Antillas Neerlandesas (Bonaire, Curazao, San Eustaquio, Saba)

Los hablantes nativos de Curazao son los que figuran en el censo como hablantes de español en el hogar (Central Bureau of Statistics Curaçao, 2023), proyectados a 2025. Los hablantes con competencia limitada son los que el censo da como hablantes de papiamento, proyectados a 2025. Los hablantes de español y de papiamento en los demás territorios se calculan sobre los porcentajes que ofrece Statista (2024a), que se calculan sobre la población total. Los porcentajes son los siguientes: Bonaire (hablantes de español: 76 %; hablantes de papiamento: 88,4 %); San Eustaquio (hablantes de español: 31,5 %; hablantes de papiamento: 31,5%); Saba (hablantes de español: 31,8 %; hablantes de papiamento: 97,6 %). Estas cifras revelan que una buena parte de la población es bilingüe o trilingüe. Los datos generales de población se obtienen de Central Bureau of Statistics Curaçao (s. f.).

Argelia

En los hablantes nativos, incluimos a los españoles residentes en Argelia (INE España, 2024b), así como a los mexicanos residentes (Instituto de los Mexicanos en el Exterior, 2020). Entre los hablantes con competencia limitada, incluimos a los refugiados saharauis en territorio argelino, donde el español es cooficial y tiene una presencia notable en la enseñanza. Se estiman los hablantes de español sobre la población mayor de 50 años: la población total es de unas 173.000 personas y las mayores de 50 años son alrededor de 30.000, a partir de datos de la Agencia de Refugiados de la ONU (2018). Los hablantes de español del Sáhara Occidental se contabilizan dentro de las cifras correspondientes a Marruecos.

Aruba

Datos del Central Bureau of Statistics Aruba (2021, 2023). Se incluyen los residentes con origen en países hispanohablantes. Como hablantes con competencia limitada, se incluyen los conocedores del papiamento, que son aproximadamente un 70 % de la población (Sato Center, 2012).

Australia

A partir del censo de 2021 (Australian Bureau of Statistics, 2021), con proyección a 2025. La cifra de hablantes de español en casa sería de 171.370. Para los hablantes del Grupo de Competencia Limitada, se ha actualizado a 2025 la cantidad estimada para 2024.

Belice

Según el censo de 2022, el 56,55 % de los beliceños habla español (Statistical Institute of Belize, 2022). Se ha realizado proyección al 2025. El 80 % de los inmigrantes nacidos en el extranjero procede de países hispanohablantes según la Organización Internacional para las Migraciones (IOM, 2022). Los hablantes de competencia limitada se obtienen del total de hablantes de español, menos los nacidos en países hispanohablantes. Es posible que muchos de los beliceños que hablan español lo hagan con dominio nativo, pero aquí no se presupone esa realidad.

Brasil

Los hablantes con dominio nativo, procedentes de países hispanohablantes, se recuentan a partir del Registro Nacional Migratorio (inmigrantes internacionales registrados por la Policía Federal) (UNICAMP, 2024). En 2025, el número de venezolanos sobresale claramente sobre las personas procedentes de otros países. Por información de la asociación de escuelas brasileñas Cultura Inglesa y de Babbel, Juliana Pazetti estima que menos de un 1 % de los brasileños hablan español (Nuadda, s. f.). Esa proporción supondría en la actualidad un total de más de 2 millones de hablantes de español con dominio nativo. Los hablantes con competencia limitada se estiman, a partir del registro de fronterizos del Instituto Brasileiro de Geografia e Estatística (IBGE, 2024) y de un estudio de conjunto de las fronteras terrestres de Brasil (Moura, 2024). Dado que la población fronteriza ronda los 10 millones de habitantes, proponemos una estimación conservadora de un grupo de hablantes con unas prácticas lingüísticas que podrían denominarse portuñol de frontera. A ese grupo poblacional de competencia limitada en español habría que unir el de los estudiantes ya egresados del sistema educativo (reglado y no reglado) y el de los hablantes de herencia que hayan podido conservar la lengua y que, a partir de datos del IBGE, podrían estimarse entre 500.000 y 1 millón de personas. Asimismo, podrían incluirse en este grupo los hablantes sefardíes que, sobre todo en las generaciones de mayor edad, han podido mantener el español en su variedad de judeoespañol.

Canadá

A partir de datos de hablantes de español en casa: 1,76 % (Statistics Canada, 2025). Porcentaje proyectado a la población de 2025 (41.206.022). Para los hablantes de competencia limitada, al porcentaje de población de origen hispano (3,3 %), que se ha incrementado notablemente en los últimos años, se le ha restado el porcentaje de hablantes de español en casa. La estimación para Canadá de hablantes de judeoespañol, en diferentes niveles, es de alrededor de 5.000.

China

Los datos de hispanohablantes en China son escasos y de difícil acceso. Contamos con registros para 2020 de peruanos (10.455), basados en datos de la ONU (Datosmacro, s. f.-a), y mexicanos (3.261) (Instituto de los Mexicanos en el Exterior, 2020). Para residentes españoles en China, contamos con los registros del INE (4.675) (INE España 2024b). Estas cifras tienen el inconveniente de no poder proyectarse con un mínimo de fiabilidad y de que se basan en registros consulares, a los que no acuden todos los extranjeros que se instalan en un país. Para el GDN, se incluyen los que han conseguido un certificado DELE de nivel C antes de 2019. Para el GCL, se incluyen los que han conseguido un certificado DELE nivel B antes de 2019.

Corea del Sur

Los datos de hispanohablantes de Corea del Sur proceden del Gobierno coreano. La información publicada por el National Geographic Information Institute de Corea del Sur (NGII, 2021) para los países hispánicos no es exacta, sino que se presenta por rangos. Hemos tomado los valores medios de tales rangos para cada uno de los países. Los residentes extranjeros hispánicos más numerosos proceden

de México, Perú y España. El GDN incluye personas con certificado DELE de nivel C anterior a 2019; el GCL incluye personas con certificado DELE de nivel B anterior a 2019.

Egipto

Los datos de hablantes con dominio nativo corresponden a residentes españoles (INE España, 2024b) y mexicanos (Instituto de los Mexicanos en el Exterior, 2020), a los que se suman los que consiguieron un certificado DELE de nivel C antes de 2019. El GCL corresponde a las personas que consiguieron un certificado DELE de nivel B antes de 2019.

Estados Unidos

La población de Estados Unidos a 1 de enero 2025 es de 341.178.551 (Oficina del Censo de los Estados Unidos, 2025). La proyección a 2025 de la población hispana arroja una proporción del 19,72 %. De los hispanos, el 67,6 % habla español en casa (Oficina del Censo de los Estados Unidos, 2023). No obstante, el Pew Research Center ofrece los resultados de una Encuesta Nacional de Latinos en la que se obtiene que el 75 % de los hispanos adultos dicen ser capaces de mantener una conversación, comprendiendo y hablando muy bien, como mínimo (Mora y López, 2023). En cualquier caso, este porcentaje se ha ido reduciendo paulatinamente a lo largo del siglo ^{XXI}, por la dinámica de las migraciones y por la alta proporción de hispanos ya nacidos en Estados Unidos. Para los hablantes de competencia limitada, se ofrece una estimación que incluye a los hispanos que conservarían habilidades comunicativas por haber adquirido el español como lengua heredada, así como a los que aprendieron español desde los sistemas educativos y conservan el español aprendido, y los que puedan haber conservado el judeoespañol, sobre todo los sefardíes de edad más avanzada. Al ir debilitándose la identificación entre hispano y hablante de español, optamos en esta ocasión por dar una cifra redondeada para el GCL. La población hispana o latina no autorizada o indocumentada, que en la actualidad asciende a unos 11 millones (Passel y Krogstad, 2024), debe estar incluida en el censo. La actual situación política no es favorable para presuponer la estabilidad de esta situación.

Filipinas

Datos actualizados a partir del Censo de Población y Hogares de 2020 y proyectados a 2025 (Philippine Statistics Authority, 2024). Para los hablantes con dominio nativo en Filipinas, se han tenido en cuenta las cifras de españoles (INE España, 2024b) y mexicanos residentes en el extranjero (Instituto de los Mexicanos en el Exterior, 2020), a los que se añaden los residentes procedentes de Argentina, Bolivia, Chile, Colombia, Costa Rica, Ecuador, Perú y Puerto Rico (ONU, 2020). Se incluyen los componentes de los 127 hogares en los que se habla español, según el censo de 2020 (Badillo Matos, 2024): 4,1 miembros por hogar. No se incluyen, en cambio, los residentes de origen español (4.952 en 2020), ya que pueden coincidir con la población española residente en Filipinas, ya contabilizada. Para el Grupo de Competencia Limitada, se han tenido en cuenta las personas registradas en el censo de 2020 como chabacanos caviteños, cotabateños, davaeños y zamboangueños (Philippine Statistics Authority, 2023), con proyección a 2025. Los hablantes de chabacano procedentes de esas etnias se incluyen en el grupo de competencia limitada por sus previsibles limitaciones para la comunicación con hispanohablantes de otras procedencias, no porque sus variedades muestren carencias lingüísticas internas de ningún tipo. No se incluyen en este grupo los 1.730.000 hablantes egresados de la enseñanza superior antes de 1986 por no tener constancia de la entidad cuantitativa de este grupo después de 40 años (Madrid Álvarez-Piñer, 2018).

Guam e Islas Marianas del Norte

La población hispana en Guam es de 4.522 personas (Oficina del Censo de los Estados Unidos, 2020). No se incluyen en el recuento por estar incluidos como latinos en el total de Estados Unidos. El grupo

de hablantes con competencia limitada está integrado principalmente por hablantes de chamorro. Se estima que un 17,8 % de la población puede hablar chamorro (The Bureau of Statistics and Plans, 2020). Los datos proceden de *Ethnologue* (Eberhard, Simons y Fening, 2025). Puede ponerse en duda la inclusión de los hablantes de chamorro dentro del grupo de competencia limitada en español, pero optamos por hacerlo debido a la fuerte impronta del español de esta lengua criolla, sobre todo en el léxico, y por el contacto de esta población con hablantes de origen hispano.

India

Los datos de hispanohablantes en la India son escasos y de difícil acceso. Estimación realizada a partir de la suma de 910 españoles (INE España, 2024b), 3.300 argentinos (seguidores de «Argentinos en India» en Facebook); 77 mexicanos (Instituto de los Mexicanos en el Exterior, 2020); 670 colombianos (según datos proporcionados por Registro Electoral de la Embajada de Colombia en India); 20 chilenos (INE Chile, 2017b); 329 venezolanos (seguidores de «Venezolanos en India» en Facebook); 1.700 peruanos (seguidores de «Peruanos en India» en Facebook). Para el GDN, se incluyen los que han conseguido un certificado DELE de nivel C antes de 2019. Para el GCL, se incluyen los que han conseguido un certificado DELE nivel B antes de 2019. El censo de la India, muy detallado en cuanto al número de hablantes de L1, L2 y L3, no ofrece información sobre el español, incluido en la categoría de «otras lenguas».

Islandia

Datos provisionales procedentes del estudio en curso sobre la demolingüística del español en los países nórdicos, incluido en el proyecto «El español en Europa» de la Universidad de Heidelberg y la Universidad de Zúrich.

Israel

Los datos de hispanohablantes con dominio nativo corresponden a residentes extranjeros (Datosmacro, s. f.-b). Los datos están basados en fuentes de la ONU. El 73,5 % de la población israelí es judía (Central Bureau of Statistics Israel, 2011, 2022a, 2022 b) y de ella el 50 % aproximadamente son sefardíes. Para el cálculo del grupo de competencia limitada, se estima como hablante de judeoespañol un 10 % de la población sefardí (Central Bureau of Statistics Israel, 2023). En 2011, el Central Bureau of Statistics Israel publicó cifras de hablantes por lengua materna en relación con su conocimiento del hebreo. Entonces se daban 81.300 personas mayores de 20 años con el español como lengua materna. Pueden existir grupos de hispanohablantes formados por descendientes de los pobladores llegados a partir de 1948 principalmente desde Argentina y México. Dado que muchos de ellos están nacionalizados israelíes, no es accesible la información sobre su origen.

Jamaica

El Joshua Project (2024) afirma que hay 8.500 hispanos en Jamaica, principalmente cubanos. El GDN incluye personas con certificado DELE de nivel C anterior a 2019; el GCL incluye personas con certificado DELE de nivel B anterior a 2019.

Japón

Registros de gobiernos de países hispanohablantes. La cantidad viene a coincidir con la aportada en Badillo Matos (2021). No se ha publicado información actualizada y fidedigna sobre hablantes con competencia limitada. El GDN incluye personas con certificado DELE de nivel C anterior a 2019; el GCL incluye personas con certificado DELE de nivel B anterior a 2019.

Marruecos

Hablantes nativos a partir de población hispanohablante en Marruecos, que incluye españoles (INE España, 2024b) y una estimación de población extranjera de origen americano no estadounidense (Haut-Commissariat au Plan, 2014). Dado que, en 2014, el Sáhara Occidental se excluía de los recuentos demográficos de Marruecos, se añade ahora una estimación de hablantes de español de esta región. Población estimada para el Sáhara Occidental en 2025: 673.266, de la cual serían hablantes de español los mayores de 50 años (18 % de la población) (Gómez y Garrido, 2020). A falta de fuentes actualizadas, los hablantes con competencia limitada en español se calculan a partir de Fernández Vitores y Benlabbah (2014). Hemos mantenido la proporción del 4,6 % de marroquíes hablantes de español para no crear una disrupción en la serie temporal de estos informes, aunque las cifras están desactualizadas, a la espera de un estudio más completo y reciente. Otras fuentes apuntan unas cifras muy distintas: la embajadora de Colombia en Marruecos, Farida Loudaya, apuntaba 7 millones en 2017 (Rojas, 2017). En nuestras estimaciones, se prescinde del Sáhara Occidental como entidad política autónoma, si bien se incluyen en Marruecos y Argelia las cantidades correspondientes.

Moldavia

Datos procedentes de la obra *Demolingüística del español en Rumanía, Bulgaria y Moldavia* (Bleorțu et al., 2024).

Noruega

Datos estimados para 2025 a partir de información provisional de un estudio en curso sobre la demolingüística del español en los países nórdicos, incluido en el proyecto «El español en Europa» de la Universidad de Heidelberg y la Universidad de Zúrich. La proyección de población a 2025 se ha realizado a partir de datos del Gobierno noruego (Statistics Norway, 2021).

Nueva Zelanda

Para los hablantes con dominio nativo, se parte de información oficial sobre el grupo de población latinoamericana (Stats New Zealand, 2018), proyectada a 2025, excluidos los brasileños. Se añaden los residentes españoles en Nueva Zelanda (INE España, 2024b). El GDN incluye personas con certificado DELE de nivel C anterior a 2019; el GCL incluye personas con certificado DELE de nivel B anterior a 2019.

Palestina

Datos de españoles en el extranjero (INE España, 2024b). Como es sabido, no todos los que se trasladan a otros países se registran en los consulados pertinentes. No contamos con registros de otros países hispanohablantes sobre personas trasladadas a Palestina. El GDN incluye personas con certificado DELE de nivel C anterior a 2019; el GCL incluye personas con certificado DELE de nivel B anterior a 2019. La situación actual de Palestina hace, por otro lado, imposible ofrecer estimaciones o datos contrastados, más allá de lo aquí considerado.

Reino Unido

La información sobre hablantes de español se obtiene a partir del censo (Office for National Statistics, 2022), con porcentajes proyectados a 2025. En cuanto al grupo de competencia limitada, el *Eurobarómetro especial 386* (Eurostat, 2012) daba un 3 % de británicos capaces de mantener una conversación en español, mientras que el British Council daba un 4 % para 2014, porcentaje que aumentó hasta el 8 % en 2015 para las personas entre 18 y 34 años (British Council, 2017). Teniendo

en cuenta el avance del español en el Reino Unido, a pesar del Brexit, por razones migratorias, de turismo y de enseñanza, se estima que el porcentaje en 2025 puede ser de un 12 % de la población, especialmente entre los jóvenes. El *Eurobarómetro especial 540* (Eurostat, 2024a) ya no incluye datos sobre el Reino Unido.

Rusia, Bielorrusia, Kazajistán y Kirguistán

Datos provisionales procedentes del estudio en curso sobre la demolingüística del español en los países con el ruso como lengua oficial o estatal, incluido en el proyecto «El español en Europa» de la Universidad de Heidelberg y la Universidad de Zúrich. En Rusia, los cubanos son los más numerosos, junto a los ecuatorianos y los españoles.

Suiza

A partir de datos de Kabatek (2022). Proyección de crecimiento sobre población total: 9.058.300 (Federal Statistical Office, 2020).

Trinidad y Tobago

En 2023, llegaron a Trinidad Tobago alrededor de 44.800 inmigrantes de origen venezolano, que se sumaron a la población residente procedente de países hispanohablantes. Los hablantes con competencia limitada se calculan a partir de una estimación del 5 % de hablantes de español sobre la población total del país (Grau Perejoan y Gea Monera, 2006, p. 209).

Turquía

Entre los hablantes nativos, incluimos los de origen español (INE España, 2024b), junto a un millar de colombianos y mexicanos. El GDN incluye personas con certificado DELE de nivel C anterior a 2019; el grupo de competencia limitada incluye personas con certificado DELE de nivel B anterior a 2019. Asimismo, el GCL incluye hablantes de judeoespañol. Las comunidades más relevantes están en Estambul y en Esmirna. Según las autoridades sefardíes, en Estambul la mayoría de los sefardíes saben leer, pero solamente un 20 % aproximadamente son capaces de comprenderlo y algo menos, de hablarlo. En Esmirna, dos tercios de los 1.200 miembros de la comunidad son hablantes. Agradecemos la información que nos han proporcionado Karen Sharhon y Jak Sigura, líderes de las comunidades sefardíes de Estambul y Esmirna. Agradecemos también la ayuda de Fernando Vara de Rey, director del Instituto Cervantes de Estambul.

Unión Europea-27

Información del *Eurobarómetro especial 540* (Eurostat, 2024a), que incluye datos sobre población mayor de 15 años en 2023. Los datos se refieren a este sector de población. Según el Eurobarómetro, tienen el español como lengua materna (excluida España) 2.207.036 personas y lo hablan muy bien el 3 % (11.433.891). Se incluyen en el Grupo de Competencia Limitada los que tienen un español bueno o básico: 9 % (34.301.672). No se tiene en cuenta la población menor de 15 años, que está escolarizada y se contabiliza aparte.

Otros países

Datos obtenidos de registros oficiales de ciudadanos hispanohablantes en el extranjero. Se han tenido en cuenta los países con mayor cantidad de residentes en el exterior: españoles (en todos los países no incluidos en este cuadro); mexicanos (en Brasil y en Asia, excluidos los países de este cuadro); cubanos (en Haití). La mayor parte de los países hispanohablantes cuentan con residentes en otros países hispanohablantes, por lo que no se han tenido en cuenta para la elaboración de esa

tabla. Algunos de los países más poblados del mundo tienen una población nativa de español muy escasa, que en buena medida queda incluida en la cantidad global. Merecen, en todo caso, destacarse Pakistán (con una estimación de 3.000 hablantes nativos) y Tailandia (con una estimación de 1.750 hablantes nativos). Indonesia, cuarto país más poblado del mundo, cuenta con una población nativa de español estimada en 900 hablantes. La población flotante, principalmente por razones de turismo (p. ej. en Tailandia o Indonesia), no se contabiliza en nuestras estimaciones.

El cuadro 2 presenta las cifras de hablantes de español con dominio nativo y con competencia limitada en países y territorios no hispanohablantes. Las notas que acompañan a las cifras de cada país explican la razón de algunas interpretaciones y decisiones metodológicas de naturaleza demolingüística. Para los recuentos de ambos grupos y la confección del cuadro se han tenido en cuenta los criterios que se explican en los siguientes puntos:

1. La nómina de países ha sido revisada, corregida y actualizada de acuerdo con la situación geopolítica actual (por ejemplo, respecto al Sáhara) y de acuerdo con los nuevos datos de que vamos disponiendo (por ejemplo, gracias al proyecto «El español en Europa» de la Universidad de Heidelberg y la Universidad de Zúrich).
2. La población está actualizada a 2025 de acuerdo con la información ofrecida por las respectivas oficinas del censo o institutos de estadísticas. Cuando no se cuenta con cifras oficiales para 2025, se adopta la proyección realizada por los mismos organismos oficiales o por organismos internacionales. En algunos casos, no se han obtenido datos suficientemente fiables. Además, la proyección no se ha realizado cuando la situación demográfica o sociopolítica pueda haber modificado seriamente las tendencias, por lo que simplemente se han realizado estimaciones. Según la fecha de los datos, se ha tenido en cuenta la posible incidencia de la pandemia del coronavirus.
3. La cantidad de hablantes integrantes del Grupo de Dominio Nativo se calcula fundamentalmente a partir de la población inmigrante de primera generación de origen hispánico. Como hemos explicado desde una perspectiva metodológica, el dominio nativo de la lengua española no es óbice para tener un dominio nativo de otra u otras lenguas. Para los países donde hay muy

escasa población hispanohablante, se han considerado las personas que han conseguido el certificado DELE en su nivel C hasta 2019, por existir poco riesgo de doble contabilidad con otras bolsas de hablantes, especialmente residentes extranjeros.

4. La cantidad de hablantes integrantes del Grupo de Competencia Limitada se obtiene generalmente como resultado de la suma de los hablantes reconocidos como descendientes de hispanohablantes o continuadores de una población, una cultura o una tradición hispánicas. En este sentido, los grupos de más delicada valoración son los hablantes de judeoespañol y de lenguas criollas como el chabacano o el chamorro. Para los países donde hay poca población hispanohablante, se han considerado las personas que han conseguido el certificado DELE en su nivel B antes de 2019, por existir poco riesgo de doble contabilidad con otras bolsas de hablantes, especialmente con posibles estudiantes en la enseñanza secundaria, cuando existe, o universitaria.

El conjunto de datos presentados merece un comentario general y varias consideraciones específicas. El comentario general se refiere a la cantidad total de hablantes potenciales del español fuera del dominio hispánico, que sobrepasa los 127 millones, un 28 % por encima de los datos ofrecidos para 2024. Más allá del peso de la cifra misma, debe especificarse que la diferencia también se debe a razones metodológicas: algunos censos y datos se han actualizado o han ofrecido datos más completos; contamos con estudios más sólidos y pormenorizados, como los derivados del proyecto «El español en Europa»; los movimientos poblacionales pueden haberse intensificado, como en el caso de los venezolanos hacia Brasil; además, se han considerado poblaciones migrantes a nivel mundial.

En cuanto a la cifra en sí, los 122 millones de hablantes potenciales de español remiten directamente a uno de los ámbitos en que el español podría tener una mayor probabilidad de crecimiento en los próximos años: el de su conocimiento y uso como lengua segunda y extranjera (generalmente, uso con competencia o habilidades limitadas). A ello habría de sumarse el potencial

incremento del número mundial de estudiantes, a lo que más adelante se hará referencia. Esos 122 millones, consecuentemente, permiten estimar un progresivo crecimiento del español como lengua internacional y franca, más allá de los intercambios internos de la comunidad hispánica.

Las consideraciones específicas que requiere el cuadro 2 obligan a dirigir la mirada hacia el peso demográfico de la población hispanohablante en Estados Unidos y en la Unión Europea. En el caso de Estados Unidos, se observa una desaceleración en el incremento de hablantes de español, motivada por el número de nacimientos de hispanos o latinos en el mismo Estados Unidos, que lleva a un conocimiento y un uso más generalizado del inglés entre las generaciones más jóvenes y, en muchos casos, a la no adquisición del español. Por lo demás, las políticas migratorias y socioeconómicas del país determinarán la incidencia que puedan tener en el futuro tanto los flujos migratorios como las tendencias lingüísticas de los grupos hispanos ya residentes. En el caso de Europa, además de la incidencia de la aplicación de una metodología más refinada, el *Eurobarómetro especial 540* (Eurostat, 2024a) indica con claridad que, si en este momento solamente el 7 % del total de los europeos mayores de 15 años son capaces de mantener una conversación en español, esa proporción se eleva al 13 % entre los jóvenes de 15 a 24 años, con tendencia al alza. Este último porcentaje sitúa al español a 1 punto porcentual del francés en el mismo segmento de edad y 1 punto por encima del alemán, con ambas lenguas mostrando tendencia a la baja entre los jóvenes.

Este informe de 2025 revela asimismo que la movilidad de los hispanohablantes sigue siendo causa directa de la presencia del español en espacios del globo alejados lingüística y culturalmente del espacio hispánico. En estos casos contamos con datos de los residentes extranjeros de origen hispánico como hablantes con dominio nativo de la lengua, pero no contamos con suficiente información sobre su manejo con competencia o habilidades limitadas, bien porque no existen grupos asentados de origen hispano², bien

2. Además de las incluidas en el cuadro, existen otras comunidades sefardíes en diversos países del mundo (p. ej. Bulgaria, Serbia, Singapur, Egipto...), pero tienen una dimensión muy reducida y no se han tenido en cuenta por la dificultad de conseguir información sobre sus repertorios y hábitos lingüísticos.

porque no se conocen datos de los egresados de la enseñanza con conocimientos de español. En casos como los de Guam, Islas Marianas o Marruecos sí han existido comunidades hispánicas anteriormente que pueden justificar un conocimiento y uso del español con competencia limitada.

2.3 Los aprendices de español

En 2025, alrededor de 24,6 millones de personas aprenden español como lengua extranjera en distintos contextos educativos. La cifra (24.560.143 personas) supone un crecimiento del 1,5 % con respecto a 2024 y continúa la tendencia de los últimos años. Desde 2012, año del primer cálculo exhaustivo ofrecido por el informe del Instituto Cervantes, se contabiliza un aumento del número de aprendices de español en el mundo de más de un tercio, aunque el aprendizaje está condicionado por las muy distintas realidades de la geografía educativa.

El cuadro 3 presenta la cuantificación y distribución mundial de los aprendices de español en la enseñanza reglada (primaria, secundaria y universitaria), así como en otros espacios de enseñanza no reglada como los centros de idiomas o las aulas del Instituto Cervantes de todo el mundo. Los datos apuntan directamente a dos claves: por un lado, la enseñanza de español depende altamente de su grado de institucionalización como lengua extranjera en contextos educativos de enseñanza reglada no universitaria; por otro lado, se dibuja un mapa de la enseñanza en torno a tres territorios, que delimitan sus principales espacios de influencia: Estados Unidos, Brasil y la Unión Europea.

La importancia cuantitativa de los tres espacios como vectores del aprendizaje de español en el mundo resulta evidente, pues en ellos se concentran cuatro de cada cinco aprendices de español. El español es, tras el inglés, la primera lengua de Estados Unidos, donde la aprende el 76 % de los estudiantes de

En la mayoría de los casos, el uso del judeoespañol responde a una competencia limitada, por restringirse a la liturgia. En otros casos, no es previsible el mantenimiento del judeoespañol porque la comunidad sefardí convive con grupos judíos de origen no hispano, aparte de su convivencia con la población no judía de sus países de residencia. No se identifican específicamente las comunidades sefardíes en países o territorios que cuentan con estadísticas de hablantes de español en casa.

lenguas extranjeras de primaria y secundaria, y el 50 % en la universidad. En Brasil, estudia español un 90 % de los alumnos que cursan una segunda lengua extranjera, alrededor de 4 millones de alumnos, si bien la legislación aprobada recientemente está modificando esa cifra a la baja. Por último, en la Unión Europea el español es la segunda o la tercera lengua extranjera con más alumnos, aunque su institucionalización varía notablemente en función del país, la etapa educativa o el tipo de itinerario educativo que se considere.

El mapa de aprendices de español no estaría completo si no se consideraran otras regiones, sobre todo del continente africano, donde se encuentra uno de cada diez aprendices de español en el mundo. Las cifras en estos países se basan fundamentalmente en la institucionalización del español como segunda lengua extranjera dentro de los respectivos sistemas educativos en las etapas de secundaria en cinco territorios de África occidental y central: Costa de Marfil (798.095), Benín (412.515), Camerún (403.000), Senegal (356.224) y Gabón (200.000). En Costa de Marfil, por ejemplo, alrededor de dos tercios de los alumnos que cursan una segunda lengua extranjera estudian español (Djandué, 2021). El español también se aprende en otros países del norte de África como Argelia, Marruecos o Túnez, aunque su presencia en la enseñanza reglada sea más bien limitada. Por fortuna, estos datos contarán con una actualización a lo largo de 2025, por lo que el informe del año siguiente podrá recogerlos y poner al día la información sobre esta importante región del mundo.

Más allá de la enseñanza secundaria, merece atención la enseñanza de español en universidades y centros de enseñanza superior. En algunos países, la educación terciaria compensa de algún modo la falta de posibilidades de estudiar español en etapas educativas anteriores. Esto sucede en China o Argelia, por ejemplo, donde más de la mitad de los aprendices de español lo cursa en centros de educación superior. La base documental para conocer el aprendizaje de la lengua española en las universidades es, sin embargo, considerablemente limitada. Ello se debe no solo a la ausencia de recuentos estadísticos sistemáticos de los estudiantes de español y de las filologías relacionadas con el español, sino sobre todo a que los países generalmente no disponen de una

cuantificación oficial del número de universitarios que aprenden la lengua como extranjera o para especialistas de otras disciplinas.

Comentario aparte merecen los espacios de enseñanza no reglada. La enseñanza de idiomas es un sector poco transparente en lo que se refiere a la información sobre su actividad y cobertura. Desde el punto de vista de la demolingüística del español, el número de hablantes que lo aprenden en contextos informales es una población oculta estimada en otras ediciones del informe en 5 millones de personas. El guarismo podría ser incluso mayor si se considera el interés que despierta su aprendizaje en plataformas como Duolingo, como más adelante comentaremos. Otro indicador importante es su valor de empleabilidad, observado a partir de los datos de certificación oficial. En los últimos años pospandemia, el número de candidatos a los Diplomas de Español como Lengua Extranjera (DELE) ha superado los 100.000 examinados. Buena parte de esos candidatos son alumnos de alguno de los cursos que ofrece el Instituto Cervantes en sus sedes internacionales, que acogen en 2025 a cerca de 68.000 alumnos.

CUADRO 3 NÚMERO APROXIMADO DE ESTUDIANTES DE ESPAÑOL EN EL MUNDO

País	Enseñanza primaria, secundaria y formación profesional	Enseñanza universitaria	Otros	Centros del Instituto Cervantes y AVE 2023-2024	Total
Estados Unidos	8.358.465	584.453		3.738	8.946.656
Brasil	4.001.530	43.517		3.371	4.048.418
Francia	3.357.221	244.080	5.978	2.037	3.609.316
Reino Unido	2.017.796	21.505		3.378	2.042.679
Italia	862.243	61.000	41.685	5.813	970.741
Costa de Marfil	795.488	1.752	855		798.095
Alemania	518.021	52.563	178.255	5.442	754.281
Benín					412.515

País	Enseñanza primaria, secundaria y formación profesional	Enseñanza universitaria	Otros	Centros del Instituto Cervantes y AVE 2023-2024	Total
Camerún	400.000	3.000			403.000
Senegal				224	356.224
Suecia	261.598	7.612	7.455	283	276.948
Polonia	245.983	18.680	3.928	2.510	271.101
Gabón	200.000				200.000
Canadá	156.126	10.992			167.118
Portugal	101.694	5.600		594	107.888
Noruega	95.863	587			96.450
Argelia	40.290	45.409		3.024	88.723
Austria	56.507	4.000	9.131	774	70.412
Australia	52.726	9.300		803	62.829
Marruecos	43.261	1.772	4.839	11.031	60.903
Japón				498	60.498
Rusia	42.022	8.398	4.418	1.277	56.115
Irlanda	49.451	3.600		1.013	54.064
China	8.874	34.823	8.866	1.477	54.040
República Checa	44.219	8.317		1.264	53.800
Bélgica	14.320	7.799	24.338	735	47.192
India	25.000	7.681	11.001	2.670	46.352
Túnez	39.500			899	40.399
Nueva Zelanda	34.275	2.098			36.373
Países Bajos	19.097	14.160		457	33.714
Dinamarca	28.813	248			29.061
Otros países					304.238
TOTAL					24.560.143

FUENTE: elaboración propia. Dando continuidad al criterio seguido en informes anteriores, para no alterar radicalmente la proyección estadística, el cuadro solo desglosa los países donde hay más de 30.000 alumnos, aunque la cifra total sí agrega los datos de los países con un número inferior de estudiantes. El número de estudiantes que se proporciona corresponde a 112 países, por lo que el número total en el mundo ha de ser superior al indicado en este cuadro. Si en alguna celda no aparecen cifras, se debe a la ausencia de datos específicos o a que se ofrecen los datos agregados.

FUENTES E INFORMACIÓN ADICIONAL DEL CUADRO 3

Alemania

Estudiantes en enseñanza primaria, secundaria y formación profesional, datos de la Oficina Federal de Estadística correspondientes al año escolar 2023-2024 (Statistisches Bundesamt, s. f.). Estudiantes universitarios, en el año escolar 2019-2020 había 5.393 alumnos inscritos en cursos de Filología Hispánica, más una estimación de 47.170 alumnos que estudian español como lengua extranjera en 106 universidades alemanas (Loureda Lamas *et al.*, 2020, 2021, 2022). Estudiantes en enseñanza no reglada inscritos en universidades populares (*Volkshochschulen*) en 2019 (Loureda Lamas *et al.*, 2022). Para evitar una doble contabilización, se ha restado de la cifra global proporcionada el número de alumnos del Instituto Cervantes en Alemania para el curso 2022-2023.

Argelia

A partir de datos de Jiménez Caballero, Aous y Salhi (2021). Estudiantes en enseñanza universitaria incluyen 3.409 alumnos inscritos en departamentos universitarios de español en el curso 2020-2021 y 42.000 estudiantes universitarios en otras facultades como traducción e interpretación, así como aquellos que estudian español como lengua extranjera, cifra correspondiente a 2016.

Australia

A partir de datos del Ministerio de Educación y Formación Profesional (MEFP). El ministerio reporta 35.306 estudiantes de español en primaria y 17.420 en secundaria durante el ciclo escolar 2021 (MEFP, 2022). Las cifras corresponden solamente a centros públicos. El número de estudiantes universitarios corresponde al año 2017 (MEFP, 2018).

Austria

Estudiantes en enseñanza primaria, secundaria y formación profesional, datos de la Oficina Federal de Estadística correspondientes al año escolar 2022-2023 (Statistik Austria, 2024). Estudiantes universitarios, estimación realizada a partir del número de estudiantes Erasmus procedentes de cada país matriculados en universidades españolas en 2019, al que se ha aplicado un factor multiplicador uniforme calculado a la baja. Estudiantes en enseñanza no reglada, datos de alumnos inscritos en cursos de español durante el año 2019 suministrados por la Asociación de Universidades Populares (*Volkshochschulen*).

Bélgica

Estudiantes en enseñanza primaria, secundaria y formación profesional, datos de Eurostat (2024b) correspondientes al año escolar 2022-2023. Estudiantes universitarios, estimación realizada a partir del número de estudiantes Erasmus procedentes de cada país matriculados en universidades españolas en 2019, al que se ha aplicado un factor multiplicador uniforme calculado a la baja. Estudiantes en enseñanza no reglada, el Ministerio de Educación y Formación Profesional reporta 5.270 alumnos de español en la comunidad francófona, 18.958 en la comunidad flamenca y 110 en la comunidad germanófona durante el curso escolar 2019-2020 (MEFP, 2022).

Benín

A partir de datos de Serrano Avilés (2014). Estos datos se están revisando para un nuevo estudio de la enseñanza de español en África subsahariana.

Brasil

Estudiantes en enseñanza primaria, secundaria y formación profesional, dato correspondiente a 2021 facilitado por el Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP, 2020) a través de la Consejería de Educación de España en Brasil. Para los estudiantes universitarios, acudimos al INEP (2020). Estas cifras, sin embargo, deben tomarse con cautela, ya que la situación puede estar modificándose como consecuencia de la sanción de una reforma educativa de la educación secundaria que incluye como «opcional» la enseñanza de la lengua española, que había sido de oferta obligatoria entre 2005 y 2017. Está en elaboración un estudio de la situación actual que podría reducir a la mitad el alumnado de español. Por el momento, trabajamos con los últimos cálculos oficiales.

Camerún

Estimaciones suministradas en febrero de 2024 por la Embajada de España en Camerún. La cifra de estudiantes universitarios incluye tanto estudios de lengua española como de traducción.

Canadá

A falta de datos representativos a escala nacional (MEFP, 2018, 2020, 2022), se ha hecho una estimación de la población escolar de la provincia de Quebec a partir de datos de Gay (2014) para el curso 2022-2023 (Statistics Canada 2024a, 2024b). El número total de estudiantes será muy superior al reportado. En el momento de redactar este informe, se está completando un estudio sobre el español en las universidades canadienses por parte de Lauren Beck y Yolanda Iglesias. En él podría estimarse el alumnado de español, en este nivel educativo, en una cifra superior a los 15.000. estudiantes. El informe «El español: lengua para el mundo. Informe 2026» recogerá las cifras definitivas.

China

Enseñanza primaria, secundaria y formación profesional, a partir de Instituto Cervantes (2023). Estudiantes universitarios, a partir de González Puy (2018).

Costa de Marfil

A partir de datos suministrados en febrero de 2024 por el Instituto Cervantes en Costa de Marfil, consultados al Ministerio de Educación Secundaria y Alfabetización, y la Embajada de España. Estudiantes en enseñanza primaria, secundaria y formación profesional, 528.648 en primaria y 266.840 en secundaria durante el ciclo escolar 2021-2022. Estudiantes en enseñanza no reglada, corresponden a profesores que han recibido formación de español por parte de la Embajada de 2015 a 2023.

Estados Unidos

Estudiantes en enseñanza primaria, secundaria y formación profesional, estimación basada en el porcentaje de alumnos de K-12 que estudia una lengua extranjera (20 %) y la proporción de estos que elige español (76 %) según el informe de American Councils for International Education (2017, p. 7) y la población escolar (National Center for Education Statistics US, 2023a, 2023b, 2024). Para los estudiantes universitarios, los datos proceden de Lusin *et al.* (2023, p. 6).

Francia

Estudiantes en enseñanza primaria, secundaria y formación profesional, datos de Eurostat (2024b) correspondientes al año escolar 2022-2023. Estudiantes universitarios, estimación de Tano (2023, p. 157). Estudiantes en enseñanza no reglada, datos a partir de Instituto Cervantes (2023).

Gabón

Estimación para estudiantes de enseñanza secundaria suministrada en febrero de 2024 por la Embajada de España.

India

Estimación a la baja a partir de cifras de Pujol Riembau (2020). La cifra real podría ser muy superior.

Irlanda

Estudiantes en enseñanza primaria, secundaria y formación profesional, datos de Eurostat (2024b) correspondientes al año escolar 2022-2023. Estudiantes universitarios, estimación realizada a partir del número de estudiantes Erasmus procedentes de cada país matriculados en universidades españolas en 2019, al que se ha aplicado un factor multiplicador uniforme calculado a la baja.

Italia

Estudiantes en enseñanza primaria, secundaria y formación profesional, datos de Eurostat (2024b) correspondientes al año escolar 2022-2023. Estudiantes universitarios, estimación realizada a partir del número de estudiantes Erasmus procedentes de cada país matriculados en universidades españolas en 2019, al que se ha aplicado un factor multiplicador uniforme calculado a la baja. Estudiantes en la educación no reglada, a partir de Bonomi, Calvi y Uberti-Bona (2024). Para evitar una doble contabilización, se ha restado de la cifra global proporcionada por las autoras el número de alumnos del Instituto Cervantes en Italia para el curso 2020-2021.

Japón

A partir de Ugarte Farrerons (2012) y Ministerio del Interior y de Comunicación de Japón (2023). En el momento de redactar de este informe, se está completando un exhaustivo estudio sobre el español en las universidades japonesas por parte de Juan Carlos Moyano, Yoshimi Hiroyasu y Ángela Yamaura. En él podría estimarse el alumnado de español, en este nivel educativo, en una cifra superior a los 60.000 estudiantes. «El español: lengua para el mundo. Informe 2026» recogerá las cifras definitivas.

Marruecos

Estudiantes en enseñanza primaria, secundaria y formación profesional, a partir de Instituto Cervantes (2023). Estudiantes universitarios, datos del Ministerio de Educación y Formación Profesional (2020). Estudiantes en enseñanza no reglada, datos suministrados en febrero de 2024 por la Embajada de España, correspondientes al alumnado en centros de titularidad del Estado español dependientes de la Consejería de Educación.

Noruega

Estudiantes en enseñanza primaria, secundaria y formación profesional, datos de Eurostat (2024b) correspondientes al año escolar 2022-2023. Estudiantes universitarios, estimación realizada a partir del número de estudiantes Erasmus procedentes de cada país matriculados en universidades españolas en 2019, al que se ha aplicado un factor multiplicador uniforme calculado a la baja.

Nueva Zelanda

A partir de datos del Ministerio de Educación y Formación Profesional (2020).

Países Bajos

Estudiantes en enseñanza primaria, secundaria y formación profesional, datos de Eurostat (2024b) correspondientes al año escolar 2022-2023. Estudiantes universitarios, estimación realizada a partir del número de estudiantes Erasmus procedentes de cada país matriculados en universidades españolas en 2019, al que se ha aplicado un factor multiplicador uniforme calculado a la baja.

Polonia

Estudiantes en enseñanza primaria, secundaria y formación profesional, datos de Eurostat (2024b) correspondientes al año escolar 2022-2023. Estudiantes universitarios, datos a partir de Zieliński *et al.* (en prensa). Estudiantes en enseñanza no reglada, a partir de Instituto Cervantes (2023).

Portugal

Estudiantes en enseñanza primaria, secundaria y formación profesional, datos de Eurostat (2024b) correspondientes al año escolar 2022-2023. Estudiantes universitarios, dato a partir de Santos Rovira y Serrano Lucas (2022).

Reino Unido

Estudiantes en enseñanza primaria, secundaria y formación profesional, datos de Eurostat (2024b) correspondientes al año escolar 2018-2019. Estudiantes universitarios, estimación realizada a partir del número de estudiantes Erasmus procedentes de cada país matriculados en universidades españolas en 2019, al que se ha aplicado un factor multiplicador uniforme calculado a la baja.

República Checa

Estudiantes en enseñanza primaria, secundaria y formación profesional, datos de Eurostat (2024b) correspondientes al año escolar 2022-2023. Estudiantes universitarios, estimación realizada a partir del número de estudiantes Erasmus procedentes de cada país matriculados en universidades españolas en 2019, al que se ha aplicado un factor multiplicador uniforme calculado a la baja.

Rusia

A partir de los datos provisionales procedentes del estudio en curso sobre la demolingüística del español en los países con el ruso como lengua oficial o estatal, incluido en el proyecto «El español en Europa» de la Universidad de Heidelberg y la Universidad de Zúrich. Estudiantes en enseñanza no reglada, datos a partir de Instituto Cervantes (2023).

Senegal

A partir de cifras oficiales de la Dirección de la Planificación de la Reforma de la Educación de Senegal, facilitadas al Instituto Cervantes en noviembre de 2020.

Suecia

Estudiantes en enseñanza primaria, secundaria y formación profesional, datos de Eurostat (2024b) correspondientes al año escolar 2022-2023. Estudiantes universitarios, estimación realizada a partir del número de estudiantes Erasmus procedentes de cada país matriculados en universidades españolas en 2019, al que se ha aplicado un factor multiplicador uniforme calculado a la baja. Para estudiantes en enseñanza no reglada, datos de Statistics Sweden proporcionados en el estudio en curso sobre la demolingüística del español en los países nórdicos, incluido en el proyecto «El español en Europa» de la Universidad de Heidelberg y la Universidad de Zúrich.

Túnez

Datos suministrados en febrero de 2024 por la Embajada de España, facilitados por el Ministerio de Educación tunecino para el ciclo escolar 2022-2023.

Otros países

Incluye cifras de países con menos de 30.000 alumnos en África y Asia (Catar, Emiratos Árabes Unidos, Filipinas, Ghana, Guinea-Bissau, Irán, Irak, Israel, Libia, Malawi, Mali, Togo, Vietnam y Zimbabue), cuyos datos han sido proporcionados por las embajadas de España en los respectivos países en febrero de 2024; datos de Brunéi, Indonesia, Malasia y Tailandia a partir de Badillo Matos (2023); datos de Kazajistán y Kirguistán a partir de los datos provisionales procedentes del estudio en curso sobre la demolingüística del español en los países con el ruso como lengua oficial o estatal, incluido en el proyecto «El español en Europa» de la Universidad de Heidelberg y la Universidad de Zúrich; así como datos no desglosados de otros países de Europa: para Albania, Bosnia y Herzegovina, Macedonia del Norte, Montenegro y Serbia, véase Loureda Lamas *et al.* (2025); datos de Bielorrusia a partir de los datos provisionales procedentes del mismo estudio de la Universidad de Heidelberg y la Universidad de Zúrich; para la Unión Europea-27 (Bulgaria, Chipre, Croacia, Eslovaquia, Eslovenia, Estonia, Finlandia, Grecia, Hungría, Letonia, Lituania, Luxemburgo, Malta y Rumania), véase Eurostat (2024b).

El recuento para los países de la UE-27 recoge estudiantes en enseñanza primaria, secundaria y formación profesional correspondientes al año escolar 2022-2023. Para los estudiantes universitarios, con el fin de ofrecer cierta continuidad en la estadística, seguimos el criterio aplicado en «El español: una lengua viva. Informe 2023» (Instituto Cervantes, 2023): estimación realizada a partir del número de estudiantes Erasmus procedentes de cada país matriculados en universidades españolas, al que se ha aplicado un factor multiplicador uniforme calculado a la baja. En cuanto a los estudiantes de cursos de español en la red de centros del Instituto Cervantes y en la plataforma AVE, se excluyen datos de estudiantes de aulas del Instituto Cervantes situadas en países en los que no hay centros propiamente dichos.

2.4 Los hablantes potenciales de lengua española

La suma de los totales ofrecidos en los cuadros anteriores queda recogida en el cuadro 4. El total corresponde al grupo de hablantes potenciales de español dentro y fuera de los países mayoritariamente hispanohablantes. Recordamos que el uso potencial del español incluye tanto el que de él hacen los hablantes que tienen un dominio nativo de la lengua, como aquellos cuyo dominio o competencia encuentran eventualmente limitaciones de naturaleza estilística o sociolingüística y aquellos que lo están aprendiendo dentro de un sistema educativo.

CUADRO 4 HABLANTES DE ESPAÑOL DE DOMINIO NATIVO, CON COMPETENCIA LIMITADA Y APRENDICES DE LENGUA, POR TIPOS DE PAÍSES Y CON TOTALES

	GDN	GCL	GALE ³	Total
Países hispanohablantes	459.172.613	24.856.989	-	484.029.602
Países no hispanohablantes	59.942.645	67.211.254	24.560.143	151.714.042
TOTAL	519,115,258	92,068,243	24,560,143	635,743,644

FUENTE: elaboración propia..

El cuadro 4 consolida algunas dinámicas que merecen resaltarse. En primer lugar, la comunidad hispanohablante continúa su dinámica de crecimiento demográfico y supera los 500 millones de hablantes de dominio nativo. En segundo lugar, se observa el crecimiento de los hablantes potenciales de español fuera del espacio mayoritariamente hispanohablante, lo que se debe en buena medida a la población de trasfondo migratorio, aunque también a la actualización de las fuentes y de los datos, así como al refinamiento del análisis. En cuanto a los aprendices de español, se constata el crecimiento sostenido de este grupo, con un incremento del 1,5 % el último año. En conjunto, los hablantes potenciales de lengua española en el mundo superan ampliamente los 630 millones, lo que mantiene al español entre las lenguas con mayor potencial demográfico del planeta.

3. El español entre las lenguas internacionales

3.1 El español como lengua nativa, primera o materna

Este informe destaca como información relevante el hecho de que la comunidad hispanohablante con un dominio nativo de la lengua ha superado con claridad los 500 millones de hablantes. Este dato la sitúa por debajo de la comunidad nativa de chino mandarín, como siempre ha ocurrido, pero

3. En el GALE de los países hispanohablantes sería susceptible incluir aprendices que ya son contados como tales en sus respectivos países de origen; por este motivo no se incluyen en este cuadro, de modo que se minimice la posibilidad de incurrir en doble contabilidad.

en estos momentos también por detrás de la comunidad nativa de hindi, al tiempo que la cifra se mantiene por encima de la correspondiente a los hablantes nativos de inglés. El cuadro 5 recoge los guarismos que nos permiten establecer esta jerarquía.

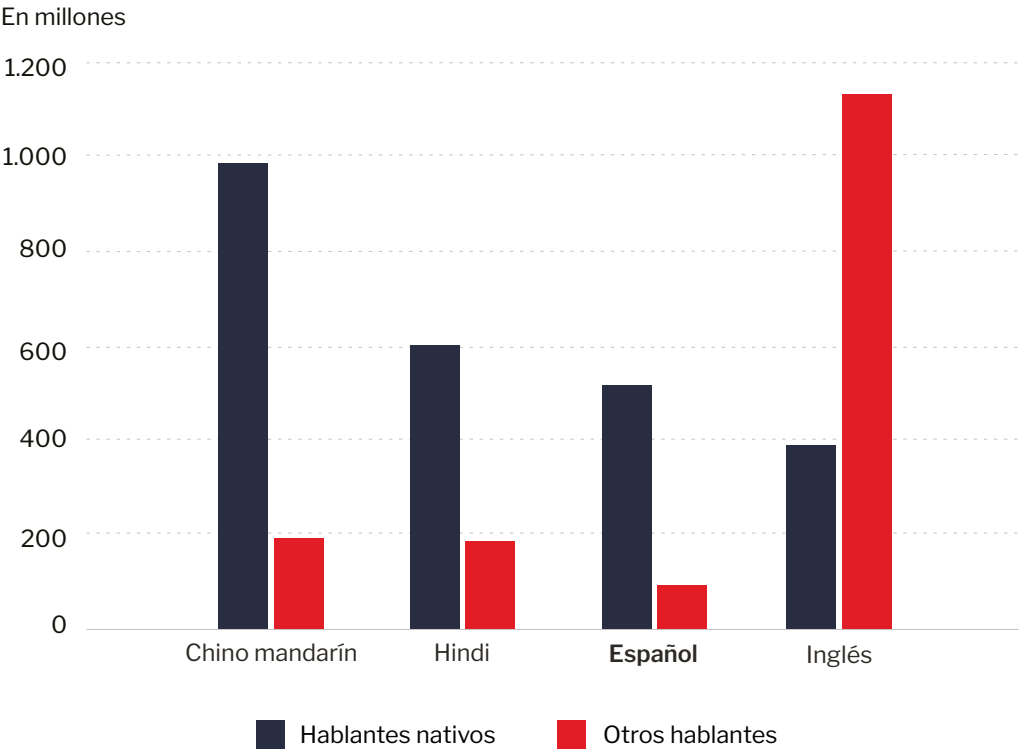
CUADRO 5 LENGUAS MÁS HABLADAS DEL MUNDO COMO NATIVAS Y COMO SEGUNDAS Y/O EXTRANJERAS

Lengua	Hablantes nativos	Otros hablantes
Chino mandarín	989.874.880	193.927.500
Hindi	603.372.494	186.393.633
Español	519.115.258	92.068.243
Inglés	390.348.200	1.137.573.260

FUENTE: los datos del chino mandarín y del inglés proceden de *Ethnologue* (Eberhard, Simons y Fenning, 2025); los datos del hindi son una proyección a 2025 a partir del Censo de 2011 (Office of the Registrar General & Census Commissioner, 2021, 2022; National Commission on Population, 2019) en el caso de los nativos y de *Ethnologue* en el caso de «otros hablantes»; los datos del español son los presentados en este informe.

Para la interpretación de este cuadro, hay que tener en cuenta que la caracterización como lengua nativa (y de hablantes con dominio nativo) alterna, según las fuentes, con la de lengua materna o lengua primera (L1), lo que impide garantizar que la referencia sea exactamente la misma para todos los recuentos (Moreno Fernández, 2024). Además, una adecuada interpretación de esas cantidades de hablantes requiere manejar y valorar informaciones adicionales referidas a la naturaleza social y lingüística de cada una de las lenguas consideradas.

GRÁFICO 1 LENGUAS MÁS HABLADAS DEL MUNDO COMO NATIVAS (L1) Y COMO SEGUNDAS Y/O EXTRANJERAS (L2)



FUENTE: los datos del chino mandarín y del inglés proceden de *Ethnologue* (Eberhard, Simons y Fenning, 2025); los datos del hindi son una proyección a 2025 a partir del Censo de 2011 (Office of the Registrar General & Census Commissioner, 2021, 2022; National Commission on Population, 2019) en el caso de los nativos y de *Ethnologue* en el caso de «otros hablantes»; los datos del español son los presentados en este informe.

En primer lugar, el chino mandarín, desarrollado como dialecto del grupo étnico han del norte de China, es una macrolengua que agrupa una docena de variedades que comparten una escritura y una tradición escrita, y que no son inteligibles para otras variedades de China (min, hakka), aunque compartan la misma escritura. Geográficamente, el chino mandarín es más hablado en la provincia de Pekín, así como al norte, centro y suroeste de China (Eberhard, Simons y Fenning, 2025).

El chino estándar o putonghua es un tipo de chino mandarín considerado como lengua oficial en China continental, cuyas normas no se aplican a Macao ni a Hong Kong. Alrededor del 70 % de los chinos hablan alguna versión del mandarín como primera lengua o lengua nativa. En cuanto

a la escritura del chino, existen dos versiones: una tradicional, utilizada oficialmente en Taiwán, Hong Kong y Macao, y una simplificada (pinyin), creada en la segunda mitad del siglo ^{xx} para hacer más accesible la escritura y consecuentemente la alfabetización. La escritura simplificada es la oficial en China continental y la más popular entre las generaciones jóvenes de los hablantes internacionales de chino.

Existen serias dificultades para conocer el número exacto de hablantes de chino mandarín estándar, dado que los censos lingüísticos, si existen, no suelen ser accesibles internacionalmente. En cuanto a la cifra que ofrecemos como «Otros hablantes», corresponde en su mayor parte a los hablantes en Taiwán, Macao y Singapur, así como en Estados Unidos o Canadá. En cualquier caso, el chino mandarín es lengua de la República Popular China (con sus distintas regiones), Singapur y Taiwán, de modo que su uso internacional corresponde primordialmente a los grupos migrantes que se distribuyen por todo el mundo, sin que suelen constituir una comunidad idiomática integrada socialmente en sus respectivos países de acogida.

En segundo lugar, la lengua hindi deriva del sánscrito, con la incorporación de elementos y voces del árabe y el persa (Dipshikha, 2023). En la actualidad es lengua oficial en la India, si bien el inglés funciona *de facto* como lengua de gobierno, por las reticencias de varios estados a aceptar el uso del hindi (Department of Official Language, 2016). Debe tenerse en cuenta que la Constitución de la India cataloga 22 lenguas a las que da reconocimiento oficial. Todo ello en un entorno nacional en el que se hablan 122 lenguas principales y se clasifican 1.369 lenguas y variedades más, según el Censo de 2011 (Office of the Registrar General & Census Commissioner, 2011).

Geográficamente, los hablantes nativos de hindi se asientan principalmente en el centro-norte del país, en los estados de Rajasthan, Uttar Pradesh, Madhya Pradesh, Bihar, Jharkhand, Chhattisgarh y Dehli. En el sur, son otras las lenguas oficiales más extendidas (marathi, telugu, kannada,

tamil)⁴. Como se ha indicado, el cuadro 5 presenta una estimación de la cifra de hablantes nativos de hindi en India proyectadas a 2025 a partir de los datos de población del Censo de 2011 (Office of the Registrar General & Census Commissioner, 2022). Esa fuente, además de los hablantes de L1, aporta también las cifras de hablantes de segundas (L2) y terceras (L3) lenguas en India (Office of the Registrar General & Census Commissioner, 2021). De este modo, la comunidad de hablantes de hindi adquiere unas dimensiones mucho mayores porque son 158,6 millones los que lo tienen como L2 y otros 27,7 millones los que lo tienen como L3. Esto no resulta sorprendente en uno de los países más poblados del mundo y donde el hindi es lengua oficial. Además, con toda probabilidad muchos de los hablantes de hindi como L2 han de dominarla como lengua nativa, lo que conferiría aún mayor entidad a su comunidad de dominio nativo. Con todo, el hindi no es estrictamente una lengua internacional, por lo que podría decirse que se trata de un gigante con pies de barro en el terreno de la geopolítica.

Esta descripción de las lenguas china e hindi es necesaria para comprender y perfilar convenientemente la comunidad con dominio nativo de español. Las diferencias entre estas comunidades nativas son más que notables en muchos sentidos, como lo son las diferencias respecto de la comunidad anglófona nativa. El cuadro 6 refleja las principales particularidades y coincidencias entre las cuatro comunidades de hablantes nativos más grandes del mundo.

4. Asimismo, en los informes demolingüísticos es posible encontrar referencias al indostánico o hindi-urdu. Esto se debe a la proximidad lingüística de ambas variedades (urdu e hindi), pero entre ellas existen diferencias significativas: el urdu, lengua nacional en Pakistán, aunque no mayoritaria, se escribe mediante una adaptación del alfabeto persa y es utilizado principalmente por musulmanes, mientras que el hindi se escribe en alfabeto devanagari y es utilizado mayoritariamente por hinduistas.

CUADRO 6 COMPARACIÓN DE COMUNIDADES NATIVAS DE CHINO MANDARÍN, HINDI, ESPAÑOL E INGLÉS

	Familia	Geografía	Oficial ONU	Escritura	Comunidad L1	Comunidad L2
Chino	Sino-tibetana	1 continente	Sí	china	creciente	estable
Hindi	IE-Indoirania	1 región	No	devanagari	creciente	estable
Español	IE-Romance	3 continentes	Sí	latina	creciente ralentizada	creciente
Inglés	IE-Germánica	5 continentes	Sí	latina	creciente ralentizada	creciente

FUENTE: elaboración propia. IE = Indoeuropeo.

La comparación entre estas cuatro comunidades de hablantes y sus respectivas lenguas arroja diferencias sustanciales, más allá de las cifras. Tres de las lenguas pertenecen a la gran familia indoeuropea, aunque se vinculen a ramas diferentes: el hindi a la indoirania, el español a la romance y el inglés a la germánica. La geografía marca unas distancias considerables: mientras el hindi y el chino mandarín son lenguas instaladas fundamentalmente como nativas en regiones de sus respectivos países, el español es una lengua nativa geográficamente internacional (21 países de Europa, América y África), como en mayor medida lo es el inglés (57 países y 28 entidades soberanas de Europa, América, África, Asia y Oceanía). El hindi, por otro lado, es la única lengua nativa que no es oficial en el sistema de la ONU. En lo que se refiere a la escritura, el inglés y el español comparten el uso del alfabeto latino, mientras que para el chino y el hindi se utilizan alfabetos de menor difusión internacional.

Finalmente, las dinámicas demolingüísticas del chino y el hindi, por un lado, y del español y el inglés, por otro, son opuestas: mientras la demografía de hablantes nativos crece de modo galopante en los casos de China y la India, especialmente en este último caso, la de los países hispanófonos y anglófonos se ralentiza, con la excepción tal vez de los países africanos, donde el porcentaje de hablantes nativos de inglés es relativamente modesto. Por otro lado, los hablantes de español y de inglés como L2 muestran una dinámica

creciente, sobre todo en el caso del inglés, mientras que el hindi como L2 solo crece al ritmo de la demografía india; por su parte, el chino mandarín como L2, si bien su aprendizaje interesa cada vez más, no llega a alcanzar las cifras de las lenguas internacionales, tanto por la distancia lingüística y escritural, como por la capacidad creciente de los hablantes de chino para manejarse internacionalmente en inglés y en español.

De esta comparación no se debe deducir que haya lenguas nativas (o maternas) más importantes que otras, ya que todas son cognitivamente, lingüística, cultural y socialmente esenciales para sus hablantes. La comparativa, en cambio, permite apreciar con claridad cuáles son las diferencias derivadas de factores externos a las lenguas mismas, como la geografía o la oficialidad. Esta comparación resulta de gran utilidad para conocer las particularidades del español como lengua nativa en contraste con las de otras lenguas. Con todo, el conocimiento bilingüe de las lenguas española e inglesa es la combinación más prometedora con vistas a la comunicación internacional, tanto por el peso conjunto de sus comunidades nativas (alrededor de 1.000 millones de hablantes), como por el número de países en los que una u otra es oficial (más de 75 en varios continentes), la difusión de su alfabeto compartido o el prestigio de las culturas y los conocimientos a los que sirven de expresión.

3.2 El español y otras lenguas de la Organización de Naciones Unidas

La comparación de la comunidad hispanohablante nativa con las comunidades hablantes de chino mandarín, hindi e inglés tiene un gran interés lingüístico y cultural, al tiempo que permite ubicar al español en su dimensión demográfica. Sin embargo, desde una perspectiva sociopolítica internacional, resulta muy revelador comparar la demografía del español, del inglés y del chino mandarín con la de las demás lenguas oficiales de la ONU, a saber: francés, ruso y árabe.

Un primer escollo que encuentra tal tarea es la de establecer unas cifras de hablantes fiables para cada una de las seis lenguas oficiales de la ONU. Para ello, resulta imprescindible distinguir entre lenguas consideradas como nativas o

maternas (L1) y lenguas segundas y adicionales (L2; dominio no nativo). En este caso, procederemos de igual modo que se hizo con el hindi: atendiendo a las lenguas como nativas o L1. Con este fin, sin duda la fuente más buscada, citada y reproducida es Ethnologue. Otros proveedores de información, como Wikipedia o Statista, no dudan en recurrir al proyecto Ethnologue como referencia. Sin embargo, desde el Observatorio Global del Español hemos abordado un recuento de los hablantes de las seis lenguas en cuestión, partiendo del listado de países que las tienen como oficiales o cooficiales y ampliándolo al resto de países, acudiendo fundamentalmente a fuentes censales. Las cifras buscadas se refieren a hablantes de cada lengua como L1 (nativa) en todos los países del mundo. Los resultados obtenidos por el Observatorio, como era previsible, no coinciden plenamente con los datos que ofrece Ethnologue fundamentalmente porque ni las fechas de los datos manejados ni las fuentes suelen coincidir plenamente en todos los análisis.

CUADRO 7 NÚMERO DE MIEMBROS DE COMUNIDADES NATIVAS (L1) DE LAS LENGUAS OFICIALES DE LA ONU EN TODO EL MUNDO, SEGÚN EL OBSERVATORIO GLOBAL DEL ESPAÑOL Y SEGÚN ETHNOLOGUE

Lengua	Observatorio Global del Español	Ethnologue
Chino mandarín	954.412.674	989.874.880
Español	519.115.258	483.979.360
Inglés	443.268.546	390.348.200
Árabe	224.263.923	334.765.000
Ruso	172.383.133	145.166.390
Francés	139.138.190	74.170.080

FUENTE: elaboración propia con datos del Observatorio Global del Español y de Eberhard, Simons y Fennig (2025).

Como ya se ha detallado, la cifra de hablantes de español con dominio nativo es superior a los 515 millones. El cuadro 7 ofrece las cifras totales de hablantes de español en todos los países en los que se han podido contabilizar los hablantes de cada lengua con dominio nativo. Estas cifras, una vez más no redondeadas por respeto a los datos manejados, ofrecen muchos puntos

críticos: por ejemplo, el recuento de hablantes en países de multilingüismo extremo, como ocurre en muchos territorios de África. En este caso, el francés o el inglés ofrecen un porcentaje de hablantes en el hogar o como primera lengua muy bajo, frecuentemente por debajo del 10 % de la población total, y, entre la población que lo tienen como segunda lengua, resultad difícil saber cuántas personas lo manejan con dominio nativo. Pensemos que en estas situaciones no solo influye la multiplicidad de lenguas, que convierte a cada territorio en un mosaico idiomático, sino que en muchos países es frecuente el conocimiento y el uso de lenguas francas africanas, como el suajili, el wolof, el hausa o el lingala, o incluso de lenguas pidgin (pidgin nigeriano), que recortan espacio comunicativo a las lenguas europeas. Por otro lado, el cuadro 7 no ofrece cifras de hablantes con dominio limitado o como segundas lenguas⁵. Con todo entre las cifras del Observatorio y las de Ethnologue no se observa una total falta de correspondencia en cuanto al volumen relativo de cada comunidad de hablantes, probablemente por la base poblacional total de la que se parte, que en el caso de Ethnologue no está tan actualizada como en la del Observatorio Global del Español.

Más allá de las cifras exactas, así como de los comentarios ya realizados sobre el chino mandarín y el inglés, conviene llamar la atención sobre las notables diferencias existentes entre estas comunidades lingüísticas. Así, para el árabe, por ejemplo, se ofrecen cifras referidas, en el caso de Ethnologue, a un árabe estándar, que no tiene una presencia equiparable en todas las naciones arabófonas. Por otro lado, el árabe se manifiesta en diversas variedades que pueden presentar dificultades notables de intercomprensión, especialmente entre las modalidades occidentales (magrebíes) y las orientales (sudanés, peninsulares...). Entre todas ellas, la variedad egipcia es la de mejor

5. Este es el primer intento de cuantificación de las lenguas internacionales realizado desde el Instituto Cervantes, en este caso desde el Observatorio Global del Español, mediante investigación propia realizada por un equipo de investigación situado en San Millán de la Cogolla (Amalia Ovín, Alejandro Bocigas, Pol Pérez y Enrique Ramírez, bajo la dirección de Francisco Moreno Fernández). La metodología empleada es paralela a la empleada con la lengua española, aunque no ha alcanzado el mismo grado de exhaustividad, lo que pretende paliarse en futuras investigaciones. En cualquier caso, los recuentos se han hecho sobre datos más actualizados que los que manejan otras fuentes y aplicando a todas las lenguas unos mismos criterios.

comprensión en el espacio arabófono, por su presencia en los medios y las artes. En cualquier caso, el mundo árabe se caracteriza por la existencia de situaciones diglósicas, en las que las variedades dialectales ocupan los espacios regionales, coloquiales y familiares, mientras el estándar o el egipcio ocupan las esferas internacionales, públicas y académicas. Así pues, las diferencias respecto a las características del ámbito hispánico son evidentes.

El caso de la lengua rusa remite a un territorio de amplia geografía, pero proporcionalmente de menor población. En el momento actual, la oficialidad del ruso se reconoce apenas en media docena de países (bien como lengua oficial, bien como lengua con carácter estatal), dado que las antiguas repúblicas vinculadas a la URSS han ido prescindiendo de ella conforme se han ido produciendo conflictos con la Federación de Rusia. Esto no ha supuesto la desaparición radical de las comunidades rusohablantes en esas repúblicas, pero sí la postergación pública del ruso en la vida nacional. En el caso de los países hispanohablantes, no todos ellos tienen el español como lengua oficial (así ocurre en Argentina, México o Uruguay), pero este sí conserva su valor vehicular, general o nacional *de facto*.

Finalmente, en lo que se refiere al francés, podrían establecerse paralelismos con el español, fundamentados en las trayectorias históricas de España y de Francia, pero la realidad es que los estilos coloniales, en un caso y en otro, fueron muy diferentes, como ocurrió con las colonias británicas, cuyo resultado geopolítico derivó en un conjunto de países, principalmente en África, en los que el francés es L1 de una relativamente baja proporción de la población, mientras que el español es L1 de la mayoría de los habitantes de los países hispánicos.

3.3 El español y el desarrollo humano

No resulta conveniente, cuando de comparar situaciones idiomáticas se trata, limitarse a la información cuantitativa que aporta la demografía de las lenguas. Ciertamente, las comunidades lingüísticas se definen por las características de las propias lenguas, así como por los rasgos esenciales de su

demografía, pero existen otros factores que permiten completar la imagen de tales comunidades y que tienen que ver con la política, la sociología o la economía. Por tal motivo, interesa en este momento comparar los perfiles de las comunidades hablantes de las lenguas oficiales de la ONU en lo que se refiere a su nivel de desarrollo humano y a otros aspectos coadyuvantes, como son el nivel de pobreza o el nivel de igualdad de género.

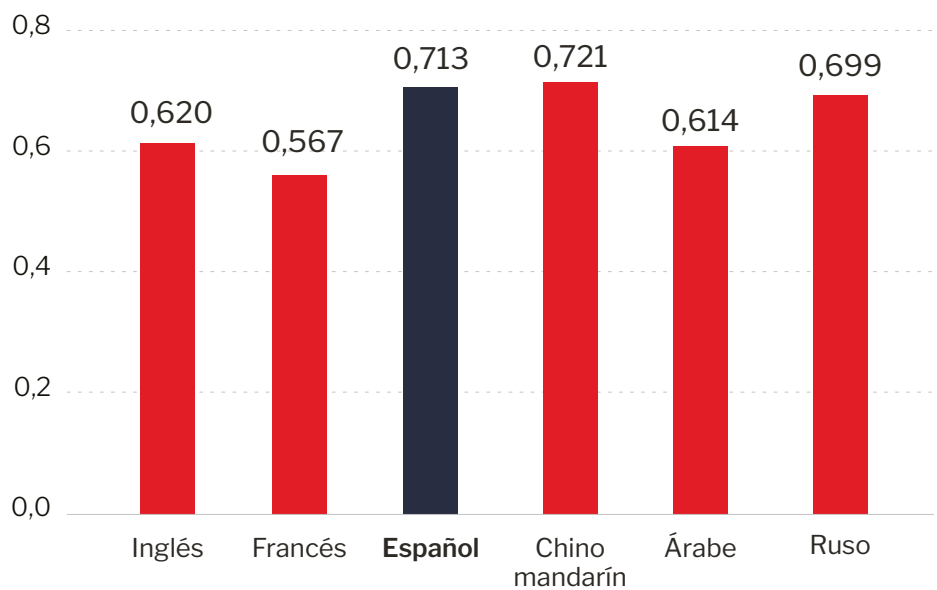
El más reciente *Informe sobre desarrollo humano 2023/2024* publicado por el Programa de Naciones Unidas para el Desarrollo (PNUD, 2024a) ofrece un buen retrato del momento que están atravesando las naciones del mundo. Parte del estado en 2021-2022, cuando la incertidumbre de la pandemia parecía frenar el desarrollo humano y quebrar la evolución de las naciones. En aquel momento, el valor global del Índice de Desarrollo Humano (IDH) conoció un descenso por primera vez en su historia. Como es sabido, el IDH es un índice compuesto que mide el resultado promedio de tres dimensiones básicas: longevidad y salud, conocimiento y nivel de vida (PNUD, 2024b). El informe 2023-2024 muestra un repunte del IDH que en 2023 ha alcanzado un máximo histórico, con alza de todos sus componentes.

Asimismo, el informe nos alerta de dos hechos relevantes que afectan a las comunidades lingüísticas transcontinentales. Por un lado, se constata un aumento de la desigualdad entre los países con IDH muy alto y bajo, contrariamente a lo que se esperaría que debe ocurrir a largo plazo. Por otro lado, se evidencia que el cambio climático puede provocar un marcado crecimiento de las desigualdades, que se hace patente en el número de muertes provocadas por los fenómenos asociados a tal cambio. Dentro del espacio hispanohablante, mientras los países americanos ven estabilizado ese específico volumen de muertes, los países desarrollados, entre los que se encuentra España, conocerían un descenso del mismo.

Para nuestro análisis, hemos comparado el IDH promedio del conjunto de países hablantes de cada una de las lenguas oficiales en la ONU, con un resultado proporcional no muy diferente del obtenido mediante el cálculo de estos índices en informes anteriores. La comparación de las seis comunidades

idiomáticas nos muestra unas poblaciones sinohablante e hispanohablante con un promedio de IDH que lidera el desarrollo en el concierto idiomático (chino mandarín: 0,721; español: 0,713) y una comunidad francófona (0,567) cuyo promedio queda relegado al último lugar. Obviamente, son los promedios los que arrojan estos resultados, pues los valores de la francofonía dependen del bajo IDH de muchas naciones africanas, como ocurre con el inglés, mientras que el chino mandarín se ve privilegiado por los altos índices de Hong-Kong o Singapur. En el caso del mundo hispanohablante, menos desigual entre sí en cuanto a los valores del IDH, es el alto promedio general el que beneficia el IDH del conjunto, sustentado en el desarrollo humano etiquetado como «muy alto» de España, Chile, Panamá y Costa Rica.

GRÁFICO 2 PROMEDIOS EN 2022 DEL ÍNDICE DE DESARROLLO HUMANO (IDH) EN COMUNIDADES HABLANTES DE LAS LENGUAS OFICIALES DE LA ONU (VALORES ENTRE 0-1)



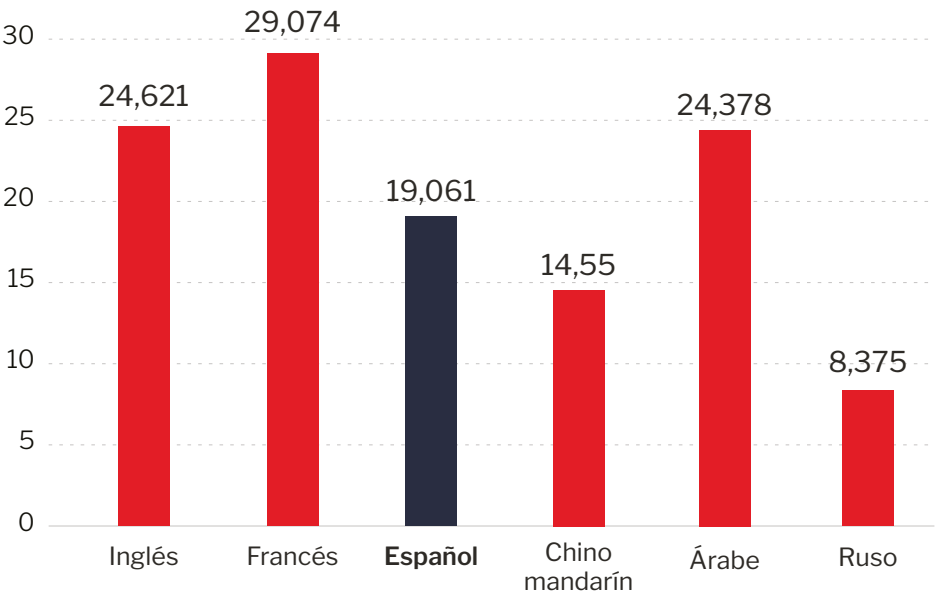
FUENTE: PNUD (2024a).

EL IDH, sin embargo, al ser un valor complejo, encierra matices que también merecen comentario, aparte del difuminado que supone su tratamiento en forma de promedios. Para demostrarlo, nos centraremos en dos de esos matices: el relativo a la pobreza multidimensional y el relativo a la desigualdad de género.

La pobreza multidimensional es aquella que no se define únicamente por la falta de dinero, sino por la privación o la dificultad de acceso a bienes y servicios, como buenos centros médicos, viviendas de calidad, empleo o educación, junto a la capacidad de empoderamiento y un adecuado estado físico (p. ej. masa corporal) (Multidimensional Poverty Peer Network [MPPN], s. f.). Son, pues, varias las dimensiones que afectan a la pobreza, de ahí su denominación. La pobreza multidimensional se cuantifica mediante un índice que expresa el porcentaje de la población que la sufre, ajustado según la intensidad de las privaciones.

Con el fin de comparar el Índice de Pobreza Multidimensional (IPM) en las comunidades hablantes nativas de las lenguas oficiales de la ONU, hemos procedido de nuevo mediante promedios, con el resultado que se observa en el gráfico 3. En líneas generales, se aprecia una proporcionalidad inversa con el IDH de estas comunidades, de modo que las comunidades hablantes de chino mandarín y de español, revelan un IPM más bajo que el de las comunidades francófona y anglófona. Sin embargo, llama la atención el proporcionalmente elevado IPM de los países arabófonos. En este caso, la comunidad rusohablante es la que presenta un índice de pobreza más bajo.

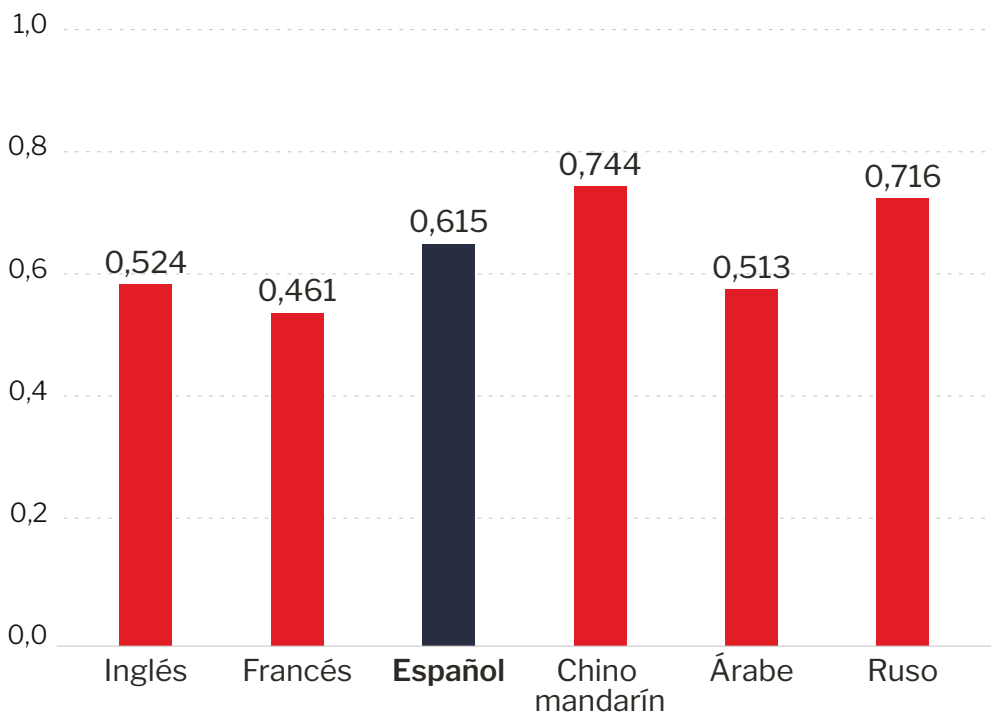
GRÁFICO 3 PROMEDIOS EN 2022 DEL ÍNDICE DE POBREZA MULTIDIMENSIONAL EN COMUNIDADES HABLANTES DE LAS LENGUAS OFICIALES EN LA ONU



FUENTE: PNUD (2024a).

Finalmente, el Índice de Desigualdad de Género (IDG) es un valor compuesto que refleja la desigualdad entre mujeres y hombres en tres dimensiones: salud reproductiva, empoderamiento y mercado de trabajo (PNUD, 2024b). Resulta muy interesante observar la falta de correlación directa entre los índices anteriores y un aspecto tan significativo en la vida social de las comunidades como la desigualdad de género. En este caso, son las comunidades hablantes de chino mandarín (0,744) y de ruso (0,716) las que revelan un promedio de mayor desigualdad, seguidas, aunque algo de lejos por la comunidad hispanohablante (0,615). El IDG es menos prominente en el caso de las comunidades anglófona (0,524) y francófona (0,461), y llama la atención, en contra de lo que generalmente suele pensarse en Occidente, que los países árabes presenten un IDG relativamente bajo e inferior al de la comunidad anglohablante. En todo caso, a este respecto aún queda un largo trecho que recorrer para todas las comunidades idiomáticas.

GRÁFICO 4 PROMEDIOS EN 2022 DEL ÍNDICE DE DESIGUALDAD DE GÉNERO EN COMUNIDADES HABLANTES DE LAS LENGUAS OFICIALES EN LA ONU (VALORES ENTRE 0-1)



FUENTE: PNUD (2024a).

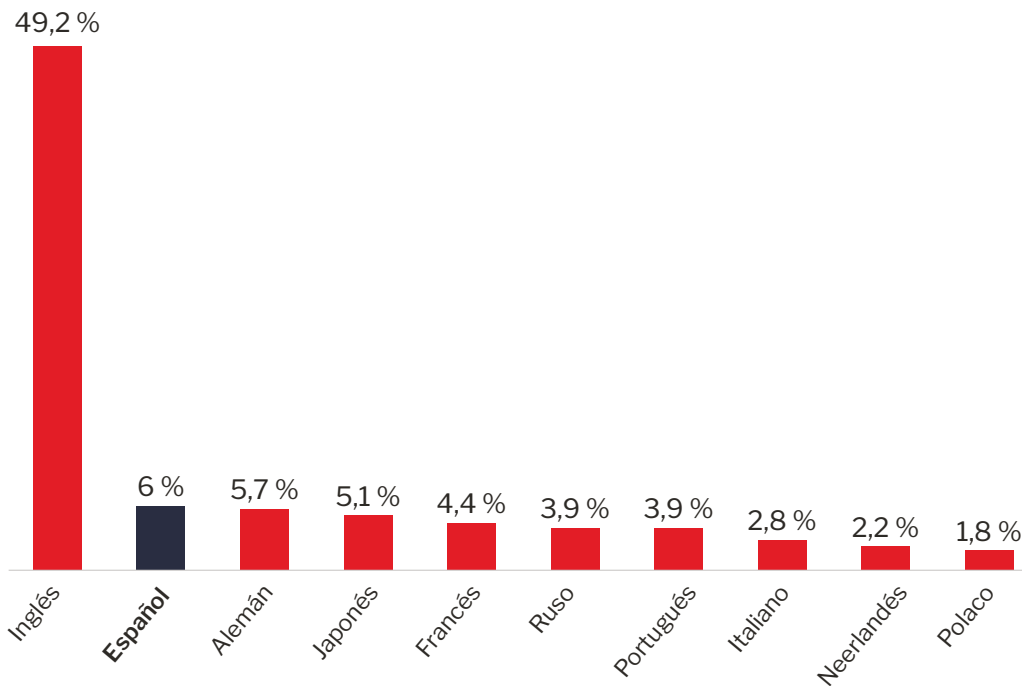
3.4 El español en las redes

El mundo virtual también concede a la lengua española un lugar de relevancia. Las páginas web que ofrecen contenidos en español suponen en la actualidad el 6 % de todas las páginas del mundo, porcentaje que sitúa a la lengua española por encima de la alemana, la japonesa o la francesa, por mencionar algunas otras. A este respecto, sin embargo, ninguna lengua puede hacerle sombra al inglés, puesto que da contenido al 49 % del total de las páginas web. Con todo, resulta interesante conocer que la tendencia de las páginas web en español es de crecimiento (1,5 % desde 2015, con 0,4 % en 2024), mientras que el peso de las páginas en inglés ha descendido aproximadamente un 6,2 % desde 2014; las páginas en ruso, un 2,2 %; y las páginas en alemán, un 0,3 %⁶.

Gráfico 5 · Página siguiente >

6. Optamos por presentar las estadísticas de W3Tech por ser datos obtenidos automáticamente de páginas web que se consideran relevantes por su contenido y funcionalidad: se excluyen los sitios duplicados, redirigidos o subdominios, así como los que solo muestran una página del servidor, por ejemplo. Otras fuentes utilizan criterios diferentes para cuantificar el peso de las lenguas en internet. Así, el Observatorio de la Diversidad Lingüística y Cultural en Internet (OBDILCI), con sede en Niza (Francia) y ligado a la Organización Internacional de la Francofonía y previamente a Unión Latina, afirma que el 20 % de las páginas en internet están en inglés, el 19 % en chino y el 7,7 % en español. Pero, al mismo tiempo, el OBDILCI ha propuesto un Índice de Ciberglobalización (CGI) que incluye como factores el número de hablantes de las lenguas, el número de países donde se hablan y el número de hablantes conectados a internet. Este índice complejo cuantifica el inglés con un 14,13 % de la ciberglobalización mundial, el francés con un 9,55 % y el español con un 2,09 % (OBDILCI, 2024). Las cifras de hablantes se toman de Ethnologue y otras fuentes, con un intervalo de confianza del 20 %.

GRÁFICO 5 PORCENTAJES DE SITIOS WEB DE CONTENIDO POR LENGUA (FEBRERO 2025)



FUENTE: WW3Techs (s. f.).

Sin duda, el desarrollo de la presencia del español en internet está ligado a la demanda por parte de los usuarios, que a su vez se correlaciona con el crecimiento del uso de la telefonía móvil. En este sentido, los usuarios de internet móvil en América Latina están experimentando un notable incremento, al pasar de 418 millones en 2023, con una tasa de penetración del 65 %, a la estimación de 485 millones en 2030, con una tasa de penetración del 72 %. Asimismo, el porcentaje de conexiones desde teléfonos inteligentes pasará del 80 % en 2023 al 92 % en 2030 (GSMA, 2024a). Por tipo de contenido, en América Latina, hay tres usos que concentran el grueso del tráfico en internet: las redes sociales, la navegación web y los directos, y en los tres casos se hace un uso intensivo de vídeos, con su posible influencia sobre la lengua hablada (GSMA, 2024b).

Por otro lado, merecen conocerse las conclusiones de un estudio sobre la complejidad lingüística de los comentarios hechos por los usuarios en múltiples plataformas digitales (Facebook, Twitter,

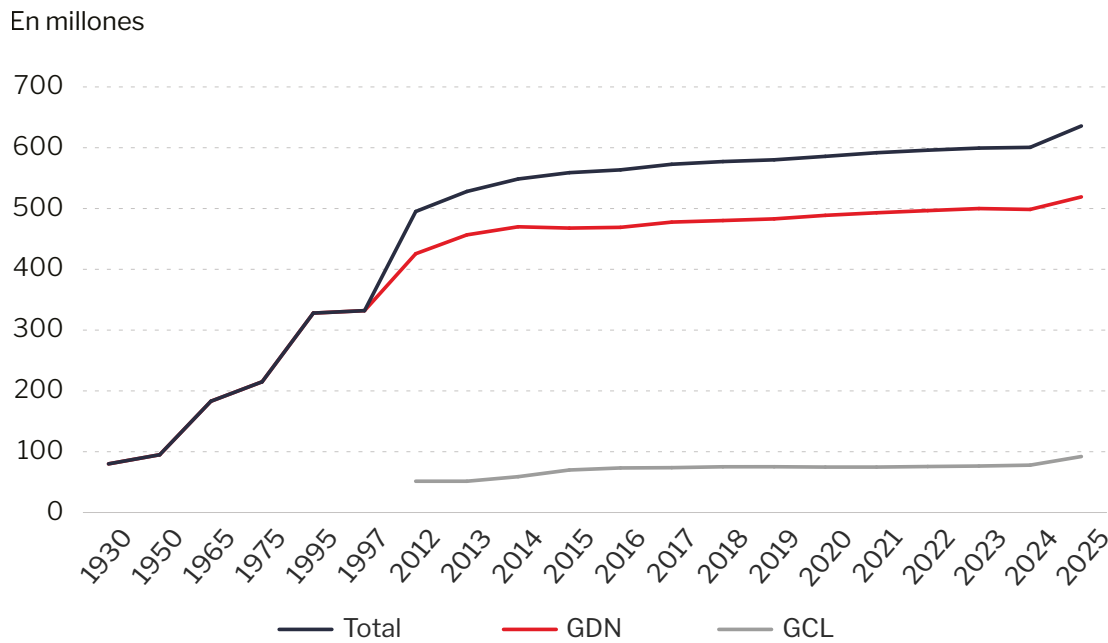
YouTube, Voat, Reddit, Usenet, Gab y Telegram), realizado sobre 300 millones de comentarios en inglés. Sus resultados revelan una tendencia constante a la disminución de la longitud de los textos y de la riqueza léxica, aunque se introducen constantemente nuevas palabras. Estas tendencias, observadas para la lengua inglesa, probablemente también podrían ser observables en el español y en otras lenguas del mundo (Di Marco *et al.*, 2024).

Finalmente, en cuanto a otros aspectos relativos al español en las redes, merece una mínima reflexión la presencia del español en la aplicación de aprendizaje de idiomas Duolingo. El informe de 2024 de esta organización revela que, durante ese año, el español ha sido el segundo idioma más popular, tras el inglés, que se está aprendiendo en 33 países (frente a los 135 del inglés y los 16 del francés) y que los países donde este idioma ha sido más demandado han sido Estados Unidos, los países nórdicos, Sudáfrica y Australia (Blanco, 2024).

4. La proyección del español

La situación actual de la lengua española y de sus comunidades de hablantes en el mundo no es producto de una coyuntura contemporánea, sino consecuencia de una evolución que tuvo en el siglo ^{xx} uno de sus más significativos impulsos. Las escasas y poco precisas referencias demolingüísticas anteriores a los pasados años noventa narran una evolución del español desde los años treinta, cuando la comunidad hispánica no alcanzaba los 100 millones de hablantes, hasta la última década del siglo ^{xx}, cuando se sobrepasaron los 300 millones. Esa evolución describe una asíntota que encuentra su inflexión en los años setenta y ochenta, cuando la demografía hispana conoció un gran despegue como consecuencia de la transición demográfica iniciada en los años sesenta (Agar Corbinos y Ferrer, 2001), al tiempo que Estados Unidos comenzó a dar visibilidad a su población hispanohablante mediante la utilización en el censo de la etiqueta «hispano».

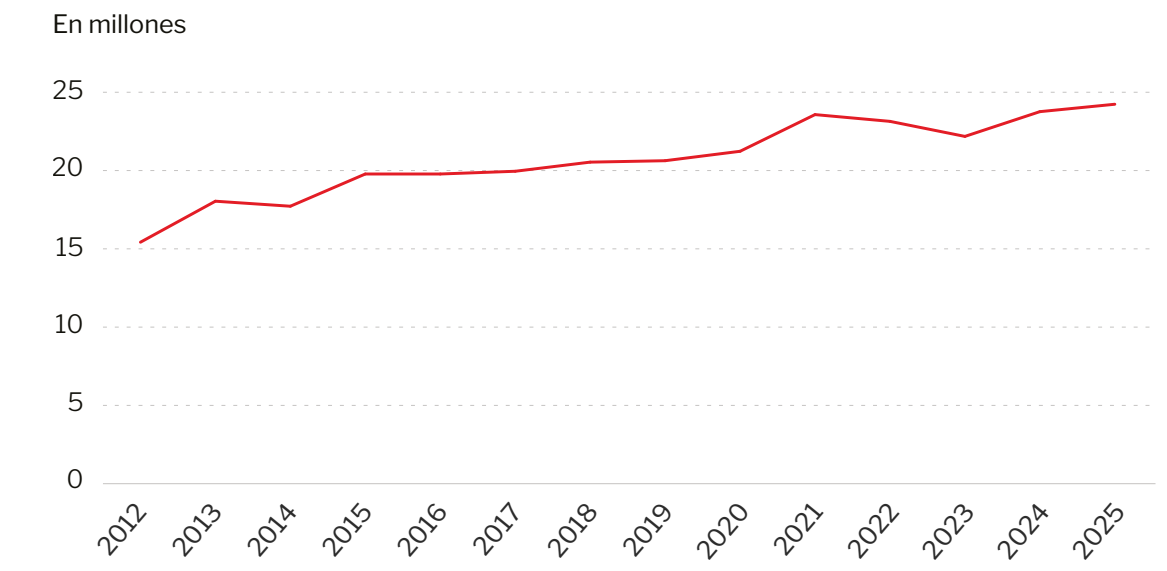
GRÁFICO 6 DESARROLLO DE LA DEMOGRAFÍA DE LOS HABLANTES POTENCIALES DE ESPAÑOL, POR DOMINIO NATIVO Y COMPETENCIA LIMITADA (1930-2025)



FUENTE: 1930 (Herrero Mayor, 1944); 1950 (Rosenblat 1977); 1965 (Rosenblat, 1977: 183.000; Pottier, 1968: 130.000); 1996 (Gordon y Grimes, 2005: 266.000; Moreno Fernández y Otero Roth, 1998: 350.000; Hernández Alonso, 1994: 450.000); 2012-2023 (Instituto Cervantes, anuarios del 2012 al 2023); 2024 (Observatorio Global del Español, 2024).

Como es sabido, desde 1992 el Instituto Cervantes viene haciendo un seguimiento de la evolución del español que nos ha proporcionado detallada y valiosa información sobre su desarrollo en cuanto al uso y al aprendizaje. En esa línea continuada de análisis, se observa el progresivo crecimiento de la población hablante de español en todo el mundo (países hispanohablantes y no hispanohablantes), con distinto dominio de la lengua, hasta superar los 630 millones de hablantes potenciales en 2024. El crecimiento de los potenciales hablantes de español insertos en procesos de aprendizaje también muestra una persistencia, observada desde los algo más de 18 millones de aprendices recontados en 2012 hasta los más de 24 millones de la actualidad. En ese proceso se produjo una disrupción severa motivada por la pandemia del COVID-19, que produjo un retroceso de prácticamente tres años respecto del estado inmediatamente anterior, y que solo en 2025 parece haber recuperado su impulso crecedero.

GRÁFICO 7 DESARROLLO DEL GRUPO DE APRENDICES DE LENGUA ESPAÑOLA (GALE) EN EL MUNDO (2012-2025)



FUENTE: 2012-2023 (Instituto Cervantes, anuarios del 2012 al 2023); 2024 (Observatorio Global del Español (2024)).

La comparación de los diferentes grupos de hablantes de español que la serie de estudios del Instituto Cervantes ha considerado nos permite apreciar que, entre sus porcentajes de crecimiento, los grupos que proporcionalmente más han aumentado en su total son los hablantes con competencia limitada (79 %) y los aprendices de español (36 %), mientras que los hablantes de dominio nativo, más pegados a la pura demografía, han crecido un 22 %.

CUADRO 8 EVOLUCIÓN DEL NÚMERO DE HABLANTES POTENCIALES CON DOMINIO NATIVO (GDN), COMPETENCIA LIMITADA (GCL) Y APRENDICES DE ESPAÑOL (GALE) ENTRE 2012 Y 2025

Año	GDN	GCL	GALE	Total
2012	425,636,709	51,442,310	18,017,913	495,096,932
2013	456,676,339	51,442,310	19,962,500	528,081,149
2014	469,949,593	58,905,216	19,724,511	548,579,320
2015	467,801,452	69,922,994	21,252,789	558,977,235

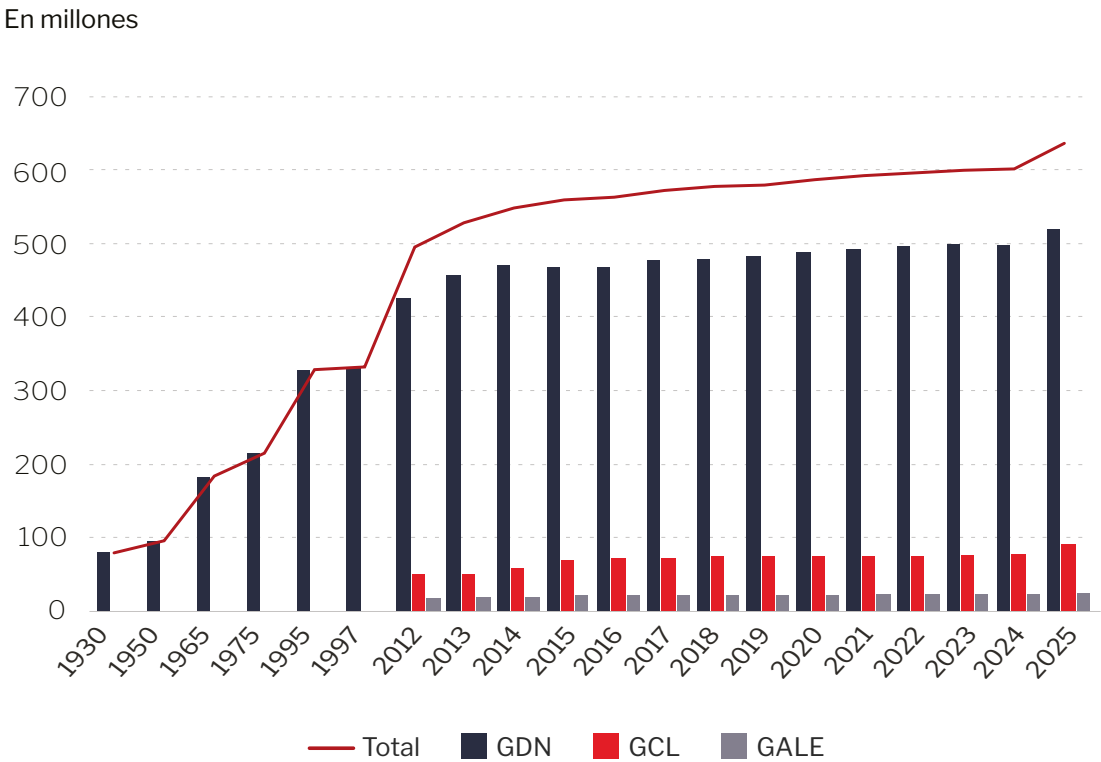
Año	GDN	GCL	GALE	Total
2016	469,037,629	73,284,395	21,252,789	563,574,813
2017	477,659,770	73,749,162	21,382,247	572,791,179
2018	480,229,642	75,201,405	21,815,280	577,246,327
2019	482,949,086	75,267,790	21,882,448	580,099,324
2020	488,767,362	74,710,449	22,329,275	585,807,086
2021	492,990,519	74,716,239	24,069,206	591,775,964
2022	496,573,842	75,625,447	23,748,298	595,947,587
2023	499,947,796	76,422,128	23,035,198	599,405,122
2024	498,497,757	77,901,236	24,208,813	600,607,806
2025	519,115,258	92,068,243	24,560,143	635,743,644
% Δ	22%	79%	36%	28%

FUENTE: 2012-2023 (Instituto Cervantes, anuarios del 2009 al 2023); (Observatorio Global del Español (2024).

A este respecto, no debe dejar de mencionarse que el cálculo anual de estas proporciones ha podido verse afectado por cuestiones metodológicas, como el refinamiento metodológico y las permanentes actualización y revisión de los datos, así como el descubrimiento progresivo de nuevas bolsas de hablantes, por mucho que las decisiones técnicas siempre se hayan inclinado hacia las soluciones más conservadoras. Tal conservadurismo responde fundamentalmente a una prudencia investigadora, siempre necesaria en el manejo de datos de naturaleza social, aunque también ha podido influir el deseo de evitar la sospecha de sesgo institucional. En todo caso, es evidente que el grupo de hablantes potenciales de español más claramente destacado en volumen es el de dominio nativo, muy por encima del número de hablantes con dominio o competencia socioestilísticamente limitada y de los aprendices de español. Del mismo modo, se hace innegable que los grupos de competencia y habilidades limitadas, y de aprendices son los que ofrecen un más amplio margen de crecimiento, debido a su propia naturaleza y dinámica. El gráfico

8 presenta el volumen de hablantes con diferentes dominios a lo largo del último siglo, de acuerdo con los datos disponibles.

GRÁFICO 8 DESARROLLO DE LA DEMOGRAFÍA DE LOS HABLANTES POTENCIALES DE ESPAÑOL (1930-2025))



FUENTE: 1930 (Herrero Mayor, 1944); 1950 (Rosenblat 1977); 1965 (Rosenblat, 1977: 183.000; Pottier, 1968: 130.000); 1996 (Gordon y Grimes, 2005: 266.000; Moreno Fernández y Otero Roth, 1998: 350.000; Hernández Alonso, 1994: 450.000); 2012-2023 (Instituto Cervantes, anuarios del 2012 al 2023); 2024 (Observatorio Global del Español, 2024).

Junto al ejercicio de análisis de la evolución de los grupos de hablantes potenciales de español, merece la pena el intento de proyectar la evolución de la comunidad hispanohablante en las próximas décadas. No es este último un ejercicio que merezca llevarse a cabo anualmente, dado que las tendencias demográficas no suelen modificarse de un año para otro, pero sí conviene presentarlo con cierta periodicidad, aunque sea a modo de recordatorio.

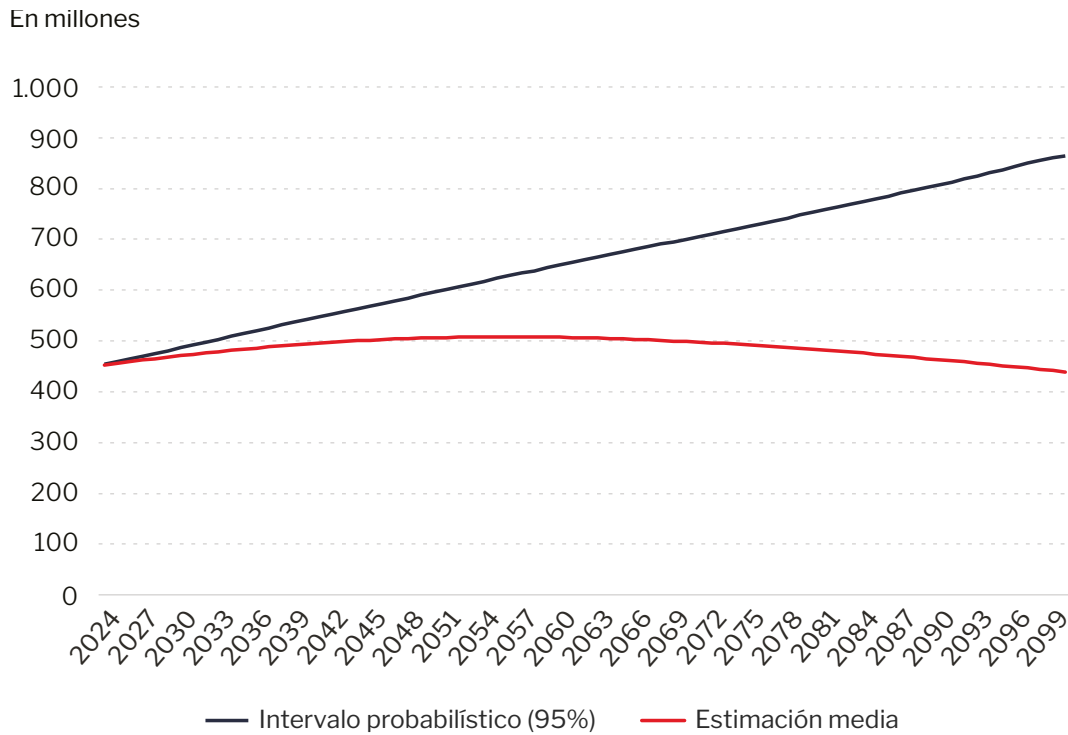
El panorama mundial entre 2025 y 2100 nos presenta una demografía en crecimiento, que comenzará a ralentizarse en 2050 y que alcanzará su cima en

2086, con 10.430 millones de personas (ONU, 2022b; Banco Mundial, 2025). A partir de ahí, la curva poblacional comenzará su descenso camino del 2100. Para entonces, el 80 % de la población mundial vivirá en África y en Asia, si bien hacia 2050 será África el continente que muestre un crecimiento más rápido y la mitad de la población mundial vivirá en los siguientes nueve países: India, Nigeria, Pakistán, República Democrática del Congo, Etiopía, Estados Unidos, Indonesia y Uganda. Si trasladamos la demografía al ámbito idiomático, obtendríamos como resultado que, entre las lenguas oficiales de la ONU, la inglesa sería la más beneficiada de tal crecimiento, como L1 y como L2, al gozar de oficialidad en cinco de esos nueve países (India, Nigeria, Pakistán, Estados Unidos y Uganda), mientras que el francés es oficial solamente en uno de ellos (República Democrática del Congo). Dentro de este paisaje demolingüístico, el español no disfrutaría de las ventajas de una demografía sobresaliente.

Para obtener una proyección de los hablantes nativos de español (L1) hasta el año 2100, este informe ha partido de la proyección poblacional de cada uno de los países hispanohablantes, calculada por la ONU, con el resultado que se presenta en el gráfico 9. En él se dibujan dos curvas que representan, en primer lugar, una estimación media de la comunidad hispanohablante nativa: esa curva alcanzaría su punto máximo a mediados de siglo y experimentaría un paulatino descenso, de la mano de la demografía, en el último tercio de siglo. En segundo lugar, el gráfico presenta una línea de posible crecimiento máximo que, aunque estadísticamente posible, no parece demográficamente probable: quizá sea más previsible una asíntota que un crecimiento totalmente lineal. En cualquier caso, esta proyección debe tratarse con sumo cuidado porque podría subestimar el impacto de las migraciones.

Gráfico 9 · Página siguiente >

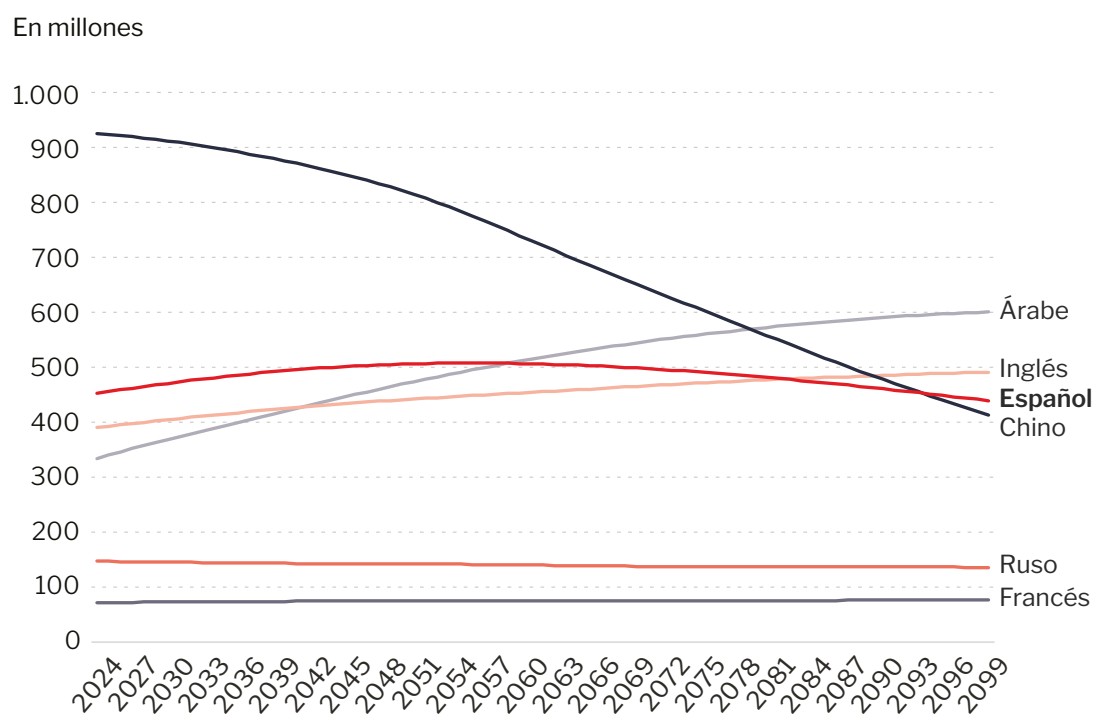
GRÁFICO 9 PROYECCIÓN DE LA COMUNIDAD DE HABLANTES NATIVOS DE ESPAÑOL (L1) HASTA 2100



FUENTE: elaboración propia a partir de datos de la ONU (2022b) y Eberhard, Simons y Fennig (2025).

Si comparamos la proyección de la comunidad hispanohablante con la correspondiente a las otras cinco lenguas oficiales de la ONU, observamos comportamientos muy dispares en su proyección, que se muestran ligados a la evolución demográfica de cada comunidad lingüística.

El gráfico 10 ofrece unas proyecciones basadas en números absolutos de población estimada. Para estimar el número de hablantes nativos se considera el porcentaje de hablantes de cada lengua como primera lengua (L1), según los datos de Ethnologue, en los países donde es lengua oficial en 2025. De este modo, la actual situación demolingüística de las lenguas internacionales como lenguas nativas para cada país se proyecta sobre las previsiones demográficas de la ONU para cada uno de ellos. Así, el escenario proyectado considera, por ejemplo, que la proporción de hablantes de francés en Canadá se mantiene constante (alrededor del 20 %) a lo largo de la serie temporal.

GRÁFICO 10 PROYECCIÓN DE LAS COMUNIDADES DE HABLANTES NATIVOS (L1) DE LAS LENGUAS OFICIALES DE LA ONU HASTA 2100

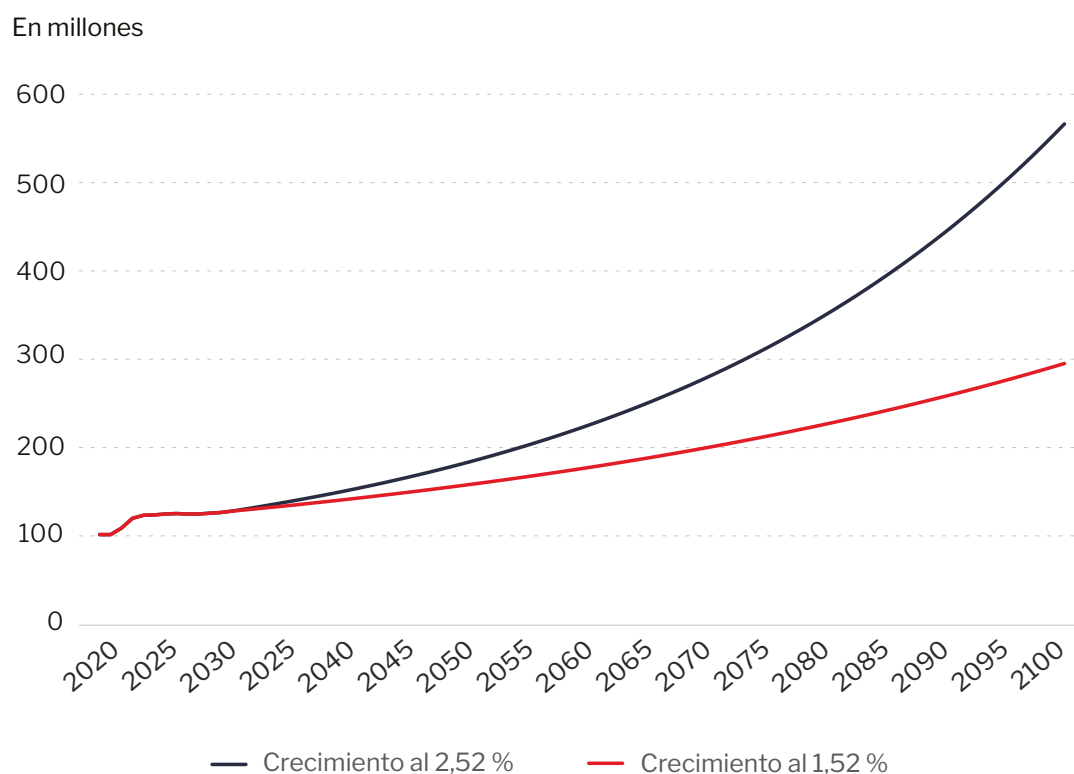
FUENTE: elaboración propia a partir de datos de la ONU (2022b) y Eberhard, Simons y Fennig (2025).

En lo que se refiere al resultado de las proyecciones, llama la atención tanto la reducción de hablantes de chino mandarín como el crecimiento de las comunidades arabófonas, ambos casos ligados a los porcentajes poblacionales previsibles para cada comunidad. Asimismo, podría plantearse la duda sobre la proyección obtenida para la lengua francesa: si África ha de tener el mayor desarrollo demográfico mundial, buena parte del continente es francófono. Para desentrañar esta paradoja, habría que tener en cuenta, por un lado, que la proporción de hablantes nativos de francés (L1) es relativamente baja en las naciones africanas; por otro, sin embargo, cabe suponer que, si la alfabetización y la escolarización se generalizan en estos países, el porcentaje de hablantes de francés con dominio nativo —aunque no lo tengan como L1— habría de crecer necesariamente, aspecto que no queda reflejado en el gráfico. Con todo, el informe *La lengua francesa en el mundo* de 2022 señalaba que, desde 2010, ningún país del espacio africano había experimentado

un cambio significativo en el porcentaje de su población francófona, en su mayoría por debajo del 50 %, en relación con su población total, lo que podría considerarse como una señal débil que presagia un escenario poco optimista (Organización Internacional de la Francofonía, 2022).

Volviendo a las proyecciones relativas específicamente al español, uno de los aspectos más interesantes sería el que afecta al posible futuro de los hablantes que aquí se consideran de dominio o competencia limitada (GCL), en términos de habilidades socioestilísticas y de fluidez, dado que es uno de los grupos de hablantes con mayor potencial de crecimiento. En este sentido, hay que tener en cuenta que durante los últimos 13 años el crecimiento anual del GCL ha sido del 3,52 %. Si se mantuviera este mismo ritmo de crecimiento, en 2100 se alcanzarían los 1.000 millones de personas hablantes potenciales de español; es decir, un 11 % de la población mundial. Podría pensarse que tal proyección sobrepasa unos límites razonables, dados los condicionamientos a que se ve sometida su base empírica, pero podría tenerse en cuenta igualmente que, por ejemplo, en la Unión Europea, según el *Eurobarómetro especial 540* (2024a) el 7 % de la población que habla español como lengua extranjera sobrepasa el 12 % en la población menor de 15 años, por lo que en una sola generación podrían darse saltos significativos.

En cualquier caso, dado que el GCL depende parcialmente de los hablantes egresados de los sistemas educativos, así como de los descendientes de inmigrantes, y dado que ambos factores se supeditan a elementos coyunturales, hemos optado por ofrecer dos proyecciones a la baja, donde el GCL crece un punto porcentual menos en cada una de ellas; es decir, un 2,52 % y un 1,52 % anualmente.

GRÁFICO 11 PROYECCIÓN DE CRECIMIENTO DEL GRUPO DE COMPETENCIA LIMITADA HASTA 2100

FUENTE: elaboración propia a partir de datos de la ONU (2022b).

La proyección que ofrece el gráfico 11 para el español dibuja un panorama con más de 200 millones de hablantes con dominio limitado, en su estimación más conservadora, y de más de 500 millones en su proyección menos conservadora. Si se contrasta con las previsiones demográficas del grupo de hablantes nativos, se evidencia un escenario futuro en que el peso del grupo de hablantes de dominio o competencia limitada será mucho mayor que en la actualidad (1 de cada 3 hispanohablantes). Esto tiene varias implicaciones. La más relevante sería que se prevé que la dimensión internacional del español aumente, ya que la mayor parte de estos hablantes con dominio limitado de la lengua son hablantes de español como lengua extranjera que residen en países no hispanohablantes.

Las cautelas expresadas en la proyección de hablantes de dominio limitado se aprecian e incluso se acrecientan en la posible proyección de los

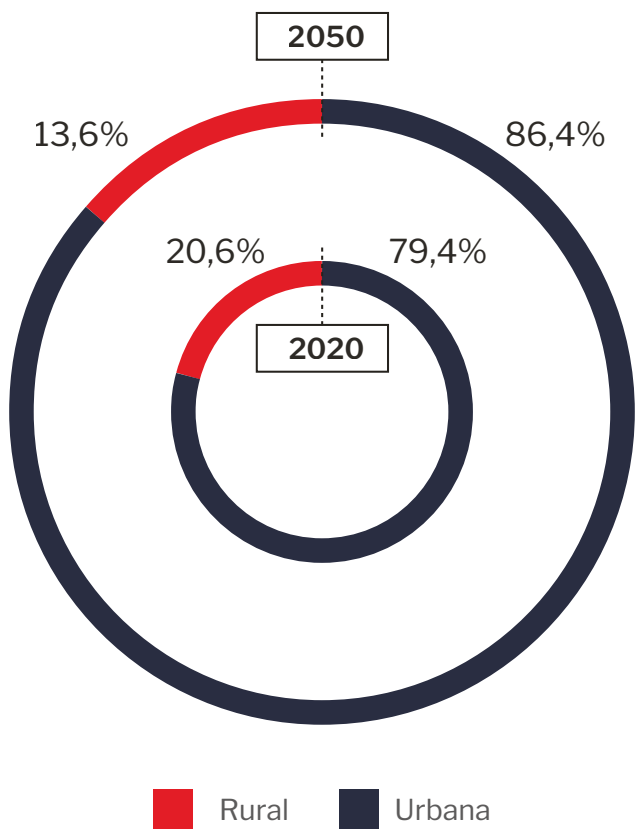
aprendices de español. Por eso no se ofrece aquí una proyección del número de aprendices de español como lengua extranjera. De modo general, realizar proyecciones demolingüísticas sobre el número de aprendices de una lengua resulta una tarea muy insegura por varias razones. Una de las más importantes es la dificultad de controlar las variables que influyen en el interés por las lenguas extranjeras; otra es su institucionalización en los sistemas de enseñanza. Si bien se sabe que el número de personas que aprenden una lengua depende de varios factores sociales asociados con las relaciones internacionales, sobre todo con las económicas (Ginsburgh, Melitz y Toubal, 2014), lo cierto es que existen otros factores decisivos, como las políticas educativas relativas a la enseñanza de lenguas extranjeras, que son difíciles de predecir. Además, en el caso del español encontramos una dificultad metodológica adicional: la base de datos disponible para las proyecciones, construida desde 2012 para la elaboración de los informes del Instituto Cervantes, ha ido ampliando, ajustando y mejorando su base documental en sus sucesivas ediciones, de modo que no sería posible identificar los cambios debidos a la evolución misma del número de aprendices y los debidos a la incorporación de nuevas bolsas de aprendices o al ajuste de los criterios de cuantificación.

Un último ejercicio de proyección nos lleva a comparar la evolución de la población rural y urbana en los países hispanohablantes entre 2020 y 2050. El interés lingüístico de esta comparación reside, por un lado, en que las ciudades son el principal hábitat de uso de las lenguas oficiales, nacionales o vehiculares (en nuestro caso, el español), mientras que el mantenimiento de las lenguas indígenas, originarias o minoritarias, cuando se produce, lo hace generalmente en las áreas rurales.

Las poblaciones totales de los países hispanohablantes muestran con claridad un predominio de la población urbana sobre la rural, tanto en 2020 como en 2050: en la primera referencia temporal, la proporción de población rural apenas supera el 20 %; en la segunda referencia, la población rural desciende al 13,6 %. Esto significa que, en las décadas más próximas, la urbanización seguirá siendo un proceso intenso en el espacio

hispanohablante, que irá en detrimento del mantenimiento de las lenguas originarias y minoritarias y que acrecentará el volumen proporcional de hablantes de español, especialmente de los que lo hablan con un dominio limitado, como podría ser el de la primera generación de migrantes internos desde las poblaciones rurales originarias a las ciudades.

GRÁFICO 12 POBLACIÓN URBANA Y RURAL DE LOS PAÍSES HISPANOHABLANTES EN 2020 Y 2050

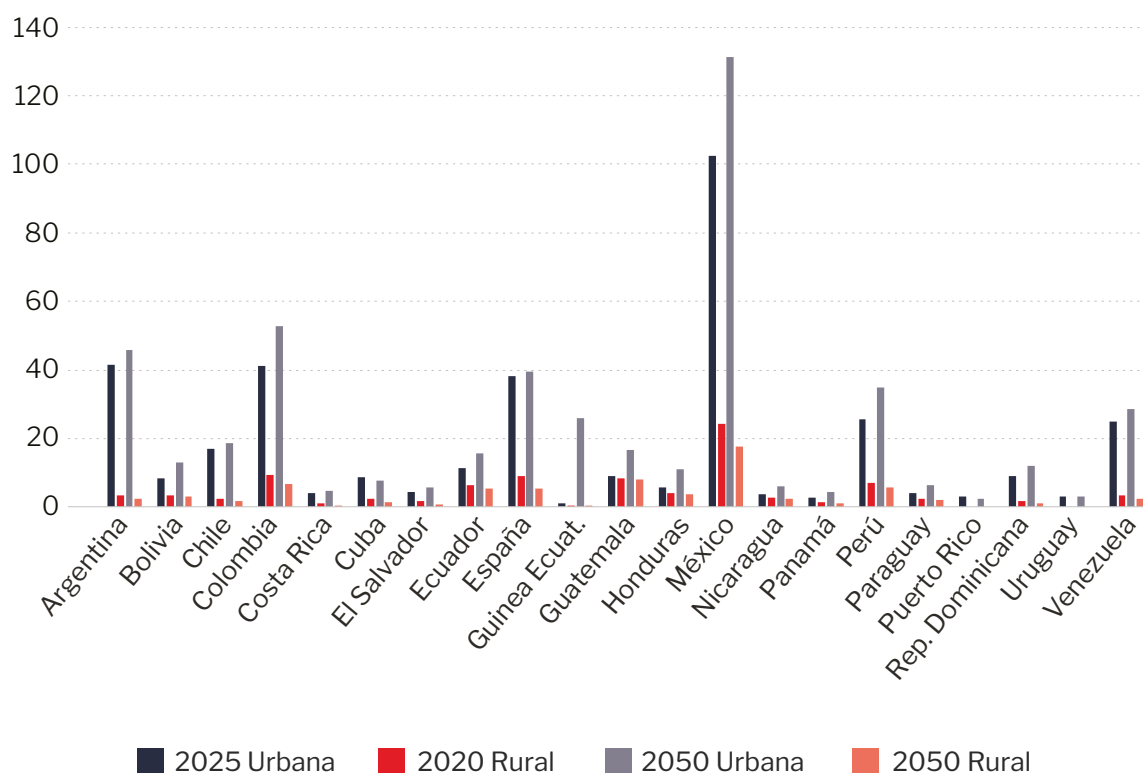


FUENTE: Banco Mundial (s.f.-b).

Si prestamos atención a la evolución de la población rural y urbana en cada uno de los países hispanohablantes, observamos exactamente el mismo proceso de desruralización o urbanización que señalaban las cifras totales. Prácticamente en todos los países, entre 2020 y 2050, la población urbana crecerá en números absolutos y la rural decrecerá, si bien será México el país en que el despegue urbano y el hundimiento rural se hará más evidente.

GRÁFICO 13 POBLACIÓN URBANA Y RURAL DE LOS PAÍSES HISPANOHABLANTES EN 2020 Y 2050 (POR PAÍS)

En millones

**FUENTE:** Banco Mundial (s.f.-b).

Estos datos apuntan, por tanto, a un mayor fortalecimiento del español en los países hispanohablantes y a una minoración demográfica de las áreas rurales, tradicional repositorio de las hablas más tradicionales. En aquellos países que ya cuentan con una población rural pequeña, como Puerto Rico o Uruguay, estos hechos tendrán poca incidencia sobre las variedades de su español.

Referencias bibliográficas

- AGAR CORBINOS, L. y FERRER, M. (2001). Démographie et contexte socio-économique en Amérique Latine et dans les Caraïbes. En F. Gendreau (Dir.), *Les transitions démographiques des pays du sud* (pp. 25-40). Éditions Stem. https://horizon.documentation.ird.fr/exl-doc/pleins_textes/divers17-09/010027446.pdf
- AGENCIA CUBANA DE NOTICIAS (2024). Actualizan cifras sobre la población efectiva en Cuba. *Tribuna de La Habana*. <https://www.tribuna.cu/cuba/2024-07-19/actualizan-cifras-sobre-la-poblacion-efectiva-en-cuba>
- ÁLVAREZ QUIÑONES, R. (30 de mayo de 2023). Cuba, sin «revolución», tendría ahora 17 millones de habitantes. *Diario de Cuba*. https://diariodecuba.com/cuba/1685444611_47487.html#:~:text=Son
- AMERICAN COUNCILS FOR INTERNATIONAL EDUCATION (2017). *The National K-12 Foreign Language Enrollment Survey Report*. <https://www.americancouncils.org/sites/default/files/FLE-report-June17.pdf>
- ARGENTINOS EN INDIA (s. f.). Inicio [Página de Facebook]. Facebook. Recuperado el 30 de marzo de 2025. <https://www.facebook.com/groups/277316635769716/>
- AUSTRALIAN BUREAU OF STATISTICS (2021). *Census of Population and Housing Language spoken at home*. <https://www.abs.gov.au/ausstats/abs@.nsf/Lookup/2900.0main+features100622016>
- BADILLO MATOS, Á. (2021). *Lengua y cultura en español en el Japón de la era Reiwa*. Real Instituto Elcano. <https://www.realinstitutoelcano.org/wp-content/uploads/2021/01/badillo-lengua-y-cultura-en-espanol-en-japon-era-reiwa-1.pdf>
- BADILLO MATOS, Á. (2024). *Español para tigres sudasiáticos. Lengua y cultura en español en Filipinas y el sudeste asiático*. Instituto Cervantes y Real Instituto Elcano.
- BANCO CENTRAL DE RESERVA DE EL SALVADOR (BCR) (2022). *Estimaciones y Proyecciones de Indicadores Demográficos*. <https://www.bcr.gob.sv/documental/Inicio/busqueda/192>
- BANCO MUNDIAL (s. f.-a). *Datos de libre acceso del Banco Mundial*. Consultado el 30 de marzo de 2025. <https://datos.bancomundial.org/>
- BANCO MUNDIAL (s. f.-b). *Population estimates and projections*. Consultado el 30 de marzo de 2025. <https://databank.worldbank.org/Population/id/622a9444#>
- BANCO MUNDIAL (2024). *Informe anual 2024*. <https://www.bancomundial.org/es/about/annual-report>
- BANCO MUNDIAL (2025). *Desarrollo urbano*. Grupo Banco Mundial. <https://www.bancomundial.org/es/topic/urbandevelopment/overview#3>

- BLANCO, C. (2024). *The 2024 Duolingo Language Report*. Duolingo. <https://blog.duolingo.com/2024-duolingo-language-report/>
- BLEORȚU, C., BUZILĂ, P., ILIAN, I., ȚIȚEI-ĂVĂDANEI, A., LOUREDA LAMAS, Ó., ÁLVAREZ MELLA, H., ROȘCA, A. y MANCHEVA, D. (2024). *Demolingüística del español en Rumanía, Bulgaria y Moldavia*. Instituto Cervantes, Universidad de Heidelberg y Universidad de Zúrich. https://cvc.cervantes.es/lengua/espanol_europa/espanol_rumania/
- BONOMI, M., CALVI, M. V. y UBERTI-BONA, N. (2024). *Demolingüística del español en Italia. Con un anexo sobre el español en Malta, San Marino y Ciudad del Vaticano*. Instituto Cervantes, Universidad de Heidelberg y Universidad de Zúrich. https://cvc.cervantes.es/lengua/espanol_europa/espanol_italia/default.htm
- BRAVO GARCÍA, E. (2008). *El español internacional*. Arco Libros.
- BRITISH COUNCIL (2017). *Languages for the future*. https://www.britishcouncil.org/sites/default/files/languages_for_the_future_2017.pdf
- CENTRAL BUREAU OF STATISTICS ARUBA (2021). *Census 2020 Housing Tables*. <https://cbs.aw/wp/index.php/2021/11/30/census-2020-housing-tables/>
- CENTRAL BUREAU OF STATISTICS ARUBA (2023). *Population projections*. <https://cbs.aw/wp/index.php/2023/05/04/population-projections/>
- CENTRAL BUREAU OF STATISTICS CURAÇAO (s. f.). *Population*. Consultado el 30 de marzo de 2025. <https://www.cbs.cw/population>
- CENTRAL BUREAU OF STATISTICS CURAÇAO (2023). *Population by most spoken language (in private households)*. <https://senso.cbs.cw/education-characteristics>
- CENTRAL BUREAU OF STATISTICS ISRAEL (2011). *Judíos y otros de 20 años y más, según lengua materna y uso del hebreo*. https://www.cbs.gov.il/he/publications/doclib/2013/seker_hevratil1/pdf/t28.pdf
- CENTRAL BUREAU OF STATISTICS ISRAEL (2022a). *Population of Israel on the Eve of 2023*. https://www.cbs.gov.il/he/mediarelease/DocLib/2022/426/11_22_426e.pdf
- CENTRAL BUREAU OF STATISTICS ISRAEL (2022b). *Selected Data from the 2021 Social Survey on Languages*. <https://www.cbs.gov.il/en/mediarelease/Pages/2022/Selected-Data-from-the-2021-Social-Survey%20on%20Languages.aspx>
- CENTRAL BUREAU OF STATISTICS ISRAEL (2023). *Population Statistical Abstract of Israel 2023*. <https://www.cbs.gov.il/en/publications/Pages/2023/Population-Statistical-Abstract-of-Israel-202-No-74.aspx>
- CENTRAL INTELLIGENCE AGENCY (CIA). (s. f.). *The World Factbook*. Consultado el 30 de marzo de 2025). <https://www.cia.gov/the-world-factbook/>

- COMISIÓN ECONÓMICA PARA AMÉRICA LATINA Y EL CARIBE (CEPAL) (s. f.-a). *América Latina y el Caribe: Estimaciones y proyecciones de población*. Consultado el 30 de marzo de 2025. <https://www.cepal.org/es/temas/proyecciones-demograficas/america-latina-caribe-estimaciones-proyecciones-poblacion>
- COMISIÓN ECONÓMICA PARA AMÉRICA LATINA Y EL CARIBE (CEPAL) (s. f.-b). REDATAM. *Procesamiento, análisis y diseminación de Estadísticas y Censos a lo largo del mundo*. Consultado el 30 de marzo de 2025. <https://redatam.org/es>
- COMISIÓN ECONÓMICA PARA AMÉRICA LATINA Y EL CARIBE (CEPAL) (1983). *Venezuela: Estimaciones de proyecciones de población, 1950-2025*. Centro Latinoamericano de Demografía - Oficina Central de Estadística e Informática. <https://hdl.handle.net/11362/8005>
- COMISIÓN ECONÓMICA PARA AMÉRICA LATINA Y EL CARIBE (CEPAL) (2022). *Panorama Social de América Latina y el Caribe*. <https://repositorio.cepal.org/server/api/core/bitstreams/3ca376cf-edd4-4815-b392-b2a1f80ae05a/content>
- CONSEJO DEL SECTOR PRIVADO PARA LA ASISTENCIA EDUCACIONAL (CEPAL) (2025). *Programas y proyectos educativos*. <https://cospae.org/>
- CONSEJO NACIONAL DE POBLACIÓN DE MÉXICO (CONAPO) (2024). *Indicadores demográficos de la República Mexicana, en el año 2024*. https://portales.segob.gob.mx/es/Derechos_Humanos/Contexto_Demografico#:~:text=Se%20estima%20que%20en%202024,a%C3%B1os%2C%20representa%2029.4%20por%20ciento
- CONSEJO NACIONAL DE POBLACIÓN DE MÉXICO (CONAPO) y SECRETARÍA DE MEDIO AMBIENTE Y RECURSOS NATURALES (SEMARNAT) (2020). *Proyecciones de población*. https://apps1.semarnat.gob.mx:8443/dgeia/compendio_2020/dgeiawf.semarnat.gob.mx_8080/ibi_apps/WFServletOeae.html
- COUNTRYMETERS (s. f.). *Población de Venezuela*. Consultado el 30 de marzo de 2025. <https://countrymeters.info/es/Venezuela>
- DATOSMACRO (s. f.-a). *China Inmigración*. Consultado el 30 de marzo de 2025. <https://datosmacro.expansion.com/demografia/migracion/inmigracion/china#:~:text=China%20tiene%2C%20seg%C3%BAn%20los%20%C3%BAltimos,%2C%20que%20son%20el%2038.63%25>
- DATOSMACRO (s. f.-b). *Israel Inmigración*. Consultado el 30 de marzo de 2025. <https://datosmacro.expansion.com/demografia/poblacion/israel#:~:text=Solo%20el%2021%2C%25%20de,442%20habitantes%20por%20Km2>
- DEPARTAMENTO ADMINISTRATIVO NACIONAL DE ESTADÍSTICA DE COLOMBIA (DANE) (2025). *Censos de población*. <https://telencuestas.com/censos-de-poblacion/colombia/2025>

- DEPARTMENT OF OFFICIAL LANGUAGE (2016). *Constitutional Provisions*. Ministry of Home Affairs India. <https://rajbhasha.gov.in/en/constitutional-provisions>
- DI MARCO, N., LORU, E., BONETTI, A., GRAZIA SERRA, A. O., CINELLI, M. y QUATTROCIOCCHI, W. (2024). Patterns of linguistic simplification on social media platforms over time. *PNAS*, 121 (50), 1-11. <https://doi.org/10.1073/pnas.2412105121>
- DIPSHIKHA, B. (2023). Hindi. En *Linguistic Survey of India*. Language Division. Office of the Registrar General. <https://language.census.gov.in>
- DJANDUÉ, B. D. (2021). El español lengua extranjera en Costa de Marfil: ¿por qué y para qué se aprende? *redELE*, 33, 1-11. <https://www.educacionfpydeportes.gob.es/dam/jcr:5fc-ce4ed-51a6-48b5-9e63-26f24b153766/ele-marfil.pdf>
- DOMÍNGUEZ, Y. (30 de agosto de 2024). Población dominicana crecerá casi un 1 % para el 2025. *Listín Diario*. https://listindiario.com/la-republica/20240830/poblacion-dominicana-crecera-1-2025-one_823598.html
- EBERHARD, D., SIMONS, G. y FENNIG, C. (2025). *Ethnologue: Languages of the World*. (28.^a ed.). SIL International. <https://www.ethnologue.com/>
- EUROSTAT (2012). *Europeans and their Languages*. Special Eurobarometer 386. <https://op.europa.eu/en/publication-detail/-/publication/f551bd64-8615-4781-9be1-c592217dad83>
- EUROSTAT (2024a). *Europeans and their Languages*. Special Eurobarometer 540. <https://europa.eu/eurobarometer/surveys/detail/2979>
- EUROSTAT (2024b). *Pupils by education level and modern foreign language studied*. https://ec.europa.eu/eurostat/web/products-datasets/-/educ_uae_lang01
- FEDERAL STATISTICAL OFFICE (2020). *Evolution of the permanent resident population according to the 3 basic scenarios, 1990-2050*. <https://www.bfs.admin.ch/bfs/en/home/statistics/population/population-projections/national-projections.assetdetail.12947897.html>
- FERNÁNDEZ VÍTORES, D. y BENLABBAH, F. (2014). *La lengua española en Marruecos*. AECID e Instituto de Estudios Hispano-Lusos.
- FIGUEREDO REINALDO, Ó., HÉCTOR RODRÍGUEZ, Y., EXTREMERA PEREGRÍN, D. A., EXTREMERA SAN MARTÍN, D., CARMONA TAMAYO, E. y PADRÓN PADILLA, A. (19 de julio de 2024). Hoy somos menos de 10 millones de cubanos residentes. *Cubadebate*. <http://www.cubadebate.cu/noticias/2024/07/19/hoy-somos-menos-de-10-millones-de-cubanos-residentes/>
- GAY, M. L. (25-28 de junio de 2014). *La situación del español en Canadá y la enseñanza de ELE en Quebec* [comunicación]. V Congreso internacional FIAPE: ¿Qué español enseñar y cómo? Variedades del español y su enseñanza. Cuenca, España. <https://www.educacionfpydeportes.gob.es/dam/jcr:e88c4c04-15ff-45ad-8eae-235703f0c114/13--la>

-situacion-del-espanol-en-canada-y-la-ensenanza-de-ele-en-quebec--gaylourdes-pdf.pdf

GINSBURGH, V., MELITZ, J. y TOUBAL, F. (2014). Foreign Language Learning: An Econometric Analysis. CESifo Working Papers, 4923. <https://www.cesifo.org/en/publications/2014/working-paper/foreign-language-learning-econometric-analysis>

GÓMEZ, E. y GARRIDO, A. (2020). *Sáhara Occidental. Preguntas frecuentes*. Una mirada al Sáhara Occidental. <https://Saharaoccidental.es/faqs/Sahara-occidental/>

GONZÁLEZ PUY, I. (2018). El español, un valor en alza en China. En VV. AA., *El español en el mundo. Anuario del Instituto Cervantes 2018* (pp. 287-303). Instituto Cervantes y AEBOE. https://cvc.cervantes.es/lengua/anuario/anuario_18/gonzalez/p01.htm

GORDON, R. G. y GRIMES, B. F. (2005). *Ethnologue: Languages of the World* (15.^a ed.). Summer Institute of Linguistics. <https://www.sil.org/resources/archives/6143>

GOVERN D'ANDORRA (2022). *Coneixements i usos lingüístics de la població d'Andorra. Situació actual i evolució 1995-2022*. <https://www.ari.ad/images/pdf/llengua.pdf>

GRAU PEREJOAN, M. y GEA MONERA, M. P. (2006). El español en Trinidad y Tobago. En VV. AA., *Enciclopedia del español en el mundo. Anuario del Instituto Cervantes 2006-2007* (pp. 209-2011). Instituto Cervantes, Plaza & Janés y Círculo de Lectores. https://cvc.cervantes.es/lengua/anuario/anuario_06-07/pdf/paises_41.pdf

GSMA (2024a). *La economía móvil en América Latina 2024*. <https://www.gsma.com/solutions-and-impact/connectivity-for-good/mobile-economy/wp-content/uploads/2024/06/La-economia-movil-en-America-Latina-2024.pdf>

GSMA (2024b). *Uso de redes móviles en América Latina. Tráfico de datos en la actualidad y proyecciones a 2030*. <https://www.gsma.com/about-us/regions/latin-america/wp-content/uploads/2024/10/Uso-de-redes-moviles-en-America-Latina-SPA.pdf>

HAUT-COMMISSARIAT AU PLAN (2014). *Les residentes étrangères au Maroc*. <https://www.hcp.ma/file/231891/>

HERNÁNDEZ ALONSO, C. (1994). Política lingüística española en América. En VV. AA., *Los derechos humanos en América. Una perspectiva de cinco siglos* (pp. 97-121). Cortes de Castilla y León.

HERRERO MAYOR, A. (1944). *Presente y futuro de la lengua española en América*. El Ateneo.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE) (2024). *Sistema de Registro Nacional Migratório*. <https://portaldeimigracao.mj.gov.br/pt/dados/microdados/1733-obmigra/dados/microdados/401205-sismigra>

INSTITUTO CERVANTES (2023). *El español: Una lengua viva. Informe 2023*. En VV. AA., *El español en el mundo 2023. Anuario del Instituto Cervantes* (pp. 25-140). Instituto Cervantes y

McGraw-Hill. https://cvc.cervantes.es/Lengua/anuario/anuario_23/el_espanol_en_el_mundo_anuario_instituto_cervantes_2023.pdf

INSTITUTO DE LOS MEXICANOS EN EL EXTERIOR (2020). *Población mexicana en el exterior*. <https://ime.gob.mx/estadisticas>

INSTITUTO NACIONAL DE ESTADÍSTICA DE BOLIVIA (INE BOLIVIA) (2024). *Censo de Población y Vivienda 2024*. <https://censo.ine.gob.bo/>

INSTITUTO NACIONAL DE ESTADÍSTICA DE ESPAÑA (INE ESPAÑA) (2021). *Encuesta de características esenciales de la población y las viviendas. Año 2021*. <https://www.ine.es/dynt3/inebase/index.htm?padre=8981&capsel=9008>

INSTITUTO NACIONAL DE ESTADÍSTICA DE ESPAÑA (INE ESPAÑA) (2024a). *Estadística continua de población*. https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736177095&menu=ultiDatos&idp=1254735572981

INSTITUTO NACIONAL DE ESTADÍSTICA DE ESPAÑA (INE ESPAÑA) (2024b). *Estadística del Padrón de españoles residentes en el extranjero*. <https://www.ine.es/dynt3/inebase/index.htm?padre=3669&capsel=4553>

INSTITUTO NACIONAL DE ESTADÍSTICA DE ESPAÑA (INE ESPAÑA) (2024c). *Proyecciones de Población. Años 2024-2074*. <https://www.ine.es/dyngs/Prensa/es/PROP20242074.htm>

INSTITUTO NACIONAL DE ESTADÍSTICA DE GUATEMALA (INE GUATEMALA) (2020). *Estimaciones y proyecciones de población a nivel departamental y municipal. Resultados*. https://www.ine.gob.gt/ine/wp-content/uploads/2020/08/Presentacion_19_08-2020-resultados.pdf

INSTITUTO NACIONAL DE ESTADÍSTICA DE GUATEMALA (INE GUATEMALA) (2024). *Encuesta Nacional de Calidad y Bienestar de los Hogares – ENCABIH*. INE. https://www.ine.gob.gt/wp-content/uploads/2024/03/INFORME_ENCABIH.pdf

INSTITUTO NACIONAL DE ESTADÍSTICA DE GUATEMALA, OFICINA NACIONAL DE ESTADÍSTICA DE LA REPÚBLICA DOMINICANA (ONE) y FUNDACIÓN PARA EL ANÁLISIS ESTRATÉGICO Y DESARROLLO DE LA PEQUEÑA Y MEDIANA EMPRESA (FAEDPYME) (2024). *Informe de indicadores estadísticos empresariales*. INE, ONE y FAEDPYME. <https://www.ine.gob.gt/2024/03/21/informe-de-indicadores-estadisticos-empresariales-con-enfoque-de-genero/>

INSTITUTO NACIONAL DE ESTADÍSTICA DE GUINEA ECUATORIAL (INEGE) (2023). *Demografía y población*. <https://inege.org/wp-content/uploads/2024/01/Guinea-Ecuatorial-en-Cifras-2023.pdf>

INSTITUTO NACIONAL DE ESTADÍSTICA DE GUINEA ECUATORIAL (INEGE) (2024). *Anuario Estadístico de Guinea Ecuatorial 2024*. INEGE. <https://inege.org/wp-content/uploads/2024/11/ANUARIO-ESTADISTICO-DE-GUINEA-ECUATORIAL-2024.pdf>

INSTITUTO NACIONAL DE ESTADÍSTICA DE GUINEA ECUATORIAL (INEGE) (2024b). *II Encuesta Nacional de Hogares de Guinea Ecuatorial. Informe de resultados definitivos*. INEGE. <https://inege.org/wp-content/uploads/2024/12/Informe-Definitivo-de-la-ENH2.pdf>

INSTITUTO NACIONAL DE ESTADÍSTICA DE HONDURAS (INE HONDURAS) (2017). *Encuesta Permanente de Hogares de Propósitos Múltiples (EPHPM)*. Instituto Nacional de Estadística. <https://ine.gob.hn/encuesta-de-hogares>

INSTITUTO NACIONAL DE ESTADÍSTICA DE HONDURAS (INE HONDURAS) (2024a). *Sistema Estadístico Nacional*. <https://sen.ine.gob.hn/Home/Index>

INSTITUTO NACIONAL DE ESTADÍSTICA DE HONDURAS (INE HONDURAS) (2024b). *BASEINE. Base de datos en línea*. <https://temp.ine.gob.hn/bases-de-datos/>

INSTITUTO NACIONAL DE ESTADÍSTICA DE PARAGUAY (INE PARAGUAY) (2022). *Primeros resultados finales del IV Censo Nacional de Población y Vivienda para Pueblos Indígenas*. <https://www.ine.gov.py/censo2022/documentos/indigena/Resultados-Finales-Censo-Indigena.pdf>

INSTITUTO NACIONAL DE ESTADÍSTICA DE PARAGUAY (INE PARAGUAY) (2024a). *Población*. <https://www.ine.gov.py/microdatos/index.php?cant=99&tema=TODO>

INSTITUTO NACIONAL DE ESTADÍSTICA DE PARAGUAY (INE PARAGUAY) (2024b). *Proyecciones de la población nacional por sexo y edad al 2050*. INE Paraguay. <https://www.ine.gov.py/publication-single.php?codec=266>

INSTITUTO NACIONAL DE ESTADÍSTICA DE URUGUAY (INE URUGUAY) (2023). *Censo de Población, Hogares y Vivienda*. <https://www.gub.uy/instituto-nacional-estadistica/comunicacion/noticias/poblacion-preliminar-3444263-habitantes>

INSTITUTO NACIONAL DE ESTADÍSTICA E INFORMÁTICA DE PERÚ (INEI) (2024). *Situación de la Población Peruana 2024: Una mirada de la diversidad étnica*. <https://cdn.www.gob.pe/uploads/document/file/6616587/5751291-situacion-de-la-poblacion-peruana-2024-una-mirada-de-la-diversidad-etnica.pdf>

INSTITUTO NACIONAL DE ESTADÍSTICA Y CENSO DE PANAMÁ (INEC PANAMÁ) (s. f.). *XII CENSO NACIONAL DE POBLACIÓN Y VIII DE VIVIENDA 2023*. Consultado el 30 de marzo de 2025. https://www.inec.gob.pa/publicaciones/Default3.aspx?ID_PUBLICACION=1199&ID_CATEGORIA=19&ID_SUBCATEGORIA=71

INSTITUTO NACIONAL DE ESTADÍSTICA Y CENSOS DE COSTA RICA (INEC COSTA RICA) (s. f.). *ESTADÍSTICAS DEMOGRÁFICAS*. Consultado el 30 de marzo de 2025. <https://inec.cr/estadisticas-fuentes/estadisticas-demograficas>

INSTITUTO NACIONAL DE ESTADÍSTICA Y CENSOS DE COSTA RICA (INEC COSTA RICA) (2024). *Estimaciones y proyecciones nacionales de población 1950-2100*. <https://admin.inec.cr/sites/default/files/2024-09/mepoblacpepnac%201950%20-%202100.pdf>

INSTITUTO NACIONAL DE ESTADÍSTICA Y CENSOS DE ECUADOR (INEC ECUADOR) (2022). *Censo Ecuador*. INEC. https://www.censoecuador.gob.ec/wp-content/uploads/2024/05/Presentacion_Nacional_2da_entrega.pdf

INSTITUTO NACIONAL DE ESTADÍSTICA Y CENSOS DE ECUADOR (INEC ECUADOR) (2024). *Estimaciones y proyecciones de la población de Ecuador, Revisión 2024*. https://www.ecuadorencifras.gob.ec/documentos/web-inec/Poblacion_y_Demografia/Proyecciones_Poblacionales/censo_2022/revision_2024/Boletin_tecnico_proy_rev2024.pdf

INSTITUTO NACIONAL DE ESTADÍSTICA Y CENSOS DE LA REPÚBLICA ARGENTINA (INDEC) (2023). *Censo Nacional de Población y Viviendas 2022 por sexo y edad. Indicadores demográficos*. INDEC. https://www.indec.gob.ar/ftp/cuadros/poblacion/censo2022_indicadores_demograficos.pdf

INSTITUTO NACIONAL DE ESTADÍSTICA Y CENSOS DE LA REPÚBLICA ARGENTINA (INDEC) (2024). *Proyecciones*. <https://www.indec.gob.ar/>

INSTITUTO NACIONAL DE ESTADÍSTICA Y GEOGRAFÍA DE MÉXICO (INEGI) (s. f.). CONSULTADO EL 30 de marzo de 2025. *Indicadores sociodemográficos*. <https://www.inegi.org.mx/>

INSTITUTO NACIONAL DE ESTADÍSTICA Y GEOGRAFÍA DE MÉXICO (INEGI) (2020). *Censo General de Población y Vivienda*. <https://www.inegi.org.mx/programas/ccpv/2020/>

INSTITUTO NACIONAL DE ESTADÍSTICA Y GEOGRAFÍA DE MÉXICO (INEGI) (2022). *Estadísticas a propósito del Día Internacional de los Pueblos Indígenas*. <https://inegi.org.mx/app/saladeprensa/noticia.html?id=7519#:~:text=De>

INSTITUTO NACIONAL DE ESTADÍSTICAS DE CHILE (INE CHILE) (2017a). *Censo de Población y Vivienda*. <https://www.ine.gob.cl/estadisticas/sociales/censos-de-poblacion-y-vivienda/censo-de-poblacion-y-vivienda>

INSTITUTO NACIONAL DE ESTADÍSTICAS DE CHILE (INE CHILE) (2017b). *Segundo Registro de Chilenos en el Exterior*. Instituto Nacional de Estadísticas y Ministerio de Relaciones Exteriores.

INSTITUTO NACIONAL DE ESTADÍSTICAS DE CHILE (INE CHILE) (13 de agosto de 2024). *Sala de Prensa*. <https://www.ine.gob.cl/sala-de-prensa/prensa/general/noticia/2024/08/13/en-marzo-de-2025-se-entregan-los-primeros-resultados-del-censo#:~:text=Marzo%202025%3A%20entrega%20del%20n%C3%BAmero,de%20indicadores%20del%20Censo%202024>

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA (INEP) (2020). *Censo da Educação Superior 2019*. <https://www.gov.br/inep/pt-br/aceso-a-informacao/dados-abertos/microdados/censo-da-educacao-superior>

INSTITUTO NACIONAL DE INFORMACIÓN DE DESARROLLO DE NICARAGUA (INIDE) (s. f.). *ANUARIOS ESTADÍSTICOS*. CONSULTADO EL 30 de marzo de 2025. <https://www.inide.gob.ni/Home/Anuarios>

- INSTITUTO NACIONAL DE INFORMACIÓN DE DESARROLLO DE NICARAGUA (INIDE) (2005). VIII Censo de Población y IV de Vivienda 2005. <https://www.inide.gob.ni/Estadisticas/censoCEPOV2005>
- INSTITUTO NACIONAL DE INFORMACIÓN DE DESARROLLO DE NICARAGUA (INIDE) (2025). Estadísticas Sociales y Demográficos. <https://www.inide.gob.ni/>
- JIMÉNEZ CABALLERO, I., AOUS, I. Y SALHI, S. E. (2021). El español en Argelia. En VV. AA., *El español en el mundo 2021. Anuario del Instituto Cervantes* (pp. 527-538). Instituto Cervantes y Bala Perdida S. L. https://cvc.cervantes.es/lengua/anuario/anuario_21/africa/argelia.htm
- JIMÉNEZ SALCEDO, J. (2021). De lengua extranjera a lengua dominante: Representaciones epilingüísticas del castellano en el Principado de Andorra. *Revista de Investigación Lingüística*, 24, 119-145. <https://doi.org/10.6018/ril.480081>
- JOSHUA PROJECT (2024). Country: Jamaica. <https://joshuaproject.net/countries/JM>
- KABATEK, J. (2022). El español en Suiza, con un apunte sobre su presencia en Liechtenstein. En VV. AA., *El español en el mundo 2022. Anuario del Instituto Cervantes*. Instituto Cervantes y McGraw Hill (pp. 245-266). https://cvc.cervantes.es/lengua/anuario/anuario_22/kabatek/p01.htm
- LA CIUTAT (17 de enero de 2025). El 16,7 % de la población de Andorra es de otras nacionalidades. <https://laciutat.cat/es/laciutatandorra-es/el-167-de-la-poblacio-dandorra-es-daltres-nacionalitats-duplicant-se-en-els-darrers-anys>
- LOUREDA LAMAS, Ó., MORENO FERNÁNDEZ, F., ÁLVAREZ MELLA, H. Y SCHEFFLER, D. (2020). Alemania y sus hablantes de español. En VV. AA., *El español en el mundo 2020. Anuario del Instituto Cervantes* (pp. 321-354). Instituto Cervantes y Bala Perdida, S. L. https://cvc.cervantes.es/lengua/anuario/anuario_20/loureda_moreno/p01.htm
- LOUREDA LAMAS, Ó., MORENO FERNÁNDEZ, F., ÁLVAREZ MELLA, H. Y SCHEFFLER, D. (2021). *Demolingüística del español en Alemania*. Instituto Cervantes, Universidad de Heidelberg y Universidad de Zúrich. https://cvc.cervantes.es/lengua/espanol_europa/espanol_alemania/default.htm
- LOUREDA LAMAS, Ó., MORENO FERNÁNDEZ, F., ÁLVAREZ MELLA, H. Y SCHEFFLER, D. (2022). *El español en la Alemania precovid-19*. Instituto Cervantes, Universidad de Heidelberg y Universidad de Zúrich. https://cvc.cervantes.es/lengua/espanol_europa/espanol_alemania_precovid/default.htm
- LOUREDA LAMAS, Ó., PIHLER CIGLIČ, B., KAPOVIĆ, M., BLEORTU, C., VESELKO, V., PALAPANIDI, K., SPAHIĆ, E., USTAMUJIĆ, I., MIHAJLOVIKJ-KOSTADINOVSKA, S., JOVANOVIĆ, A., ŠIFRAR KALAN, M., POPOVSKI, I., KOVAČ BARETT, I., VUČINA SIMOVIĆ, I., KOVAČ, J., GEORGIJEV, I., KABA, F. Y VALERO FERNÁNDEZ, P. (2025). *Demolingüística del español en el sureste europeo*. Instituto Cervantes, Universidad de Heidelberg y Universidad de Zúrich. https://cvc.cervantes.es/lengua/espanol_europa/espanol_sureste/

- LUSIN, N., PETERSON, T., SULEWSKI, C. y ZAFER, R. (2023). *Enrollments in Languages Other Than English in US Institutions of Higher Education, Fall 2021*. Modern Language Association of America. <https://www.mla.org/Resources/Guidelines-and-Data/Reports-and-Professional-Guidelines/Enrollments-in-Languages-Other-Than-English-in-United-States-Institutions-of-Higher-Education>
- MADRID ÁLVAREZ-PIÑER, C. (2018). La evolución de la lengua y la cultura en español en Filipinas. En VV. AA., *El español en el mundo. Anuario del Instituto Cervantes 2018* (pp. 305-312). Instituto Cervantes y AEBOE. https://cvc.cervantes.es/lengua/anuario/anuario_18/madrid/p01.htm
- MASUDA, K. (2020). *Ideología del estándar y realidad plurinormativa de la lengua española: El caso de E/LE en Japón* [Tesis de doctorado, Universitat de Barcelona]. <https://www.tdx.cat/handle/10803/668728#page=1>
- MINISTERIO DE EDUCACIÓN Y FORMACIÓN PROFESIONAL (MEFP) (2018). *El mundo estudia español 2018*. Secretaría General Técnica del Ministerio de Educación. <https://www.educacionyfp.gob.es/mc/redele/el-mundo-estudia-espa-ol/2018.html>
- MINISTERIO DE EDUCACIÓN Y FORMACIÓN PROFESIONAL (MEFP) (2020). *El mundo estudia español 2020*. Secretaría General Técnica del Ministerio de Educación. <https://www.educacionyfp.gob.es/mc/redele/el-mundo-estudia-espa-ol/2020.html>
- MINISTERIO DE EDUCACIÓN Y FORMACIÓN PROFESIONAL (MEFP) (2022). *El mundo estudia español 2022*. Secretaría General Técnica del Ministerio de Educación. <https://www.educacionyfp.gob.es/mc/redele/el-mundo-estudia-espa-ol/2022.html>
- MINISTERIO DE SALUD DEL PERÚ (MINSA) (2024). *Población Estimada 2025*. REUNIS Repositorio Única Nacional de Información y Salud. <https://www.minsa.gob.pe/reunis/?op=1&niv=5&tbl=1>
- MINISTERIO DEL INTERIOR Y DE COMUNICACIÓN DE JAPÓN (2023). *Statistical Handbook of Japan*. Oficina Estadística del Ministerio del Interior y de Comunicación. <https://www.stat.go.jp/english/data/handbook/pdf/2023all.pdf>
- MORA, L. y LÓPEZ, M. H. (20 de septiembre de 2023). *Latinos' Views of and Experiences With the Spanish Language*. Pew Research Center. <https://www.pewresearch.org/race-and-ethnicity/2023/09/20/latinos-views-of-and-experiences-with-the-spanish-language/>
- MORENO FERNÁNDEZ, F. (2024). *Demografía de las lenguas*. Iberoamericana Vervuert.
- MORENO FERNÁNDEZ, F. y OTERO ROTH, J. (1998). Demografía de la lengua española. En VV. AA., *El español en el mundo. Anuario del Instituto Cervantes 1998* (pp. 59-86). Instituto Cervantes y Arco Libros.
- MORENO FERNÁNDEZ, F. y OTERO ROTH, J. (2006). *Demografía de la lengua española*. Universidad Complutense de Madrid. <https://eprints.ucm.es/8936/1/DT03-06.pdf>

- MOURA, R. (2024). Aspectos populacionais da fronteira terrestre segundo o censo 2022. *Boletim regional urbano e ambiental*, 31. <http://dx.doi.org/10.38116/brua31art3>
- MULTIDIMENSIONAL POVERTY PEER NETWORK (MPPN) (S. F.). ¿QUÉ ES POBREZA MULTIDIMENSIONAL? UNIVERSIDAD DE OXFORD. OXFORD DEPARTMENT OF INTERNATIONAL DEVELOPMENT (OPHI). CONSULTADO EL 30 de marzo de 2025. <https://www.mppn.org/es/pobreza-multidimensional/que-es-el-ipm/>
- NATIONAL CENTER FOR EDUCATION STATISTICS US (2023a). *Private elementary and secondary school enrollment, percentage distribution of private school enrollment, and private enrollment as a percentage of combined enrollment in public and private schools, by school orientation and grade: Selected years, fall 1999 through fall 2021*. Digest of Education Statistics. https://nces.ed.gov/programs/digest/d23/tables/dt23_205.15.asp?current=yes
- NATIONAL CENTER FOR EDUCATION STATISTICS US (2023b). *Statistics U. S.* <https://nces.ed.gov/factsfats/display.asp?id=372#PK12-enrollment>
- NATIONAL CENTER FOR EDUCATION STATISTICS US (2024). *Enrollment in public elementary and secondary schools, by region, state, and jurisdiction: Selected years, fall 1990 through fall 2023*. Digest of Education Statistics. https://nces.ed.gov/programs/digest/d24/tables/dt24_203.20.asp
- NATIONAL COMMISSION ON POPULATION (2019). *Population Projections for India and States 2011-2036. Report of the Technical Group on Population Projections*. https://geographicinsights.iq.harvard.edu/files/geographicinsights2/files/report_population_projection_2019.pdf
- NATIONAL GEOGRAPHIC INFORMATION INSTITUTE (NGII) (2021). *National Atlas of Korea III. (International Migration of Foreign Nationals, p. 2.683)*. http://nationalatlas.ngii.go.kr/pages/page_2683.php
- NUADDA (S. F.). *MENOS DEL 1 % de los brasileños hablan español*. <https://www.nuadda.com/menos-del-1-de-los-brasilenos-hablan-espanol/>
- OBSERVATORIO DE LA DIVERSIDAD LINGÜÍSTICA Y CULTURAL EN LA INTERNET (OBDILCI) (2024). *Indicadores de la presencia de lenguas en la Internet (Versión 5.2 - noviembre de 2024)*. <https://www.obdilci.org/proyectos/principal/>
- OBSERVATORIO GLOBAL DEL ESPAÑOL (2024). *El español: lengua para el mundo 2024*. En VV. AA., *El español en el mundo 2024. Anuario del Instituto Cervantes* (pp. 21-164). Instituto Cervantes y McGraw-Hill. https://cvc.cervantes.es/lengua/anuario/anuario_24/el_espanol_en_el_mundo_anuario_instituto_cervantes_2024.pdf
- OFFICE FOR NATIONAL STATISTICS UNITED KINGDOM (2022). *Language, England and Wales: Census 2021*. <https://www.ons.gov.uk/peoplepopulationandcommunity/culturalidentity/language/bulletins/languageenglandandwales/census2021>

OFFICE OF THE REGISTRAR GENERAL & CENSUS COMMISSIONER INDIA (2011). *Concepts and Definitions*. <https://censusindia.gov.in/nada/index.php/catalog/42561>

OFFICE OF THE REGISTRAR GENERAL & CENSUS COMMISSIONER INDIA (2021). *Population by bilingualism and trilingualism, India-2011*. <https://censusindia.gov.in/nada/index.php/catalog/10262#metadata-description>

OFFICE OF THE REGISTRAR GENERAL & CENSUS COMMISSIONER INDIA (2022). *Population by mother tongue, India-2011*. <https://censusindia.gov.in/nada/index.php/catalog/10191>

OFICINA DE INFORMACIÓN DIPLOMÁTICA (OID) (2025). *Nicaragua*. Ministerio de Asuntos Exteriores, Unión Europea y Cooperación de España. https://www.exteriores.gob.es/documents/fichaspais/nicaragua_ficha%20pais.pdf

OFICINA DEL CENSO DE LOS ESTADOS UNIDOS (2020). *Guam Demographic Profile*. US Census Bureau Summary Tables. <https://data.census.gov/table/DECENNIALDPGU2020.DP1?g=040XX00US66&d=DECIA%20Guam%20Demographic%20Profile&tid=DECENNIALDPGU2020.DP1>

OFICINA DEL CENSO DE LOS ESTADOS UNIDOS (2023). *National Population Projections Tables: Main Series*. US Census Bureau Summary Tables. <https://www.census.gov/data/tables/2023/demo/popproj/2023-summary-tables.html>

OFICINA DEL CENSO DE LOS ESTADOS UNIDOS (2025). *U.S. and World Population Clock*. <https://www.census.gov/popclock/>

OFICINA NACIONAL DE ESTADÍSTICA DE LA REPÚBLICA DOMINICANA (ONE) (10 de agosto de 2023). *Boletín Censal*. <https://www.one.gob.do/media/0hq0f5b/boletín-censal-26.pdf>

OFICINA NACIONAL DE ESTADÍSTICA DE LA REPÚBLICA DOMINICANA (ONE) (2025). *Estimaciones y proyecciones demográficas*. <https://www.one.gob.do/datos-y-estadisticas/temas/estadisticas-demograficas/estimaciones-y-proyecciones-demograficas/>

OFICINA NACIONAL DE ESTADÍSTICA Y CENSOS DE EL SALVADOR (ONEC) (s. f.). *POBLACIÓN Y ESTADÍSTICAS DEMOGRÁFICAS*. CONSULTADO EL 30 de marzo de 2025. <https://onec.bcr.gob.sv/poblacion-y-estadisticas-demograficas/>

ORGANIZACIÓN DE NACIONES UNIDAS (ONU) (2018). *Sahrawi Refugees in Tindouf, Algeria: Total In-Camp Population*. Alto Comisionado de Naciones Unidas para los Refugiados (ACNUR). https://wsrw.org/files/dated/2018-10-29/unhcr_2018.pdf

ORGANIZACIÓN DE NACIONES UNIDAS (ONU) (2020). *International Migrant Stock*. Departamento de Asuntos Económicos y Sociales. División de Población. <https://www.un.org/development/desa/pd/content/international-migrant-stock>

ORGANIZACIÓN DE NACIONES UNIDAS (ONU) (2022a). *Población mundial*. Departamento de Asuntos Económicos y Sociales. <https://www.un.org/es/desa-es/la-poblaci%C3%B3n-mundial-llegar%C3%A1-8000-millones-en-2022>

- ORGANIZACIÓN DE NACIONES UNIDAS (ONU) (2022b). *World Population Prospects 2022. Summary of Results*. Departamento de Asuntos Económicos y Sociales. División de Población. https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/wpp2022_summary_of_results.pdf
- ORGANIZACIÓN DE NACIONES UNIDAS (ONU) (2024a). *Peace, dignity and equality on a healthy planet*. <https://www.un.org/en/global-issues/population#:~:text=The>
- ORGANIZACIÓN DE NACIONES UNIDAS (ONU) (2024b). *World Population Prospects 2024*. Departamento de Asuntos Económicos y Sociales. División de Población. <https://www.un.org/development/desa/pd/world-population-prospects-2024>
- ORGANIZACIÓN INTERNACIONAL DE LA FRANCOFONÍA (2022). *La lengua francesa en el mundo. Síntesis 2022*. Gallimard y Organización Internacional de la Francofonía. https://observatoire.francophonie.org/wp-content/uploads/2022/08/Livret_OIF2_Esp.pdf
- ORGANIZACIÓN INTERNACIONAL PARA LAS MIGRACIONES (IOM) (2022). *Migration and Belize. Results of the Migration Modules Attached to the 2021 Labor Force Survey*. <https://sib.org.bz/wp-content/uploads/MigrationandBelize.pdf>
- PASSEL, J. S. Y KROGSTAD, J. M. (22 de julio de 2024). *What we know about unauthorized immigrants living in the U.S.* Pew Research Center. <https://www.pewresearch.org/short-reads/2024/07/22/what-we-know-about-unauthorized-immigrants-living-in-the-us/#:~:text=Who are unauthorized immigrants%3F>
- PERUANOS EN INDIA (S. F.). INICIO [PÁGINA DE FACEBOOK]. FACEBOOK. CONSULTADO EL 30 de marzo de 2025. <https://www.facebook.com/peruanosenindia/>
- PHILIPPINE STATISTICS AUTHORITY (2023). *Ethnicity in the Philippines (2020 Census of Population and Housing)*. <https://psa.gov.ph/statistics/population-and-housing/node/1684059978>
- PHILIPPINE STATISTICS AUTHORITY (2024). *Population Projection Statistics*. <https://psa.gov.ph/statistics/census/projected-population>
- POTTIER, B. (1968). L'Espagnol. En A. Martinet (Ed.), *Le langage* (pp. 887-905). Gallimard.
- PROGRAMA DE NACIONES UNIDAS PARA EL DESARROLLO (PNUD) (2024a). *Human Development Report 2023/2024*. <https://hdr.undp.org/content/human-development-report-2023-24>
- PROGRAMA DE NACIONES UNIDAS PARA EL DESARROLLO (PNUD) (2024b). *Human Development Report 2023/2024 technical notes*. https://hdr.undp.org/sites/default/files/2023-24_HDR/hdr2023-24_technical_notes.pdf
- PUJOL RIEMBAU, Ò. (2020). El español en la India: Una lengua favorecida. En VV. AA., *El español en el mundo 2020. Anuario del Instituto Cervantes*. Instituto Cervantes y Bala Perdida, S. L. https://cvc.cervantes.es/lengua/anuario/anuario_20/pujol/p01.htm

- R4V. Plataforma de Coordinación Interagencial para Refugiados y Migrantes de Venezuela (2024). *R4V América Latina y el Caribe, Refugiados y Migrantes Venezolanos en la Región*. <https://www.r4v.info/es/document/r4v-america-latina-y-el-caribe-refugiados-y-migrantes-venezolanos-en-la-region-noviembre-3>
- REPÚBLICA DE GUINEA ECUATORIAL (2016). *Censo de población 2015. Resultados preliminares*. República de Guinea Ecuatorial, PNUD, UNFPA, FAO, OIT y UNICEF. https://ireda.ceped.org/inventaire/ressources/gnq-2015-rec_resultados_preliminares.pdf
- ROJAS, J. C. (26 de marzo de 2017). Entrevista con Farida Loudaya. *El Tiempo*. <https://www.eltiempo.com/mundo/africa/entrevista-con-farida-loudaya-embajadora-de-marruecos-en-colombia-71664>
- ROSENBLAT, Á. (1977). El futuro de nuestra lengua. En Á. Rosenblat, *Sentido mágico de la palabra*. Universidad Central de Venezuela.
- SANTOS ROVIRA, J. M. Y SERRANO LUCAS, C. M. (2022). *Demolingüística del español en Portugal*. Instituto Cervantes, Universidad de Heidelberg y Universidad de Zúrich. https://cvc.cervantes.es/lengua/espanol_europa/espanol_portugal/
- SATO CENTER (2012). *Language Varieties*. Papiamentu. <https://www.hawaii.edu/satocenter/langnet/definitions/papiamentu.html>
- SERRANO AVILÉS, J. (2014). La enseñanza del español en África Subsahariana: documentación y propuestas. En J. Serrano Avilés (Ed.), *La enseñanza del español en el África Subsahariana* (Capítulo 1, pp. 15-92). Embajada de España en Kenia, AECID, Instituto Cervantes, Casa África y Los Libros de la Catarata. <https://cvc.cervantes.es/lengua/eeas/default.htm>
- STATE DATA CENTER-PUERTO RICO (2023). *Estimados Anuales Poblacionales*. SDC - PR. <https://censo.estadisticas.pr/EncuestaComunidad>
- STATE DATA CENTER-PUERTO RICO (2025). *Estimados Anuales Poblacionales*. SDC - PR. <https://censo.estadisticas.pr/>
- STATISTA (2024a). *Spoken languages of Bonaire, Saint-Eustatius and Saba 2021*. <https://www.statista.com/statistics/1377799/caribbean-netherlands-spoken-languages-2021/#:~:text=Papiamento>
- STATISTA (2024b). *Total population of Venezuela*. <https://www.statista.com/statistics/370820/total-population-of-venezuela/>
- STATISTICAL INSTITUTE OF BELIZE (SIB) (2022). *Belize Population and Housing Census 2022*. <https://sib.org.bz/census/2022-census/>
- STATISTICS CANADA (2024a). *Number of students in regular programs for youth, public elementary and secondary schools, by grade and sex*. <https://doi.org/10.25318/3710000701-eng>
- STATISTICS CANADA (2024b). *Postsecondary enrolments, by registration status, institution type, status of student in Canada and gender*. <https://doi.org/10.25318/3710001801-eng>

- STATISTICS CANADA (2025). *Mother tongue by single and multiple mother tongue responses: Canada, provinces and territories, census divisions and census subdivisions*. <https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=9810018001>
- STATISTICS NORWAY (2021). *Population: Children, families and households*. <https://www.ssb.no/en/befolkning/?de=Children%2C+families+and+households&innholdstype=publikasjon-artikkel>
- STATISTIK AUSTRIA (2024). *Fremdsprachenunterricht der Schüler: innen im Schuljahr 2022/23 [Dataset]*. <https://www.statistik.at/statistiken/bevoelkerung-und-soziales/bildung/schulbesuch/schuelerinnen>
- STATISTISCHES BUNDESAMT (S. F.). SCHÜLER MIT FREMDSPRACHEN-UNTERRICHT. [HTTPS://WWW-GENESIS.DESTATIS.DE/GENESIS//ONLINE?OPERATION=TABLE&CODE=21111-0006&bypass=true&levelindex=1&levelid=1709031181542#abreadcrumb](https://www-genesis.destatis.de/genesis//online?operation=table&code=21111-0006&bypass=true&levelindex=1&levelid=1709031181542#abreadcrumb)
- STATS NEW ZEALAND (2018). *Latin American ethnic group*. <https://www.stats.govt.nz/tools/2018-census-ethnic-group-summaries/latin-american>
- TANO, M. (2023). Dinámicas de la lengua española en los espacios de la educación superior francesa. *Archiletras científica: revista de investigación de lengua y letras*, 10, 147-163.
- THE BUREAU OF STATISTICS AND PLANS. GOVERNMENT OF GUAM (2020). *The 2020 Census of Guam*. <https://bsp.guam.gov/2020-census/>
- UGARTE FARRERONS, V. (2012). El español en Japón. En VV. AA., *El español en el mundo*. Anuario del Instituto Cervantes 2012. Instituto Cervantes y AEBOE. https://cvc.cervantes.es/lengua/anuario/anuario_12/ugarte/p01.htm
- UNICAMP (2024). *Inmigrantes internacionais registrados no Brasil*. Observatório das Migrações em São Paulo. <https://www.nepo.unicamp.br/observatorio/bancointerativo/numeros-imigracao-internacional/sinere-sismigra/>
- VENEZOLANOS EN INDIA (S. F.). INICIO [PÁGINA DE FACEBOOK]. FACEBOOK. CONSULTADO EL 30 de marzo de 2025. <https://www.facebook.com/venezolanosenindia/>
- W3Techs (s. f.). *Usage statistics of content languages for websites*. Consultado en febrero de 2025. https://w3techs.com/technologies/overview/content_language
- ZIELŃSKI, A., BERTA, T., MEDVECZKÁ, M., PESKOVA, J., ŠTRBÁKOVÁ, R., TATOJ, C., VALERO FERNÁNDEZ, P., ÁLVAREZ MELLA, H. Y LOUREDA LAMAS, Ó. (en prensa). *Demolingüística del español en Europa Central*. Instituto Cervantes, Universidad de Heidelberg y Universidad de Zúrich.

II. ANÁLISIS COMPLEMENTARIOS

EL VALOR ECONÓMICO DEL ESPAÑOL ANTE UN NUEVO ESCENARIO INTERNACIONAL

Juan Carlos Jiménez Jiménez
UNIVERSITY OF ALCALÁ

1. Introducción

Las décadas recientes han registrado notables avances del español en distintos ámbitos, desde su peso demográfico a escala mundial hasta su papel como facilitador de los flujos económicos internacionales —de personas y de mercancías, de servicios y de capitales— que han acompañado al gran proceso de globalización y transformación tecnológica vivido en el tránsito del siglo xx al xxi. Una lengua internacional, esta será aquí nuestra premisa, necesita una economía global para hacer valer sus potencialidades. El español lo ha demostrado.

Sin embargo, cuando se inicia el segundo cuarto de esta ya accidentada centuria, todas las señales apuntan a un giro en el orden mundial —y no solo en lo económico— marcado por el nacionalismo y la introversión: se sustituyen aceleradamente las reglas por la discrecionalidad, las instituciones multilaterales por la ley del más fuerte y, en el seno de algunas democracias (y, sorprendentemente, en la más poderosa), flaquea el Estado de derecho, tornándose así la anterior predictibilidad en caos, también en las relaciones internacionales. Para entenderse en el intercambio es tan importante la lengua como, más si cabe, la confianza que dan el respeto a las normas, a los pactos y a las instituciones, formales e informales. Una lengua compartida, no importará

decirlo una vez más, es una suerte de moneda común: rinde sus mayores frutos con el mercado y la libre competencia.

Falta sin duda perspectiva para concretar la magnitud y velocidad de los cambios, pero no por desgracia para ver ya con claridad el sentido del péndulo. Que este giro en el orden liberal internacional hacia postulados iliberales, por más que preludiado en no pocas señales previas, venga patrocinado por quien —benefactor y beneficiado— más contribuyó a crearlo y sostenerlo durante ocho décadas, no es una buena noticia para el futuro del español. Tanto menos por cuanto es ese país, Estados Unidos, el epicentro de la sacudida, justo donde se cifraban muchas de las esperanzas del español como lengua de una clase media mundial en ascenso, y donde el español y sus hablantes, dentro de una batalla política y cultural de muchos más matices, parece convertirse en el enemigo interno a batir frente al deseado dominio imperial del inglés, dentro y fuera de sus fronteras. Lo de menos es el estatus o no de lengua oficial; lo grave es la imposición de una política de *English only* que remite a otras épocas y que hoy, además, parece dirigirse, más allá de lo lingüístico, al rechazo de los signos de identidad de una parte sustancial de la población y de una gran fracción del territorio donde ya se hablaba español antes de que apareciera el inglés.

Cualquier análisis debe contemplar esta nueva situación. Publicado a finales de 2023, los autores de *Los futuros del español. Horizonte de una lengua internacional* (Alonso Rodríguez, Jiménez Jiménez y García Delgado, 2023) tratamos de dibujar los escenarios previsibles de nuestra lengua en sendos plazos, medio (2030) y largo (2050). Pero, como acaba de indicarse, sobre la línea de ese horizonte se levanta hoy una gran ola que no cabe ignorar y que hay que surfear con la pericia necesaria para no perder el equilibrio. Estas breves páginas se proponen resumir lo que sabemos hasta la fecha acerca del valor económico del español y, sobre todo, reevaluar sus previsibles escenarios a la luz de unas realidades no tan prometedoras como podía pensarse (pero tampoco carentes de oportunidades). Cómo quedan, pues, las previsiones del español en el PIB mundial, en los flujos de comercio y en el ámbito

de la economía digital será el objeto de los apartados que siguen, antes de recobrar la visión de conjunto para ofrecer, dentro del epígrafe conclusivo, una propuesta sintética de lo que puede hacerse para mantener el vigor del español como palanca de progreso del mundo hispanohablante.

2. El poder de compra de los hispanohablantes: del escenario «conservador» al «pesimista»

Casi dos décadas atrás, las estimaciones realizadas para 2006 cifraban en torno del 9 % del PIB mundial (en realidad, en una horquilla situada entre el 8,7 y el 9,2 %, atendiendo a dos hipótesis alternativas) el poder de compra de los hispanohablantes en el mundo (Jiménez Jiménez, 2007, 2009). La metodología se basaba en atribuir a estos la renta per cápita de sus respectivos países, salvo en el caso de los hispanos norteamericanos, en que era obligado hacer algunos ajustes. Los cálculos, replicados recientemente tomando como referencia 2019-2020, ofrecían órdenes de magnitud ligeramente por encima, entre un 9,1 % y un 9,6 % del PIB mundial (Alonso Rodríguez, Jiménez Jiménez y García Delgado, 2023). Tanto aquellas como estas cifras reflejaban el empuje del colectivo hispano de Estados Unidos, tanto demográfico como en lo tocante a su creciente poder adquisitivo dentro de la primera economía del mundo.

Las proyecciones efectuadas con el horizonte de 2050 (y los umbrales intermedios de 2025 y 2030) fueron particularmente cautas, dentro de las muchas y difícilmente predecibles variables que entran en este tipo de cálculos. Pues bien, si hasta 2025 podía preverse que la evolución de la capacidad de compra de los hispanohablantes siguiera lentamente al alza, hasta situarse entre el 9,4 y el 9,9 % del PIB mundial, la perspectiva para 2050 rebajaba esos porcentajes en cerca de medio punto, en una horquilla entre el 8,8 y el 9,4 %, en una trayectoria descendente que solo se amortiguaba en la más optimista de las hipótesis respecto de los hispanos norteamericanos (Alonso Rodríguez, Jiménez Jiménez y García Delgado, 2023).

Pero es necesario que los porcentajes y sus decimales no nos impidan ver lo que están reflejando estas cifras. Primero, que la capacidad de compra

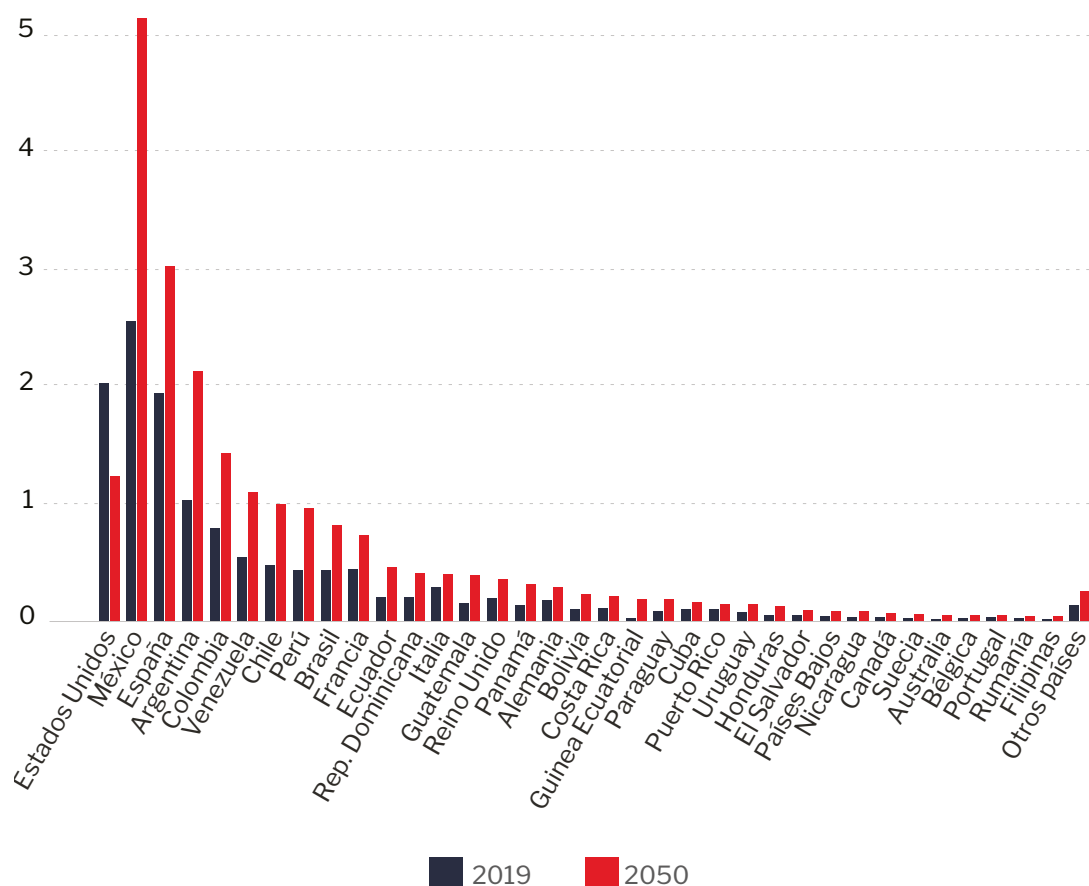
hispanohablante respecto al total mundial no ofrece abruptas oscilaciones, ni en las dos décadas previas, ni en las dos y media por venir: en torno del 9 % al comienzo y al final de la primera mitad del presente siglo, y medio punto por encima en el álgido punto medio. Segundo, que la leve caída en la cuota de hablantes de español en el mundo que tendrá lugar —la demografía es tozuda— de aquí a 2050 no va a verse previsiblemente compensada por el aumento de la renta per cápita de estos: algo que solo podría cambiar la extensión a gran escala del español fuera de sus actuales áreas de dominio nativo en países más ricos, lo que no entra dentro de lo previsible. Y, tercero, que cualquier cálculo es muy sensible a las hipótesis que se hagan acerca de los hispanos norteamericanos: de cómo evolucione su demografía, de cómo lo haga su renta y su capacidad adquisitiva, y de cómo lo haga la pervivencia del español dentro de este colectivo. En particular, de este último factor.

Los cambios en el escenario del español en Estados Unidos que hoy se entrevén como algo más que probables obligan a considerar, junto a las proyecciones que acaban de apuntarse, alguna otra aún más prudente —por no decir pesimista—, y por eso mismo tal vez más realista. La contribución de los hispanos norteamericanos al PIB *del español* se estimó inicialmente, en los cálculos de 2023 y en sus proyecciones hasta 2050, sobre dos conjuntos de hipótesis. La primera, relativa a la renta per cápita de los hispanos, admitía a su vez dos opciones: una, más optimista, suponía que la renta per cápita de estos —en realidad, el ingreso medio de los hogares hispanos— podía cifrarse en 2019 en un 81,9 % de la media nacional (a tenor de las estimaciones de la Oficina del Censo de los Estados Unidos), si bien esa proporción se iba luego corrigiendo al alza, hasta converger en 2050 con el nivel de vida promedio norteamericano (opción 1); la otra opción, más prudente, aplicaba al PIB de Estados Unidos de cada año el mismo porcentaje que representan los hispanos en el poder de compra norteamericano, según la proyección de los datos del Selig Center para 2050, lo que rebajaba los cálculos en casi una tercera parte (opción 2). La segunda hipótesis (común a las opciones 1 y 2) aplicaba un factor de corrección que desinflaba las proyecciones demográficas de los hispanos del censo de Estados Unidos de acuerdo con el porcentaje de estos que mantienen

competencias en español; proporción estimada inicialmente en el 71,6 % de los que en 2020 hablaban español en el hogar, y a partir de ahí reducida de acuerdo con la senda observada hasta ahora por la progresiva —aunque mucho más contenida que en otras lenguas— pérdida de competencias lingüísticas en español de la comunidad hispana.

Las circunstancias actuales aconsejan no obstante considerar un escenario no radical, pero sí sustancialmente alternativo a los dos que acaban de apuntarse. Sería nuestra «opción 3»¹: ¿Y si la pérdida de las competencias lingüísticas de las comunidades hispanas fuera más acusada de lo estimado hasta ahora? Supóngase que en 2050 solo el 25 % del colectivo hispano de Estados Unidos conservara un dominio fluido del español. No es un porcentaje al azar: es el que actualmente representan los hispanos de tercera generación que son bilingües en español e inglés (el resto son monolingües en inglés) (Lacomba, 2023, con los datos del Pew Research Center). En tres décadas, si se siguen reduciendo —aún más— los flujos migratorios y si todo el aumento demográfico hispano se basa en los nacidos ya en Estados Unidos, esto es, cuando la parte mayoritaria del conjunto sea de tercera generación, la política de *English only*, si vence los contrapesos actuales, podría consolidar esa menguada proporción de hablantes. Si en la más favorable de las opciones previas contempladas los hispanos norteamericanos alcanzaban una capacidad global de compra por encima de la de cualquier otro país, por delante de México, y suavizaban el descenso en la capacidad de compra hispanohablante dentro del PIB mundial, en esta más desfavorable (pero no inverosímil) opción 3, perdido ese freno, el peso del español caería en 2050 al 7,7 % del PIB mundial, es decir, un punto porcentual menos de lo que marcaba la más prudente de las estimaciones previas (opción 2) para 2050 y casi dos puntos menos de lo que promediaban las estimaciones en 2019-2020; de hecho, claramente por debajo de lo que estimábamos con los criterios más prudentes a comienzos de este siglo. Por países, el gráfico 1 muestra cómo quedaría el *ranking*, encabezado por México de forma crecientemente destacada.

1. Confróntese nuestro gráfico 1 con los dos que aparecen en Jiménez Jiménez y Narbona Moreno (2022, pp. 10-11): el más optimista (alternativa 1) y el —en aquel momento— prudente (alternativa 2).

GRÁFICO 1 CAPACIDAD DE COMPRA DE LOS HABLANTES DE ESPAÑOL EN EL MUNDO 2019-2050 (PIB EN PPA, BILLONES DE \$ A PRECIOS INTERNACIONALES)**FUENTE:** elaboración propia.

En definitiva, si los cálculos iniciales de la capacidad de compra de los hispanohablantes en el mundo reflejaban un progresivo declinar que podía ser, no obstante, compensado por el avance económico de unas comunidades de hispanos norteamericanos que conseguían conservar en buena medida sus competencias lingüísticas a lo largo de las generaciones, una hipótesis menos optimista respecto de esta evolución del español en Estados Unidos muestra cómo aquel declinar se convierte en inevitable y mucho más acusado.

Con todo, la tendencia de pérdida de dominio del español entre los hispanos ha contado hasta ahora con algunos contrapesos en cuya fortaleza hay que confiar: entre ellos, la frecuente comunicación de la

diáspora hispana en Estados Unidos con sus comunidades de origen, singularmente con México, o el vigor de las manifestaciones culturales en español (música, audiovisual...), de gran atractivo entre los más jóvenes. Pero, sobre todo, la percepción desde hace un par de décadas de que el español —rompiendo un viejo estigma como marcador social negativo— *valía* en el mercado de trabajo, y para los jóvenes hispanos un bilingüismo equilibrado inglés-español era la más rentable de las opciones, tanto en términos de acceso a los estudios superiores como, luego, de una más favorable inserción laboral y unos salarios más altos. ¿Cómo es posible —se preguntaba Gándara (2015)— que hablar solo inglés valiera lo mismo (o incluso más) que hablar inglés y español? Desde luego, desafía al sentido común —y a la economía de la educación, a la psicología y a la neurociencia— que fuera mejor para los hispanos (en términos de salarios y de acceso al empleo) ser monolingües en inglés que bilingües en inglés y español.

A partir de bases de datos longitudinales —abarcando toda la trayectoria escolar de cohortes de hispanos bilingües y monolingües en inglés hasta llegar al mercado laboral—, tres trabajos han venido a darle la vuelta a esa extraña paradoja del bilingüismo en español dentro de Estados Unidos (Agirdag, 2014; Rumbaut, 2014; Santibañez y Zárate, 2014). De ellos se deduce que hoy, en Estados Unidos: a) los bilingües equilibrados en español e inglés ganan significativamente más que los monolingües (o que los bilingües limitados) con los que comparten «raíces inmigrantes» (entre 2.000 y 3.200 dólares anuales más); b) el bilingüismo equilibrado reporta una ventaja significativa, mayor cuanto más alto es el nivel de competencia bilingüe, y tanto en términos de menor abandono escolar como de salarios más altos y ocupaciones de más prestigio; y c) los bilingües en español tienen más probabilidades de ir a la universidad.

Y no solo esto: a la «prima salarial» que obtienen en Estados Unidos —¡al fin!— los hispanos que conservan su español, se une la que perciben los no hispanos que lo aprenden (más de 7 millones). Un reciente estudio estima una diferencia entre la retribución promedio de estos y los que son solo monolingües en inglés cercana a los 2.000-3.000 dólares anuales a lo largo del

periodo 2014-2019, lo que viene a significar una prima de ingresos equivalente al 3,3 % (Martínez García y Martínez García, 2022). Y, aunque sea como simple elemento indiciario que ratifica esta percepción, en Martín Leralta, Jiménez Jiménez y Carrera Troyano (2024) hemos hallado, al examinar el valor económico de la certificación del español a través de una amplia muestra de candidatos DELE y SIELE de todo el mundo, que los estadounidenses destacan por la valoración que hacen de sus certificados de español (la más alta por nacionalidades, y que duplica la media general), en particular los que ya están insertos en el mercado de trabajo.

Las expectativas parecían buenas, y ojalá lo sigan siendo. Pero que rebrote hoy el estigma social negativo de hablar español en Estados Unidos, dentro de un empeño por forzar la oficialidad única del inglés y de una ola de reversión social mucho más amplia, deja en el aire aquellos avances y, con ello, el propio futuro del español a varias décadas vista. No es esta, en todo caso, la única de las amenazas que se cierne sobre nuestra lengua por los vientos desatados desde el otro lado del Atlántico.

3. Intercambios comerciales y proteccionismo: puertas que se cierran y otras que se abren

Difícilmente se encontrará a un economista de prestigio —y aun sin él— capaz de sostener que «arancel», en su acepción de derechos aduaneros, es la palabra más bella del diccionario. Eso sí, el proteccionismo, una vez puesto en marcha, es insaciable: al erosionar la competencia (y con ello la competitividad) de las producciones locales protegidas, incita a cubrir bajo su égida al resto de las actividades, afectadas directa o indirectamente por los mayores costes de aquellas. Poco importa que se trate solo de una estrategia transaccional (con el fin de conseguir ventajas en otros órdenes) o forme parte de una política defensiva integral; la arbitraria discrecionalidad con que se administran las amenazas arancelarias sume a los agentes en la incertidumbre y deprime el comercio más incluso que los aranceles mismos. De modo que se debilita, si no se anula, el favorable efecto que distintos factores tienen sobre los intercambios internacionales, incluido el de contar con una lengua común entre las partes.

El desafío es cierto. Desde el nacimiento del Acuerdo General sobre Aranceles Aduaneros y Comercio en 1947 el mundo no se había enfrentado como hoy al peligro de un aumento universal e indiscriminado de esta forma impositiva sobre las importaciones de otros países. Las pulsiones proteccionistas de Estados Unidos —en una política de «empobrecer al vecino» que ya se reveló catastrófica en los años treinta del siglo pasado y tampoco rindió frutos en el primer Gobierno de Trump— amenazan, de generalizarse, el comercio internacional. Un comercio que durante décadas ha propiciado el progreso económico de buena parte del mundo (sin negar algunas sombras que no es este el lugar de discutir) y que ha permitido, en el terreno que aquí nos ocupa, el de la lengua, que el español haya desempeñado un importante papel facilitador de los intercambios —y de los movimientos de capitales, no se olvide— entre los países que lo compartimos (inclúyase aquí, por cierto, al propio Estados Unidos). Un papel facilitador que hoy está seriamente comprometido.

Los modelos de gravedad han sido hasta la fecha el instrumento metodológico básico para aproximar el efecto que tiene compartir una lengua sobre el comercio internacional. Pues bien, nuestros datos nos dicen que la lengua común es una variable que, dentro de las holguras de este tipo de modelos, duplica aproximadamente los intercambios entre los países que comparten alguna de ellas, controlados los otros factores que influyen en el comercio. Efecto que al observarlo en el caso del español para el periodo 1996-2019 (por evitar la dislocación causada por la pandemia) se revela tan solo algo inferior a ese promedio; pero, sobre todo, claramente menor del que se obtenía para nuestra lengua hasta 2007, superando por entonces el promedio de otras lenguas, y también el efecto del inglés —más allá de su condición de *lingua franca* de los negocios internacionales— entre los países anglófonos (Jiménez Jiménez y Narbona Moreno, 2011, 2022).

Es decir, que el efecto multiplicador de la lengua sobre el comercio de bienes se mantiene, en el caso del español, dentro del molde general de las otras lenguas, esto es, aumentándolo en torno a un 100 % entre los países hispanohablantes, si bien contrastando con los resultados obtenidos en anteriores análisis, en particular antes de la Gran Recesión de 2008. Diríase, pues, que ese

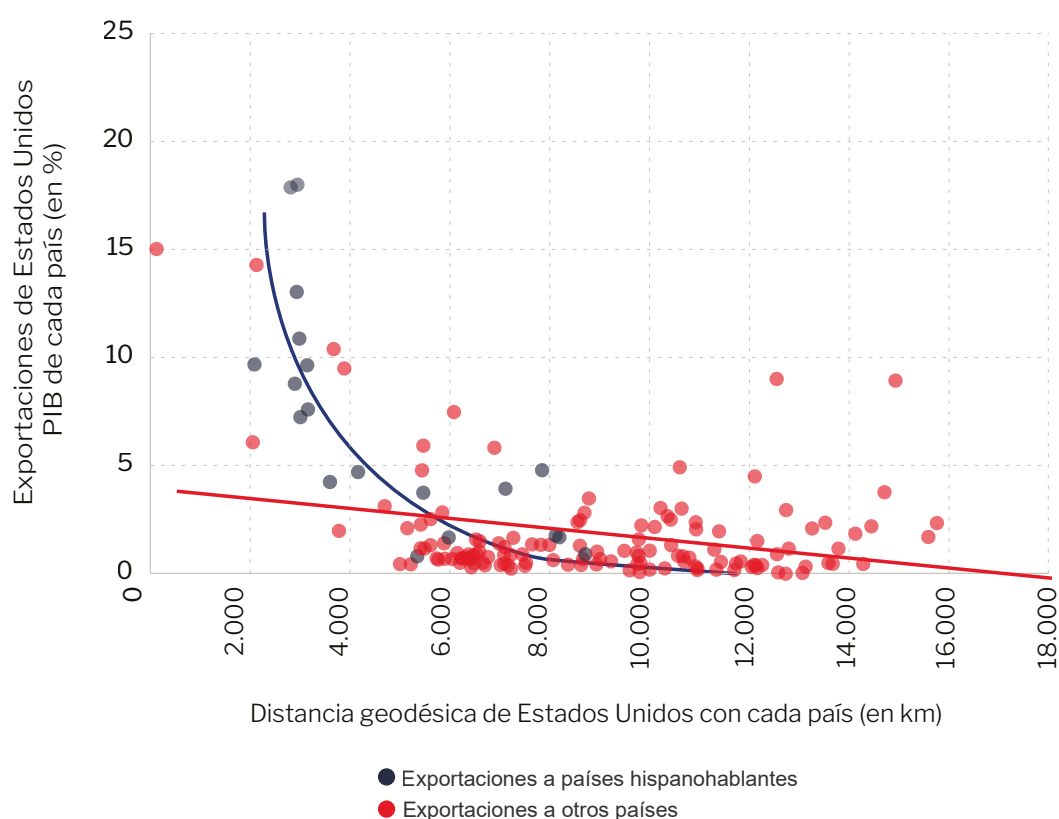
efecto facilitador del comercio del español, aunque mucho más importante en el pasado, lo sigue siendo, pero ya dentro de unos parámetros universales. Ahora bien, una de las limitaciones de estos cálculos es que la identificación de cada país (y, por tanto, del total de sus habitantes) con su lengua mayoritaria, supuesto que para el español guarda una razonable correspondencia en general, choca con el caso particular —pero importantísimo, sobre todo para México y las naciones centroamericanas— de Estados Unidos.

Por eso resulta muy interesante el intento de Martínez García y Martínez García (2022) y Martínez García *et al.* (2021) de cuantificar en qué medida las exportaciones e importaciones de Estados Unidos están condicionadas —por no decir favorecidas— por el hecho de que una amplia proporción de su población sea hispanohablante. Para ello ponen en relación la distancia geodésica de cada país con un índice relativo que mide la intensidad de las exportaciones que se dirigen desde Estados Unidos hacia ellos y de las importaciones que este país recibe de cada uno de ellos. Su conclusión principal es que el extendido conocimiento del español en Estados Unidos constituye una importante externalidad positiva que facilita sus relaciones comerciales bilaterales con los de habla hispana. En concreto, estiman que «las exportaciones de Estados Unidos son 160.000 millones de dólares más altas y las importaciones 109.000 millones más de lo que se predeciría si no tuviera una fuerte relación cultural y lingüística con los países de habla hispana».

Fijémonos en las exportaciones. El gráfico 2, con algunos retoques propios sobre los datos de base de los citados autores y el ajuste, exponencial en vez de logarítmico, permite apreciar con claridad cómo las exportaciones de Estados Unidos tienen un sesgo muy favorable hacia los países en los que se habla español —línea continua— respecto del resto —línea discontinua—. De nuestros cálculos se sigue que los países de habla hispana duplican largamente los índices de exportación hacia ellos por parte de Estados Unidos en relación con el resto de países, y casi triplican los de importación procedentes de ellos. No todo puede seguramente ser atribuido solo al español, pero es indudable que el hecho de ser la lengua más hablada en Estados Unidos tras el inglés no es ajeno a esa marcada inclinación comercial hacia los países hispanohablantes.

¿Sobrevivirá a una guerra arancelaria que cercene los intercambios en ambas direcciones (en el previsible escenario de que los países agraviados respondan)? Lo lógico es que no, lo que añade un elemento más de irracionalidad —y de inconveniencia para el papel del español— a esa política de introversión.

GRÁFICO 2 EXPORTACIONES DE ESTADOS UNIDOS EN RELACIÓN CON EL TAMAÑO DE LA ECONOMÍA, LA DISTANCIA Y LA LENGUA DE REFERENCIA DE CADA PAÍS SOCIO COMERCIAL (2015-2019)



FUENTE: elaboración propia a partir de Martínez-García et al. (2021), sobre la base de datos del FMI.

Un freno radical a los intercambios comerciales de Estados Unidos con el mundo hispánico, con México a la cabeza, tiene una segunda derivada negativa para el español en aquel país, al debilitar una fuente fundamental de su valor en el mercado laboral: el interés de las empresas por contar con hablantes bilingües que atiendan las oportunidades de negocio que abre —hasta ahora ha sido así— la interrelación económica con los países al sur de su frontera. Sería depreciar un activo de interés mutuo, por cuanto los beneficios de esa mayor actividad se extienden incluso a los que no hablan español.

De este modo, el nacionalismo económico en boga, que entiende no solo los intercambios de todo tipo con los otros países, sino también las lenguas, como parte de un juego de suma cero donde lo que obtiene otro es siempre a costa de uno mismo, despreciando así la riqueza que encierra el multilingüismo, dibuja, de extenderse, un panorama poco alentador para el futuro del español. Es una pena que una vez demostrados los beneficios cognitivos del multilingüismo —puede citarse, entre un sinfín de acreditados estudios, a Bialystok (2011)—, estos no puedan calar en las mentes más obtusas que se le oponen (y que, no por casualidad, tampoco suelen confiar en los trabajos científicos).

No obstante, la imposición de aranceles por parte del principal socio comercial de muchos de los países hispanohablantes, como es Estados Unidos, tal vez sea una ocasión de oro para reorientar, en aquellos productos que sea factible, los intercambios en la región. De hecho, el comercio intrarregional de América Latina —y esto vale para su amplia parte hispana— no alcanza las proporciones que se dan en otras grandes áreas del planeta y, de hecho, ha venido reduciéndose a lo largo de las últimas décadas, hasta situarse, a comienzos de esta, no más allá de un exiguo 12 % del total (Comisión Económica para América Latina y el Caribe [CEPAL], 2021). Dejando a un lado el caso de España dentro de su mercado regional (la Unión Europea), solo México ha estado integrado —dejémoslo de momento en pretérito perfecto compuesto— en un gran acuerdo, pero con sus dos vecinos del norte. Mercosur, la Comunidad Andina de Naciones o el Mercado Común Centroamericano han sido hasta ahora proyectos de integración limitados, lo que a China le ha servido para erigirse en pocos años en segundo socio comercial de toda la región. Si ahora el primero retrocede, se abre la oportunidad, que es a la vez un reto, de que un renovado regionalismo permita a los países hispanohablantes —más Brasil, aprovechando la proximidad lingüística— hacer de su lengua común el factor de estímulo comercial que los modelos nos dicen que tiene. Las empresas translatinas han sido quizá las primeras en apreciarlo. Cooperación e institucionalidad son las palabras clave.

En suma, que para el mundo hispánico el español —y todo el sustrato compartido que transmite— haya sido tradicionalmente una muleta con la que superar otras deficiencias y facilitar los intercambios mutuos, abre un camino con el que aliviar en parte el previsible freno al comercio mundial y, en particular, el que llega desde su gran cliente del norte: la profundización del mercado regional que compone buena parte del continente americano, al tiempo que se aprovechan las vías recientemente abiertas por el acuerdo comercial Mercosur-Unión Europea, contando con el pilar de impulso que debiera suponer la familiaridad lingüística del español y el portugués a ambos lados del Atlántico. Apuntado queda.

4. Los retos del español en la economía digital

En el apartado previo se ha hablado del comercio de bienes, históricamente el más importante a escala mundial, y también en el que se han centrado hasta ahora —por motivos añadidos de disponibilidad de datos— los análisis realizados desde la economía de la lengua con el fin de examinar la relación que guardan estas transacciones con el hecho de disfrutar de un idioma común. Pero desde hace décadas, y singularmente a raíz de la recesión posterior a 2008, el comercio de servicios viene cobrando un imparable protagonismo: si en 2008 este era apenas una cuarta parte del de mercancías, en 2023 (con los datos más recientes del Banco Mundial) se acerca ya a la tercera parte. De modo que hoy en día las exportaciones de servicios están creciendo tres veces más rápido que las de bienes. Esto es crucial desde la perspectiva de la economía de la lengua y del valor económico del español, que, por su número de hablantes (y su capacidad de compra) se sitúa en una muy favorable posición —al menos *prima facie*— para hacer suyas las externalidades de red que proporciona el tamaño.

Nos faltan aún series de datos para replicar en los servicios los análisis que se han hecho para desvelar la relación cuantitativa entre lengua y comercio (de bienes), pero sí podemos entrever que en esa parte ahora creciente de los intercambios, la de servicios, el papel facilitador de contar con una lengua común es superior a la de los bienes (primera ventaja): los

servicios requieren particularmente la interacción de los seres humanos y la «proximidad psicológica» que una lengua común procura. De igual modo que comprobamos que una parte esencial de ese comercio de servicios está muy relacionada con el propio desarrollo de la economía digital, que no solo activa el intercambio de intangibles sino que sirve también de cauce —a través del comercio electrónico y sus plataformas— a una parte creciente de los movimientos de bienes. Y aquí viene la segunda ventaja para el papel de una lengua común: los servicios (y el comercio electrónico minorista) escapan con mucha más facilidad que los bienes a los aranceles.

La economía digital es el nuevo terreno de juego donde la lengua tiene que hacer valer su carácter de «bien de club», capaz de difundir externalidades de red allí donde esta es precisamente la forma de interactuar. Y si decisivo en el aprovechamiento de esta ventaja es el número de hablantes, el español cuenta —también aquí— con un activo privilegiado. En efecto: que la economía digital salte más fácilmente sobre la geografía y las fronteras que la economía tradicional —y que la conectividad digital esté sustituyendo en muchas actividades al contacto presencial— no reduce en absoluto, sino antes al contrario, la importancia de la comunicación y de disponer de una lengua común que reduzca los costes de ponerse de acuerdo en los intercambios y de crear los lazos de confianza necesarios.

La pregunta ahora es si los hispanohablantes estamos preparados —o al menos bien posicionados— para aprovechar la ventaja de contar con una lengua hablada por más de 600 millones de personas en ese renacido comercio que abre la economía digital, deslizándose por encima de fronteras políticas, accidentes geográficos, distancias encarecedoras y hasta caprichosos aranceles. La información todavía es muy parcial, pero sí permite vislumbrar, junto a las oportunidades, si se hacen las cosas bien, algunas dificultades si estas no se hacen con la prontitud y ambición necesarias. Lo ha expresado recientemente la secretaria general de la Conferencia de las Naciones Unidas sobre Comercio y Desarrollo (UNCTAD), Rebeca Grynszpan, al referirse al papel de América Latina —sírvanos la generalización para aplicarla al

mundo hispanohablante del continente— en la digitalización del comercio: «La región puede liderar en plataformas digitales y conectividad, si invierte masivamente en tecnología, infraestructura y capacitación de la fuerza laboral» (Grynspan, 2025). La situación presente, sin embargo, está lejos de ese desiderátum.

La posición de los países hispanohablantes en el panorama mundial de la economía digital es aún manifiestamente mejorable. Muchas veces se alude al más favorable de los indicadores: el número de internautas, favorecido en nuestros países, en general, por el paso acelerado a la telefonía digital y la banda ancha desde un deficiente equipamiento en telefonía fija. Dentro de lo lábil y a veces poco coincidente de las estimaciones en este punto (seríamos el 7-8 % del total mundial, según unas u otras fuentes), sí puede decirse que la proporción de internautas hispanohablantes —la ratio de penetración o conectividad— se sitúa en torno del 75 %, esto es, unos diez puntos porcentuales por encima de los promedios mundiales (si bien persisten grandes diferencias dentro del condominio hispánico, desde la casi absoluta conectividad de los hispanos norteamericanos, españoles y hablantes del Cono Sur, hasta la mucho más limitada de centroamericanos o ecuatoguineanos). Eso significa, aplicando los mismos criterios que antes se siguieron para el conjunto de la población, que los internautas que hablan español poseen el 12,5 % de la capacidad de compra «virtual» del planeta². Porcentaje llamado a descender en parecida proporción a la capacidad de compra global, dependiendo de los escenarios antes manejados sobre la persistencia del español en Estados Unidos.

Pero ni el número de internautas —o su peso relativo— garantiza el aprovechamiento de las ventajas de la economía digital, sobre todo si su papel es más pasivo (usuarios) que activo (productores de contenidos) dentro de la red, ni el desarrollo de la sociedad de la información y de las TIC en los países de habla hispana alcanza los niveles deseables para ello. Los

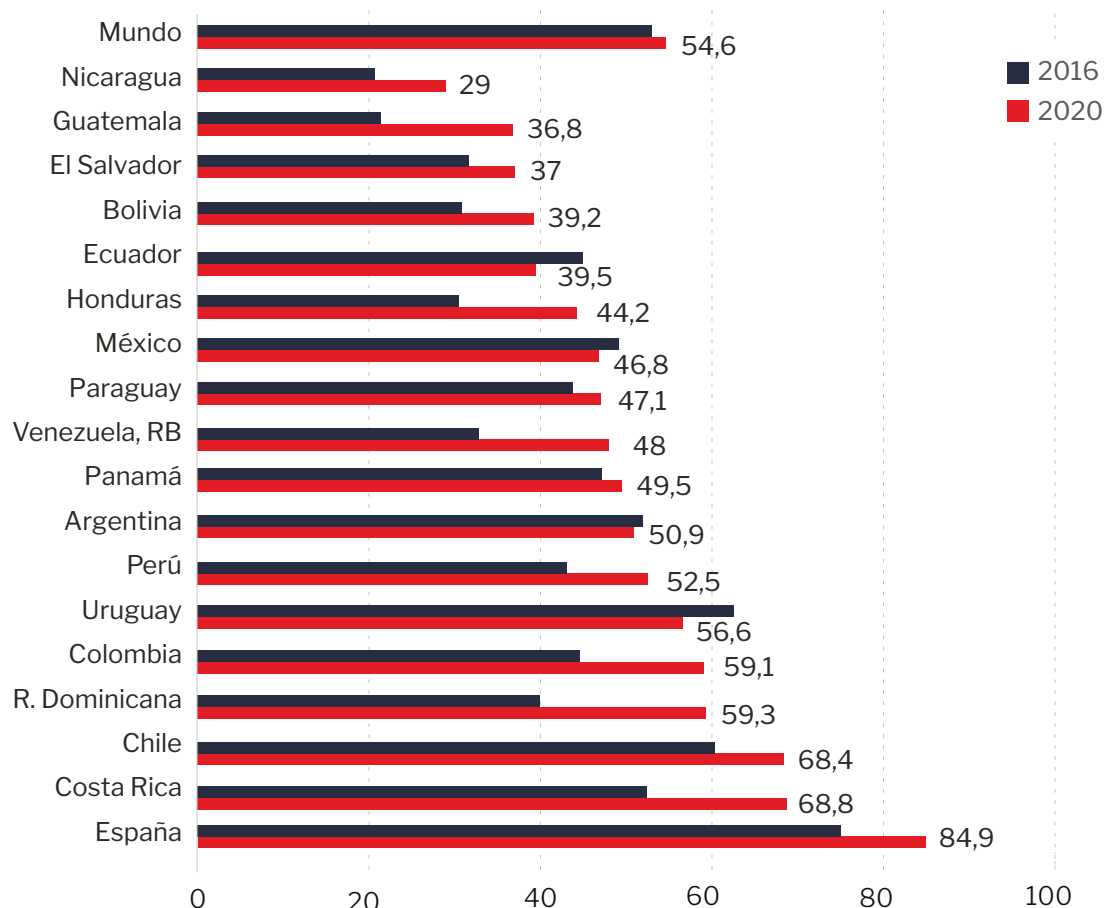
2. Teniendo en cuenta una «ratio de multilingüismo» del 43 % a escala mundial, de modo que la suma de la capacidad de compra de todas las lenguas superaría el 100 %. Véase Pimienta (2021).

principales indicadores globales de la economía digital —ICT Development Index (IDI) y Networked Readiness Index (NRI)— revelan, dentro de la heterogeneidad de los países hispanohablantes, una posición en general intermedia-baja en los respectivos *rankings*. Tomando los datos más actualizados de 2024 del NRI, solo 5 de los 16 países de habla hispana considerados dentro de la muestra de 133 naciones están por encima de la posición media de la clasificación; esto es, apenas un tercio³. Siete, en cambio, se sitúan en el tercio inferior de la escala, reproduciendo además un esquema que se repite desde 2017, lo que indica escasos avances en este terreno, al menos comparativamente con el resto del mundo.

Otro índice, este de la UNCTAD (B2C E-commerce Index), evalúa el grado de preparación de cada economía a la hora de apoyar las compras en línea —el comercio electrónico— de sus consumidores, a partir del promedio simple de cuatro indicadores (o «pilares»): (i) porcentaje de personas con titularidad de una cuenta bancaria; (ii) proporción de personas que utilizan internet; (iii) índice de fiabilidad postal; y (iv) servidores de internet seguros por cada millón de habitantes. Pues bien, los datos, llevados al gráfico 3, nos muestran de nuevo cómo solo un tercio, 6 de los 18 países hispanohablantes considerados, superan el índice promedio mundial, a pesar de una evolución, en general, positiva en este terreno. Hay, sin duda, factores estructurales que están lastrando la expansión potencial del comercio electrónico en nuestros países, a pesar de la amplia presencia de internautas de habla hispana. Pueden resumirse en tres: reducido grado de bancarización (esto es, de medios de pago utilizables electrónicamente), limitada fiabilidad postal (sistemas confiables de entrega de los bienes adquiridos) y carencia de servidores seguros de internet, una infraestructura esencial para el desarrollo de la sociedad de la información.

Gráfico 3 · Página siguiente >

3. Calculado a partir de <https://networkreadinessindex.org/>.

GRÁFICO 3 UNCTAD B2C E-COMMERCE INDEX DE LOS PAÍSES HISPANOHABLANTES 2016-2020 (0-100)

FUENTE: elaboración propia a partir de datos de la UNCTAD. Véase: https://unctad.org/system/files/official-document/tn_unctad_ict4d17_en.pdf.

En resumidas cuentas, no importa tanto estar conectado como *para qué* se está conectado. De poco sirve que seamos muchos y muy conectados a internet, si eso solo sirve para navegar pasivamente descargando contenidos, ver la televisión en línea, chatear o compartir correo electrónico, y no se traduce —por unas u otras trabas o insuficiencias— en el acceso a servicios como los financieros, contactar con la Administración pública, comprar en línea o subir contenidos propios, por ejemplo.

De nuevo aquí, cooperación e institucionalidad son las palabras clave para superar muchas de las dificultades que afronta el futuro del español en la nueva economía de redes y plataformas: es precisa una gobernanza común

del ecosistema digital en los países hispanohablantes, dando impulso y coordinación —dos propósitos siempre hazañosos en la comunidad hispánica— a algunas iniciativas que han surgido en este ámbito. Dos de ellas no pueden dejar de citarse aquí: por un lado, la Agenda Digital para América Latina y el Caribe (eLAC), promovida por la Comisión Económica para América Latina y el Caribe (CEPAL) y el Banco de Desarrollo de América Latina; por otro, en el terreno más específico de la lengua y del desarrollo de nuevas herramientas digitales para particulares y empresas, el proyecto Lengua Española e Inteligencia Artificial (LEIA), liderado por la Real Academia Española.

5. *Lingua est libertas*

NPor supuesto, no se agotan en los aspectos hasta aquí contemplados los riesgos que para la consolidación del valor económico del español encierra el nuevo escenario de relaciones internacionales y el papel de ariete que ha adoptado Estados Unidos. El caso de las migraciones es evidente en sí mismo y está, de hecho, en el germen de esa política aislacionista. Las modelizaciones que se han realizado para cuantificar qué efecto tiene el hecho de compartir un idioma —controlando el resto de variables— sobre la decisión de emigrar subrayan cómo la proximidad lingüística entre dos países, el de origen y el de destino, ejerce un papel muy positivo de estímulo sobre los flujos migratorios entre ambos (Alonso Rodríguez, Jiménez Jiménez y García Delgado, 2023).

El dominio del español, en concreto, incide con fuerza tanto en la decisión de emigrar como en la selección del destino: en el caso de la gran emigración desde la América hispana hacia España de los años 2000, triplicando las cifras de atracción y ofreciendo un «premio salarial» por el dominio de la lengua que alcanzó hasta un 30 % (Alonso Rodríguez y Gutiérrez Palacios, 2010). Y no menos importante: esa comunidad lingüística es un factor explicativo esencial del éxito migratorio, reduciendo los costes y facilitando la integración. Más allá de los flujos masivos de emigración hispana hacia Estados Unidos que se dieron en el pasado —movidos por factores socioeconómicos, pero con el español allanando el camino—, lo cierto es que cerca del 60 % de la

inmigración registrada en la gran mayoría de países hispanohablantes procede de otro país de lengua hispana. Convendría preservar esa movilidad, que no es sino fuente de progreso para los países receptores y alivio de tensiones y generación de divisas para los emisores.

Hoy más que nunca, cuando soplan vientos de frente, no cabe mantenerse al páiro: si el español quiere acrecentar su valor económico debe basarse en el progreso económico de los países que lo hablan. ¿Cómo? Invirtiendo en los factores que favorecen su inserción en la economía digital (infraestructuras, educación...), mejorando sus niveles de calidad institucional —tan importantes como la lengua común para estimular los intercambios— y salvaguardando la libertad de movimientos de factores (personas y capitales) y productos. Nada, por otro lado, sostendrá mejor el valor del español dentro de Estados Unidos como la pujanza que alcance, fuera y en todos los órdenes, el resto del mundo hispanohablante.

La lengua, parafraseando al padre de la economía, Adam Smith, es lo que hace posible el intercambio y, con ello, la división del trabajo —sin distinción de fronteras—, base de la prosperidad de las naciones. Por eso mismo, necesita de la libertad para desplegar sus potencialidades. Aplicando a la economía la reflexión de Manuel Alvar (1983), «poseer una lengua no es estar emparedados en una celda que nos aprisiona, sino disponer de una clave para descubrir a la luz». Ojalá la facultad discursiva de los seres humanos se imponga finalmente a muros, aranceles y fronteras. A las lenguas les sienta mejor la libertad.

Referencias bibliográficas

- AGIRDAG, O. (2014). The literal cost of language assimilation for the children of immigration: The effects of bilingualism on labor market outcomes. En R. M. Callahan y P. C. Gándara (Eds.), *The bilingual advantage. Language, literacy, and the U.S. labor market* (pp. 160-181). Channel View Publications.
- ALVAR LÓPEZ, M. (1983). *La lengua como libertad*. Cultura Hispánica.
- ALONSO RODRÍGUEZ, J. A. Y GUTIÉRREZ PALACIOS, R. (DIRS.) (2010). *Emigración y lengua: El papel del español en las migraciones internacionales* (Colección Fundación Telefónica). Ariel.
- ALONSO RODRÍGUEZ, J. A., JIMÉNEZ JIMÉNEZ, J. C. Y GARCÍA DELGADO, J. L. (2023). *Los futuros del español. Horizonte de una lengua internacional*. Alianza Editorial.
- BIALYSTOK, E. (2011). Reshaping the mind: The benefits of bilingualism. *Canadian Journal of Experimental Psychology / Revue canadienne de psychologie expérimentale*, 65 (4), 229-235. <https://criancabilingue.wordpress.com/wp-content/uploads/2013/09/cep-65-4-229.pdf>
- COMISIÓN ECONÓMICA PARA AMÉRICA LATINA Y EL CARIBE (CEPAL) (2021). *Perspectivas del Comercio Internacional de América Latina y el Caribe, 2021*. CEPAL. <https://repositorio.cepal.org/entities/publication/acfd934b-f47a-4e9e-ab2d-410d3ef4a9fe>
- GÁNDARA, P. C. (2015). Is There Really a Labor Market Advantage to Being Bilingual in the U.S.? *Policy Information Report and ETS Research Report Series RR-15-07*.
- GRYNSPAN, R. (7 de enero de 2025). *Por una economía global más justa: el papel de América Latina*. UNCTAD. <https://unctad.org/es/news/por-una-economia-global-mas-justa-el-papel-de-america-latina>
- JIMÉNEZ JIMÉNEZ, J. C. (2007). *El español: valor de un activo económico*. Fundación Caja de Burgos e Instituto Universitario de Investigación Ortega y Gasset.
- JIMÉNEZ JIMÉNEZ, J. C. (2009). El poder de compra del español en el mundo. *Revista de Occidente*, 335, 21-36.
- JIMÉNEZ JIMÉNEZ, J. C. Y NARBONA MORENO, A. (2011). *El español en los flujos económicos internacionales. Un instrumento de la internacionalización empresarial*. (Colección Fundación Telefónica). Ariel.
- JIMÉNEZ JIMÉNEZ, J. C. Y NARBONA MORENO, A. (2022). *El español en los flujos de comercio internacional: horizonte 2030-2050*. Observatorio Nebrija del Español. <https://www.nebrija.com/catedras/observatorio-nebrija-espanol/pdf/espanol-flujos-comercio-intern.pdf>
- LACOMBA, C. (2023). *Mapa hispano de los Estados Unidos 2022* (serie Estudios del Observatorio/Observatorio Studies, n.º 84). Instituto Cervantes at the Faculty of Arts and Sciences of Harvard University. https://cervantesobservatorio.fas.harvard.edu/sites/default/files/84_sp_mapa_hispano_2022_1.pdf

- MARTÍN LERALTA, S., JIMÉNEZ JIMÉNEZ, J. C. Y CARRERA TROYANO, M. (2024). *La certificación en español: un universo valioso. Cartografía y valor económico*. Instituto Cervantes y Fundación Nebrija.
- MARTÍNEZ GARCÍA, E. Y MARTÍNEZ GARCÍA, M. T. (2022). *El valor económico del español en Estados Unidos: Oportunidades y retos para el futuro* (serie Estudios del Observatorio/Observatorio Studies, n.º 78). Instituto Cervantes at the Faculty of Arts and Sciences of Harvard University. https://cervantesobservatorio.fas.harvard.edu/sites/default/files/78_sp_martinez_y_martinez_junio_28.pdf
- MARTÍNEZ GARCÍA, E., MARTÍNEZ GARCÍA, M. T., COULTER, J. Y GROSSMAN, V. (13 de julio de 2021). Se habla español: U.S. yet to realize many benefits of a growing bilingual population, *Dallas Fed Economics Blog*. <https://www.dallasfed.org/research/economics/2021/0713#:~:text=Growing%20Spanish%E2%80%93English%20bilingualism%20in,compared%20with%20monolingual%20English%20speakers>.
- PIMIENTA, D. (2021). *Versión nueva y mejorada de un enfoque alternativo para la producción de indicadores lingüísticos en la Internet*. Observatorio de la diversidad lingüística y cultural en la Internet. <https://funredes.org/lc2021/ALI%20V2-ES.pdf>
- RUMBAUT, R. G. (2014). English plus: Exploring the socioeconomic benefits of bilingualism in Southern California. En R. M. Callahan y P. C. Gándara (Eds.), *The bilingual advantage. Language, literacy, and the U.S. labor market* (pp. 182-205). Channel View Publications.
- SANTIBAÑEZ, L. Y ZÁRATE, M. E. (2014). Bilinguals in the U.S. and college enrollment. En R. M. Callahan y P. C. Gándara (Eds.), *The bilingual advantage. Language, literacy, and the U.S. labor market* (pp. 211-233). Channel View Publications.

INTELIGENCIA ARTIFICIAL Y LENGUA ESPAÑOLA: PANORAMA ACTUAL Y PERSPECTIVAS

Elena González-Blanco García
MICROSOFT

1. Introducción

La rápida evolución de la inteligencia artificial, especialmente tras la llegada de la IA generativa, está transformando el procesamiento del lenguaje natural en todo el mundo, y la lengua española no es una excepción. Con más de 600 millones de hispanohablantes a escala global, el español representa un patrimonio lingüístico de enorme importancia cuyo futuro en la era digital ofrece grandes promesas, aunque también retos significativos. En el presente artículo se analiza el estado de la IA aplicada al español y al mundo hispanohablante, abarcando las estrategias públicas que han impulsado esta área, los recursos lingüísticos disponibles, el ecosistema empresarial emergente, los grupos de investigación más destacados y el impacto de estas tecnologías en el español y sus variantes.

2. La inteligencia artificial y el lenguaje: historia de una disciplina científica

La historia de la inteligencia artificial (IA) y su aplicación al lenguaje natural es una de las más fascinantes transformaciones del mundo contemporáneo. A lo largo de las últimas siete décadas, esta disciplina ha pasado de ser una utopía filosófica a convertirse en una realidad tecnológica omnipresente, que afecta tanto a la enseñanza de lenguas como a la traducción automática, pasando por la asistencia conversacional y la generación textual.

Los primeros pasos de la IA se remontan a la década de 1950, cuando el matemático Alan Turing propuso su célebre test para determinar si una máquina

podía «pensar» mediante una conversación en lenguaje natural (Turing, 1950). En paralelo, lingüistas como Noam Chomsky sentaron las bases formales de la sintaxis generativa, lo que permitió que los primeros sistemas informáticos intentaran procesar el lenguaje humano con reglas explícitas (Chomsky, 1957). En 1954, un experimento en la Universidad de Georgetown logró traducir frases del ruso al inglés mediante un sistema de reglas (Hutchins, 2014), lo que dio lugar a grandes expectativas sobre la traducción automática. Sin embargo, en 1966 un informe del Gobierno estadounidense concluyó que esos avances eran modestos y la financiación para el procesamiento del lenguaje natural (PLN) se redujo drásticamente, iniciando una etapa conocida como el «invierno de la IA» (Automatic Language Processing Advisory Committee [ALPAC], 1966).

Durante los años setenta y ochenta, el paradigma dominante fue el de los sistemas simbólicos: programas que seguían reglas programadas por humanos para entender o generar frases. Un ejemplo famoso fue ELIZA, un rudimentario *chatbot* diseñado para simular a un terapeuta (Weizenbaum, 1966). Más ambicioso fue SHRDLU, que permitía dar órdenes en inglés a un mundo virtual de bloques (Winograd, 1972). Sin embargo, estos sistemas solo funcionaban bien en contextos muy limitados. El gran cambio comenzó hacia finales de los años ochenta y noventa, cuando el aumento del poder computacional y la disponibilidad de grandes corpus bilingües (gracias, por ejemplo, al Parlamento Europeo) permitieron abandonar los sistemas de reglas y adoptar métodos estadísticos (Koehn, 2010). Por primera vez, las máquinas comenzaron a «aprender» patrones lingüísticos a partir de grandes cantidades de texto, sin intervención humana directa. Esto dio lugar a los sistemas de traducción estadística y a los primeros intentos de usar redes neuronales para el lenguaje (Hochreiter y Schmidhuber, 1997).

El siglo ^{XXI} trajo consigo avances fundamentales. En 2013 se popularizó la idea de representar palabras como vectores numéricos, capturando sus significados mediante proximidad matemática y mediante la representación vectorial (Mikolov *et al.*, 2013). Poco después, Google introdujo la traducción automática neuronal, que superó con creces a los sistemas anteriores (Wu *et al.*, 2016). Pero el punto de inflexión llegó en 2017 con la publicación del

artículo «Attention Is All You Need», que proponía una nueva arquitectura informática llamada Transformer (Vaswani *et al.*, 2017). Este modelo revolucionó el procesamiento del lenguaje al permitir que las máquinas entendieran contextos completos de texto de manera mucho más eficiente, abandonando los enfoques anteriores basados en secuencias.

Desde entonces, el progreso ha sido vertiginoso. En 2018, Google presentó BERT, un modelo basado en vectores que anticipaba la comprensión profunda del texto (Devlin *et al.*, 2018). OpenAI, por su parte, lanzó su serie de modelos GPT (Generative Pre-trained Transformer), entrenados con cantidades masivas de texto en múltiples idiomas (Radford *et al.*, 2018). En noviembre de 2022, el lanzamiento de ChatGPT, basado en GPT-3.5, supuso un antes y un después: en solo cinco días superó el millón de usuarios y dos meses después alcanzaba los 100 millones (OpenAI, 2022). Este modelo permitía mantener conversaciones fluidas, redactar textos, resolver dudas o escribir código, todo en lenguaje natural. En 2023, OpenAI presentó GPT-4, con capacidades aún más sofisticadas, incluyendo la interpretación de imágenes y gráficos (OpenAI, 2023).

La popularización de estos sistemas generativos provocó una carrera tecnológica global. Google lanzó su serie Gemini (antes Bard), Meta publicó los modelos LLaMA¹ de código abierto y surgieron otros como Claude², desarrollado por Anthropic, o la serie Phi, especializada en modelos pequeños pero eficientes. A escala global, la carrera es imparable, como han demostrado países como China, que ha desarrollado a un coste menor y gran velocidad sus propios modelos (DeepSeek, Yi-34B, Moonshot, SparkDesk).

En paralelo, la región de Oriente Medio ha emergido como un actor relevante en el desarrollo de inteligencia artificial, con iniciativas lideradas principalmente desde los Emiratos Árabes Unidos. El Instituto de Innovación en Inteligencia Artificial (IIA)³ de Abu Dhabi, en colaboración con el grupo tecnológico G42, desarrolló la familia de modelos Falcon, destacando Falcon

1. Véase: <https://www.llama.com/>.

2. Véase: <https://claude.ai/>.

3. Véase: <https://www.tii.ae/ai-and-digital-science>.

40B y Falcon 180B, este último lanzado en 2023 y reconocido como uno de los primeros modelos de código abierto más potentes. Asimismo, en 2024 se presentó la empresa AI71⁴, impulsada también por G42 e IIA, con el objetivo de desarrollar soluciones seguras y personalizadas de IA para gobiernos y grandes instituciones. Esta apuesta estratégica sitúa a Oriente Medio en la vanguardia del desarrollo de IA generativa, promoviendo una soberanía tecnológica regional y ampliando el mapa global de innovación más allá de los polos tradicionales de Estados Unidos, Europa y China.

3. El panorama de la IA en español: corpus, recursos y modelos

El desarrollo y entrenamiento de los modelos de inteligencia artificial requiere gran cantidad de datos y, por ende, de recursos lingüísticos: corpus textuales y orales, léxicos computacionales o bases terminológicas, entre otros. La lengua española cuenta con una larga tradición en la compilación de corpus, iniciada por proyectos pioneros de la Real Academia Española (RAE) y las academias asociadas, y continuada hoy por iniciativas abiertas de diverso tipo.

Entre los corpus de referencia clásicos destacan los creados bajo los auspicios de la Real Academia Española (RAE) para sus labores lexicográficas. El Corpus de Referencia del Español Actual (CREA)⁵ recopiló versiones electrónicas de textos en español producidos entre 1975 y 2004, con el fin de documentar el uso real contemporáneo del idioma tanto en España como en América. El CREA, lanzado en los años noventa, alcanzó varios cientos de millones de palabras, abarcando diversidad de géneros (un 90 % de textos escritos —libros, prensa, blogs— y un 10 % de transcripciones orales) y asegurando un equilibrio geográfico del 50 % de materiales de España y 50 % de Hispanoamérica. Por su parte, el Corpus Diacrónico del Español (CORDE)⁶, concebido para fines históricos, reúne

4. Véase: <https://ai71.ai/>.

5. Véase: <https://corpus.rae.es/creanet.html>.

6. Véase: <https://www.rae.es/banco-de-datos/corde>.

unos 300 millones de palabras de textos representativos de todas las épocas del español, desde sus orígenes medievales hasta el siglo ^{xx}. El CORDE, publicado a inicios de los 2000, integra unas 5.500 obras completas (literarias, documentales, científicas, etc.) cuidadosamente seleccionadas para reflejar la evolución del español peninsular e hispanoamericano a lo largo de más de mil años. Estas colosales bases de datos —CREA y CORDE— han sido herramientas fundamentales para la investigación lingüística en español y sirvieron de fundamento para obras como el *Nuevo diccionario histórico del español*.

En el siglo ^{xxi}, la RAE y la Asociación de Academias de la Lengua Española (ASALE) emprendieron un nuevo corpus sincronizado de gran escala: el Corpus del Español del Siglo XXI (CORPES XXI)⁷. Iniciado en 2013, el CORPES se plantea recoger 300 millones de formas ortográficas producidas entre el año 2001 y 2030, cubriendo todos los países hispanohablantes y variedad de registros. Las actualizaciones periódicas del CORPES dan cuenta de su crecimiento: la versión 0.92 (publicada en junio de 2020) alcanzó 312 millones de formas ortográficas procedentes de más de 300.000 documentos (incluyendo por primera vez una porción significativa de transcripciones orales).

Asimismo, la disponibilidad de corpus en español se ha visto potenciada por iniciativas internacionales de *web crawling* y datos abiertos. Por ejemplo, el corpus esTenTen (parte de la serie TenTen de corpus de internet) alcanzó los 8.000 millones de palabras en español, extraídas de la web e incluyendo textos de 19 variedades nacionales del idioma (Jakubíček *et al.*, 2013). Este tipo de corpus masivo, compilado por universidades europeas, es valioso para entrenar sistemas de *machine learning* que requieran cantidades ingentes de texto. En el ámbito multilingüe, abundan también los «corpus paralelos» que incluyen al español, usados sobre todo para traducción automática: desde el clásico Europarl⁸ (procedente del Parlamento Europeo) hasta colecciones como OpenSubtitles, ParaCrawl o WikiMatrix, con cientos de millones de segmentos bilingües español-inglés, español-francés, etc.

7. Véase: <https://www.rae.es/corpes/>.

8. Véase: <https://www.statmt.org/europarl/>.

En efecto, la comunidad de PLN ha ido construyendo repositorios abiertos donde el español está bien representado. Un reflejo de esto es la plataforma Hugging Face Datasets⁹, que a fecha de 2025 ofrece docenas de conjuntos de datos en español: desde colecciones para *benchmarking* (p. ej., Semantic Textual Similarity en español, traducciones del cuestionario de conocimiento general MMLU al español o el set ArguAna de consultas argumentativas traducidas) hasta corpus especializados como MilaNLProc HONEST¹⁰ (diseñado para evaluar sesgos en generación de texto en español). Muchos de estos datos han sido recopilados gracias a competiciones y foros como el IberLEF (Iberian Languages Evaluation Forum)¹¹, donde equipos de España y Latinoamérica comparten datos para tareas de análisis de sentimientos, detección de lenguaje ofensivo o resumen automático, entre otras.

Un recurso fundamental, probablemente menos reconocido aún en el ámbito de la IA, son los corpus orales y fonéticos en español, que han cobrado relevancia con el auge del reconocimiento de voz y la síntesis de habla por IA. Iniciativas globales como Common Voice¹² de Mozilla incluyen aportaciones masivas de audio de voluntarios en español (y también en catalán, gallego, euskera, etc.), proporcionando decenas de miles de grabaciones que sirven para entrenar algoritmos de *speech-to-text* y *text-to-speech*. Del mismo modo, diversas universidades han compilado corpus de habla espontánea en distintas variedades (por ejemplo, corpus de diálogos telefónicos, de entrevistas o de radio en español de América), aunque estos recursos suelen estar menos centralizados que los textuales, pero la ventaja es que, por su distribución desigual y su desconocimiento, en la mayor parte de los casos no han sido utilizados para entrenar LLMs ni algoritmos de IA.

9. Véase: <https://huggingface.co/>.

10. Véase: <https://github.com/MilaNLProc/honest>.

11. Véase: <https://sites.google.com/view/iberlef-2025>.

12. Véase: <https://commonvoice.mozilla.org/>.

En suma, el ecosistema de recursos lingüísticos en español combina «grandes corpus generales» (RAE, BNE, web) con «*datasets* específicos» modernos, accesibles en repositorios abiertos. Esta abundancia de datos —inimaginable hace apenas dos décadas— constituye el combustible que alimenta los avances actuales en IA del lenguaje. No obstante, siguen existiendo áreas que reforzar, como la disponibilidad de datos etiquetados de calidad para tareas complejas (análisis semántico profundo, co-referencia, etc.) o corpus representativos de registros coloquiales y de redes sociales. La colaboración entre instituciones (académicas, gubernamentales y empresariales) está siendo crucial para identificar vacíos y crear nuevos recursos que continúen enriqueciendo el panorama del español en la era de la IA.

En el ámbito hispanohablante, en los últimos años han surgido modelos de lenguaje entrenados específicamente en español que marcan un salto cualitativo en la aplicación de inteligencia artificial al ecosistema hispanohablante. Uno de los primeros fue el modelo MarIA, desarrollado por el Barcelona Supercomputing Center — Centro Nacional de Supercomputación (BSC-CNS), considerado el primer modelo de lenguaje masivo entrenado íntegramente en español, utilizando como base los fondos digitales de la Biblioteca Nacional de España (Gutiérrez Fandiño *et al.*, 2021). En su momento, mostró un rendimiento notable en tareas como comprensión lectora, clasificación textual y generación de contenidos, siendo especialmente útil en contextos que requieren sensibilidad cultural o normativa. Tras la aparición de ChatGPT, el primer modelo generativo español, creado por la *startup* española Clibrain, fue LINCE¹³, una familia de modelos concebida específicamente para abarcar el español y todas sus variantes que obtuvo excelentes resultados para español en los *benchmarks* de evaluación. LINCE, lanzado comercialmente en 2023 y disponible en Huggingface en una versión abierta de 7.000 millones de parámetros, logró resultados comparables en calidad a GPT-3 pero optimizados para el español y destaca por su capacidad de manejar eficazmente las variantes dialectales del español.

13. Véase: <https://techcrunch.com/2023/07/12/lince-llm/>.

GRÁFICO 1 LINCE FAMILY MODELS, BY CLIBRAIN

Models 19

clibrain/gemma-2b-coder

Updated Nov 28, 2024

clibrain/mamba-2.8b-ft-synthia-v1.3

Updated Dec 31, 2023 • 11

clibrain/Llama-2-7b-ft-instruct-es-sharded-bf16

Text Generation • Updated Nov 13, 2023 • 3

clibrain/lince-zero

Text Generation • Updated Nov 7, 2023 • 133 • 48

clibrain/Llama-2-7b-ft-instruct-es-gguf

Updated Oct 23, 2023 • 3

clibrain/Llama-2-7b-ft-instruct-es-gptq-3bit

Text Generation • Updated Sep 20, 2023 • 3

clibrain/Llama-2-13b-ft-instruct-es-gptq-4bit

Text Generation • Updated Sep 4, 2023 • 13 • 3

clibrain/lince-zero-f16-ggml-q4_0

Updated Aug 30, 2023 • 5

clibrain/idefics-9b-ft-describe-diffusion

Updated Aug 24, 2023 • 2

clibrain/Llama-2-7b-ft-instruct-es

Text Generation • Updated Aug 9, 2023 • 1.26k • 25

clibrain/mamba-2.8b-chat-no_robots

Text Generation • Updated Jan 23, 2024 • 4 • 15

clibrain/mamba-2.8b-instruct-openhermes

Text Generation • Updated Dec 31, 2023 • 140 • 72

clibrain/lince-mistral-7b-it-es

Text Generation • Updated Nov 7, 2023 • 93 • 27

clibrain/Llama-2-13b-ft-instruct-es-gguf

Updated Oct 23, 2023

clibrain/Llama-2-13b-ft-instruct-es-gptq-3bit

Text Generation • Updated Sep 20, 2023 • 1 • 3

clibrain/legal-longformer-base-8192-spanish

Fill-Mask • Updated Sep 12, 2023

clibrain/Llama-2-7b-ft-instruct-es-gptq-4bit

Text Generation • Updated Sep 1, 2023 • 14 • 9

clibrain/Llama-2-13b-ft-instruct-es

Text Generation • Updated Aug 30, 2023 • 1.28k • 9

clibrain/Llama-2-ft-instruct-es

Text Generation • Updated Aug 9, 2023 • 1.24k • 18

FUENTE: <https://huggingface.co/clibrain>.

Asimismo, en foros de evaluación como IberLEF, se han puesto a prueba modelos específicamente optimizados para el español en tareas como detección de sentimientos, lenguaje ofensivo o análisis semántico. Entre ellos destacan BETO, una versión adaptada de BERT al español, y RoBERTuito (Pérez *et al.*, 2022), una variante de RoBERTa enfocada en el español hispanoamericano, que han obtenido resultados sobresalientes en múltiples *benchmarks* (Cañete *et al.*, 2020).

Ahora bien, este impulso reciente no habría sido posible sin una base sólida de herramientas y recursos computacionales desarrollados a lo largo de las

dos últimas décadas. Como ejemplo de algunos de estos proyectos pioneros, cabe citar FreeLing (Carreras *et al.*, 2004), una suite de análisis lingüístico desarrollada por la Universitat Politècnica de Catalunya, que proporciona funcionalidades como tokenización, lematización, análisis morfosintáctico, etiquetado gramatical, análisis sintáctico y reconocimiento de entidades, y que fue utilizada ampliamente en entornos industriales y académicos multilingües. En el campo de la traducción automática, destaca Apertium (Forcada *et al.*, 2011), una plataforma de código abierto especializada en pares de lenguas emparentadas (como español-catalán o español-gallego), ampliamente utilizada en gobiernos autonómicos y en la enseñanza del procesamiento lingüístico.

El análisis sintáctico profundo en español también ha contado con corpus anotados como AnCora (Taulé Delor, Martí Antonín y Recasens Potau, 2008), desarrollado por la Universitat de Barcelona y la Universitat Autònoma de Barcelona, que incluye anotaciones sintácticas y semánticas tanto en español como en catalán. La herramienta AnCoraPipe permite su anotación automática y ha sido ampliamente empleada para el entrenamiento y evaluación de *parsers*. Por su parte, el grupo IXA de la Universidad del País Vasco ha desarrollado el conjunto de herramientas IXA Pipes, centradas en el análisis morfosintáctico, la desambiguación y la anotación semántica en español y euskera, dentro de una visión multilingüe del procesamiento lingüístico (Agerri Gascón, Bermúdez y Rigau Claramunt, 2014).

En América Latina, el Grupo de Ingeniería Lingüística de la Universidad Nacional Autónoma de México (UNAM) ha producido una serie de herramientas específicas para el análisis del español mexicano, conocidas como NLP Tools del grupo GIL, aplicadas tanto a proyectos de PLN como a estudios de lingüística computacional (Galicia Haro, Gelbukh y Bolshakov, 2002). Asimismo, el grupo TALP de la Universitat Politècnica de Catalunya ha impulsado herramientas de traducción automática y extracción de información bajo el nombre TALP Tools, con especial enfoque en métodos estadísticos y de aprendizaje automático (Crego Clemente, 2008). Otro recurso fundamental ha sido UDPipe, un conjunto de modelos sintácticos entrenados para

múltiples idiomas (incluido el español) conforme a los estándares de Universal Dependencies que proporciona análisis de dependencias gramaticales, lematización y etiquetado POS (Straka y Strakova, 2017). Finalmente, cabe mencionar LinguaKit (Gamallo et al., 2018), una herramienta desarrollada por el grupo CITIUS de la Universidad de Santiago de Compostela, que combina análisis sintáctico, extracción de entidades y clasificación de sentimientos, y que ha sido utilizada tanto en la investigación académica como en prototipos empresariales; y, finalmente, los recursos desarrollados por la Universidad de Stanford, que si bien estuvieron principalmente enfocados en el inglés, contaron también con librerías específicas que se desarrollaron en español a partir de los sistemas de NLTK¹⁴.

Estos recursos, creados en su mayoría en universidades e instituciones públicas, fueron decisivos para el desarrollo inicial del procesamiento automatizado del español. Aunque muchos han sido superados en precisión por modelos de IA generativa, siguen siendo relevantes en tareas especializadas, en la creación de corpus lingüísticos controlados, en la formación de estudiantes y en la evaluación comparada de nuevos sistemas.

4. Grupos de investigación en IA y lenguaje en el mundo hispanohablante

El avance de la IA aplicada al español ha sido posible gracias a décadas de trabajo de grupos académicos y de investigación especializados en procesamiento del lenguaje natural, inteligencia artificial y *machine learning*. En el mundo hispanohablante existe una comunidad científica activa y consolidada, con centros de excelencia que han contribuido tanto al desarrollo de recursos y algoritmos como a la formación de nuevos investigadores.

En España, múltiples comunidades autónomas cuentan con al menos un equipo destacado en tecnologías del lenguaje, y los más prolíficos han sido aquellos que se encuentran vinculados a la existencia de lenguas cooficiales.

14. Véase: <https://www.nltk.org/index.html>.

Uno de los grupos más prominentes es la Unidad de Tecnologías del Lenguaje del BSC-CNS, en Barcelona. Este grupo ha estado al frente de proyectos nacionales como Plan TL, MarIA o los modelos ALIA (Barcelona Supercomputing Center, 2025) y otros europeos, combinando la potencia de supercomputación con la lingüística computacional.

Las universidades también han desempeñado un papel fundamental en este ecosistema. El grupo IXA de la Universidad del País Vasco es reconocido internacionalmente por sus trabajos en traducción automática multilingüe (especialmente entre euskera, español e inglés) y en procesamiento del lenguaje biomédico. IXA ha contribuido a herramientas de código abierto (por ejemplo, al motor de traducción Apertium para pares como español-euskera) y ha liderado proyectos de PLN clínico en colaboración con sistemas de salud regionales. En Cataluña, la Universitat Politècnica de Catalunya cuenta con el grupo TALP (Tecnologies del Llenguatge i la Parla) que lleva décadas investigando en traducción automática estadística y neuronal, así como en *speech recognition*; y la Universitat Pompeu Fabra alberga el grupo TALN, con aportes en resumen automático y análisis semántico. En Madrid, la Universidad Politécnica de Madrid y la Universidad Nacional de Educación a Distancia (UNED) han contado con grupos activos en recuperación de información y PLN: la primera de ellas, por ejemplo, fue pionera en España en temas de QA (pregunta-respuesta) y la UNED organiza desde hace años competiciones internacionales de búsqueda de respuestas en español (tareas CLEF). Otro referente es el PRHLT Research Center de la Universitat Politècnica de València, que se especializa en minería de textos, detección de *hate speech* y plagio, habiendo coordinado retos en análisis de opiniones en español. El Instituto de Ingeniería del Conocimiento (IIC), ligado a la Universidad Autónoma de Madrid, lleva años aplicando PLN a industria (bancos, encuestas, etc.); el Consejo Superior de Investigaciones Científicas (CSIC) ha albergado proyectos en el Instituto de Lengua, Literatura y Antropología orientados a lingüística de corpus y a dialectología que, si bien no estaban inicialmente orientados a PLN, son una fuente valiosísima de información a través de sus *Atlas Lingüísticos*.

Asimismo, la reciente apertura de la unidad ELLIS Alicante (Laboratorio Europeo de Aprendizaje y Sistemas Inteligentes) trae un nodo internacional enfocado en IA que incluye investigación en modelos de lenguaje y aprendizaje profundo, lo que podría generar sinergias con el ámbito lingüístico.

En cuanto a asociaciones científicas, cabe resaltar la Sociedad Española para el Procesamiento del Lenguaje Natural (SEPLN), fundada en los años ochenta, que ha sostenido una comunidad académica cohesionada y que organiza un congreso anual con publicaciones de relevancia científica importante en el sector.

Hispanoamérica cuenta también con grupos de investigación notables en PLN, a pesar de disponer de recursos más limitados en algunos casos. México es uno de los países más activos: el Laboratorio de Lenguaje Natural del Instituto Politécnico Nacional (IPN) ha producido numerosos trabajos en semántica y lingüística computacional; la UNAM alberga el Grupo de Ingeniería Lingüística (GIL) que desarrolla desde diccionarios electrónicos hasta sistemas de extracción de información enfocados en el español de México. En América del Sur, Argentina tiene la Sociedad Argentina de Tecnologías del Lenguaje Humano, que reúne a investigadores locales; en la Universidad de Chile y la Pontificia Universidad Católica de Chile existen equipos que colaboraron en proyectos regionales (p. ej. desarrollo de *spell-checkers* para español de Chile, análisis de corpus del español de redes sociales chilenas). Países como Colombia y Perú han incrementado su presencia: por ejemplo, la Universidad de los Andes (Colombia) ha participado en competiciones de traducción automática al español y en Perú se han empezado a explorar herramientas para traducción entre castellano y quechua usando IA. Muchos de estos esfuerzos están interconectados a través de eventos iberoamericanos (el congreso CICLing en México o la conferencia IBERAMIA de IA, que suele incluir sesiones de PLN, etc.).

Mención especial merecen las iniciativas comunitarias internacionales impulsadas por jóvenes investigadores hispanohablantes, que han surgido para democratizar el PLN en nuestro idioma. Un caso reciente es SomosNLP, una

comunidad nacida en 2021 que agrupa a estudiantes y profesionales de diversos países con la misión de lograr una representación justa del español y sus variantes en la era de la IA. Entre sus proyectos figura #Somos600M, cuyo objetivo es crear corpus y modelos que reflejen la rica diversidad dialectal del español hablado por más de 600 millones de personas en todo el mundo. Iniciativas así, al margen de instituciones formales, indican la pasión de las nuevas generaciones por impulsar el español en la IA global, con enfoques abiertos y colaborativos (p. ej.: «hackatones» internacionales organizados por SomosNLP con participantes de decenas de países, centrados en desarrollar recursos lingüísticos libres).

5. Ecosistema empresarial de la IA lingüística en español

El ecosistema de *startups* se ha convertido en la gran apuesta de innovación, dada la velocidad que requiere este campo. No obstante, aunque ha habido una notable disponibilidad de fondos europeos y se ha incidido en la necesidad de fomentar este tipo de emprendimientos, lo cierto es que pocas de estas *startups* han alcanzado un tamaño o una relevancia internacional equiparables al de referentes europeos como Mistral.ai, *startup* francesa considerada «unicornio» por haber alcanzado ya una valoración superior a 1.000 millones de dólares, apenas creada hace tres años y que ha destacado no solo por la calidad de sus LLMs y algoritmos, sino por la velocidad para levantar rondas de inversión especialmente cuantiosas y por recibir un respaldo político firme por parte del Gobierno francés.

El ecosistema español de las dos últimas décadas nos deja hitos mucho más pequeños, pero con nombres relevantes que han dado dinamismo al sector, como Sherpa.ai (con sede en Bilbao, centrada en asistentes virtuales por voz e IA predictiva), Inbenta (originada en Barcelona, desarrolladora de *chatbots* y búsquedas semánticas en múltiples lenguas, con fuerte enfoque en español y que logró expandirse a Silicon Valley) o Narrativa (*startup* de Madrid que aplica generación de lenguaje natural para periodismo automatizado produciendo noticias financieras de forma autónoma). También los correctores

textuales han contado con larga tradición, con muchos ejemplos, como Escribelo.ai, un asistente de escritura en español optimizado para SEO.

La traducción automática, con especial hincapié en las lenguas peninsulares, ha sido otra de las áreas donde han surgido proyectos empresariales interesantes. Por ejemplo, Pangeanic (con sede en Valencia) ha destacado en traducción automática y anonimización textual; esta compañía participó en proyectos europeos y estuvo presente en la presentación del Plan TL en 2015 como empresa referente del sector. Otras firmas establecidas serían Indra (y su filial Minsait), que ha desarrollado soluciones de procesamiento de lenguaje para administraciones públicas; Eleka/Elhuyar y Vicomtech, con base en el País Vasco, enfocadas en sistemas multilingües que integran el euskera; o Verbio, originaria de Barcelona, especialista en tecnologías de voz (reconocimiento y síntesis en español y otros idiomas) desde inicios de los 2000. En el terreno de los asistentes virtuales, Telefónica impulsó su asistente Aura con capacidades en español, lo que motivó colaboraciones con la RAE para asegurar un uso correcto del idioma. Asimismo, grandes consultoras tecnológicas (IBM, NTT Data, Atos, etc.) han establecido en España unidades de PLN para desarrollar *chatbots* y análisis semántico para sus clientes locales, adaptando tecnologías globales al español.

La nueva oleada de la IA generativa, unida a la creciente financiación y el foco en la creación de empresas de base tecnológica, ha provocado que comiencen a surgir empresas de IA aplicada a determinados ámbitos (como SoyLuzia para la educación o MAIte para el ámbito legal), pero aún es pronto para determinar cuál de estas empresas alcanzará el suficiente tamaño para sobrevivir en esta carrera tecnológica. Un ejemplo paradigmático de las dificultades del ecosistema es Clibrain, una prometedora empresa fundada en Madrid en 2022 que lanzó el primer modelo de lenguaje de gran tamaño enteramente entrenado en español, LINCE, en 2023. A pesar de su ambición y de la calidad técnica de sus modelos –que ofrecían resultados comparables a GPT-3 en español y destacaban por su cobertura dialectal–, la compañía cesó su actividad menos de un año después de su lanzamiento. Entre las buenas noticias, en cambio, tenemos el caso de Magnific.ai, *startup* murciana

pionera en el tratamiento de mejora de imágenes con IA que fue vendida a Freepik en un tiempo récord, o Multiverse, con sede en Euskadi, que ofrece un interesante sistema de compresión de LLMs.

Uno de los aspectos positivos en todo este entorno es que ha aumentado considerablemente el número de *venture capital* que está invirtiendo en *startups* españolas en sus diferentes fases, desde época temprana hasta grandes rondas donde compiten con empresas europeas e internacionales.

6. Estrategias públicas en España y marco europeo

En la última década, el Gobierno de España ha desarrollado una serie de planes estratégicos para impulsar las tecnologías del lenguaje y la IA aplicada al español, en consonancia con políticas europeas más amplias. Un hito inicial fue el lanzamiento, en octubre de 2015, del Plan de Impulso de las Tecnologías del Lenguaje (Plan TL). Este programa quinquenal (2016-2020) supuso una inversión pública de 90 millones de euros para dinamizar la industria del procesamiento del lenguaje natural (PLN) en español.

Tras la culminación del Plan TL en 2020, la estrategia nacional se integró en iniciativas más amplias de IA. En diciembre de 2020 el Gobierno aprobó la Estrategia Nacional de Inteligencia Artificial (ENIA), que reconocía la lengua española como un activo estratégico en el ecosistema digital. Poco después, en 2021, se lanzó el ambicioso Proyecto Estratégico para la Recuperación y Transformación Económica (PERTE) denominado «Nueva economía de la lengua», en el marco del plan Next Generation EU. Este PERTE, presentado oficialmente en 2022, movilizó 1.100 millones de euros de inversión pública (provenientes en parte de fondos europeos) con el objetivo de potenciar el español y las lenguas cooficiales como factor de crecimiento económico.

Paralelamente, España ha participado activamente en el marco regulatorio y de financiación europeo relacionado con la IA y el lenguaje, apoyando la propuesta del Reglamento Europeo de IA (AI Act). Una cuestión crítica en el contexto hispánico es cómo la revolución de la IA lingüística afecta a las lenguas

cooficiales de España –catalán, gallego, euskera– así como a las numerosas variantes dialectales del español tanto en la península como en América. Estas lenguas y variedades representan la riqueza plurilingüe y cultural de España e Hispanoamérica; garantizar su presencia en la era de la IA es fundamental para evitar brechas digitales y pérdida de patrimonio lingüístico.

En el caso de las lenguas cooficiales españolas, históricamente su desarrollo tecnológico ha sido desigual. El catalán, con más de 10 millones de hablantes, cuenta desde hace años con traductores automáticos, correctores ortográficos y asistentes de voz (por ejemplo, versiones catalanas de Siri o Alexa) impulsados en parte por su similitud con otras lenguas romances y por el apoyo institucional de la Generalitat de Catalunya. Iniciativas recientes, como el Projecte AINA (lanzado en 2022), buscan recopilar grandes corpus orales y escritos catalanes para entrenar modelos de IA específicos y asegurar la supervivencia digital del catalán en el futuro. De forma similar, el euskera (lengua de raíces no indoeuropeas, hablada por alrededor de 750.000 personas) ha requerido desarrollos propios: instituciones vascas y empresas como Elhuyar o Vicomtech han creado traductores neuronales e incluso asistentes (p. ej., Zurekin para voz en euskera), partiendo de corpus bilingües euskera-español cuidadosamente elaborados. El gallego, cercano al portugués y hablado por unos 2 millones de personas, también dispone de herramientas como correctores (la Xunta de Galicia liberó uno integrado en LibreOffice) y trabaja en sintetizadores de voz para gallego.

Una de las iniciativas más ambiciosas en el ámbito de la IA lingüística en español es el desarrollo del modelo ALIA-40B, liderado por el BSC-CNS. Este modelo cuenta con 40.000 millones de parámetros y ha sido entrenado sobre 2,5 TB de datos textuales en español y lenguas cooficiales. Su arquitectura de tipo Mixture-of-Experts permite activar únicamente las partes del modelo necesarias para cada tarea, optimizando así los recursos computacionales. El proyecto, en fase activa en 2025, ha liberado parte de sus versiones y pesos a través de Hugging Face¹⁵. Por su parte, la versión redu-

15. Véase: <https://huggingface.co/BSC-LT/ALIA-40b>.

cida, Salamandra-7B¹⁶, diseñada para entornos de baja latencia, ha mostrado resultados especialmente competitivos en lenguas minoritarias. Ambos modelos se encuentran en evolución constante y están siendo aplicados en dominios sensibles como el jurídico y el médico, mediante técnicas de aprendizaje reforzado.

En paralelo, el Centro Vasco de Tecnología de la Lengua (HiTZ), integrado por los grupos IXA y Aholab de la Universidad del País Vasco (UPV/EHU), ha destacado por sus avances en tecnologías lingüísticas aplicadas. Entre sus desarrollos recientes se encuentran sistemas de síntesis de voz adaptados a las variantes dialectales del euskera (con un 94 % de inteligibilidad media), herramientas de desambiguación semántica para textos jurídicos en español y modelos lingüísticos destinados a la detección temprana de enfermedades neurodegenerativas mediante el análisis de patrones de habla. Su colaboración con Osakidetza (el Servicio Vasco de Salud) ha permitido implementar sistemas de triaje automático con una precisión del 92 % en la clasificación de urgencias médicas, lo que representa un ejemplo pionero de transferencia tecnológica en sanidad pública.

Por otro lado, el proyecto TAN-IBE (Traducción Automática Neuronal para las Lenguas Ibéricas), coordinado por la Universitat Oberta de Catalunya (UOC), apuesta por una estrategia multilingüe inclusiva en el terreno de la traducción automática. Utilizando técnicas de *transfer learning* y redes neuronales de tipo Transformer-XL, el consorcio ha trabajado con corpus paralelos de más de 85 millones de pares oracionales para siete lenguas románicas peninsulares: español, catalán, gallego, portugués, asturiano, aragonés y aranés. La combinación de alineamiento semántico mediante grafos de conocimiento y entrenamiento cruzado ha permitido mejorar en un 37 % la puntuación BLEU para traducciones entre lenguas minoritarias, reduciendo significativamente la dependencia de grandes corpus bilingües (Oliver González, 2024). Este enfoque representa un avance relevante en la accesibilidad y preservación digital de las lenguas regionales de España.

16. Véase: <https://huggingface.co/BSC-LT/salamandra-7b-instruct>.

7. Conclusiones

El español cuenta con una ventaja competitiva indiscutible en el campo de la IA lingüística, como es la enorme disponibilidad de datos y corpus en español y en sus múltiples variantes geográficas y sociales. Sin embargo, esta riqueza se ve limitada por una dificultad estructural clave: el acceso a dichos recursos sigue siendo restringido en muchos casos, ya sea por barreras legales, técnicas o de interoperabilidad. Aunque existen diversas fuentes de financiación y múltiples iniciativas paralelas, la fragmentación del ecosistema limita el impacto de los esfuerzos actuales. Sería deseable una mayor colaboración interinstitucional y público-privada que permitiera lanzar proyectos de gran escala, capaces de competir tanto en financiación como en velocidad y calidad con los grandes desarrollos internacionales, especialmente con los promovidos en Estados Unidos y China. Tenemos, además, una oportunidad única derivada de la riqueza y variedad de nuestro idioma, una diversidad que solo las comunidades hispanohablantes pueden abordar de manera integral. La aplicación de la IA a los distintos ámbitos industriales y, en particular, la especialización lingüística y cultural a través de las variantes diatópicas, diastráticas y diafásicas sigue siendo una asignatura pendiente en el desarrollo actual de la inteligencia artificial. La generación de un ecosistema donde el impulso político fomente la colaboración empresarial, académica y de las instituciones culturales es la única forma para poder llevarlo a cabo.

Referencias bibliográficas

- AGERRI GASCÓN, R., BERMÚDEZ, J. y RIGAU CLARAMUNT, G. (2014). IXA pipes: Efficient and ready to use multilingual NLP tools. En N. Calzolari, K. Choukri, T. Declerck, H. Loftsson, B. Maegaard, J. Mariani, A. Moreno, J. Odijk, y S. Piperidis (Eds.), *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)* (pp. 3.823-3.828) UPV/EHU. European Language Resources Association (ELRA). http://www.lrec-conf.org/proceedings/lrec2014/pdf/775_Paper.pdf
- AUTOMATIC LANGUAGE PROCESSING ADVISORY COMMITTEE (ALPAC) (1966). *Languages and Machines. Computers in Translation and Linguistics*. National Academy of Sciences.
- BARCELONA SUPERCOMPUTING CENTER (21 de enero de 2025). ALIA, Europe's first public, open and multilingual AI infrastructure. <https://www.bsc.es/news/bsc-news/alia-europes-first-public-open-and-multilingual-ai-infrastructure>
- CAÑETE, J., CHAPERON, G., FUENTES, R., HO, J.-H., KANG, H. y PÉREZ, J. (2020). Spanish Pre-Trained BERT Model and Evaluation Data. <https://arxiv.org/abs/2308.02976>
- CARRERAS X., CHAO, I., PADRÓ, L. y PADRÓ, M. (2004). FreeLing: An open-source suite of language analyzers. En M. T. Lino, M. F. Xavier, F. Ferreira, R. Costa y R. Silva (Eds.), *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'04)*. European Language Resources Association (ELRA). <http://www.lrec-conf.org/proceedings/lrec2004/pdf/271.pdf>
- CREGO CLEMENTE, J. M. (2008). *Architecture and Modeling for N-gram-based Statistical Machine Translation* [Tesis doctoral, Universitat Politècnica de Catalunya]. https://nlp.lsi.upc.edu/papers/phd_jmcrego.pdf
- CHOMSKY, N. (1957). *Syntactic Structures*. De Gruyter Mouton.
- DEVLIN, J., CHANG, M-W., LEE, K. y TOUTANOVA, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers. arXiv:1810.04805
- FORCADA, M. L., GINESTÍ-ROSELL, M., NORDFALK, J., O'REGAN, J., ORTIZ ROJAS, S., PÉREZ ORTIZ, J. A., SÁNCHEZ MARTÍNEZ, F., RAMÍREZ SÁNCHEZ, G. y TYERS, F. M. (2011). Apertium: A free/open-source platform for rule-based machine translation. *Machine Translation*, 25 (2), 127-144.
- GALICIA HARO, S. N., GELBUKH, A. BOLSHAKOV, I. A. (2002). Compilation of a Mexican Spanish text corpora. WSEAS Press. <https://www.gelbukh.com/CV/Publications/2002/WSEAS-2002-Corpora.pdf>
- GAMALLO OTERO, P., GARCÍA GONZÁLEZ, M., PIÑEIRO POMAR, C. A., MARTÍNEZ CASTAÑO, R y PICHEL CAMPOS, J. C. (2018). LinguaKit: A Big Data-Based Multilingual tool for linguistic annotation and information extraction. IEEE Xplore. <https://ieeexplore.ieee.org/document/8554689>

- GUTIÉRREZ FANDIÑO, A., ARMENGOL ESTAPÉ, J., PÀMIES MASSIP, M., LLOP PALAO, J., SILVEIRA OCAMPO, J., PÍO CARRINO, C., ARMENTANO OLLER, C., RODRÍGUEZ PENAGOS, C., GONZÁLEZ AGIRRE, A. y VILLEGAS MONTSERRAT, M. (2022). MarIA: Spanish Language Models. *Procesamiento del Lenguaje Natural*, 68, 39-60. <http://journal.sepln.org/sepln/ojs/ojs/index.php/pln/article/view/6405/3820>
- HITZ CENTER UPV/EHU (2025). *Dominios médico y legal*. <https://www.hitz.eus/en/domains>
- HOCHREITER, S. y SCHMIDHUBER, J. (1997). Long Short-Term Memory. *Neural Computation*, 9 (8), 1.735-1.780.
- HUTCHINS, J. (2014). The history of machine translation in a nutshell. <https://aclanthology.org/www.mt-archive.info/10/Hutchins-2014.pdf>
- KOEHN, P. (2010). *Statistical Machine Translation*. Cambridge University Press.
- MIKOLOV, T., CHEN, K., CORRADO, G. y DEAN, J. (2013). Efficient Estimation of Word Representations in Vector Space. arXiv:1301.3781
- OLIVER GONZÁLEZ, A. (2024). TAN-IBE Participation in the Shared Task: Translation into Low-Resource Languages of Spain. En B. Haddow, T. Kocmi, P. Koehn y C. Monz (Eds.), *Proceedings of the Ninth Conference on Machine Translation* (pp. 871-877). <https://aclanthology.org/2024.wmt-1.83/>
- OPEN AI (14 de marzo de 2023). GPT-4. <https://openai.com/research/gpt-4>
- OPEN AI (30 de noviembre de 2022). Presentamos ChatGPT. <https://openai.com/es-ES/index/chatgpt/>
- PÉREZ, J. M., FURMAN, D. A., ALONSO ALEMANY, L. y LUQUE, F. M. (2022). RoBERTuito: a pre-trained language model for social media text in Spanish. En N. Calzolari, F. Béchet, P. Blache, K. Choukri, C. Cieri, T. Declerck, S. Goggi, H. Isahara, B. Maegaard, J. Mariani, H. Mazo, J. Odijk y S. Piperidis (Eds.), *Proceedings of the Thirteenth Language Resources and Evaluation Conference (LREC)* (pp. 7.235-7.243). <http://www.lrec-conf.org/proceedings/lrec2022/pdf/2022.lrec-1.785.pdf>
- RADFORD, A., NARASHIMAN, K., SALIMANS, T. y SUTSKEVER, I. (2018). Improving Language Understanding by Generative Pre-Training. OpenAI. https://cdn.openai.com/research-covers/language-unsupervised/language_understanding_paper.pdf
- STRAKA, M. y STRAKOVÁ, J. (2017). Tokenizing, POS Tagging, Lemmatizing and Parsing UD 2.0 with UDPipe. En J. Hajic (Ed.), *Proceedings of the CoNLL 2017 Shared Task. Multilingual Parsing from Raw Text to Universal Dependencies* (pp. 88-89). The Association for Computational Linguistics. <https://universaldependencies.org/conll17/proceedings/K17-3.pdf>
- TAULÉ DELOR, M., MARTÍ ANTONÍN, M. A. y RECASENS POTAU, M. (2008). AnCora: Multilevel annotated corpora for Catalan and Spanish. En N. Calzolari, K. Choukri, B. Maegaard, J. Mariani,

- J. Odijk, S. Piperidis y D. Tapias (Eds.), *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08)*. European Language Resources Association (ELRA). http://www.lrec-conf.org/proceedings/lrec2008/pdf/35_paper.pdf
- TURING, A. M. (1950). Computing Machinery and Intelligence. *Mind*, 59 (236), 433-460. <https://phil415.pbworks.com/f/TuringComputing.pdf>
- V. JAKUBÍČEK, M. KILGARRIFF, A., KOVÁŘ, V, RYCHLÝ, P. y SUCHOMEL, V. (23-26 de junio de 2013). *The TenTen Corpus Family*. 7th International Corpus Linguistics Conference (CL-2013), Lancaster, Reino Unido. https://www.sketchengine.eu/wp-content/uploads/The_TenTen_Corpus_2013.pdf
- VASWANI, A., SHAZEER, N., PARMAR, N., USZKOREIT, J., JONES, L., GOMEZ, A. N., KAISER, L. y POLOSUKHIN, I. (2017). Attention Is All You Need. NIPS 2017. <https://arxiv.org/pdf/1706.03762>
- WEIZENBAUM, J. (1966). ELIZA-A Computer Program for the Study of Natural Language Communication Between Man and Machine. *Communications of the ACM*, 9 (1), 36-45.
- WINOGRAD, T. (1972). *Understanding Natural Language*. Academic Press.
- WU, Y., SCHUSTER, M., CHEN, Z., LE, Q. V., NOROUZI, M., MACHEREY, W., KRIKUN, M., CAO, Y., GAO, Q., MACHEREY, K., KLINGNER, J., SHAH, A., JOHNSON, M., LIU, X., KAISER, L., GOUWS, S., KATO, Y., KUDO, T., KAZAWA, H., ... DEAN, J. (2016). Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation. arXiv:1609.08144

TERMINOLOGÍA EN ESPAÑOL Y MULTILINGÜISMO DIGITAL

Elea Giménez Toledo

CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS
(INSTITUTO DE LENGUA, LITERATURA Y ANTROPOLOGÍA)

1. Introducción

Las lenguas encuentran hoy en el entorno digital el territorio más fértil y estimulante en el que crecer, reivindicarse e impulsar multitud de desarrollos gracias a las tecnologías del lenguaje. Sin embargo, no todas las lenguas están en disposición de aprovechar esa enorme oportunidad que representa el espacio digital, pues no todas tienen las mismas condiciones para adentrarse en ese territorio. El concepto de igualdad lingüística digital ha sido definido como «el estado en el que las lenguas se encuentran en cuanto a apoyo tecnológico y contexto para continuar existiendo y prosperar como lenguas vivas en la era digital» (Gaspari *et al.*, 2021).

El multilingüismo digital viene siendo defendido desde el ámbito de la investigación y también desde algunas políticas públicas desde hace algún tiempo, a partir de la constatación de los sesgos lingüísticos derivados del monolingüismo imperante del inglés (Spence, 2021).

Para lograr un ecosistema digital multilingüe se requiere, por un lado, una visión no estrictamente tecnológica, sino integradora de las perspectivas sociales y culturales que representa el multilingüismo (Spence, 2021). Es necesario pensar en lo que significa la presencia de múltiples lenguas, a todos los niveles, en el entorno digital. El multilingüismo trae consigo la diversidad de perspectivas, visiones y análisis. Por ello, como apuntan Horvath *et al.* (2024), a través del trabajo colaborativo y del compromiso con la diversidad y la inclusión, podemos esforzarnos en «lograr infraestructuras que

desafíen el pensamiento convencional mediante su capacidad para apoyar y enriquecer la naturaleza multilingüe y polifacética de la academia contemporánea». Añaden que la existencia de infraestructuras multilingües no debería dejarse solo en manos de las grandes compañías –reflexión compartida en el informe *Disrupting Digital Monolingualism* (Spence, 2021)– sino que habrían de ser abordadas con las soluciones que aporten los humanistas digitales.

Al convencimiento de que se necesita un ecosistema digital multilingüe, le han de seguir la visión, la estrategia, los recursos y mucho trabajo desde cada una de las lenguas. El multilingüismo digital solo puede darse si apuesta por el desarrollo de cada lengua en el entorno digital, abarcando todas las cuestiones que el concepto engloba (Giménez Toledo, 2024): políticas públicas y acciones de diplomacia científica; circulación de contenidos en la red en los distintos idiomas, correctamente etiquetados (p. ej. Céspedes *et al.*, 2025); disponibilidad de corpus y datos en formatos adecuados para que sean legibles por las máquinas; modelos de lenguaje; recursos léxicos que, en el caso de los datos abiertos, deberían seguir los principios FAIR, para que se puedan encontrar, sean accesibles, interoperables y reutilizables; existencia de infraestructuras, redes y espacios de datos; y existencia de plataformas digitales, herramientas y algoritmos que trabajen en y con los distintos idiomas.

Este capítulo está dedicado específicamente a analizar el peso del español en los recursos lingüísticos recogidos en infraestructuras digitales europeas, así como a profundizar en la relevancia y utilidad de las terminologías en español, como un tipo de recurso lingüístico clave para el desarrollo de distintas tecnologías del lenguaje.

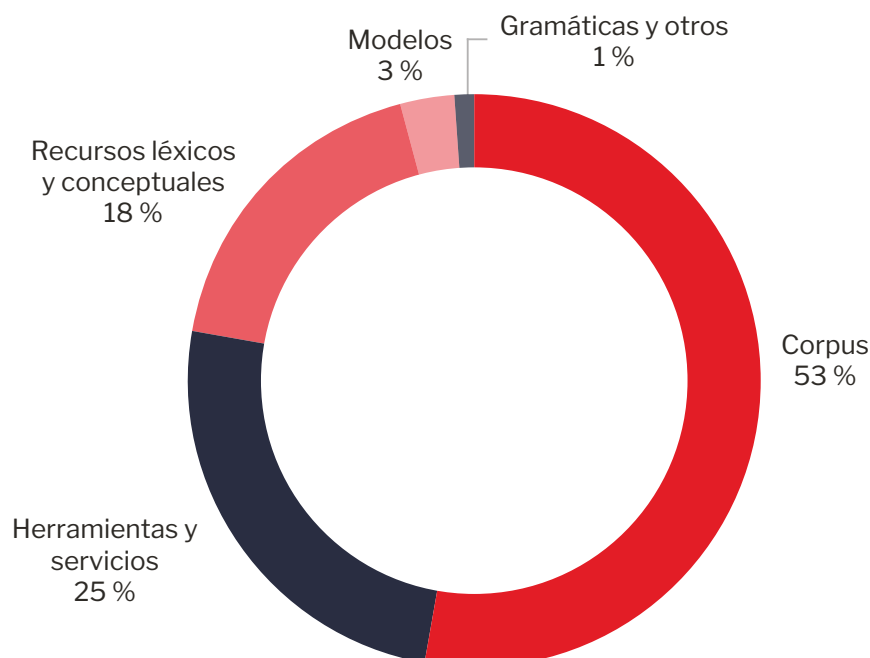
2. Los recursos del lenguaje en español en infraestructuras digitales de datos

Siendo tan críticos para el desarrollo de un idioma en el entorno digital, los recursos y las tecnologías del lenguaje han de ser conocidos, contabilizados y analizados como el valioso activo que son.

Todos ellos, desde los corpus hasta los datos lingüísticos, pasando por las herramientas de transcripción o traducción, son cruciales para el desarrollo de aplicaciones de inteligencia artificial como los asistentes virtuales, las búsquedas de voz o los sistemas conversacionales. Por esa razón, la Comisión Europea puso en marcha el Espacio Europeo de Datos Lingüísticos, con el objetivo de intercambiar recursos del lenguaje en el sector público y en el sector privado, haciendo posible así el desarrollo de herramientas y servicios. En ese espacio de datos, se comparten datos lingüísticos, se conectan con suministradores de datos y usuarios y, además, todo ello sucede en un entorno en el que se cumplen los derechos europeos. De la motivación para crear este espacio de datos, por naturaleza multilingüe, se deriva la necesidad de que sea alimentado con los datos de las distintas lenguas.

Dentro del Espacio Europeo de Datos Lingüísticos se ha desarrollado European Language Grid (ELG), en cuyo catálogo se recogen los recursos y las tecnologías del lenguaje que quieran ser visibilizados, compartidos y reutilizados. Entre estos recursos se encuentran corpus textuales y multimodales, modelos de lenguaje, sistemas de transcripción, sistemas de traducción, recursos léxicos, colecciones de datos anotados que sirven para entrenar modelos, plataformas para la síntesis o el reconocimiento de voz, sistemas de reconocimiento de entidades nombradas, analizadores morfológicos o lematizadores, entre otros muchos. El catálogo ELG es una fuente valiosa para analizar el estado las lenguas en lo que se refiere a recursos y tecnologías del lenguaje, pues además de los datos recopilados, gracias en parte a convocatorias financiadas, toma datos de distintas fuentes como CLARIN, ELRA o los repositorios Zenodo o Hugging Face.

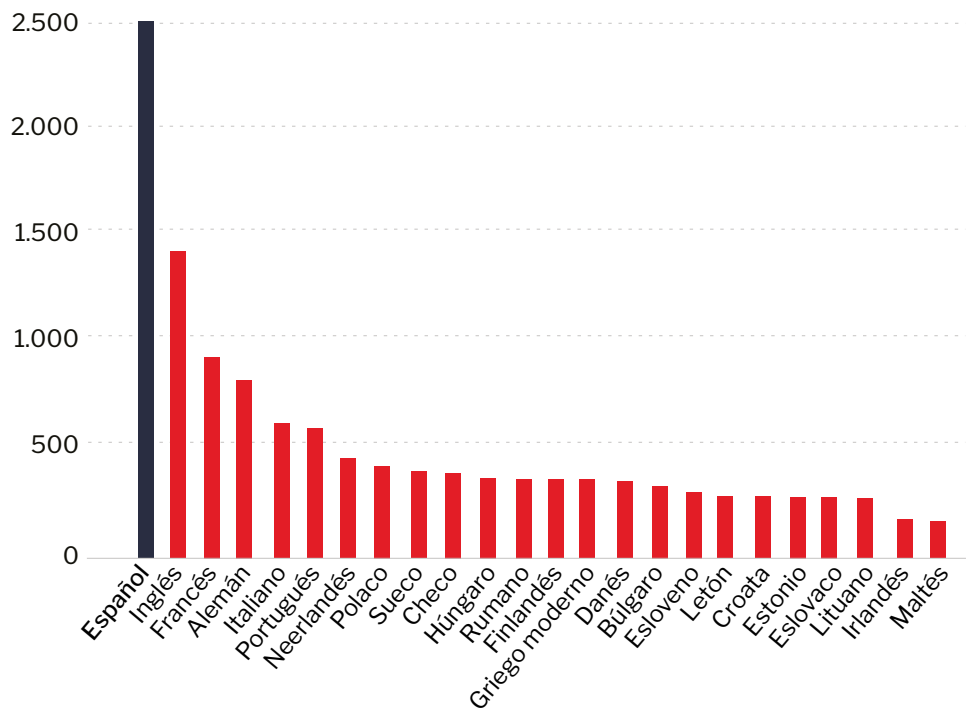
Gráfico 1 · Página siguiente >

GRÁFICO 1 TIPOLOGÍA DE RECURSOS Y TECNOLOGÍAS DEL LENGUAJE (MULTILINGÜES) EN ELG (MARZO 2025)

FUENTE: elaboración propia a partir de datos de ELG (2025).

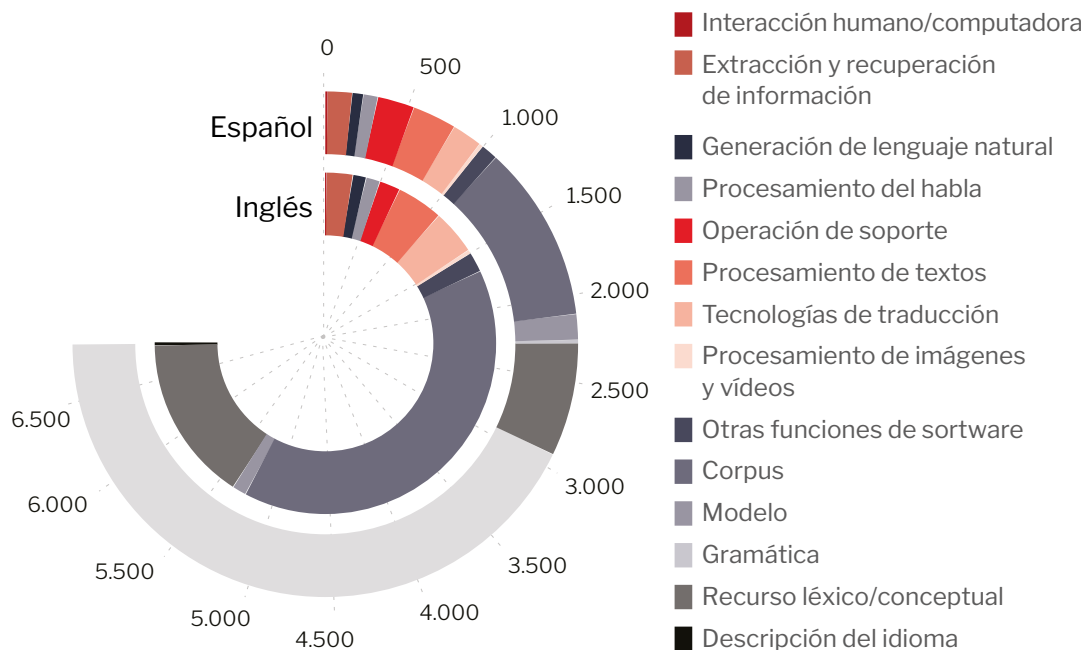
El español es la lengua con más recursos en este catálogo (2.567), por delante de inglés (1.466), francés (963), alemán (852) e italiano (644), lo que representa no solo un desarrollo nutrido de recursos y tecnologías en nuestra lengua sino también un adecuado posicionamiento en el catálogo de recursos de referencia en Europa (gráfico 2). Sin embargo, la cifra en bruto ha de ser matizada. Por un lado, la posición del español con respecto a las otras lenguas europeas varía en función del tipo de recurso lingüístico. Así, atendiendo al número de recursos léxicos o conceptuales que tiene cada lengua en ELG, es decir, al peso de cada lengua en el total de recursos léxicos (cuadro 1, última columna), el español ocupa la segunda posición, aportando un 23 % de los recursos, pero alejado del 50 % de los aportados por el inglés. El francés y el alemán les siguen, aportando un 18 % aproximadamente de recursos léxicos. Por otro lado, entre los 2.567 recursos del lenguaje en español, solo un 25,6 % son recursos léxicos, muy por debajo del peso relativo en otras lenguas. Este porcentaje señala un área que habría que desarrollar o fortalecer.

GRÁFICO 2 DISTRIBUCIÓN DE RECURSOS Y TECNOLOGÍAS DEL LENGUAJE POR IDIOMA EN ELG (MARZO 2025)



FUENTE: elaboración propia a partir de datos de ELG (2025).

GRÁFICO 3 COMPARATIVA ESPAÑOL/INGLÉS DE RECURSOS Y TECNOLOGÍAS DEL LENGUAJE POR TIPOLOGÍA EN ELG



FUENTE: elaboración propia a partir de datos de ELG (2025).

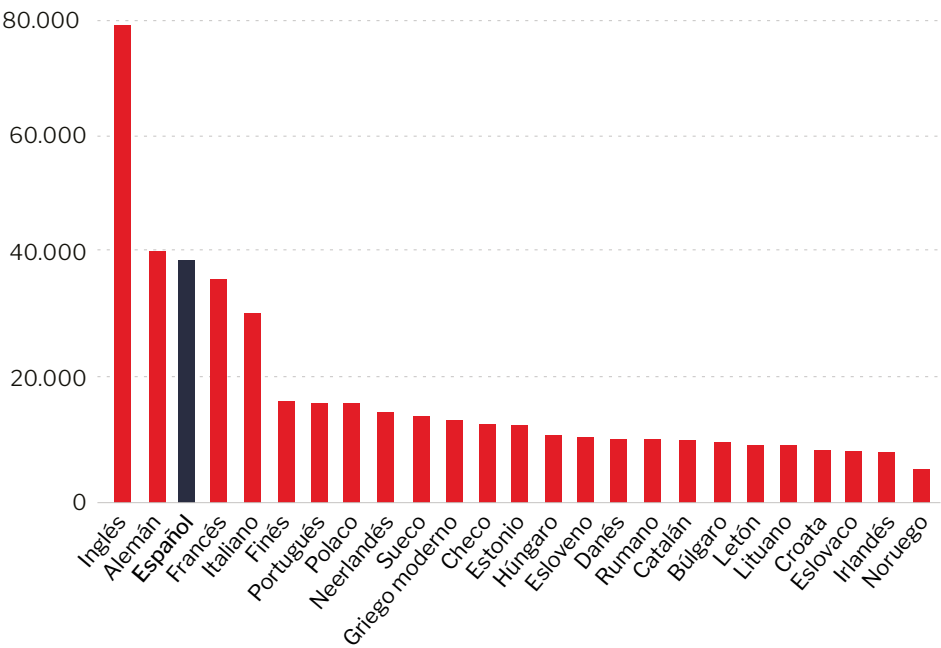
CUADRO 1 RECURSOS LÉXICOS EN EUROPEAN LANGUAGE GRID

Lenguas europeas	N.º de recursos en catálogo ELG	N.º de recursos léxicos en ELG	% de recursos léxicos (relativo)	% de recursos léxicos (absoluto)
Inglés	1.466	1.428	97,41	50,05
Estonio	292	243	83,22	8,52
Esloveno	317	236	74,45	8,27
Lituano	284	208	73,24	7,29
Alemán	852	518	60,80	18,16
Letón	294	166	56,46	5,82
Francés	963	541	56,18	18,96
Finlandés	377	202	53,58	7,08
Sueco	415	193	46,51	6,76
Griego moderno	377	164	43,50	5,75
Croata	294	105	35,71	3,68
Checo	409	141	34,47	4,94
Italiano	644	222	34,47	7,78
Irlandés	184	55	29,89	1,93
Búlgaro	347	101	29,11	3,54
Neerlandés	476	134	28,15	4,70
Polaco	440	123	27,95	4,31
Maltés	176	48	27,27	1,68
Danés	365	98	26,85	3,43
Español	2.567	657	25,59	23,03
Portugués	620	150	24,19	5,26
Rumano	378	85	22,49	2,98
Húngaro	383	81	21,15	2,84
Eslovaco	292	55	18,84	1,93

FUENTE: elaboración propia a partir de datos de ELG (2025).

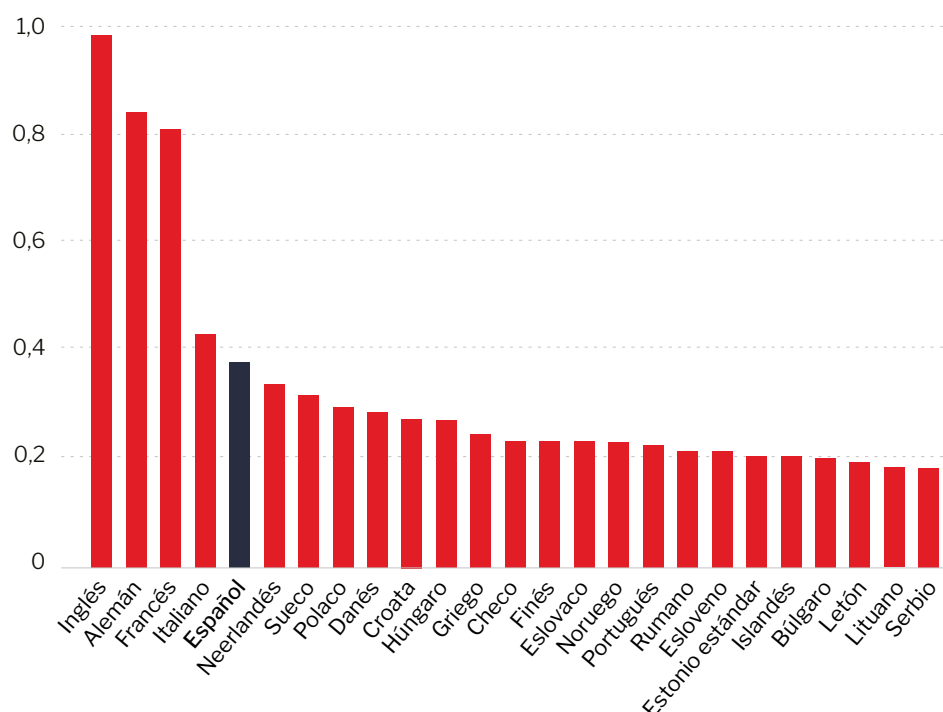
También se ve relativizado el peso del español, considerando los indicadores DLE (Digital Language Equality). Con ellos se mide el peso de cada lengua según factores tecnológicos y contextuales. En el caso de los factores tecnológicos, el indicador utilizado no solo considera el número de recursos lingüísticos que hay de cada idioma en ELG, sino que les asigna un peso específico en función de algunas variables como si son recursos anotados o no, si son especializados en una materia o no, si son de código abierto o propietario, etc. La fórmula utilizada para calcularlo viene a expresar lo legibles y reutilizables que son los recursos lingüísticos en cada idioma. El gráfico 4 muestra cómo el español es la tercera lengua, según este indicador, tras el inglés (a más de 40.000 puntos de diferencia) y el alemán.

GRÁFICO 4 INDICADORES DE IGUALDAD LINGÜÍSTICA DIGITAL SEGÚN FACTORES TECNOLÓGICOS



FUENTE: elaboración propia a partir de datos de ELG (2025).

Gráfico 5 · Página siguiente >

GRÁFICO 5 INDICADORES DE IGUALDAD LINGÜÍSTICA DIGITAL SEGÚN FACTORES CONTEXTUALES

FUENTE: elaboración propia a partir de datos de ELG (2025).

Entre las grandes infraestructuras de investigación europeas está CLARIN (Common Language Resources and Technology Infrastructure), enmarcada plenamente en la ciencia abierta y dedicada a reunir y ofrecer datos, herramientas y servicios para apoyar la investigación basada en recursos lingüísticos. Funciona con consorcios nacionales de 24 países europeos, entre ellos España, que conforman una red de instituciones productoras de recursos lingüísticos, tanto datos como herramientas. La sección de recursos reúne corpus, herramientas y recursos léxicos; entre estos últimos escasean los que cubren el español. Solo hay un inventario léxico monolingüe en español. En otros veintiún recursos léxicos multilingües (modelos de lenguaje, lexicones, diccionarios o glosarios) el español es una de las lenguas cubiertas. Se detecta una notable diferencia entre los recursos monolingües aportados desde otras lenguas como el estonio, el esloveno o el sueco en relación con el español o el francés que, siendo lenguas destacadas en ELG, tienen una presencia casi testimonial de estos recursos en CLARIN.

Esto se puede relacionar al menos con dos aspectos. Por un lado, con el desconocimiento de la infraestructura por parte de quienes producen corpus, recursos terminológicos u otro tipo de recursos y herramientas lingüísticas, lo que sugiere que debería producirse un acercamiento de las infraestructuras a los especialistas y viceversa, así como hacer una mayor difusión de lo que representan: espacios construidos entre todos para un fin común. Por otro lado, la falta de recursos en español podría estar relacionada más con una falta de legibilidad de estos que con su inexistencia.

El Virtual Language Observatory puesto en marcha por CLARIN permite rastrear recursos lingüísticos en cada idioma procedentes de distintos catálogos. La búsqueda de recursos léxicos en español no aporta más de una veintena de resultados. De nuevo, el escueto resultado puede estar relacionado con la falta de publicación de los recursos lingüísticos en español en los distintos catálogos o repositorios, pero también con cuestiones técnicas como la falta de nombres, palabras clave, enlaces rotos o metadatos incompletos (tamaño del corpus, anotación lingüística, etc.) (Fišer, Lenardič y Erjavec, 2018). Estas carencias han de ser tenidas en cuenta por productores de recursos lingüísticos.

3. Terminologías en español

Dentro de la diversidad de recursos y tecnologías del lenguaje que existen, los recursos léxicos y, más concretamente, las terminologías, han sido objeto de especial atención en los análisis cuantitativos mostrados en los anteriores apartados. La razón es que los términos, palabras especializadas de los distintos ámbitos científicos y profesionales, adecuadamente descritas por especialistas en cada dominio y cuidadas desde el punto de vista lingüístico, son la base de multitud de herramientas digitales basadas en el lenguaje. Tesoros, vocabularios, ontologías, lexicones, listas de términos o entidades nombradas (nombres propios) forman parte de lo que se denominan genéricamente terminologías. Las aplicaciones más relevantes de las terminologías en el entorno digital son:

- la recuperación eficiente y la descubribilidad de información en sistemas y buscadores;
- la indización o clasificación automática de grandes volúmenes de documentos, tal y como sucede en la mayor base de datos de producción científica médica (National Library of Medicine, 2021; Chen, Bullard y Giustini, 2023; Moore, Yaqub y Sampat, 2024);
- su conexión con otras terminologías en otros idiomas gracias a la tecnología de datos lingüísticos enlazados, que permite, entre otras cosas, buscar un término en un idioma y recuperar documentos sobre ese tema en otros idiomas distintos (p. ej. GoTriple);
- los sistemas de traducción automática, pues la propia existencia de términos cuidados y/o normalizados facilita el aprendizaje de estos sistemas y mejora la calidad de las traducciones;
- el entrenamiento de algoritmos para que las máquinas puedan entender el lenguaje científico en español, generar textos que contengan tecnicismos y «hablar» un español correcto;
- el apoyo a sistemas de lenguaje claro para permitir que los textos difíciles de comprender por su complejidad técnica (médicos, legislativos, normativos, etc.) puedan utilizar expresiones alternativas que hagan más comprensible su contenido a sus destinatarios.

Fuera del entorno digital, las terminologías son un soporte fundamental para la correcta enseñanza del español científico y profesional, para la comunicación médico/paciente a través de la medicina gráfica o de infografías, y en el ámbito de la traducción y la interpretación, donde se han creado recursos específicos para cubrir unas necesidades profesionales muy exigentes, que requieren precisión e inmediatez.

4. El español en los recursos terminológicos multilingües

El multilingüismo por el que apuestan muchas instituciones del ámbito internacional, desde la UNESCO a la ONU pasando por la Unión Europea, se traduce no solo en el impulso y la protección, a través de sus políticas, de las distintas lenguas, sino también en el establecimiento de lenguas de trabajo dentro de las propias instituciones y lenguas de publicación de los distintos documentos que generan. La defensa del multilingüismo implica respeto a la identidad de cada pueblo, pero también protección jurídica para la ciudadanía. Así, por ejemplo, las leyes y normativas que nos afectan a quienes vivimos en países de la Unión Europea se publican en cada una de las 24 lenguas oficiales de la UE para garantizar que llegan y son comprensibles para todos. Esto comporta un ingente volumen de traducciones caracterizadas por su precisión técnica; en la normativa que nos afecta no puede haber incorrecciones o inequidades entre lo que se dice en una lengua o en otra. Esta necesidad de precisión afecta, sobre todo, a la terminología. El adjetivo *green* en inglés, puede traducirse como «verde» o «ecológico». Se han manejado casi como sinónimos, pero tienen alcances diferentes, de tal manera que cada vez más, en las traducciones institucionales, se traduce «verde» por *green* y «ecológico» por *organic* (Arias Rodríguez, 2024).

Estas disquisiciones sobre los vocablos en busca de la mayor precisión y también de mayor consistencia requieren de un notable trabajo con las palabras, especialmente con los tecnicismos. Por ello se han venido construyendo bancos de datos terminológicos multilingües, críticos para la labor que desarrollan los traductores e intérpretes institucionales. El volumen de términos que contienen crece de manera constante; el trabajo en torno a la validación por expertos y a su normalización es incesante.

La ONU, que tiene 6 lenguas oficiales (árabe, chino, ruso, español, inglés y francés), desarrolla la base de datos UNTERM. Este recurso recoge la terminología necesaria –cientos de miles de palabras– para el trabajo que se desarrolla dentro de la organización y de otros organismos internacionales que dependen de la ONU, como la OMS o la UNESCO. Los términos

científicos forman parte de esta colección pues, a menudo, los temas que se abordan tienen conexiones directas con el ámbito de la ciencia.

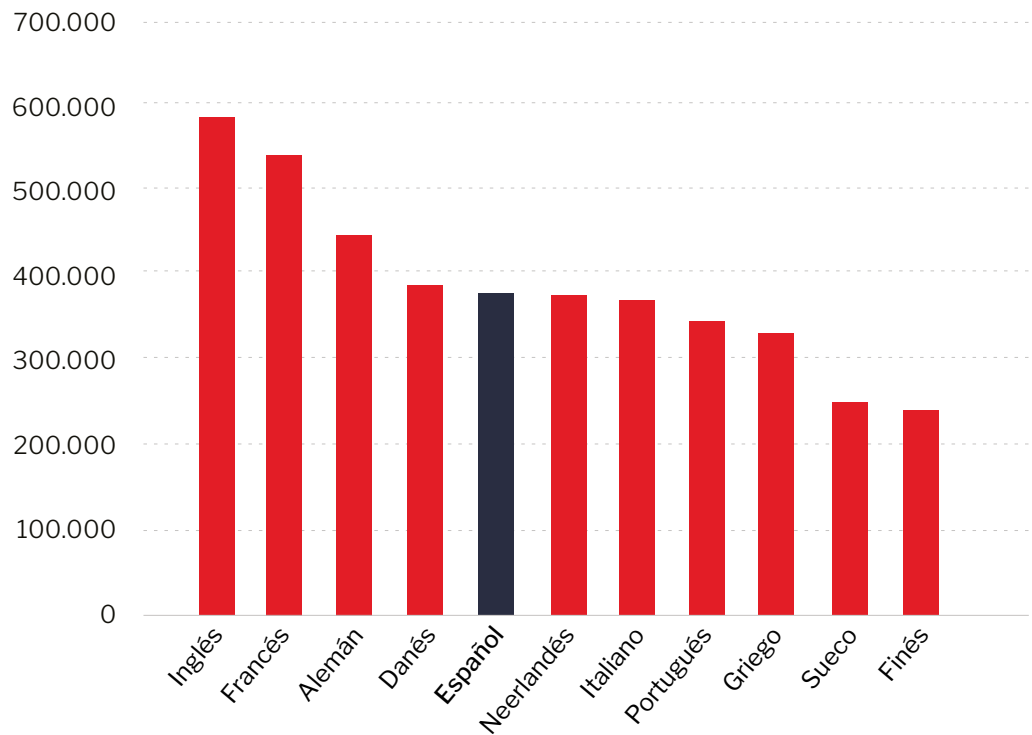
Por su parte, IATE European Union Terminology (IATE, s. f.) es la base de datos terminológica multilingüe más grande del mundo, creada en 2004 para «recopilar, divulgar y gestionar la terminología específica de la Unión Europea». Contiene casi 7 millones de términos y trabaja con las 24 lenguas oficiales de la Unión Europea, entre ellas, el español. Es un recurso en abierto, disponible para cualquier tipo de usuario y esencial para los profesionales de la lengua, especialmente para traductores e intérpretes.

A pesar de ser una base de datos multilingüe, el nivel de cobertura para cada idioma es distinto; esto se explica, por un lado, con la propia historia de la Unión Europea: a medida que se han ido incorporando países, ha crecido el número de lenguas oficiales y, por tanto, se han generado términos en esas lenguas en IATE. Por otro, tiene que ver con las necesidades de los traductores que van alimentando la base de datos en función de aquello que necesitan traducir.

Las estadísticas de IATE muestran algunos indicadores relevantes para observar el peso de los idiomas. La cobertura lingüística, por ejemplo. Se refiere al número de entradas o registros que tiene cada idioma (gráfico 6). El español es la quinta lengua con mayor cobertura. El inglés, el francés y el alemán son lenguas de trabajo en la Unión Europea, por lo que no es de extrañar que ocupen las primeras posiciones. El danés y el neerlandés tienen también muy buena cobertura por su antigüedad como lenguas oficiales de la Unión Europea.

Gráfico 6 · Página siguiente >

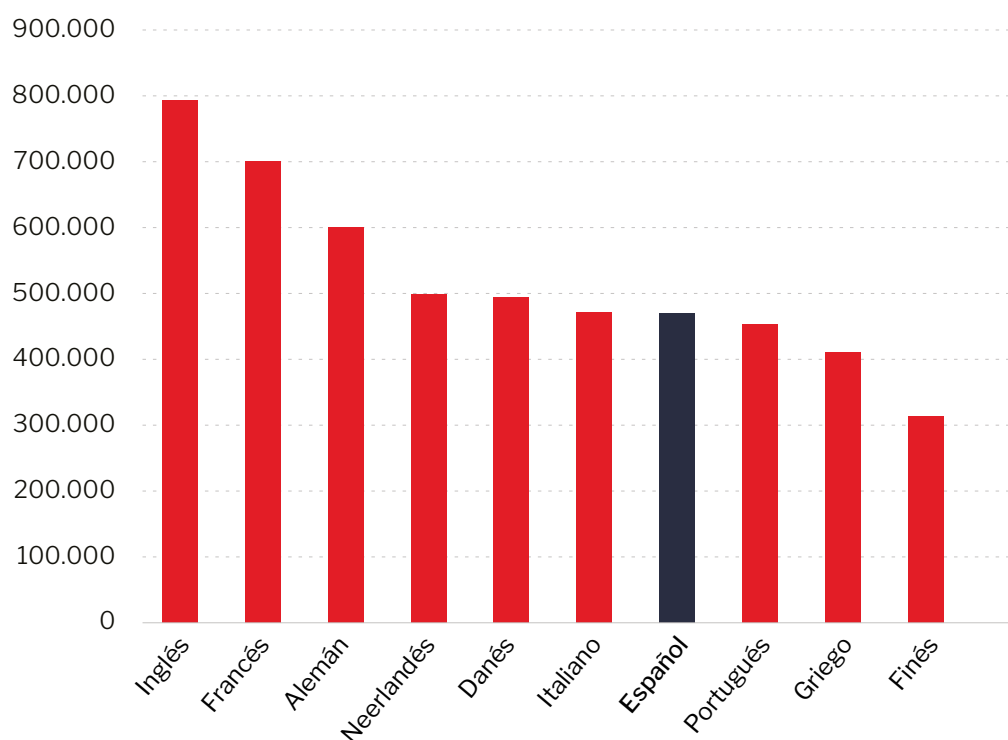
GRÁFICO 6 COBERTURA LINGÜÍSTICA EN IATE (MARZO 2025)



FUENTE: elaboración propia a partir de datos de IATE (s. f.).

Otro indicador para medir la presencia del español en esta base de datos terminológica es el número de términos. Como puede observarse en el gráfico 7, el español se sitúa en la séptima posición, con una cifra considerable de 470.000 términos. El inglés, con cerca de 800.000 términos, y el francés, con 700.000 términos, encabezan la lista.

Graph 7 · Next page >

GRÁFICO 7 NÚMERO DE TÉRMINOS POR LENGUA EN IATE (MARZO 2025)

FUENTE: elaboración propia a partir de datos de IATE (s. f.).

Es preciso contextualizar estas cifras. El volumen de datos terminológicos está relacionado con el número de años que llevan los distintos países en la UE, y también con las necesidades que van teniendo los traductores, que implican, a veces, mayor trabajo en unas lenguas que en otras. Conviene mirar los datos con otra perspectiva: el trabajo terminológico ha ido evolucionando y la calidad de la información que se ofrece para cada término es mayor cuanto más recientes son los registros. Las lenguas que se han incorporado más tarde a IATE, como el eslovaco, el croata, el maltés, el estonio o el rumano tienen registros para cada término más completos, con definiciones, notas o contextos. En este sentido, se puede observar (cuadro 2), cómo las lenguas de incorporación más reciente añaden más valor a la base de datos (registros más completos), mientras que las lenguas con mayor cobertura —el español entre ellas— tienen porcentajes de registros «ricos» menores.

CUADRO 2 NIVEL DE RIQUEZA DE LA INFORMACIÓN TERMINOLÓGICA EN IATE
(MARZO 2025)

Lengua	Valor añadido por lengua (%)
Eslovaco	80,15 %
Croata	78,20 %
Maltés	78,13 %
Estonio	75,68 %
Rumano	73,41 %
Húngaro	72,85 %
Checo	72,42 %
Búlgaro	71,97 %
Polaco	69,69 %
Lituano	68,14 %
Esloveno	67,04 %
Letón	65,63 %
Inglés	44,26 %
Alemán	38,16 %
Francés	36,68 %
Finés	33,57 %
Sueco	32,58 %
Italiano	32,47 %
Neerlandés	32,03 %
Griego	25,47 %
Portugués	25,40 %
Danés	24,57 %
Español	24,22 %
Gaélico	21,12 %

FUENTE: elaboración propia a partir de los datos de IATE (s. f.).

Los recientes desarrollos de IATE, que cuenta ya con una API REST, permiten una mayor accesibilidad y una mayor reutilización de los datos terminológicos, lo que representa un beneficio claro para profesionales de la lengua como traductores o intérpretes.

5. La terminología en español: TeresIA

La transformación digital de los recursos lingüísticos resulta esencial para que puedan interconectarse mejor las terminologías existentes, que se encuentran dispersas y que, conjuntamente, ofrecerían información muy valiosa. Precisamente este es uno de los objetivos de TeresIA, proyecto de investigación que persigue la construcción de un metabuscador para recuperar términos presentes en distintos recursos terminológicos en español. Uno de los primeros pasos que se dieron en el proyecto fue la elaboración de un censo de recursos terminológicos (TeresIA, 2024), que permitió identificar unos 400, bien en español (monolingües) bien multilingües incluyendo el español. Cubren distintas especialidades, son producidos por distintos tipos de instituciones y están en distintos formatos. El siguiente paso está siendo acordar con los productores su incorporación al metabuscador de TeresIA, analizar la estructura de cada recurso y proceder a la conversión de los datos para hacerlos legibles por el metabuscador. Muchos de ellos no son interoperables. La tarea ingente de conversión se justifica por la necesidad de hacer legible y recuperable el enorme acervo de datos terminológicos que se encuentran en distintas fuentes y que muchas veces no son fácilmente accesibles. La conversión supondrá, además, crear conjuntos de datos terminológicos que podrán estar en espacios de datos lingüísticos, como un capital del español para contribuir al desarrollo de las tecnologías del lenguaje en nuestra lengua y como un activo de la ciencia abierta, porque implica compartir datos, hacerlos accesibles y reutilizables.

Por otro lado, los neologismos terminológicos no dejan de aparecer. Las innovaciones o los descubrimientos científicos generan nuevas palabras que hay que añadir a las colecciones de términos en español y que es necesario validar con la ayuda de expertos, sancionar

lingüísticamente y convertir a formatos legibles por las máquinas. Para detectar los neologismos terminológicos que van surgiendo, se ha desarrollado en el marco de TeresIA un algoritmo de extracción de terminología (KeyCare), que permite detectarlos a partir de corpus de textos científicos. Estos nuevos datos terminológicos serán incorporados a los espacios de datos lingüísticos, generándose así la dinámica de producir información en español en abierto y disponible para que sea reutilizada por distintos tipos de profesionales de la lengua, por empresas o por desarrolladores de tecnologías del lenguaje. Las tecnologías lingüísticas necesitan de infraestructuras y de datos. En la UE se ha apostado por la creación de las primeras. Investigadores, técnicos, profesionales e instituciones hemos de apostar por la generación de grandes conjuntos de datos en español.

6. Conclusiones

Sabiendo que las tecnologías del lenguaje –omnipresentes en la vida de hoy– requieren de recursos lingüísticos en todos los idiomas y que, por otra parte, más el 80 % de todo lo que se hace en inteligencia artificial está basado en el lenguaje, el trabajo con y sobre los recursos lingüísticos en español se hace necesario si se quiere fortalecer el idioma en el entorno digital, contribuyendo así al multilingüismo.

Con los datos mostrados en este capítulo se puede concluir, generalmente hablando, que existe una buena cobertura de los datos lingüísticos en español tanto en espacios de datos como en bases de datos terminológicas. Sin embargo, esa cobertura es desigual tanto en la tipología de recursos como en las infraestructuras. Si en ELG el español tiene un peso destacado, en CLARIN la presencia de recursos en nuestro idioma es escasa. Por otro lado, en ELG se aprecia que algunos tipos de recursos lingüísticos en español predominan sobre otros. La presencia de recursos terminológicos podría mejorar, algo que tratará de lograrse con la conversión de datos que está afrontando TeresIA.

Por un lado, sería necesario difundir mejor entre productores de terminologías e instituciones académicas las infraestructuras y los espacios de datos

existentes, entre ellos también los de ámbito nacional como INESdata. Por otro lado, se requieren datos lingüísticos de todos los ámbitos, pero los terminológicos –es decir, los relativos a disciplinas científicas o profesiones– han de ser producidos y cuidados especialmente si se quiere apostar por el español en la ciencia y como lengua de trabajo en organismos internacionales. En paralelo, es necesario impulsar la digitalización de contenidos existentes en nuestro idioma y constituir corpus de textos académicos en español abiertos y con licencias Creative Commons.

Contar con terminologías de calidad en español que, además, estén convertidas a formatos de la web de datos, sigan los principios FAIR y puedan enlazarse (Linked Open Data) se convierte en una línea de trabajo crítica para que las tecnologías del lenguaje puedan desarrollarse también en español.

Agradecimientos

Estas páginas se enmarcan en TeresIA, proyecto de investigación financiado con fondos de la Unión Europea Next GenerationEU/PRTR a través del Ministerio para la Transformación Digital y de la Función Pública.

La autora agradece a la comunidad lingüística española de la Dirección General de Traducción de la Comisión Europea su colaboración para el análisis del español en IATE.

Referencias bibliográficas

- ARIAS RODRÍGUEZ, I. (2024). *Green: ¿verde o ecológico?* *Puntoycoma*, 184, 5-6. https://ec.europa.eu/translation/spanish/magazine/es_magazine_es.htm
- CÉSPÉDES, L., KOZŁOWSKI, D., PRADIER, C., SAINTE-MARIE, M. H., SHOKIDA, N. S., BENZ, P., POITRAS, C., BOUDREAU NINKOV, A., EBRAHIMY, S., AYENI, P., FILALI, S., LI, B. Y LARIVIÈRE, V. (2025). Evaluating the linguistic coverage of OpenAlex: An assessment of metadata accuracy and completeness. *Journal of the Association for Information Science and Technology*, 1-12. <https://doi.org/10.1002/asi.24979>
- CHEN, E., BULLARD, J. Y GIUSTINI, D. (2023). Automated indexing using NLM's Medical Text Indexer (MTI) compared to human indexing in Medline: a pilot study. *Journal of the Medical Library Association*, 111 (3), 684-694. <https://doi.org/10.5195/jmla.2023.1588>
- EUROPEAN LANGUAGE GRID (ELG) (2025). European Language Grid. Catalogue. ELG Consortium. <https://live.european-language-grid.eu/catalogue/dashboard>
- FIŠER, D., LENARDIČ, J. Y ERJAVEC, T. (2018). CLARIN's key resource families. En N. Calzolari, K. Choukri, C. Cieri, T. Declerck, S. Goggi, K. Hasida, H. Isahara, B. Maegaard, J. Mariani, H. Mazom, A. Moreno, J. Odijk, S. Piperidis y T. Tokunaga (Eds.), *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*. European Language Resources Association (ELRA). <https://aclanthology.org/L18-1210.pdf>
- GASPARI, F., WAY, A., DUNNE, J., REHM, G., PIPERIDIS, S. Y GIAGKOU, M. (2021). *European Language Equality. Digital Language Equality (preliminary definition)* (Documento de trabajo). ELE Consortium. https://european-language-equality.eu/wp-content/uploads/2021/05/ELE_Deliverable_D1_1.pdf
- GIMÉNEZ TOLEDO, E. (8 de marzo de 2024). El valor del multilingüismo en la comunicación de la ciencia. *The Conversation*. <https://theconversation.com/el-valor-del-multilinguismo-en-la-comunicacion-de-la-ciencia-225363>
- HORVÁTH, A., WAGNER, C., WRISLEY, D., BERNARDI, J., CHADHA, A., GARCIA, D., GRALLERT, T., GHOSH, S., ISHIDA, Y., KAYE, A., MEDEN, A., NAGASAKI, K., PALMER, D., SCHEITHAUER, H., ŚWIETLIK, M., THARSEN, J. Y WANG, Y. (2024). Multilinguality in Action: Towards Linguistic Diversity and Inclusion in Digital Humanities. *Magazén*, 5 (2), 179-196. <http://doi.org/10.30687/mag/2724-3923/2024/02/001>
- IATE EUROPEAN UNION TERMINOLOGY (S. F.). IATE EUROPEAN UNION TERMINOLOGY. CONSULTADO EN MARZO DE 2025. <https://iate.europa.eu/home>
- NATIONAL LIBRARY OF MEDICINE (1 de diciembre de 2021). MEDLINE 2022 Initiative: Transition to Automated Indexing. *NLM Technical Bulletin*, 443, e5. https://www.nlm.nih.gov/pubs/techbull/nd21/nd21_medline_2022.html

- MOORE, D. A. Q., YAQUB, O. Y SAMPAT, B. N. (2024). Manual versus machine: How accurately does the Medical Text Indexer (MTI) classify different document types into disease areas? *PLOS One*, 19 (3). <https://doi.org/10.1371/journal.pone.0297526>
- SPENCE, P. (2021). *Disrupting Digital Monolingualism: A report on multilingualism in digital theory and practice*. Language Acts and Worldmaking. <https://zenodo.org/records/5743283>
- TERESIA (2024). *Informe del censo de recursos terminológicos de Teresia*. <https://proyectoteresia.org/wp-content/uploads/2024/09/TeresIA-ICRTT-1.0.0-20240405.pdf>

LA ORIENTACIÓN DEL VOTO HISPANO EN LAS ELECCIONES PRESIDENCIALES DE ESTADOS UNIDOS EN 2024

Francisco Javier Pueyo Mena

OBSERVATORIO DEL INSTITUTO CERVANTES
EN LA UNIVERSIDAD DE HARVARD

1. Introducción

No cabe duda de que la celebración de elecciones periódicas es uno de los pilares fundamentales de la democracia, pues estas garantizan que los candidatos políticos elegidos representen la voluntad de los ciudadanos llamados a ejercer el voto. La transmisión de la información –o de la desinformación– política puede afectar, por otro lado, a la percepción de los ciudadanos sobre la realidad de las comunidades en las que recae el foco de los medios de comunicación en unas elecciones. Las últimas elecciones presidenciales estadounidenses, celebradas el 5 de noviembre del año 2024, no son una excepción. Tanto en la precampaña electoral como después de conocerse los resultados, la orientación del voto hispano ha sido uno de los principales centros de atención mediática. No faltan ejemplos de esta cobertura en los medios tanto de Estados Unidos como de Hispanoamérica y España:

Lo más sorprendente puede haber sido el aumento de 14 puntos porcentuales de Trump entre los votantes que se identificaron como hispanos, según una encuesta a pie de urna realizada por Edison Research. Alrededor del 46 % de los votantes hispanos votaron a Trump, frente al 32 % en las elecciones de 2020 (Lange, Erickson y Heath, 2024, Reuters)¹.

1. En el caso de las citas extraídas de textos publicados originalmente en inglés, se ofrece la traducción al español elaborada por el autor.

Trump obtuvo un asombroso 46 % del voto latino en estas elecciones. Es la cifra más alta jamás obtenida por un republicano, incluso más alta que la de George W. Bush en 2004 (Montanaro, 2024, NPR).

Trump consiguió un avance histórico entre los votantes de color, los inmigrantes y los latinos (Brammer, 2024, *Los Angeles Times*).

Un análisis de la encuesta de salida que aplicó *The Washington Post* revela que los hombres latinos impulsaron a Trump a la Casa Blanca, con 54 % que votó por él y 44 % por Harris (*El Universal*, 2024).

Los votantes latinos aumentaron en 13 puntos su apoyo al candidato y ahora presidente electo Donald Trump en estas elecciones, hasta alcanzar el 45 % de ese electorado, por encima del récord entre candidatos presidenciales republicanos, el de George W. Bush en 2004 (*El Tiempo*, 2024).

Trump obtuvo el apoyo del 45 % de los votantes latinos a nivel nacional, en comparación con el 53 % que logró Harris, según encuestas boca de urna de NBC News. Ese apoyo a Trump es mucho mayor que en 2020, cuando obtuvo el 32 % frente al 65 % de Joe Biden (Lugones, 2024, *Clarín*).

Trump ha alcanzado la victoria con un mayor margen que en la elección de 2020, al reducir la ventaja demócrata sobre el voto latino de 33 a 8 puntos. Eso ha sucedido a pesar de la dura agenda migratoria del magnate, quien este jueves reiteró que pondrá en marcha un plan de deportaciones masivas (Herrerías, 2024, *El País*).

El voto hispano: el seísmo electoral que amenaza el futuro del Partido Demócrata (Ansorena, 2024, ABC).

Como se desprende de la lectura de los ejemplos mediáticos citados, el mensaje que se ha transmitido en los medios de comunicación a escala global es el de un «sorprendente» cambio del voto hispano hacia el candidato republicano Donald Trump en estas elecciones, en detrimento de la candidata demócrata Kamala Harris, que rompe con la orientación tradicional del voto de la comunidad hispana en las elecciones previas desde que hay registros. Un factor clave ha determinado el estado de opinión público transmitido a los ciudadanos por los medios de comunicación: el uso general por todos los periódicos, las televisiones, las radios y los medios digitales a escala global de

un único sondeo a pie de urna, el National Election Pool (NEP)², como fuente de información de la orientación del voto hispano, tanto en el ámbito nacional como en los decisivos *swing states* («estados péndulo o bisagra»). Ningún medio tuvo en cuenta otros sondeos importantes, como el 2024 American Electorate Voter Poll (2024-AEVP) (African American Research Collaborative y BSP Research, 2024a, 2024b)³, cuyos datos para las poblaciones minoritarias difieren sustancialmente de los aportados por el citado estudio de Edison Research para los grandes medios. No se tuvieron en cuenta los posibles problemas metodológicos que presenta la encuesta del NEP a la hora de medir la tendencia del voto en comunidades minoritarias como los hispanos, los indígenas, los asiáticos o los afroamericanos, y los pocos que después llamaron la atención sobre este factor difícilmente pudieron influir en el estado de opinión que ya se había propagado de forma global:

Desafortunadamente, los sondeos a pie de urna no cuentan con un buen historial y pueden contribuir a crear narrativas erróneas durante la noche de las elecciones. En particular, los sondeos a pie de urna presentan su peor historial de acierto cuando analizan a los votantes negros o latinos y a los jóvenes (Rogers, 2024, *The New York Times*).

Las organizaciones latinas, como UnidosUS, Voto latino, La Brega y la Fuerza Fund y la Federación Hispana se han unido a las representantes de otras minorías, como Asian American Advancing Justice y AARC para denunciar que las encuestas que se hacen a pie de urna no representan bien a las minorías

2. El NEP es un consorcio de medios estadounidenses creado en 2003 para proporcionar información de sondeos a pie de urna para las elecciones estadounidenses. El sistema de muestreo produjo resultados sesgados en las elecciones presidenciales estadounidenses de 2004 y de 2016. A partir de 2018, el consorcio de empresas ABC News, CBS News, CNN y NBC News contrataron a Edison Research para realizar los sondeos a pie de urna. A partir de 2020, Reuters ha utilizado el NEP para comunicar los resultados de las elecciones presidenciales de Estados Unidos. Fox News y Associated Press formaban parte anteriormente del NEP, pero lo abandonaron en 2017 y usan desde entonces los datos de AP VoteCast.

3. La encuesta 2024 American Electorate Voter Poll es un sondeo a gran escala entre votantes afroamericanos, latinos, asiático-americanos, nativo-americanos y blancos de todo el país (en los 50 estados), así como en los estados clave que determinaron las elecciones presidenciales de 2024. La encuesta fue realizada por científicos y expertos en sondeos de BSP Research, African American Research Collaborative y la Universidad de Harvard. La encuesta se realizó entre el 18 de octubre de 2024 y el 4 de noviembre de 2024 y cuenta con una muestra de más de 9.000 votantes, de los cuales 3.750 eran hispanos.

y ofrecen una imagen desfigurada de la realidad de estas comunidades (Caro, 2024, *El País*).

¿El voto latino le dio la victoria a Trump? Un nuevo estudio asegura que no (Telemundo, 2024).

Conviene, por lo tanto, dar una visión más ponderada haciendo uso tanto de las estimaciones previas al día de las elecciones como del seguimiento del voto una vez emitido por los votantes, ya que sus metodologías, al contrario que la del NEP, son más cuidadosas a la hora de otorgar pesos demográficos y lingüísticos a las comunidades minoritarias:

Nuestra encuesta sigue ofreciendo una ventaja considerable sobre las encuestas a pie de urna nacionales, ya que cada grupo racial se pondera para coincidir con la demografía de su grupo específico dentro de cada estado. Las encuestas nacionales solo proporcionan pesos demográficos a escala estatal o nacional y no aplican ponderaciones a los grupos raciales. Esto ha generado un sesgo socioeconómico en las muestras de las comunidades afroamericanas y latinas (African American Research Collaborative y BSP Research, 2024b).

Un factor determinante sobre la metodología empleada por las dos principales encuestas a pie de urna es el idioma en el que se desarrollaron las entrevistas. Mientras 2024-AEVP ofrecía una amplia elección lingüística a los encuestados «en inglés, español, chino, coreano, tagalo o vietnamita, según la preferencia del encuestado» (African American Research Collaborative y BSP Research, 2024b), la encuesta del NEP se desarrolló exclusivamente en inglés. Como veremos más adelante, la lengua hablada en el hogar –sea el inglés o el español– es una variable significativa en la orientación del voto de las comunidades latinas en Estados Unidos.

Los datos manejados en este capítulo han sido compilados a través de un amplio y detallado informe que hemos desarrollado en el Observatorio de la lengua española y las culturas hispánicas en los Estados Unidos del Instituto Cervantes en la Universidad de Harvard (Vega Hübner y Pueyo Mena, 2025). Las conclusiones de dicho estudio indican que, si bien se observa un progresivo aumento del apoyo latino al Partido Republicano, lo cierto es que no ha ocurrido ningún vuelco sorprendente y mucho menos decisivo en la

orientación tradicional del voto en las comunidades latinas. Esta tendencia se alinea con la observada en comicios previos y sigue los mismos patrones que en las otras comunidades de votantes: cuando en el público general hay un cambio hacia una opción política determinada –sea hacia el lado republicano o el demócrata–, se observa también una tendencia similar en las comunidades latinas, particularmente en los sectores de la población hispana más integrados en la cultura mediática y lingüística del país, es decir, aquellos que hablan exclusivamente inglés en casa y que, por lo tanto, reciben la información política a través de los mismos medios y redes sociales que sus conciudadanos de la comunidad mayoritaria:

El aumento en el uso de fuentes de noticias en inglés se ha visto impulsado por un incremento en la proporción de hispanos que dicen obtener sus noticias exclusivamente en dicha lengua. Según la encuesta de 2021, así lo declara un tercio (32 %) de los adultos hispanos, frente al 22 % en 2006 (Lopez y Gonzalez-Barreda, 2013).

La generación milenial hispana utiliza las fuentes de noticias en inglés más que las generaciones de mayor edad, con un 91 % en 2016 que declara obtener al menos algunas de sus noticias en inglés, en comparación con el 68 % que dijo consumir al menos algunas de sus noticias en español (Flores y Lopez, 2018).

La población latina que habla principalmente español en casa recibe la información de fuentes más variadas, tanto de los medios estadounidenses en español y en inglés, como de los medios de sus países de origen, que lógicamente informan en español:

Los latinos nacidos en el extranjero, por el contrario, prefieren las fuentes de noticias en español: en 2016, el 89 % declaró que obtenía al menos algunas de sus noticias en español y el 70 % dijo que obtenía al menos algunas de ellas en inglés (Flores y Lopez, 2018).

2. Dimensión y distribución geográfica del electorado hispano

La población hispana supone la minoría más numerosa dentro de Estados Unidos. Según las estimaciones de la Oficina del Censo de los Estados Unidos, supera los 65 millones y representa cerca del 20 % de la población total

del país (Oficina del Censo de los Estados Unidos, 2023a, 2024). Por lo tanto, el voto de este sector poblacional es potencialmente determinante para el resultado de las elecciones presidenciales.

La cifra de hispanos con posibilidad de votar en las elecciones de 2024 superó los 36 millones (Krogstad *et al.*, 2024). De acuerdo con el Pew Research Center, unos 36,2 millones de hispanos tenían derecho a ejercer el voto en 2024, lo cual representa un 14,7 % del electorado total –una cifra histórica–. En las elecciones de 2020, fueron en torno a 32,3 millones⁴ (Krogstad *et al.*, 2024). Este incremento de prácticamente 4 millones de personas es sintomático del crecimiento de la población hispana, que en 2023 aumentó un 1,8 % con respecto al año anterior (Oficina del Censo de los Estados Unidos, 2024). Esto se debe principalmente al crecimiento natural de la población –nacen más personas hispanas de las que mueren–, aunque la inmigración sigue siendo un factor de peso.

Las estimaciones del Pew Research Center con respecto al número de hispanos con derecho al voto en las elecciones de 2020 y 2024 son coherentes con el aumento en el número de ciudadanos hispanos en edad de votar que recoge la Oficina del Censo –32,7 millones en 2021 frente a 34 millones en 2023– (Oficina del Censo de los Estados Unidos, 2021, 2023b).

Los estados en los que la cifra del electorado hispano supera el millón de personas son California (8,5 millones), Texas (6,5 millones), Florida (3,5 millones), Nueva York (2,2 millones), Arizona (1,3 millones) y Nueva Jersey (1,06 millones). En estos seis estados viven más de dos tercios del electorado hispano (Krogstad *et al.*, 2024; Oficina del Censo de los Estados Unidos, 2023b).

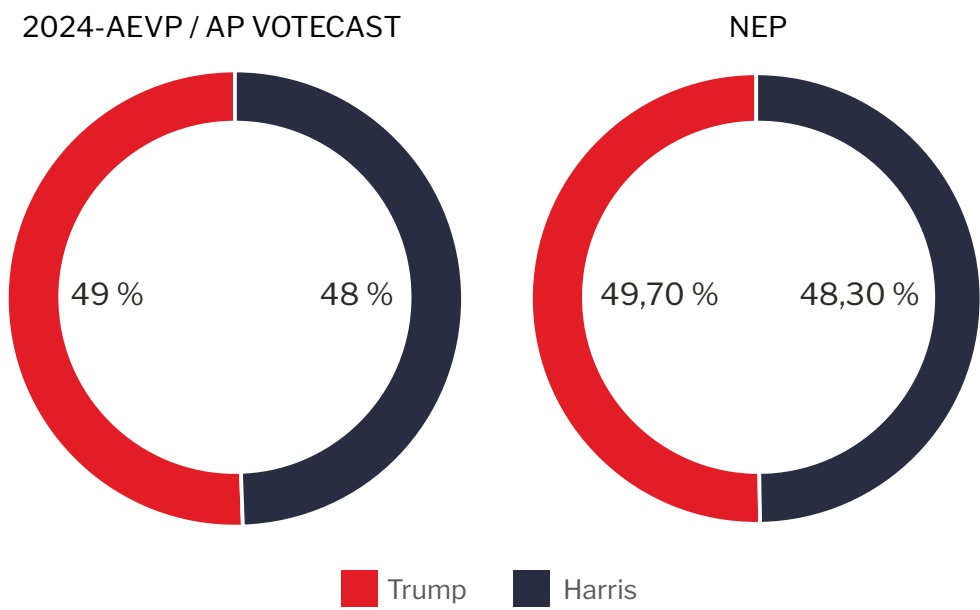
3. Resultados generales en la comunidad hispana

Los resultados globales de las elecciones presidenciales en Estados Unidos de 2024, sin distinguir entre grupos poblacionales, otorgaron un 49,8 % de apoyo al candidato republicano Donald Trump y un 48,3 % a la candidata

4. El electorado hispano aumenta aproximadamente en 1,4 millones de personas cada año y los hispanos suponen el 50 % del crecimiento del electorado estadounidense entre 2020 y 2024 (Krogstad *et al.*, 2024).

demócrata Kamala Harris (Federal Election Commission, 2025). Los resultados recogidos a pie de urna que hemos mencionado anticiparon estos resultados oficiales de forma muy acertada y dentro de sus márgenes de error estadísticos, como muestra el gráfico 1.

GRÁFICO 1 RESULTADOS GENERALES SEGÚN LOS SONDEOS A PIE DE URNA

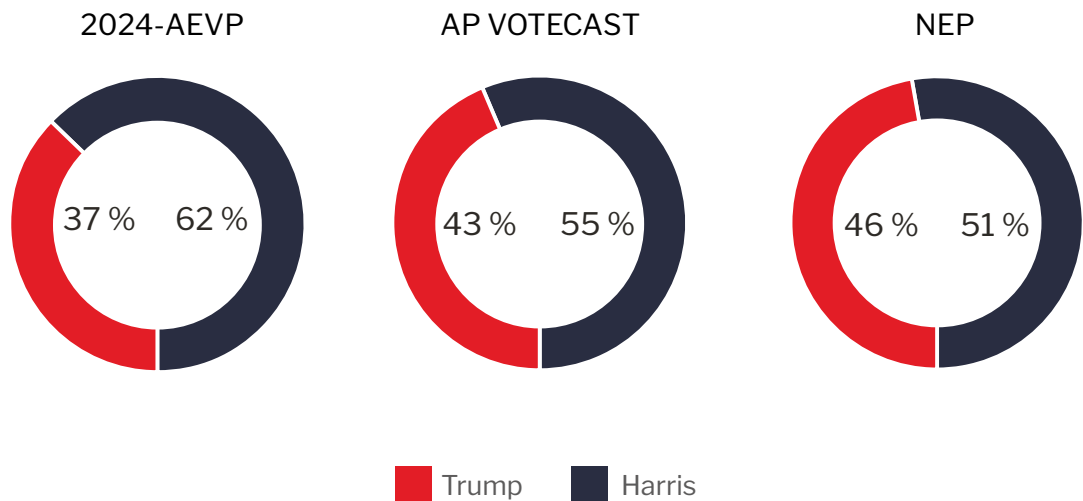


FUENTE: elaboración propia a partir de datos de 2024-AEVP (African American Research Collaborative y BSP Research, 2024a), AP VoteCast⁵ (Lohdi et al., 2024) y NEP (CNN, 2024).

Sin embargo, si nos centramos en los datos de las encuestas cuando desagregan sus resultados por comunidades de votantes, las divergencias a la hora de medir el voto en la comunidad hispana no pueden ser más llamativas. Como se observa en el gráfico 2, el respaldo a Kamala Harris oscila entre el 62 % reportado por 2024-AEVP y el 51 % que recoge NEP, es decir, una diferencia de 11 puntos porcentuales. El apoyo al candidato republicano Donald Trump varía entre el 37 % de 2024-AEVP y el 46 % de NEP, es decir, 9 puntos de diferencia.

5. AP VoteCast es una encuesta realizada por NORC en la Universidad de Chicago para Fox News, PBS NewsHour, The Wall Street Journal y Associated Press. Se llevó a cabo entre el 28 de octubre y el 5 de noviembre de 2024 con una muestra de cerca de 120.000 votantes, de los cuales el 10 % eran hispanos (Associated Press, 2024).

GRÁFICO 2 RESULTADOS GENERALES SEGÚN LOS SONDEOS A PIE DE URNA



FUENTE: elaboración propia a partir de datos de 2024-AEVP (African American Research Collaborative y BSP Research, 2024a), AP VoteCast (Lohdi et al., 2024) y NEP (CNN, 2024).

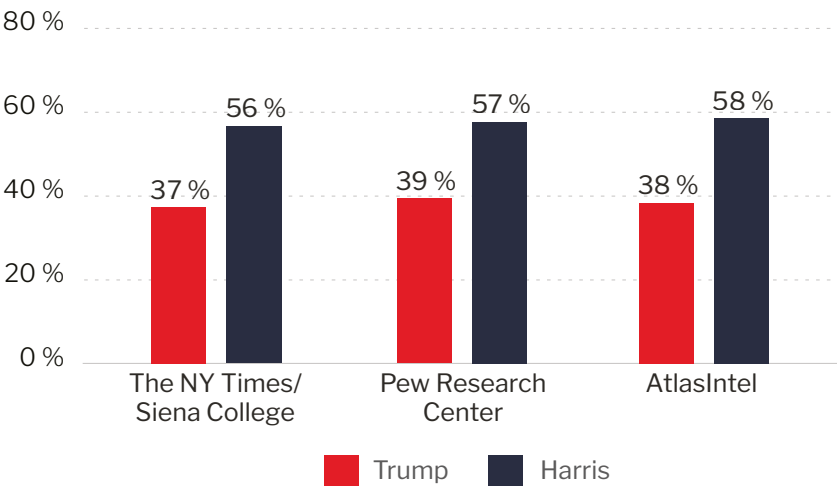
También debemos tener en cuenta las tres principales encuestas de intención de voto recogidas durante las semanas previas a las elecciones, realizadas por el Pew Research Center⁶, *The New York Times*/Siena College⁷ y AtlasIntel⁸, que ofrecen una intención de voto a medio camino entre los resultados finales reportados por 2024-AEVP y los datos ofrecidos por el NEP, tal y como se muestra en el gráfico 3. La media entre las tres encuestas de intención sería del 57 % de apoyo a la candidatura demócrata y del 39 % para la republicana.

6. Esta encuesta se realizó a estadounidenses de todo el país entre el 26 de agosto y el 2 de septiembre de 2024. La muestra incluía 691 hispanos registrados para votar (Lopez y Noe-Bustamante, 2024).

7. La encuesta de *The New York Times* y Siena College (*The New York Times*, 2024) se realizó entre el 29 de septiembre y el 6 de octubre 2024 con una muestra de 2.516 votantes de todo el país; 310 de ellos, hispanos.

8. Esta encuesta nacional se realizó entre el 1 y el 2 de noviembre de 2024 con una muestra de 2.463 votantes; el 12,1 % de ellos, hispanos (AtlasIntel, 2024).

GRÁFICO 3 APOYO DEL ELECTORADO HISPANO A LOS CANDIDATOS SEGÚN LAS ENCUESTAS PREVIAS DE INTENCIÓN DE VOTO

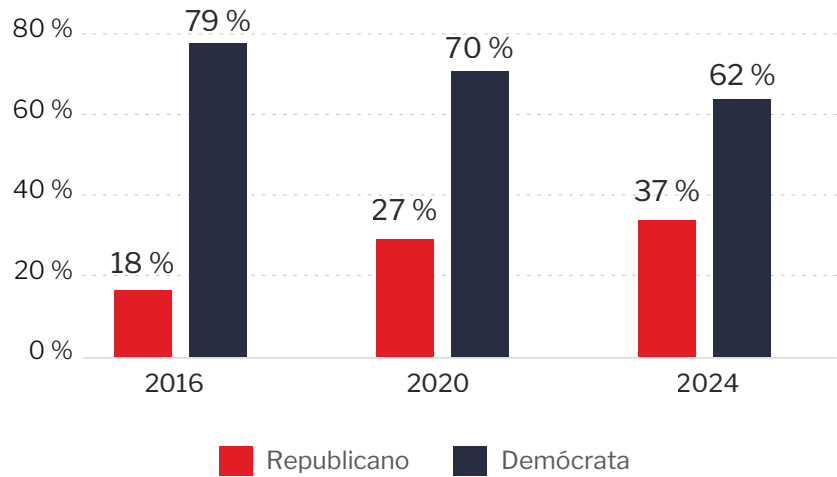


FUENTE: elaboración propia a partir de datos de *The New York Times* / Siena College (*The New York Times*, 2024), Pew Research Center (Lopez y Noe-Bustamante, 2024) y AtlasIntel (2024).

El dato ofrecido por 2024-AEVP muestra un progresivo aumento en el apoyo de la comunidad latina hacia Donald Trump respecto a los resultados de esta misma encuesta en las elecciones de 2016 y 2020. De acuerdo con esta fuente, el apoyo a Donald Trump del 37 % de los votantes supone un aumento de 10 puntos en los cuatro últimos años, aunque queda muy lejos del histórico 44 % obtenido por George W. Bush en 2004⁹.

9. Estudios posteriores del Pew Research Center indican que probablemente el apoyo del electorado hispano al Partido Republicano en las elecciones de 2004 no fuera tan alto como indicaban las estimaciones del 44 % basadas en encuestas a pie de urna, pues habría ocurrido una sobrerrepresentación de votantes cubanos prorrepúblicanos y otros desajustes demográficos (Suro, Fry y Passel, 2005).

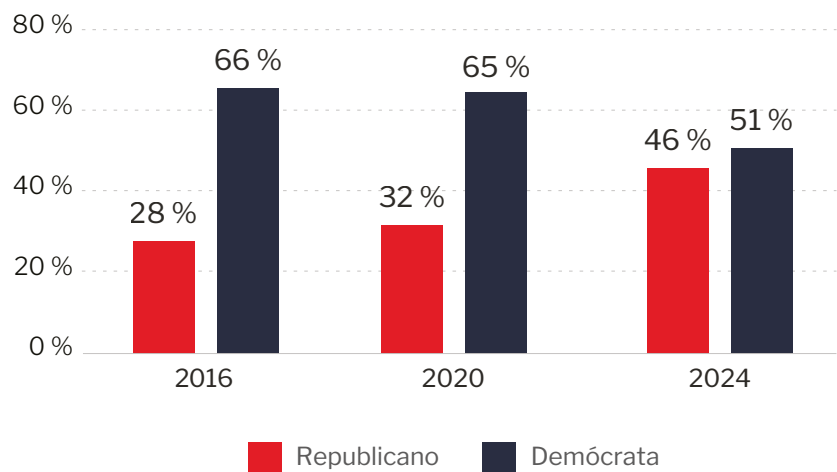
GRÁFICO 4 EVOLUCIÓN DEL VOTO HISPANO ENTRE 2016 Y 2024, SEGÚN LA AEVP



FUENTE: elaboración propia a partir de datos de AEVP.

Si nos fijamos en los datos ofrecidos por el NEP para las tres últimas elecciones presidenciales, el apoyo a Trump habría pasado del 28 % en 2016 al 32 % en 2020 y, finalmente, al 46 % en 2024 (CNN, 2016, 2020, 2024). Esto conllevaría un aumento de 14 puntos porcentuales entre 2020 y 2024.

GRÁFICO 5 EVOLUCIÓN DEL VOTO HISPANO ENTRE 2016 Y 2024, SEGÚN EL NEP



FUENTE: elaboración propia a partir de datos del NEP (CNN, 2016, 2020, 2024).

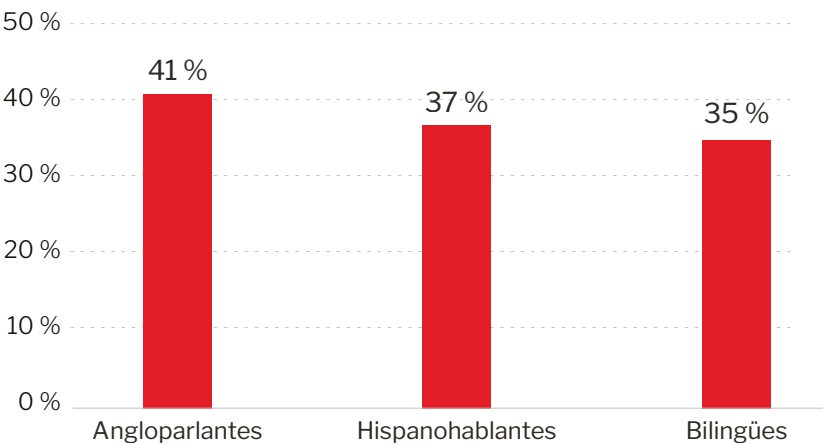
El resultado de 2024 supondría, por lo tanto, una cifra récord, pues habría superado por 2 puntos el histórico 44 % del voto hispano obtenido por George W. Bush en las elecciones presidenciales de 2004.

3.1 Orientación del voto hispano en función de la lengua hablada del hogar

Como indicábamos en la introducción, las dos encuestas a pie de urna difieren en su metodología a la hora de usar (2024-AEVP) o de no usar (NEP) el español durante la realización de las entrevistas, de acuerdo con las preferencias de los votantes. Igualmente, ambas difieren en incluir o no la pregunta de cuál es la lengua principal hablada en el hogar. La encuesta del 2024-AEVP es la única que considera este factor y clasifica la muestra en tres grupos: aquellos que solo hablan español en casa, los bilingües de español e inglés, y los que solamente utilizan el inglés en el hogar.

El gráfico 6 ilustra que el apoyo a Donald Trump es mayor entre los hispanos de habla inglesa (41 %) que entre los bilingües (35 %) y los hispanohablantes (37 %). Se observa una diferencia de 4 puntos porcentuales entre los angloparlantes y los hispanohablantes, y de 6 puntos entre los hogares de habla inglesa y los de los hablantes bilingües (African American Research Collaborative y BSP Research, 2024a).

GRÁFICO 6 APOYO DEL ELECTORADO HISPANO A DONALD TRUMP SEGÚN EL IDIOMA HABLADO EN EL HOGAR



FUENTE: elaboración propia a partir de datos de 2024-AEVP (African American Research Collaborative y BSP Research, 2024a).

Los tres grupos presentan tasas similares de votantes noveles (13-14 %), lo que sugiere que el idioma no influye significativamente en el debut electoral. Sin embargo, el historial electoral revela que el 67 % de los angloparlantes

votó por primera vez antes de 2018, frente al 60 % de los bilingües y al 64 % de los hispanohablantes, lo que sugiere una mayor consolidación del hábito electoral entre los primeros.

Un 56 % de los hispanohablantes y bilingües confían más en los demócratas para abordar sus prioridades, frente al 51 % de los angloparlantes. Estos últimos favorecen en mayor proporción a los republicanos (36 %) en comparación con los que hablan español en casa y con los bilingües (ambos con un 32 %), lo que sugiere una inclinación más conservadora entre los hogares de habla inglesa.

El coste de vida y la inflación, seguidos de la economía y el empleo, son claramente los temas más importantes para los tres grupos, aunque la salud preocupa más a los hispanohablantes (32 %) que a los bilingües (26 %) y a los angloparlantes (20 %).

La diferencia más notable se da en la prioridad otorgada a los derechos reproductivos: 29 % entre los angloparlantes, 19 % entre los bilingües y 15 % en los hispanohablantes.

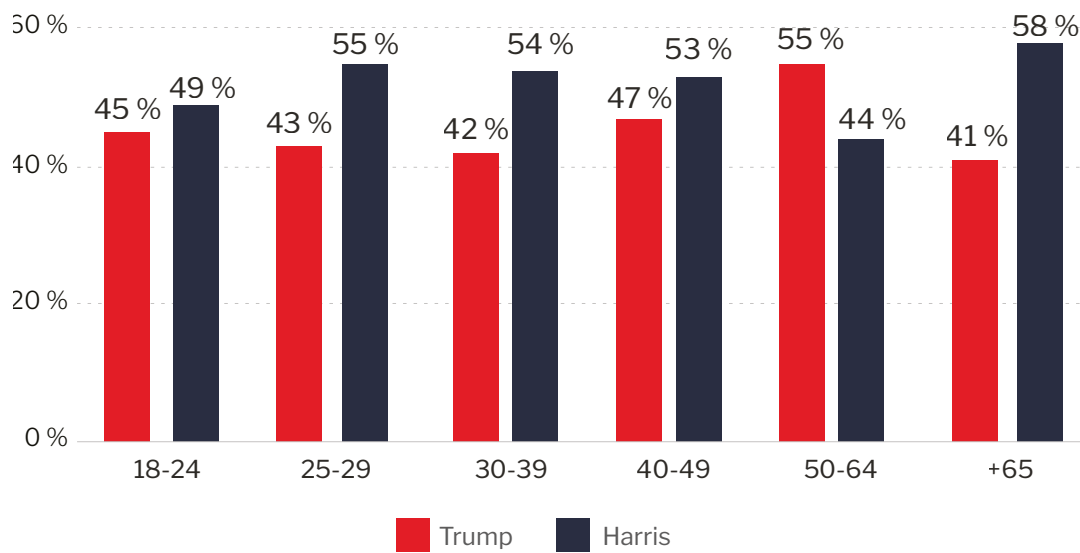
En políticas económicas y sociales, el apoyo a la construcción de viviendas y subsidios es alto en todos los grupos (del 85 % al 91 %), aunque los angloparlantes muestran una mayor oposición (15 %). En cuanto a la reforma migratoria, el respaldo a la regularización de indocumentados es mayor entre los hispanohablantes (92 %) y los bilingües (90 %) que entre los angloparlantes (85 %).

Estos datos evidencian que el idioma hablado en el hogar refleja diferencias sociales que influyen en la postura electoral: los angloparlantes son más críticos con ciertas medidas sociales y confían menos en los demócratas, mientras que los hispanohablantes y bilingües tienden a priorizar políticas económicas y sociales inclusivas. Estas observaciones subrayan la importancia de considerar las diferencias lingüísticas al formular estrategias políticas dirigidas al electorado hispano y también a la hora de analizar el comportamiento electoral de la heterogénea comunidad hispana.

3.2 Orientación del voto hispano en función de la edad

Los datos de la encuesta del NEP (CNN, 2024), que divide a los participantes en seis franjas etarias, muestra que el apoyo a Trump habría sido especialmente alto entre los hispanos del grupo de 50 a 64 años. Como muestra el gráfico 7, solo en dicha franja de edad fueron más los que apoyaron al candidato republicano que los que votaron a su contrincante (55 % a favor de Trump; 44 % a favor de Harris)

GRÁFICO 7 APOYO DEL ELECTORADO HISPANO A LOS DOS PRINCIPALES CANDIDATOS A LA PRESIDENCIA EN FUNCIÓN DE LA EDAD, SEGÚN LA ENCUESTA NEP



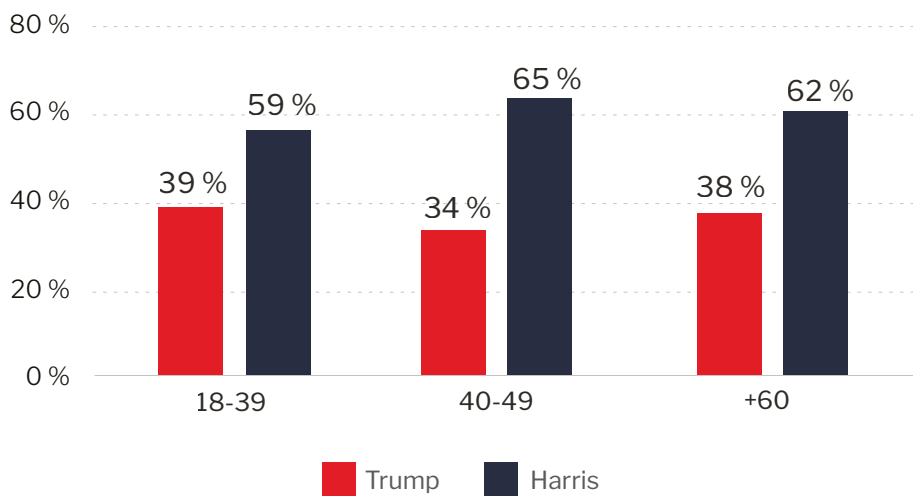
FUENTE: elaboración propia a partir de datos de NEP (CNN, 2024).

Además, esta encuesta refleja un mayor apoyo al candidato republicano en la franja etaria de 18 a 24 (45 % a favor de Trump) que en el rango etario 25-29 (43 %), si bien en ambos casos el apoyo a Harris es mayoritario. Esto podría indicar una mayor tendencia al voto republicano por parte de los votantes hispanos más jóvenes, especialmente entre los varones, como veremos en los siguientes apartados.

Sin embargo, cuando consultamos la encuesta 2024-AEVP (African American Research Collaborative y BSP Research, 2024a), que igualmente ofrece datos sobre el voto hispano en función de la edad, de nuevo nos encontramos con

diferencias sustanciales en los datos reportados. Como muestra el gráfico 8, según esta encuesta, no se registra mayoría republicana en ninguna de las tres franjas etarias estudiadas:

GRÁFICO 8 APOYO DEL ELECTORADO HISPANO A LOS DOS PRINCIPALES CANDIDATOS A LA PRESIDENCIA EN FUNCIÓN DE LA EDAD, SEGÚN LA ENCUESTA 2024-AEVP



FUENTE: elaboración propia a partir de datos de 2024-AEVP (African American Research Collaborative y BSP Research, 2024a).

El gráfico refleja que los hispanos con mayor tendencia al voto republicano pertenecerían al rango etario de 18 a 39 años, con el 39 % a favor de Trump y un 59 % a favor de Harris. Llama la atención que, en contraste con la encuesta del NEP, no se registran porcentajes de apoyo al candidato republicano superiores al 40 % en ninguno de los rangos de edad contemplados.

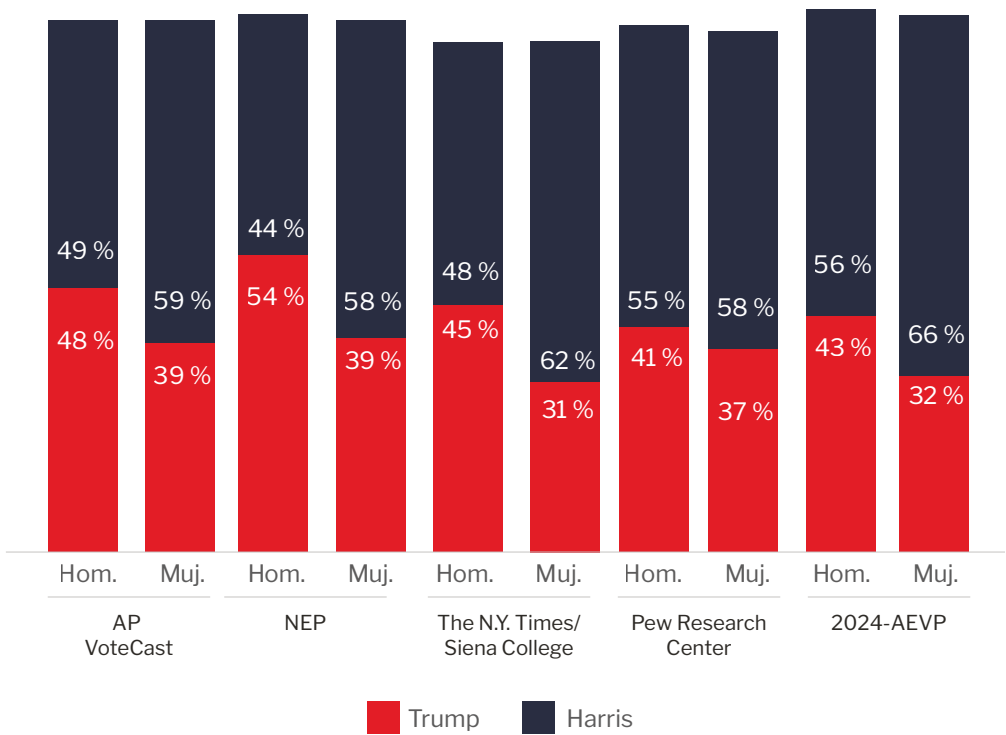
3.3 Orientación del voto hispano en función del sexo

Según los datos del Pew Research Center, Harris es la candidata preferida tanto por las hispanas (58 %) como por los hispanos (55 %). Donald Trump, en cambio, es el candidato favorito del 41 % de los hombres y del 37 % de las mujeres. Estos datos reflejan un mayor apoyo a Harris entre las mujeres que entre los hombres (Lopez y Noe-Bustamante, 2024).

Esta diferencia entre las tendencias electorales de hispanos e hispanas se acentúa al examinar los datos de las demás encuestas consultadas para este

informe. Todas ellas indican que el apoyo al candidato republicano es significativamente mayor por parte de los varones que por parte de las mujeres. Como se deduce del gráfico 9, el porcentaje de hombres hispanos que apoya a Trump supera siempre el 40 %, mientras que el porcentaje de mujeres a favor del candidato republicano no alcanza esa cifra en ningún caso, oscilando entre el 31 % y el 39 %. Asimismo, el porcentaje de hombres a favor de Harris solo superaría el 55 % en una de las estimaciones y por tan solo un punto porcentual, mientras que el voto de las hispanas ronda siempre el 60 % e incluso lo supera según dos de las encuestas.

GRÁFICO 9 APOYO DEL ELECTORADO HISPANO A LOS DOS PRINCIPALES CANDIDATOS A LA PRESIDENCIA EN FUNCIÓN DEL SEXO SEGÚN LAS DIFERENTES ENCUESTAS



FUENTE: elaboración propia a partir de los datos de AP VoteCast (Lohdi *et al.*, 2024), NEP (CNN, 2024), *The New York Times*/Siena College (*The New York Times*, 2024), Pew Research Center (Lopez y Noe-Bustamante, 2024) y 2024-AEVP (African American Research Collaborative y BSP Research, 2024a).

Según la encuesta del NEP, Trump incluso habría conseguido la mayoría del voto hispano masculino con un 54 %. Los datos del 2024-AEVP muestran de nuevo una realidad muy diferente, con una diferencia de 11 puntos porcentuales: Trump no habría alcanzado la mayoría del voto hispano masculino,

quedándose en un 43 % frente al 56 % de hombres hispanos que apoyó a Harris. Aunque se observa, sin duda, un mayor apoyo de las mujeres a Harris en ambas encuestas, los porcentajes de dicho apoyo difieren también en 8 puntos (del 58 % reportado por el NEP al 66 % que indica el 2024-AEVP).

En este punto es interesante mencionar que la tasa de participación de mujeres en las elecciones es tradicionalmente más alta que la de los hombres. Esto es extensible a todos los subgrupos raciales y étnicos, así como a la minoría hispana. En las elecciones de 2020, las mujeres supusieron el 53,7 % de los votantes hispanos, y los hombres, el 46,3 % (Fabina y Scherer, 2022).

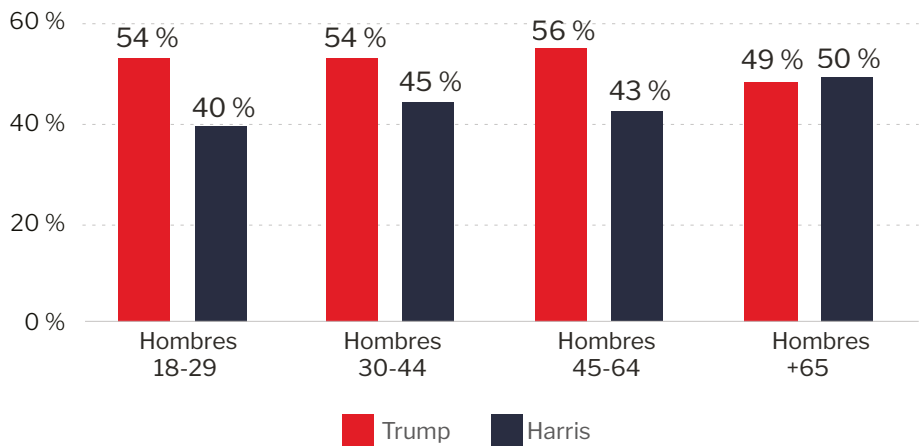
3.4 Orientación del voto hispano en función de la edad y el sexo

Tanto el NEP como la 2024-AEVP proporcionan datos sobre el voto por franjas de edad en función del sexo. Estos indican que el voto femenino es más favorable al Partido Demócrata que el masculino en todas las franjas etarias (CNN, 2024; African American Research Collaborative y BSP Research, 2024). Sin embargo, de nuevo encontramos discrepancias sustanciales entre ambas encuestas en los porcentajes que ofrecen.

Los datos del NEP indicarían un apoyo mayoritario a Trump entre los hombres en tres de las cuatro franjas etarias: 18-29, 30-44 y 45-65 (54 %, 54 % y 56 %, respectivamente), como recoge el gráfico 10. Solo los hombres de más de 65 años habrían votado en un 50 % a Harris y en un 49 % a Trump.

Gráfico 10 · Página siguiente >

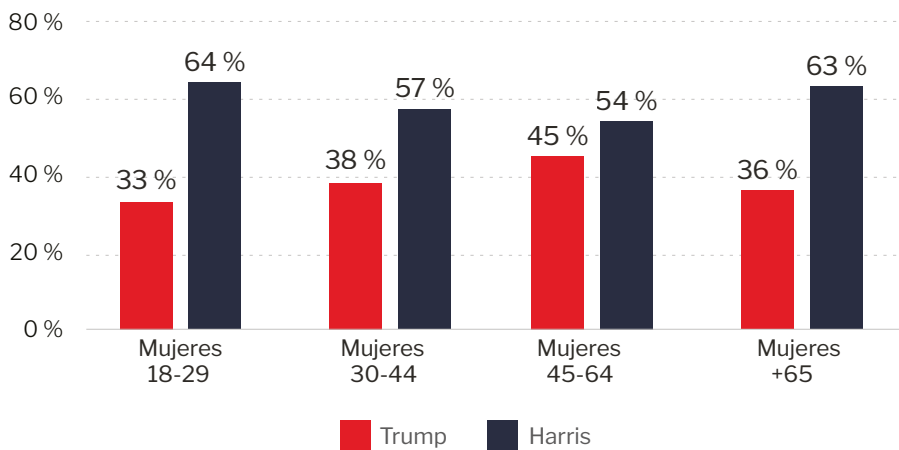
GRÁFICO 10 APOYO DE LOS HOMBRES HISPANOS A LOS DOS PRINCIPALES CANDIDATOS A LA PRESIDENCIA EN FUNCIÓN DE LA EDAD Y EL SEXO, SEGÚN LA ENCUESTA NEP



FUENTE: elaboración propia a partir de datos de NEP (CNN, 2024).

En cuanto a las mujeres, el gráfico 11, indica que Harris habría obtenido la mayoría del voto femenino en todas las categorías de edad; el apoyo al candidato republicano solo habría superado el 40 % en el grupo de hispanas entre 45 y 64 años (45 % a favor de Trump y 54 % a favor de Harris).

GRÁFICO 11 APOYO DE LAS MUJERES HISPANAS A LOS DOS PRINCIPALES CANDIDATOS A LA PRESIDENCIA EN FUNCIÓN DE LA EDAD Y EL SEXO, SEGÚN LA ENCUESTA NEP

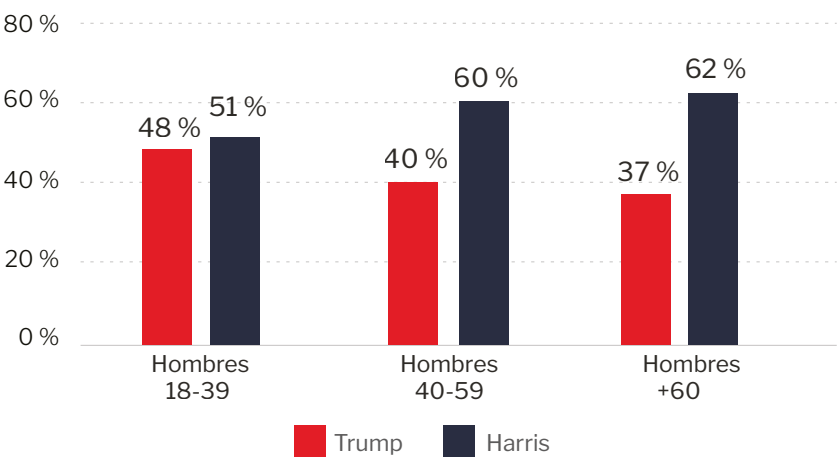


FUENTE: elaboración propia a partir de datos de NEP (CNN, 2024).

Por otro lado, la encuesta del 2024-AEVP, como muestra el gráfico 12, reporta que ningún grupo de edad entre los hombres habría apoyado a Trump

mayoritariamente. El mayor apoyo al candidato republicano estaría entre los varones de 18 a 39 años (el 48 % a favor de Trump y el 51 % para Harris).

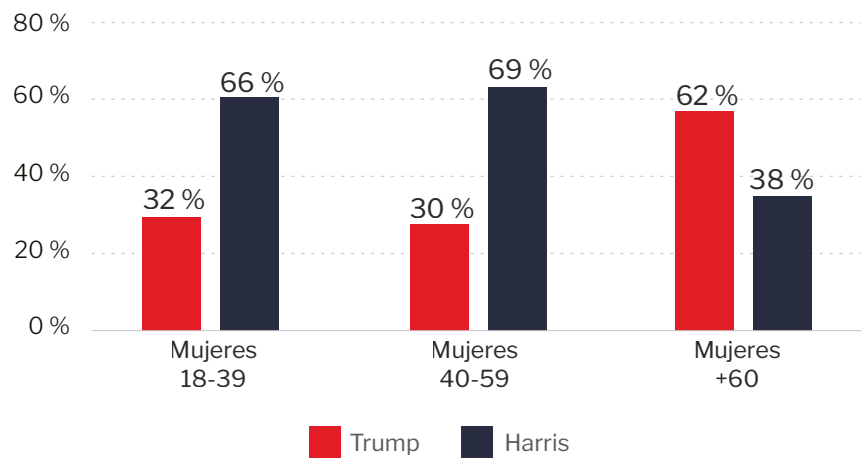
GRÁFICO 12 APOYO DE LOS HOMBRES HISPANOS A LOS DOS PRINCIPALES CANDIDATOS A LA PRESIDENCIA EN FUNCIÓN DE LA EDAD Y EL SEXO, SEGÚN LA ENCUESTA 2024-AEVP



FUENTE: elaboración propia a partir de datos de 2024-AEVP (African American Research Collaborative y BSP Research, 2024a).

El apoyo de las mujeres a Kamala Harris en todas las franjas etarias supera siempre al de los hombres, particularmente entre las hispanas más jóvenes, donde hay diferencias de entre 9 y 15 puntos porcentuales en su apoyo a Harris en comparación con los hombres, como indica el gráfico 13.

GRÁFICO 13 APOYO DE LAS MUJERES HISPANAS A LOS DOS PRINCIPALES CANDIDATOS A LA PRESIDENCIA EN FUNCIÓN DE LA EDAD Y EL SEXO, SEGÚN LA ENCUESTA 2024-AEVP



FUENTE: elaboración propia a partir de datos de 2024-AEVP (African American Research Collaborative y BSP Research, 2024a).

Aunque las cifras varíen significativamente de encuesta a encuesta, apuntan a lo que podría ser una tendencia: entre los jóvenes hispanos, que son tradicional y prevalentemente demócratas, parece estar aumentando el voto republicano, especialmente entre jóvenes varones.

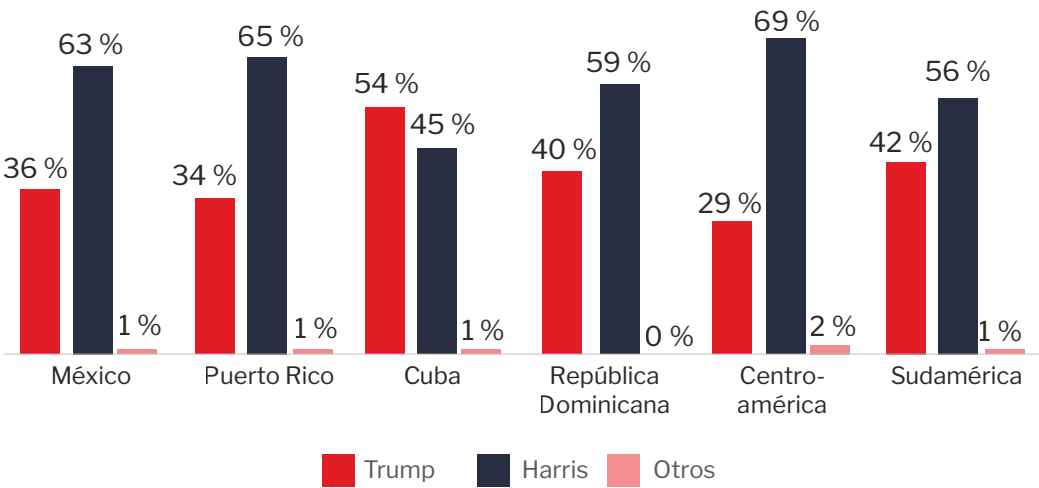
3.5 Orientación del voto en función de la ascendencia o procedencia de los votantes

En la introducción se ha descrito a la comunidad hispana como heterogénea, y lo es en función de diversas variables sociodemográficas. Una de las principales es seguramente la procedencia o ascendencia de los diferentes hispanos que conforman este colectivo. En un informe sobre el voto hispano en las presidenciales de 2020 para The Hispanic Council, Güenechea Rodríguez y Ureña Uceda (2020) advierten del riesgo de considerar a la población hispana como «una comunidad monolítica» y señalan que los diferentes trasfondos culturales que comparten las personas de un mismo país de procedencia o ascendencia se reflejan en las actitudes electorales.

La única encuesta nacional que proporciona datos sobre la conducta electoral de la muestra en función del territorio de procedencia o ascendencia

dentro del dominio hispánico es la 2024-AEVP, que divide a los encuestados en seis subgrupos: mexicanos, puertorriqueños, cubanos, dominicanos, procedentes de América Central y procedentes de América del Sur.

GRÁFICO 14 APOYO DEL ELECTORADO HISPANO A LOS PRINCIPALES CANDIDATOS A LA PRESIDENCIA EN FUNCIÓN DE LA PROCEDENCIA O ASCENDENCIA



FUENTE: elaboración propia a partir de datos de 2024-AEVP (African American Research Collaborative y BSP Research, 2024a).

El dato más llamativo que se puede extraer de la gráfica se refiere al voto entre los hispanos de origen cubano. Tradicionalmente, la mayoría de los cubanos vota al Partido Republicano. Esta comunidad es más decisiva en el estado de Florida, donde los cubanos son el grupo más numeroso dentro de la comunidad hispana (Oficina del Censo de los Estados Unidos, 2023a). En el siguiente apartado sobre los estados clave, podrá observarse que en 2024 la fuerza de esta comunidad en Florida inclinó el resultado total del electorado hispano hacia el candidato Trump.

La encuesta 2024-AEVP también distingue entre nacidos en Estados Unidos y nacidos en Puerto Rico u otros países. En este punto, los datos señalan un apoyo mayor al Partido Republicano por parte de los votantes nacidos en Estados Unidos que por parte de los nacidos en el estado libre asociado o en otros países: el 38 % de los nacidos en Estados Unidos apoya a Trump; el 60 %, a Harris. Para contrastar, solo el 34 %

de los hispanos no nacidos en Estados Unidos vota a Trump, mientras que el 65 % apoyan a Harris.

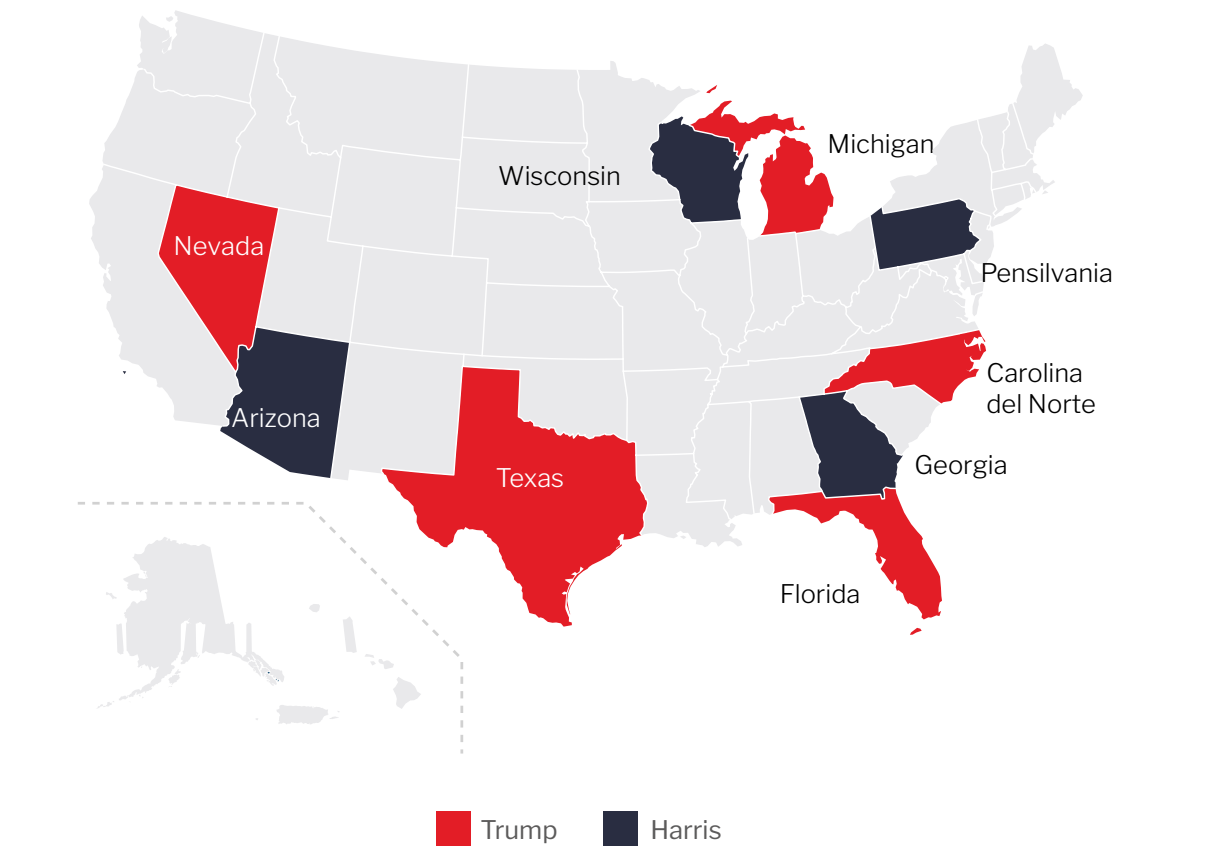
4. Resultados en los siete estados clave y en los de mayor población hispana

Solo dos encuestas ofrecen datos del voto latino en los siete estados bisagra de 2024: la elaborada por el NEP y la realizada por la 2024-AEVP. Igualmente, ambas examinan el comportamiento electoral hispano en algunos estados donde el porcentaje de la población hispana es decisivo para el resultado electoral. Según el censo de población, en California residen más de 15 millones de hispanos (el 25 % de toda la población latina del país), en el estado de Texas viven 12 millones de hispanos (que suponen el 19 % de la población latina total) y Florida cuenta con más de 6 millones de residentes latinos, es decir, casi el 10 % de todos los hispanos de Estados Unidos. (Oficina del Censo de los Estados Unidos, 2023c). Entre los tres estados reúnen, por lo tanto, más del 50 % de toda la población latina de Estados Unidos, por lo que para entender la orientación del voto hispano en 2024 se hace necesario estudiar su comportamiento en estos comicios.

Los resultados entre ambas encuestas no pueden ser más divergentes. La realizada por el NEP otorga la mayoría de los estados estudiados al republicano Donald Trump: cuatro de los siete estados clave (Michigan, Carolina del Norte, Nevada y Wisconsin) y dos de los de mayor población hispana (Florida y Texas), motivo por el que se ha difundido la idea de que el voto hispano fue decisivo para su victoria. El mapa del gráfico 15, elaborado con los datos del cuadro 1, muestra de forma visual la imagen que han transmitido los medios a nivel global.

Gráfico 15 · Página siguiente >

GRÁFICO 15 CANDIDATO CON MÁS VOTOS HISPANOS EN LOS 7 ESTADOS BISAGRA Y EN 2 DE ALTA POBLACIÓN HISPANA, SEGÚN NEP



PORCENTAJE DE APOYO

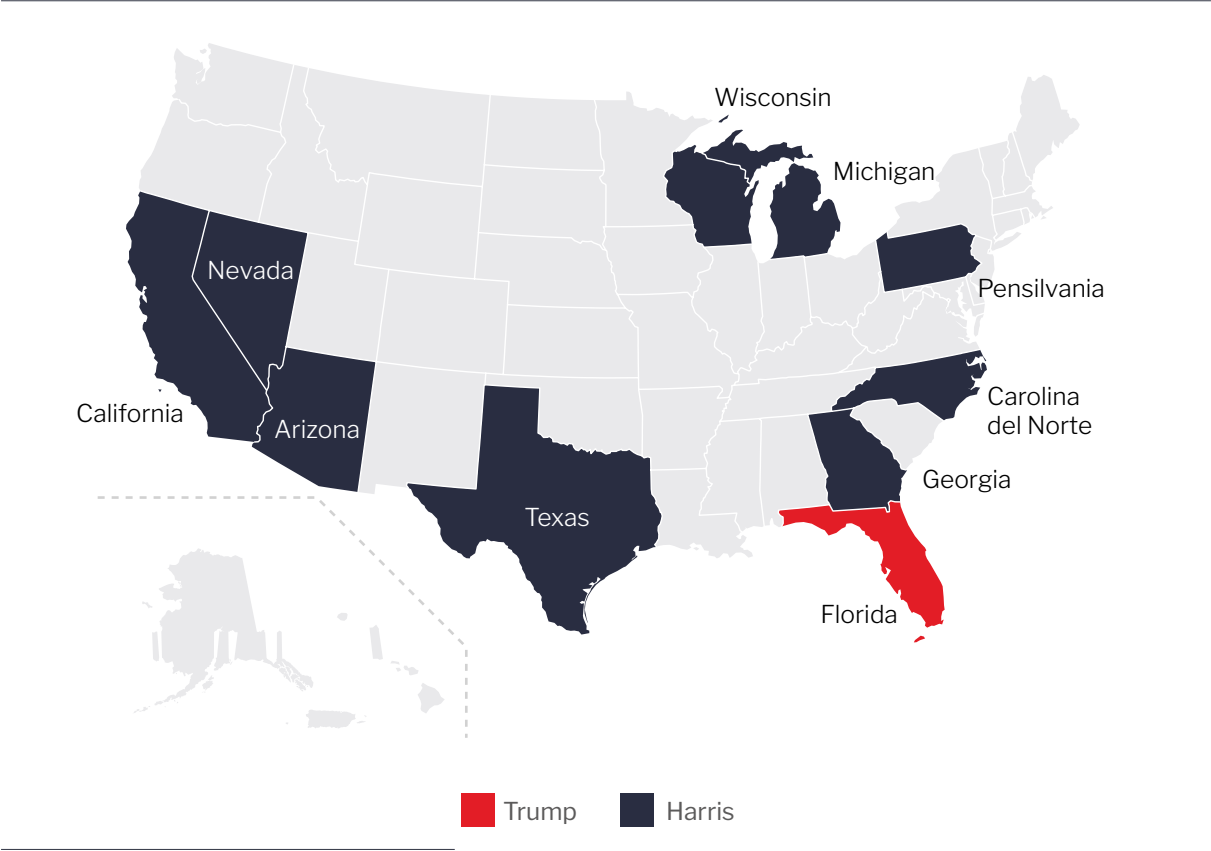
Estados bisagra	Kamala Harris	Donald Trump
Arizona	54 %	44 %
Georgia	59 %	41 %
Michigan	37 %	58 %
Carolina del Norte	49 %	50 %
Nevada	47 %	49 %
Pensilvania	59 %	41 %
Wisconsin	60 %	38 %

Estados con mayor población hispana	Kamala Harris	Donald Trump
Florida	42 %	58 %
Texas	45 %	55 %

FUENTE: elaboración propia a partir de los datos de NEP (CNN, 2024)¹⁰.

Sin embargo, la encuesta realizada por el 2024-AEVP indica que la población latina no votó mayoritariamente por Donald Trump en ninguno de los siete estados bisagra ni en dos de los de mayor población hispana (California y Texas), mostrando que solamente Florida, con un alto porcentaje de población de origen cubano, se decantó por el candidato Trump, como ha ocurrido tradicionalmente. El mapa del gráfico 16 muestra una visualización muy distinta a la anterior, la cual no ha llegado nunca a la opinión pública.

GRÁFICO 16 CANDIDATO CON MÁS VOTOS HISPANOS EN LOS 7 ESTADOS BISAGRA Y EN LOS 3 DE MAYOR POBLACIÓN HISPANA, SEGÚN 2024-AEVP



10. La encuesta del NEP no recoge datos sobre el voto hispano en el estado de California.

PORCENTAJE DE APOYO

Estados bisagra	Kamala Harris	Donald Trump
Arizona	66 %	33 %
Georgia	64 %	35 %
Michigan	64 %	33 %
Carolina del Norte	63 %	35 %
Nevada	64 %	35 %
Pensilvania	70 %	28 %
Wisconsin	64 %	34 %

Estados con mayor población hispana	Kamala Harris	Donald Trump
California	66 %	33 %
Florida	43 %	56 %
Texas	60 %	38 %

FUENTE: elaboración propia a partir de los datos de 2024-AEVP (African American Research Collaborative y BSP Research 2024a).

4. Conclusiones

La discrepancia observada a la hora de analizar la orientación del voto hispano entre el sondeo a pie de urna del NEP –ampliamente utilizado por los medios de forma global– con las encuestas previas y con otros sondeos a pie de urna menos mediáticos resulta preocupante, pues evidencia la falta de interés de los medios por reflejar con precisión la heterogeneidad y la diversidad de las minorías de Estados Unidos. Los medios de comunicación deben extremar su rigor, contrastando fuentes diversas, para mostrar la realidad social y cultural de estas comunidades minorizadas, así como a la hora de dibujar su imagen política en una sociedad tan polarizada como la estadounidense. La representación precisa de la población hispana en el campo político debería haber sido particularmente cuidadosa durante este ciclo electoral, ya que dicha comunidad ha recibido un inusitado foco de atención mediática, antes y después de las elecciones, en asuntos altamente

controvertidos como la inmigración ilegal, las deportaciones masivas, el uso de la lengua o incluso un supuesto aumento de la criminalidad.

El uso de datos sesgados y de información errónea, tanto metodológica como lingüísticamente, ha creado un estado de opinión pública que no refleja la compleja realidad de la orientación y evolución del voto hispano en las elecciones presidenciales de Estados Unidos. Nuestro análisis refuerza la necesidad de que instituciones como el Observatorio de la lengua española y de las culturas hispánicas del Instituto Cervantes en la Universidad de Harvard continúen dedicando recursos a realizar análisis rigurosos, con datos contrastados, que contemplen como base de su investigación la heterogeneidad demográfica, cultural, lingüística y política de las comunidades hispanas.

La evolución del voto hispano sigue la misma tendencia observada en ciclos electorales anteriores: la comunidad hispana se inclina mayoritariamente hacia el voto demócrata, aunque se sigue observando una tendencia creciente de apoyo al Partido Republicano.

Es fundamental analizar cómo la distribución del voto latino varía según el uso lingüístico en el hogar, considerando también la edad de los votantes. Las primeras generaciones de inmigrantes hispanos utilizan mayoritariamente el español, mientras que las segundas generaciones son bilingües. Sin embargo, en la tercera generación y las siguientes, la tendencia predominante es al monolingüismo en inglés, lo que lleva a dichas generaciones a obtener la información política de las mismas fuentes que la mayoría de sus conciudadanos estadounidenses. Como hemos mencionado anteriormente, los votantes bilingües y los dominantes de español tienden a inclinarse con más fuerza hacia el voto demócrata, mientras que entre los hispanos monolingües en inglés la tendencia hacia el voto republicano es más acentuada.

La evidencia recopilada en este estudio indica que, aunque el voto hispano en Estados Unidos sigue siendo mayoritariamente demócrata, existen variaciones significativas en función de la edad y el sexo. En términos generacionales, el mayor respaldo al Partido Republicano se da entre los mayores

de 50 años, mientras que la redistribución más notable del voto se observa entre los jóvenes. Además, se detecta una brecha de género: los hombres hispanos muestran una mayor preferencia por el candidato republicano en comparación con las mujeres, quienes mantienen un firme apoyo a Harris. Al considerar ambos factores en conjunto, el giro hacia el voto republicano es más pronunciado entre los hombres jóvenes que entre las mujeres de su misma generación. Aunque estas tendencias requieren un análisis más profundo, podrían tener implicaciones significativas en futuras elecciones presidenciales.

Referencias bibliográficas

- AFRICAN AMERICAN RESEARCH COLLABORATIVE y BSP RESEARCH (2024a). 2024 American electorate voter poll: Latino voters. <https://2024electionpoll.us/wp-content/uploads/2024/11/7.-Latino-crosstab.pdf>
- AFRICAN AMERICAN RESEARCH COLLABORATIVE y BSP RESEARCH (2024b). 2024 American electorate voter poll. Methodology. <https://2024electionpoll.us/methodology/>
- ANSORENA, J. (10 de noviembre de 2024). El voto hispano: el seísmo electoral que amenaza el futuro del Partido Demócrata. ABC. <https://www.abc.es/internacional/voto-hispano-seismo-electoral-amenaza-futuro-partido-20241109202349-nt.html>
- ASSOCIATED PRESS (4 de enero de 2024). How we survey American voters with AP VoteCast. <https://apnews.com/ap-votecast-faq>
- ATLASINTEL. (2 de noviembre de 2024). US Presidential Election 2024. <https://www.atlasintel.org/polls/general-release-polls>
- BRAMMER, J. P. (8 de noviembre de 2024). Latinx files: Reckoning with the 2024 election results. Los Angeles Times. <https://www.latimes.com/delos/newsletter/2024-11-08/latinos-trump-reckoning-deportation-right-shift-latinx-files>
- CARO, P. (12 de noviembre de 2024). Las organizaciones latinas contradicen los datos sobre el apoyo hispano a Trump. El País. <https://elpais.com/us/2024-11-13/las-organizaciones-latinas-contradicen-los-datos-sobre-el-apoyo-hispano-a-trump.html>
- CNN (2016). Exit polls. <https://www.cnn.com/election/2016/results/exit-polls>
- CNN (2020). Exit polls. <https://www.cnn.com/election/2020/exit-polls/president/national-results/7>
- CNN (2024). Election 2024: Exit polls. <https://www.cnn.com/election/2024/exit-polls/national-results/general/president/7>
- EL TIEMPO (7 de noviembre de 2024). Nuevo hito: Donald Trump rompe el récord de apoyo latino alcanzado por George W. Bush en las elecciones de Estados Unidos. <https://www.eltiempo.com/mundo/eeuu-y-canada/trump-supero-el-record-de-w-bush-en-apoyo-entre-los-votantes-latinos-en-las-elecciones-3397482>
- EL UNIVERSAL (7 de noviembre de 2024). Hispanos y mujeres, clave en el triunfo... de Trump. <https://www.eluniversal.com.mx/mundo/hispanos-y-mujeres-clave-en-el-triunfo-de-trump/>
- FABINA, J. y Z. SCHERER (febrero de 2022). Voting and Registration in the Election of November 2020, Current Population Survey Reports, P20-585. U.S. Census Bureau. <https://www.census.gov/library/publications/2022/demo/p20-585.html>

- FEDERAL ELECTION COMMISSION (16 de enero de 2025). *Official 2024 Presidential General Election Results*. <https://www.fec.gov/resources/cms-content/documents/2024presgeresults.pdf>
- FLORES, A. y LOPEZ, M. H. (11 de enero de 2018). *Among U.S. Latinos, the internet now rivals television as a source for news*. Pew Research Center. <https://www.pewresearch.org/short-reads/2018/01/11/among-u-s-latino-the-internet-now-rivals-television-as-a-source-for-news/>
- GÜENECHEA RODRÍGUEZ, J. I. y UREÑA UCEDA, D. (2020). *El papel del voto hispano en las elecciones presidenciales de EEUU 2020*. The Hispanic Council. <https://www.hispaniccouncil.org/wp-content/uploads/THC-Voto-hispano-Final.pdf>
- HERRERÍAS, A. Á. (8 de noviembre de 2024). *El gran giro de los latinos hacia Trump: del auge en los estados clave al crecimiento entre hombres jóvenes*. *El País*. <https://elpais.com/us/2024-11-08/el-gran-giro-de-los-latino-hacia-trump-del-auge-en-los-estados-clave-al-crecimiento-entre-hombres-jovenes.html>
- KROGSTAD, J. M., PASSEL, J. S., BUDIMAN, A. y NATARAJAN, A. (10 de enero de 2024). *Key facts about Hispanic eligible voters in 2024*. Pew Research Center. <https://www.pewresearch.org/short-reads/2024/01/10/key-facts-about-hispanic-eligible-voters-in-2024/>
- LANGE, J., ERICKSON, B. y HEATH, B. (7 de noviembre de 2024). *Trump's return to power fueled by Hispanic working-class voter support*. *Reuters*. <https://www.reuters.com/world/us/trumps-return-power-fueled-by-hispanic-working-class-voter-support-2024-11-06/>
- LODHI, H., CHENG, S., KAUFMANN, P., BARRIA URENDA, P. y FOX, E. J. (2024). *AP VoteCast: How America voted in 2024*. The Associated Press. <https://apnews.com/projects/election-results-2024/votecast/>
- LOPEZ, M. H. y GONZALEZ-BARRERA, A. (23 de julio de 2013). *A growing share of Latinos get their news in English*. Pew Research Center. <https://www.pewresearch.org/race-and-ethnicity/2013/07/23/a-growing-share-of-latino-get-their-news-in-english/>
- LOPEZ, M. H. y NOE-BUSTAMANTE, L. (24 de septiembre de 2024). *In Tight U.S. Presidential Race, Latino Voters' Preferences Mirror 2020*. Pew Research Center. <https://www.pewresearch.org/race-and-ethnicity/2024/09/24/in-tight-u-s-presidential-race-latino-voters-preferences-mirror-2020/>
- LUGONES, P. (6 de noviembre de 2024). *¿Por qué ganó Donald Trump? Algunas claves del contundente giro a la derecha en Estados Unidos*. *Clarín*. https://www.clarin.com/mundo/gano-donald-trump-claves-contundente-giro-derecha-estados-unidos_0_jG4ar8sTsG.html

MONTANARO, D. (8 de noviembre de 2024). Why Trump won – 9 takeaways from the 2024 election. NPR. <https://www.npr.org/2024/11/08/g-s1-33274/2024-election-how-trump-won-takeaways>

OFICINA DEL CENSO DE LOS ESTADOS UNIDOS (2021). *American Community Survey. S2901/Citizen, Voting-Age Population by Selected Characteristics*. <https://data.census.gov/table/ACSS-T1Y2021.S2901?q=S2901>

OFICINA DEL CENSO DE LOS ESTADOS UNIDOS (2023a). *American Community Survey. DP05/ACS Demographic and Housing Estimates*. <https://data.census.gov/table/ACSDPIY2023.DP05>

OFICINA DEL CENSO DE LOS ESTADOS UNIDOS (2023b). *American Community Survey. S2901/Citizen, Voting-Age Population by Selected Characteristics*. <https://data.census.gov/table/ACSST1Y2023.S2901?q=S2901>

OFICINA DEL CENSO DE LOS ESTADOS UNIDOS (2023c). *American Community Survey. S0201/Selected Population Profile in the United States*. https://data.census.gov/table/ACSSPPIY2023.S0201?t=400&g=010XX00US_040XX00US06,12,48

OFICINA DEL CENSO DE LOS ESTADOS UNIDOS (27 de junio de 2024). *New Estimates Highlight Differences in Growth Between the U.S. Hispanic and Non-Hispanic Populations, CB24-109*. <https://www.census.gov/newsroom/press-releases/2024/population-estimates-characteristics.html>

ROGERS, K. (5 de noviembre de 2024). Why Exit Polls Are Misleading. *The New York Times*. <https://www.nytimes.com/2024/11/05/us/politics/exit-polls-election.html>

SURO, R., FRY, R. y PASSEL, J. (2005). *Hispanics and the 2004 Election: Population, Electorate and Voters*. Pew Research Center. <https://www.pewresearch.org/race-and-ethnicity/2005/06/27/hispanics-and-the-2004-election/>

TELEMUNDO (13 de noviembre de 2024). ¿El voto latino le dio la victoria a Trump? Un nuevo estudio asegura que no. <https://www.telemundo.com/noticias/edicion-noticias-telemundo/elecciones-estados-unidos-2024/video/el-voto-latino-le-dio-la-victoria-trump-un-nuevo-estudio-asegura-que-no-tmvo12801389>

THE NEW YORK TIMES (13 de octubre de 2024). Cross-Tabs: October 2024 Times/Siena Poll of the Hispanic Likely Electorate. <https://www.nytimes.com/interactive/2024/10/13/us/elections/times-siena-poll-hispanic-likely-electorate.html>

VEGA HÜBNER, B. y PUEYO MENA, F. J. (DIRS.). (2025). *El voto hispano en las elecciones presidenciales de los Estados Unidos de 2024* (serie Estudios del Observatorio/Observatory Studies, n.º 95). Instituto Cervantes at the Faculty of Arts and Sciences of Harvard University. <https://cervantesobservatorio.fas.harvard.edu/es/informes/el-voto-hispano-en-las-elecciones-presidenciales-de-los-estados-unidos-de-2024>

III. CONCLUSIONES FINALES Y RECOMENDACIONES

La realidad de la lengua española en 2025, en toda su dimensión internacional, incluye datos que reflejan bien su importancia en cuanto a la demografía, el desarrollo humano, la cultura o la tecnología. El hecho de superar los 500 millones de hablantes como lengua nativa y los 630 millones de hablantes potenciales nos muestra una comunidad idiomática vigorosa y pujante en un panorama de diversidad.

Como en informes anteriores, los datos aquí reunidos reflejan de forma clara cuál es la realidad de la comunidad hispanohablante y cuáles los espacios de crecimiento y mejora. Por eso las conclusiones de este informe destacan unos datos y marcan unas líneas de acción que merecen tenerse en cuenta por parte de las autoridades, las entidades o las personas responsables de acciones que afectan a las lenguas y sus hablantes:

- a. La creación y el mantenimiento de datos demolingüísticos, así como el acopio de información sobre las dinámicas demográficas, geográficas y sociales del español y sus comunidades de hablantes, son tareas que merecen apoyarse y reforzarse. La importancia de esto es más evidente en el caso de los numerosos hablantes de español que suponen, de algún modo, poblaciones ocultas: desde los aprendices de español en la enseñanza no reglada y en espacios digitales, hasta las dinámicas de las segundas y terceras generaciones de migrantes hispanohablantes (hablantes de herencia).
- b. La comunidad hispanohablante muestra un crecimiento sostenido dentro de los propios países hispánicos, acompañando el ritmo de la demografía y sumando hablantes de otras lenguas. Ante este panorama, se hace necesario que los países presenten una oferta de enseñanza de español de calidad, compatible con programas de enseñanza bilingüe o multilingüe y orientada a políticas lingüísticas sostenibles respecto a la diversidad lingüística y cultural.

- c. La evolución del grupo de hablantes de español con habilidades limitadas, la más intensa de los grupos considerados, demanda la creación y el desarrollo de programas de refuerzo del mantenimiento de la lengua entre población migrante y de apoyo permanente al aprendizaje de español a lo largo de la vida y en comunidades donde es lengua minoritaria, especialmente en Estados Unidos, Europa y Brasil.
- d. En las próximas décadas, el peso relativo de los hablantes de competencia limitada del español aumentará dentro de la comunidad hispanohablante. Ello apunta a la necesidad de apoyar decididamente la educación lingüística, con el fin mejorar las condiciones de uso y de aprendizaje del español, sobre todo fuera de las fronteras de los territorios en que es lengua oficial, con el fin de mantener el conocimiento y uso del español entre los hablantes de herencia y de conseguir hablantes con más altas competencias.
- e. El futuro del español como lengua extranjera dependerá de su presencia en los sistemas educativos (como en Estados Unidos, Brasil, los países de Europa occidental y de África occidental y central) y de su institucionalización o consolidación en nuevos espacios educativos. Ello requiere la práctica de una diplomacia del español guiada por dos estrategias: una dirigida al sostén, la mejora y la ampliación de estructuras y acuerdos ya existentes; otra dirigida a la creación de nuevas estructuras y acuerdos bilaterales específicos para cada región.
- f. El panorama internacional del español evidencia la necesidad de reforzar la oferta de español para hacer posible su estudio allá donde tiene demanda. Para ello se hace necesario contribuir al reajuste de los sistemas educativos internacionales mediante programas de formación y actualización del profesorado de español.
- g. El desarrollo tecnológico de la lengua española requiere herramientas pensadas específicamente para el español e implementadas desde el español, para su mejor adaptación a la idiosincrasia cultural, estilística y de pensamiento de sus hablantes.

h. De nuestros análisis complementarios se obtienen conclusiones de gran interés para los sectores de referencia.

- En el campo de la economía, se hace necesario invertir en los factores que favorecen la inserción de las comunidades hispanohablantes en la economía digital (infraestructuras, educación...), mejorar los niveles de calidad institucional y salvaguardar la libertad de movimientos de personas, capitales y productos.
- En el ámbito de la inteligencia artificial, el español cuenta con una ventaja competitiva en cuanto a la disponibilidad de datos y corpus. Sin embargo, en el desarrollo actual de la IA sigue siendo una asignatura pendiente su aplicación a distintos ámbitos industriales y singularmente la adecuada atención a sus variedades sociales, estilísticas y geográficas.
- En relación con las tecnologías del lenguaje, el español es la lengua con más recursos en el catálogo European Language Grid, pero su cobertura es desigual tanto en las tipologías de recursos como en las infraestructuras.
- En cuanto al voto hispano en las últimas elecciones estadounidenses, se revela como fundamental su distribución según el uso lingüístico en el hogar, considerando también la edad de los votantes: las primeras generaciones de inmigrantes hispanos utilizan mayoritariamente el español, mientras que las segundas generaciones son bilingües; sin embargo, en la tercera generación y las siguientes, la tendencia predominante es al monolingüismo en inglés, lo que lleva a dichas generaciones a obtener la información política de las mismas fuentes que la mayoría de sus conciudadanos estadounidenses.

i. Las acciones diplomáticas siguen siendo imprescindibles para favorecer el uso institucional y público del español, especialmente en los foros internacionales, buscando su oficialidad. Son muchas las negociaciones que hay que afrontar para que el español esté al alcance de quien lo desee o lo necesite en los sistemas educativos públicos, en los organismos, acciones y gestiones internacionales, y en todo tipo de misiones de diplomacia.

- j. La creación de contenidos culturales y científicos en español, difundidos por canales especializados y por los medios de comunicación, es un factor decisivo para la valoración internacional de la lengua española y de las naciones que en ellas se comunican.

Todas estas acciones han de ser oportunamente contextualizadas y adecuadas al repertorio lingüístico y a las necesidades comunicativas de cada país o territorio. Además, el español debe presentarse, en su empleo y enseñanza, de forma responsable ante la diversidad y de modo que potencie los valores culturales que esta lengua encierra en cualquiera de sus manifestaciones. Decía Rubén Darío: «A vosotros mi lengua no debe ser extraña». No lo será si, de un modo decidido y competente, se pone al servicio de las comunidades internacionales.

