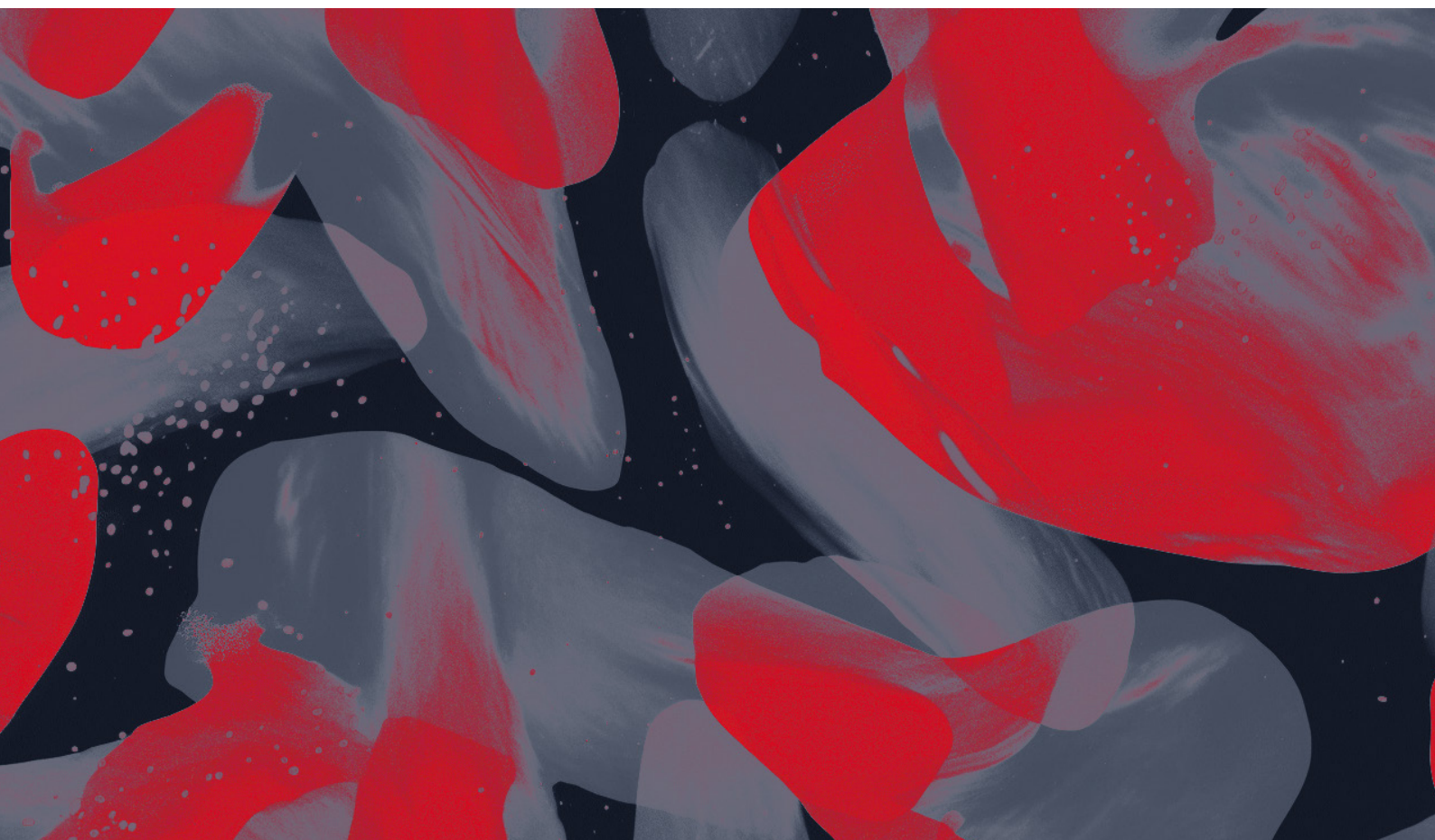


Informe 2024

# EL ESPAÑOL: LENGUA PARA EL MUNDO



**© 2024, INSTITUTO CERVANTES**

Libreros, 23. 28801 Alcalá de Henares (Madrid)

Alcalá, 49. 28014 Madrid

E-mail: [informa@cervantes.es](mailto:informa@cervantes.es)

<https://www.cervantes.org>

**CONSEJO DE REDACCIÓN:**

Luis García Montero, director del Instituto Cervantes

Carmen Noguero Galilea, secretaria general del Instituto Cervantes

Carmen Pastor Villalba, directora académica del Instituto Cervantes

Martín López-Vega, director del Gabinete de Dirección del Instituto Cervantes

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

**EL ESPAÑOL EN EL MUNDO. ANUARIO DEL INSTITUTO CERVANTES 2024:**

**DIRECTORA DE LA PUBLICACIÓN:** Carmen Pastor Villalba

**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 2024:**

**AUTOR:** Observatorio Global del Español

**EDITOR:** Francisco Moreno Fernández

**ADAPTACIÓN DE MAQUETA:** Departamento de Cultura Digital

**INSTITUTO CERVANTES**

NIPO (PDF): 110-25-074-3

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 2024

# EL ESPAÑOL: LENGUA PARA EL MUNDO



# Sumario

---

|                                                    |     |
|----------------------------------------------------|-----|
| <b>EL ESPAÑOL: LENGUA PARA EL MUNDO 2024</b>       |     |
| Observatorio Global del Español                    | 7   |
| <b>20 claves del español en el mundo</b>           |     |
| Observatorio Global del Español                    | 12  |
| <b>I. Demografía del español en el mundo 2024</b>  |     |
| Francisco Moreno Fernández y Héctor Álvarez Mella  | 15  |
| <b>II. Análisis complementarios</b>                |     |
| El español en los medios de comunicación           |     |
| Luis Miguel Pedrero Esteban                        | 81  |
| El español y la traducción en el mundo             |     |
| Ignacio Infante Fernández                          | 95  |
| El español y las industrias culturales             |     |
| Héctor Álvarez Mella                               | 104 |
| El español artificial                              |     |
| Antonio Moreno Sandoval                            | 117 |
| <b>III. Conclusiones finales y recomendaciones</b> | 141 |



# EL ESPAÑOL: LENGUA PARA EL MUNDO 2024

OBSERVATORIO GLOBAL DEL ESPAÑOL

## Preámbulo

La redacción de este informe se enmarca en los fines del Observatorio Global del Español (OGE) del Instituto Cervantes y el Gobierno de La Rioja, cuyos trabajos se han iniciado en 2024. El marco normativo del OGE señala como tareas más relevantes las de analizar la situación del español en el mundo y fomentar el conocimiento y prestigio de la lengua. Esto supone, entre otros objetivos, la elaboración de estudios sobre el español en sus distintas dimensiones (uso, conocimiento, aprendizaje, estudio), especialmente en el ámbito de la demolingüística.

El Instituto Cervantes ha venido prestando atención a la demografía de la lengua española desde el volumen fundacional de su *Anuario*, que siempre ha llevado por subtítulo «El español en el mundo» (VV. AA., 1998). El *Anuario* ha recogido año tras año información de primera mano sobre la situación del español y de la cultura en español en muy diversas áreas del planeta, sin olvidar los sectores o actividades en los que la lengua muestra un especial protagonismo. La primera etapa de los informes del Instituto Cervantes sobre el español culminó en 2006 con la publicación de la *Enciclopedia del español en el mundo* (VV. AA., 2006), que tuvo una brillante prolongación con la publicación de la *Enciclopedia del español en los Estados Unidos* (VV. AA., 2008)<sup>1</sup>.

---

1. Todos los anuarios del Instituto Cervantes aquí citados pueden consultarse en línea en la dirección: <https://cvc.cervantes.es/lengua/anuario/default.htm>.

En una segunda etapa, por iniciativa de la Dirección Académica del Instituto Cervantes y bajo sus directrices, el *Anuario* pasó a complementar sus análisis demolingüísticos regionales (VV. AA., 2009 y 2011) con un informe global que desde 2012 ha encabezado los volúmenes del *Anuario* con el título de «El español: una lengua viva» (Instituto Cervantes, 2012). La apuesta del Instituto Cervantes y de su Dirección Académica por el conocimiento detallado de la situación del español en el mundo se ve claramente reflejada en la continuidad de estos informes globales cuya ejecución material le fue encomendada a David Fernández Vitores, profesor de la Universidad de Alcalá.

Desde 2024, el OGE asume la responsabilidad de elaborar los informes demolingüísticos sobre el español en el mundo, en la línea ya consolidada del *Anuario* del Instituto Cervantes. En esta nueva etapa se dará continuidad a los contenidos propiamente demolingüísticos, cuyo valor se acrecienta al ponerse en perspectiva con los datos de años anteriores. Sin embargo, desde el OGE se introducirán algunas novedades que habrán de enriquecer el proyecto. Por un lado, los informes sobre la demolingüística y el lugar del español en el mundo irán acompañados de textos elaborados por especialistas en diferentes sectores que darán cuenta de la situación de sus respectivos ámbitos en relación con la lengua española. Además, los informes demolingüísticos incluirán desarrollos metodológicos y prestarán mayor atención a los entornos idiomáticos en que la lengua española se desenvuelve, puesto que los escenarios del español y de sus lenguas circunvecinas son indisolubles. Por otro lado, los informes elaborados desde el Observatorio pondrán un cambio de perspectiva al pasar de una mirada introspectiva (el español: una lengua viva) a una mirada de «extrospección» (el español: lengua para el mundo).

El Observatorio tendrá en los informes del *Anuario* una de las principales vías para la divulgación de los resultados de sus análisis, si bien no será la única, puesto que la complejidad y riqueza del español son tales que no pueden abarcarse en la concisión de un simple estudio anual de conjunto. Con todo, los informes han de contribuir a los fines esenciales del OGE: reforzar la identidad hispanohablante, en su comunidad y en su diversidad; proyectar

la lengua española y la cultura hispanohablante en el plano internacional; transmitir unos valores de convivencia y respeto a las identidades culturales propias y ajenas.

Decía García Márquez en 1968 que el español es un idioma fabulosamente eficaz, pero también fabulosamente olvidado (Domingo, 1968). Confiamos en que los informes «El español: lengua para el mundo» contribuyan a que no caigan en el olvido ni las lenguas del espacio hispanohablante ni las culturas y valores que en ellas se expresan.

Francisco Moreno Fernández  
Director del Observatorio Global del Español

# EL ESPAÑOL: LENGUA PARA EL MUNDO

Más de

# 600

millones de personas (600.607.806) son usuarios potenciales de español en el mundo.



# 7,5%

de la población mundial

498.497.757

personas con **dominio nativo**



77.901.236

usuarios de **competencia limitada**



24.208.813

**aprendices** de Lengua Española

El español es la **2.<sup>a</sup>** **lengua materna** del mundo tras el chino mandarín.

Fuera de los países hispanicos hay cerca de

# 100 millones

de hablantes de español.

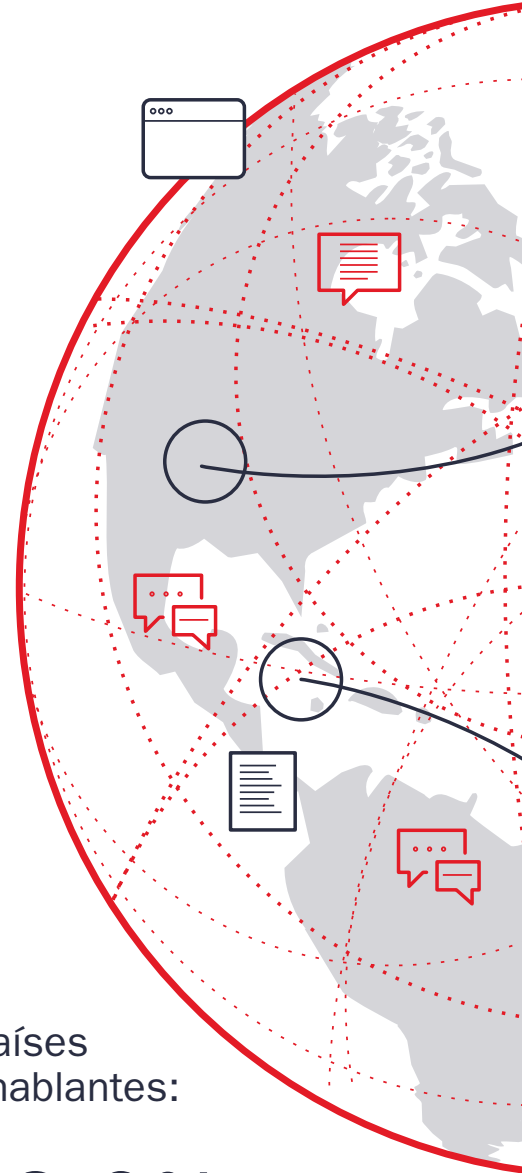
Más del **78%**

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

En los países hispanohablantes:

→ El **93,6%** de la población tiene un dominio nativo del español.

→ El **92,2%** de la población indígena u originaria puede comunicarse en español.



Hay

63,7

millones de  
hispanos en  
Estados Unidos.



El **75%** son capaces de  
mantener una conversación  
en español.

El porcentaje de votantes  
hispanos en Estados Unidos  
es del **14,7%** (el doble  
que en el año 2000).

El español es la

**6.<sup>a</sup>** lengua en cuanto a  
**producción editorial** en  
el mundo, con un 7 %  
de los libros publicados.

**2.<sup>a</sup>** lengua más importante  
de **producción  
cinematográfica** en el  
mundo, con más de  
4.300 películas  
producidas entre 2007  
y 2017.

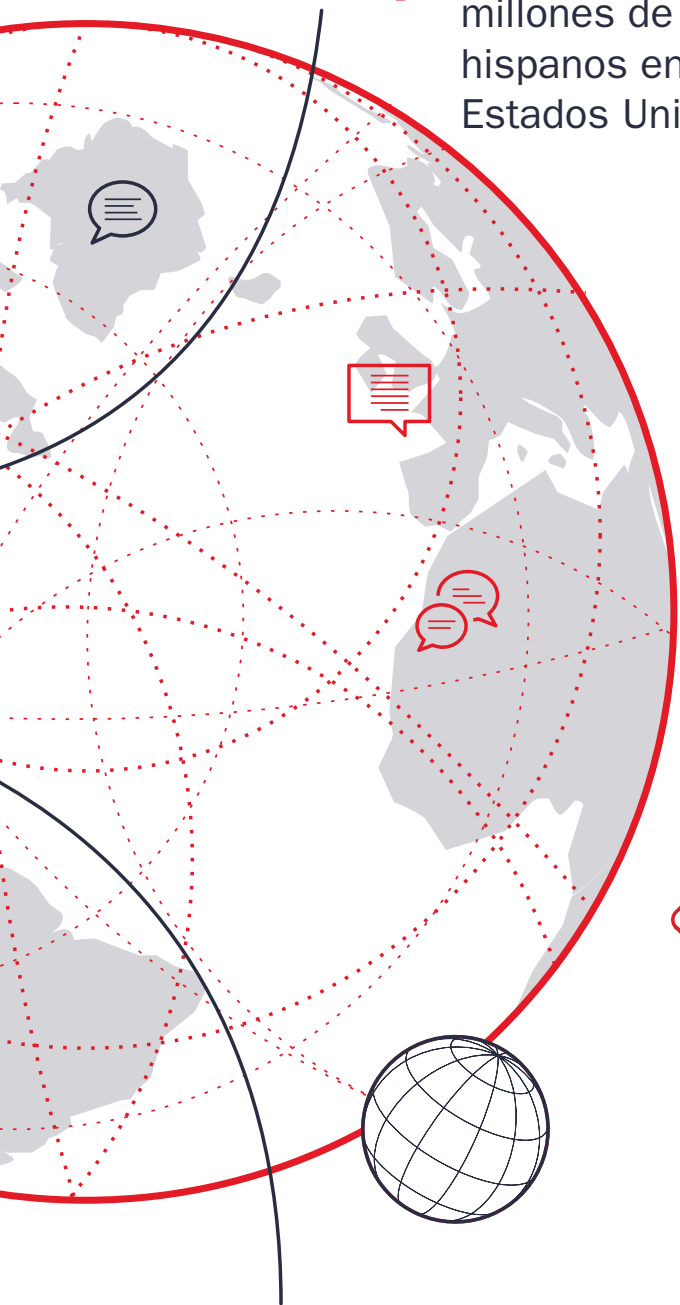
**3.<sup>a</sup>** lengua **receptora de  
traducciones**, después  
del alemán y el francés.

**6.<sup>a</sup>** lengua **desde la que se  
realizan traducciones**,  
tras el inglés, el francés,  
el alemán, el ruso  
y el italiano.

En 2023, el español  
fue la lengua del

21%

de las 500 canciones más  
reproducidas en Spotify, por  
detrás del inglés (75 %).



# EL ESPAÑOL: LENGUA PARA EL MUNDO 2024

OBSERVATORIO GLOBAL DEL ESPAÑOL

## 20 claves del español en el mundo

1. Los hablantes potenciales de español sobrepasan en 2024 por primera vez la cifra de los 600 millones en todo el mundo.
2. La comunidad hispanohablante con dominio nativo de la lengua roza los 500 millones.
3. En los países hispanohablantes, la proporción de hablantes con dominio nativo es del 93,63 % de la población.
4. Entre los países hispanohablantes, solamente Paraguay, Guatemala, Bolivia y Guinea Ecuatorial tienen una proporción de nativos de español inferior al 85 %.
5. El volumen de hablantes nativos de español fuera de los países hispánicos es superior a los 45 millones.
6. El volumen de hablantes con competencia limitada en los países hispanohablantes se sitúa cerca de los 25 millones y, en los países no hispánicos, por encima de los 50 millones.
7. La cantidad de hablantes de español fuera de los países hispánicos está cercana a los 100 millones, principalmente por su trasfondo migratorio.
8. El porcentaje de población indígena u originaria capaz de comunicarse en español es del 92,25 % y el de población monolingüe en una lengua distinta del español es del 7,75 %.

9. En 2024, más de 24 millones de personas aprenden español en el mundo.
10. Los aprendices de español se concentran fundamentalmente en Estados Unidos, la Unión Europea y Brasil, dado el peso relativo de este idioma en los sistemas de enseñanza reglada.
11. Durante los últimos 10 años, el número de aprendices de español ha venido creciendo a un ritmo medio del 2 % impulsado por su dinamismo en países de la Unión Europea y África occidental y central.
12. El volumen de hablantes de español como lengua nativa se sitúa por encima del inglés, el hindi y el árabe, y solo por debajo del chino mandarín.
13. El número de países en los que el español es oficial, vehicular o mayoritario es menor que el de países hablantes de inglés, francés o árabe, pero el índice de población hablante sitúa al español como la segunda lengua, tras el chino mandarín, entre los idiomas oficiales de la Organización de Naciones Unidas (ONU).
14. El índice de vehicularidad del español como segunda lengua es el más bajo de las lenguas oficiales de la ONU, mientras que el coeficiente de hablantes de español como primera lengua está entre los más altos.
15. El peso de la lengua española en el ámbito de la traducción muestra un alto valor como lengua meta y menor como lengua fuente, aunque, en este último caso, por encima del chino mandarín o el ruso.
16. La enseñanza universitaria internacional ofrece un espacio destacado a la lengua española, en competencia con el francés.
17. La comunidad idiomática hispanohablante muestra un nivel de desarrollo humano superior al promedio de las comunidades anglófona y francófona.
18. La proporción de sitios web relevantes cuyo contenido se expresa en español se sitúa por encima de los porcentajes de los sitios con contenido en alemán o en francés.

19. Los hispanos de Estados Unidos con capacidad de voto en las elecciones presidenciales de 2024 son más de 36 millones, lo que supone un 14,7 % del total del electorado estadounidense.
20. El porcentaje de votantes hispanos en Estados Unidos se ha duplicado en las últimas dos décadas, pasando del 7,4 % en 2000 a la previsión de un 14,7 % en 2024.

# I. DEMOGRAFÍA DEL ESPAÑOL EN EL MUNDO 2024<sup>1</sup>

## 1 | Introducción

La lengua española es, sin duda, una de las más relevantes del mundo. Esa relevancia no se fundamenta en sus caracteres lingüísticos, dado que todas las lenguas tienen en sí mismas idéntica naturaleza y valía, sino en las particularidades de su comunidad de hablantes y las condiciones de su adquisición, aprendizaje, uso y estudio, así como en los factores contextuales que las rodean. De hecho, el mayor o menor peso de las lenguas en el mundo se deriva de una multiplicidad de factores entre los cuales los puramente lingüísticos son tal vez los menos significativos.

En el análisis del conocimiento y el uso de cualquier lengua del planeta resulta fundamental la información demográfica. En el caso de la demografía del español, esta es extraordinariamente interesante por la diversidad de los espacios en que se despliega. El estudio demográfico de las lenguas requiere prestar atención a tres factores fundamentales: la geografía, la población hablante y las lenguas o variedades habladas. La geografía potencial del español es la de todos los continentes, si bien su hábitat más consolidado es el de los territorios históricamente hispanohablantes; la población hispanohablante es tanto la que habita estos territorios, como la dispersa por otros muchos espacios internacionales; la lengua, por último, es la que recibe la denominación de español o castellano, entendida como koiné resultante de la base común de todas sus variedades.

El informe «El español: lengua para el mundo 2024» pretende ofrecer una puesta al día de algunos de los caracteres esenciales de la lengua española en el panorama internacional. Entre esos caracteres destaca el

---

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

demográfico. Dentro de la demografía del español, es necesario distinguir metodológicamente tres perfiles fundamentales. El primero de ellos es el perfil de los hablantes con un dominio nativo de la lengua, que la han adquirido, han llegado a dominarla o son percibidos como hablantes nativos de español, tanto si son monolingües como bilingües. El segundo perfil es el de los hablantes con competencia limitada, quienes, siendo conocedores y usuarios de la lengua, no muestran una competencia lingüística, social o estilística como la de los nativos en las destrezas básicas. El tercer perfil es el de los aprendices de español, que pueden tener diferente nivel de proficiencia en el uso de lengua.

Este informe presenta información de los hablantes que tienen un *dominio nativo de español* dentro de cada uno de los países en que esos hablantes tienen una presencia notable. Ello incluye, por supuesto, los países históricamente considerados como hispanohablantes, pero también aquellos en los que habitan, por causas diversas, contingentes de hablantes de español. Junto a estos hablantes se tienen en cuenta aquellos que no tienen un dominio nativo, pero sí un conocimiento lo suficientemente amplio como para comunicarse en español. Este grupo de hablantes, llamado de *competencia limitada*, resulta difícil de calibrar por cuanto el conocimiento de las lenguas, así como la extensión de su uso, se difumina en la complejidad de las capacidades cognitivas y en la multiplicidad de funciones sociales e interpersonales que la lengua puede cumplir en cada contexto. El informe, finalmente, atiende a la demolingüística de los *aprendices de español* y a este respecto 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 para cada una de ellas. Como es lógico, las dificultades metodológicas solo se salvan –total o parcialmente– mediante recursos e instrumentos metodológicos, que aquí quedarán debidamente explicitados.

Al mismo tiempo, para una cabal comprensión del alcance del conocimiento y el uso del español en el mundo, es imprescindible prestar atención a las comunidades lingüísticas con las que los hispanohablantes, bilingües o no,

establecen relación de convivencia o de vecindad. Con tal fin, este informe ofrece información sobre la dimensión y la sociolingüística de los hablantes de otras lenguas, que, en el caso del espacio hispanohablante, suelen proceder mayoritariamente de *pueblos indígenas u originarios*. En este punto, las dificultades se acrecientan por la precariedad de la información existente sobre estas poblaciones y sus peculiaridades socioculturales. Con todo, se ha hecho un esfuerzo de estimación que aporta un panorama de conjunto muy revelador. En el caso de España, las estadísticas oficiales disponibles propician un conocimiento más detallado de la realidad sociolingüística del país y de sus comunidades.

Por otro lado, la proyección de la lengua española en la actualidad solo puede valorarse adecuadamente a partir de una comparación sistemática y fundamentada con otras *lenguas internacionales*. Por tal motivo, este informe analiza la posición del español respecto a otras lenguas del mundo, con especial atención a las lenguas oficiales de la ONU: el inglés, el francés, el chino mandarín, el ruso y el árabe, además del español. El análisis comparativo se hace sobre las bases de datos e informaciones referidas a varias dimensiones: una dimensión poblacional y comunicativa, una dimensión educativa y cultural, una dimensión de desarrollo humano y una dimensión digital. Obviamente, estas no son todas las variables que afectan a la vida social de las lenguas: así, la dimensión económica tiene también una gran trascendencia. Sin embargo, los datos proporcionados en los informes del Instituto Cervantes correspondientes a 2022 y 2023 no han variado sustancialmente en 2024, por lo que remitimos a ellos para mayor detalle<sup>2</sup>, con el propósito de actualizar estos aspectos en próximas entregas anuales. Asimismo, la proyección del número de hablantes para lo que resta de siglo también se ha ofrecido reiteradamente en informes anteriores, sin que las previsiones se hayan modificado de un modo significativo<sup>3</sup>.

---

2. Para obtener información sobre la dimensión económica del español en 2023, véase la segunda sección del informe elaborado por Fernández Vítóres (Instituto Cervantes, 2023).

3. Para la previsión de crecimiento de la comunidad hispanohablante en el siglo xxi, véase la primera sección del informe elaborado por Fernández Vítóres (Instituto Cervantes, 2023).

Este informe de 2024 incluye, asimismo, una nota demográfica que afecta a uno de los países más decisivos en la difusión internacional del español: Estados Unidos. En noviembre de 2024 se celebran en ese país elecciones presidenciales y nuevamente adquiere prominencia informativa todo lo relativo al voto hispano. Por tal razón, nuestro informe se hace eco de los últimos estudios referidos a la incidencia de la población hispana o latina sobre el proceso de elección presidencial.

Finalmente, este informe incluye cuatro análisis monográficos dedicados a sendos aspectos del uso y la difusión internacionales de la lengua española. En primer lugar, se analizan algunas cuestiones clave referidas al español en los medios de comunicación, con especial atención al medio radiofónico, dado que en 2024 se celebra el centenario de la implantación de la radiodifusión en España. En segundo lugar, se ofrece un estudio referido a diversos aspectos esenciales del español en las industrias culturales. En tercer lugar, se presenta un análisis de las traducciones al español y del español, que se detiene especialmente en las traducciones que afectan a la lengua inglesa. Por último, un texto sobre la inteligencia artificial y las tecnologías del lenguaje aplicadas al español ofrecerá algunos datos significativos y reflexiones esenciales sobre un ámbito que crecientemente acapara la atención de la sociedad y en el que la lengua española se juega buena parte de su relevancia internacional.

## 2 | Demolingüística fundamental del español

### 2.1 | El español en los países hispanohablantes

Las geografías de la lengua española se han caracterizado tradicionalmente por la vitalidad de su demografía, el significativo número de países y territorios que la comparten –en su mayoría contiguos–, la pujanza de las culturas que expresa y el alto índice de comunicatividad entre los hablantes de sus comunidades (Moreno Fernández y Otero Roth, 1998).

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

| País                 | Población   | Grupo de Dominio Nativo (GDN) | Hablantes nativos (%) | Grupo de Competencia Limitada (GCL) |
|----------------------|-------------|-------------------------------|-----------------------|-------------------------------------|
| México               | 132.490.415 | 124.781.373                   | 94,18 %               | 7.629.547                           |
| Colombia             | 52.695.952  | 51.684.976                    | 98,08 %               | 1.010.976                           |
| España               | 48.446.594  | 46.460.284                    | 95,90 %               | 1.647.184                           |
| Argentina            | 46.846.845  | 46.246.057                    | 98,72 %               | 600.788                             |
| Perú                 | 35.371.496  | 30.426.773                    | 86,02 %               | 1.005.682                           |
| Venezuela            | 29.395.334  | 28.632.291                    | 97,40 %               | 763.043                             |
| Chile                | 20.086.377  | 19.199.657                    | 95,59 %               | 886.720                             |
| Guatemala            | 17.843.132  | 13.542.955                    | 75,90 %               | 3.240.126                           |
| Ecuador              | 17.142.254  | 16.062.292                    | 93,70 %               | 537.552                             |
| Bolivia              | 12.332.252  | 9.927.462                     | 80,50 %               | 2.244.394                           |
| Cuba                 | 11.174.587  | 11.079.639                    | 99,15 %               | 1.117                               |
| República Dominicana | 10.792.332  | 10.241.923                    | 94,90 %               | 550.409                             |
| Honduras             | 9.892.632   | 9.409.893                     | 95,12 %               | 482.739                             |
| Paraguay             | 7.656.215   | 5.153.398                     | 67,31 %               | 2.456.048                           |
| Nicaragua            | 7.143.847   | 6.929.532                     | 97,00 %               | 214.315                             |
| El Salvador          | 6.399.471   | 6.391.792                     | 99,88 %               | 7.679                               |
| Costa Rica           | 5.309.627   | 5.254.508                     | 98,96 %               | 55.119                              |
| Panamá               | 4.217.605   | 3.643.227                     | 86,38 %               | 562.453                             |
| Uruguay              | 3.455.973   | 3.425.907                     | 99,13 %               | 30.066                              |
| Puerto Rico          | 3.264.993   | 3.105.008                     | 95,10 %               | 159.985                             |
| Guinea Ecuatorial    | 1.633.849   | 1.209.048                     | 74,00 %               | 424.801                             |
| Total                | 483.591.782 | 452.807.955                   | 93,63 %               | 24.510.743                          |

**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 (INDEC, 2023). Sobre datos de 2022, se incrementa tasa anual de crecimiento poblacional (1,04 %). En el GCL se incluyen 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 (2020). Proyección de crecimiento de población de 1,2 %. En 2024 se está realizando nuevo censo. El GCL incluye hablantes de lenguas indígenas no monolingües y de escolarización baja, así como extranjeros de origen no hispanohablante.

### Chile

Datos del Instituto Nacional de Estadísticas de Chile (2017a) estimados para 2024 con base en el Censo de Población y Vivienda de 2017. 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.

### Colombia

Datos del Departamento Administrativo Nacional de Estadística (DANE, 2018). Proyección oficial sobre censo de 2018.

### Costa Rica

Proyección del Instituto Nacional de Estadística y Censos (INEC, 2024). El último censo realizado es de 2022. El GCL incluye hablantes de lenguas indígenas y extranjeros de origen no hispanohablante.

### Cuba

Datos de la Oficina Nacional de Estadística e Información (ONEI, 2023). Estimación para 2024. Consideraciones demográficas en Álvarez Quiñones (2023).

### Ecuador

Datos del Instituto Nacional de Estadística y Censos (INEC, 2022). Proyección oficial de crecimiento poblacional. El GCL incluye hablantes de lenguas indígenas no monolingües y de escolarización baja, así como extranjeros de origen no hispanohablante.

### El Salvador

Proyección a partir de datos de la Oficina Nacional de Estadística y Censos (ONEC, 2024). Sobre el último censo, se proyecta la población con un incremento medio anual del 0,29 %. El VI Censo de Población y Vivienda se realizó en 2007. El GCL incluye extranjeros de origen no hispanohablante.

### España

Datos del Instituto Nacional de Estadística (INE, 2024). En el GDN, a la población total se le resta proporcionalmente las personas que no hablan nada de español o castellano y las que lo hablan con dificultad (INE, 2021).

### Guatemala

Proyección para 2024 del Instituto Nacional de Estadística de Guatemala (2024). El GCL incluye hablantes de lenguas indígenas no monolingües y de escolarización baja, así como extranjeros de origen no hispanohablante.

### Guinea Ecuatorial

Proyección a partir de estimación oficial de 2022 del Instituto Nacional de Estadística de Guinea Ecuatorial (INEGE, 2022), con un incremento poblacional anual de 2,4 %. El porcentaje del GDN es el manejado por Moreno Fernández y Otero (2007). El GLC incluye hablantes de lenguas nativas, criollas y europeas. El nivel de competencia lingüística es el 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 con el fin de posibilitar una valoración de la evolución de la comunidad de hablantes, considerando miembros del Grupo de Dominio Nativo a aquellos que lo son, que manejan la lengua con un dominio equivalente al de un nativo o que son considerados como tales por otros hablantes. Este criterio se fundamenta en argumentos ajenos al mero orden de adquisición. Otros análisis, a menudo sin el respaldo de pruebas empíricas sobre el dominio de la lengua, optan por no dar el tratamiento de «nativos», sino de «hablantes de L2 o de competencia limitada», a la mayoría de los hispanohablantes guineanos.

### Honduras

Proyección del Instituto Nacional de Estadística de Honduras (2013). Acceso a datos a través de BASEINE del Instituto Nacional de Estadística de Honduras (2024). El GCL incluye hablantes de lenguas indígenas y población negra anglohablante, así como extranjeros de origen no hispanohablante.

### México

Datos del Consejo Nacional de Población (CONAPO, 2024). Incremento desde datos de 2023, según el Instituto Nacional de Estadística y Geografía (INEGI, 2024). El GCL incluye hablantes indígenas con dominio precario de español y extranjeros de origen no hispanohablante (INEGI, 2022).

### Nicaragua

Proyección a partir de datos del Instituto Nacional de Información de Desarrollo (INIDE, 2024). Sobre censo de 2005, se proyecta la población con un incremento medio anual del 3 %. Consulta de datos por sistema REDATAM.

### Panamá

Proyección a partir del censo de 2023 del Instituto Nacional de Estadística y Censo (INEC, 2024), con un incremento poblacional anual del 1,3 %. 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.

### Paraguay

Proyección del Instituto Nacional de Estadística de Paraguay (2024). El GDN incluye hablantes monolingües en español y bilingües con guaraní. Datos a partir de REDATAM.

### Perú

Datos del Instituto Nacional de Estadística e Informática (INEI, 2022). Previsión oficial para 2024. El GDN incluye población indígena con estudios secundarios y universitarios en español. Acceso a

datos a través del sistema REDATAM. El GCL incluye hablantes de lenguas indígenas no monolingües y de escolarización baja, así como extranjeros de origen no hispanohablante.

#### Puerto Rico

Datos del Instituto de Estadísticas de Puerto Rico (2023). Para el cálculo del GCL se restan del GDN hablantes en casa de inglés y de otras lenguas.

#### República Dominicana

Proyección de la Oficina Nacional de Estadística (ONE, 2023). El GCL incluye extranjeros de origen no hispanohablante, principalmente haitianos.

#### Uruguay

Proyección del Instituto Nacional de Estadística de Uruguay (2023) a partir del censo de 2023, con un incremento poblacional anual del 0,34 %. El GCL incluye extranjeros de origen no hispanohablante.

#### Venezuela

Proyección de la ONU para 2024 (ONU, 2022). El censo de 2022 ofrecía una estimación de 28,3 millones (Universidad Católica Andrés Bello, 2022). El GCL incluye hablantes de lenguas indígenas no monolingües y de escolarización baja, así como extranjeros de origen no hispanohablante.

1. Los datos de poblaciones están actualizados a 2024 en la mayoría de los casos, de acuerdo con la información ofrecida por las respectivas oficinas del censo o institutos de estadísticas. Cuando no se cuenta con cifras oficiales para 2024, se adopta la proyección realizada por organismos oficiales o internacionales. En los casos en que no se cuenta con proyecciones oficiales, se ha aplicado la tasa oficial anual de crecimiento poblacional a partir de los datos del último censo realizado. Se da prioridad a las fuentes nacionales sobre las internacionales.
2. La cantidad de hablantes integrantes del Grupo de Dominio Nativo 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. El dominio nativo de la lengua española no es óbice para que también se tenga un dominio nativo de otra u otras lenguas. El bilingüismo es una significativa realidad en buena parte del territorio hispánico.
3. La proporción de hablantes nativos se calcula a partir del Grupo de Dominio Nativo en relación con la población total de cada país.

4. La cantidad de hablantes integrantes del Grupo de Competencia Limitada se obtiene generalmente como resultado de la suma de los individuos adscritos a pueblos indígenas u originarios que son analfabetos, no han sido escolarizados en lengua española o lo han sido solamente hasta un nivel primario, excluyendo los monolingües en lenguas indígenas identificados como tales. A esta 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 pueden incluirse aquí los hablantes de otras lenguas nativas que han reconocido un uso limitado de la lengua española.

La aplicación de estos criterios para 2024 ha supuesto la revisión y actualización de los porcentajes de hablantes con dominio nativo del español en cada país. En informes anteriores, los porcentajes de hablantes de dominio nativo se habían extraído del análisis demolingüístico publicado por Moreno Fernández y Otero Roth en la 3.<sup>a</sup> edición de su *Atlas* publicada en 2016. En esta ocasión, los porcentajes se han calculado *ex novo* de acuerdo con los datos actualizados que proporcionan las fuentes demográficas pertinentes. Esta actualización ha sido necesaria por el tiempo transcurrido desde 2016, fecha del último cálculo de los porcentajes, dado que tales proporciones no son estáticas, sino dinámicas, por estar sujetas a las condiciones derivadas de muy distintos procesos políticos, educativos o migratorios, entre otros.

Al comparar los datos de 2024 con los de años anteriores, se aprecia que la población del conjunto de la comunidad hispanohablante continúa creciendo y ello supone también un aumento de la cifra de hablantes nativos de español en términos absolutos, dada la condición mayoritaria de esta lengua. No obstante, el porcentaje de hablantes nativos (93,63 %) resulta algo menor respecto del informe elaborado en 2023 (94,5 %). En este caso, no es posible argumentar sobre las causas de esta diferencia interanual, dado que los porcentajes manejados en el anterior informe del Instituto Cervantes remitían a los de 2016, no a los del año 2023; y algo similar ocurre

respecto al Grupo de Competencia Limitada. El nuevo cálculo de porcentajes, basado en un pormenorizado recuento de hablantes nativos de otras lenguas, tanto monolingües como bilingües, se ajusta todo lo que ha sido posible a la realidad de 2024.

Entre los datos referidos a la proporción de hablantes nativos de español, llaman la atención los porcentajes inferiores al 85 % de Guatemala, Bolivia, Paraguay y Guinea Ecuatorial. Como es sabido, las poblaciones originarias de estos países han mantenido históricamente unas importantes proporciones demográficas: las mayas en Guatemala; las quechuas y aimaras, junto a otras, en Bolivia; las guaraníes en Paraguay; las bubi y fang, junto a otras, en Guinea Ecuatorial. Por eso resulta interesante constatar que la proporción de hablantes con un dominio nativo del español es superior al 80 % en Bolivia y Perú, y al 75 % en Guatemala. En Paraguay, el porcentaje es más bajo (67,31 %), pero no debe olvidarse que el guaraní yopará, muy extendido en el país, es una mezcla de lenguas bilingüe en la que el componente español es muy significativo.

En cuanto al Grupo de Competencia Limitada, llaman la atención por sus valores extremos las bajas cifras registradas en Cuba, que corresponden mayormente a una reducida población extranjera, y los altos guarismos de países como México, Guatemala, Bolivia o Paraguay, debidos precisamente a la dimensión de sus comunidades indígenas bilingües, o de España, con una población inmigrante de origen no hispanohablante que ha experimentado un gran crecimiento en las dos últimas décadas.

## 2.2 | El español en coexistencia con otras lenguas

Como se deduce del cuadro 1, la demografía del español en los países hispanohablantes responde de un modo directo a la demografía de las poblaciones que tienen como nativa o materna una lengua diferente, por más que también puedan conocer y usar la española. Es evidente que existe una alta correspondencia entre las poblaciones hispánicas y los hablantes con dominio nativo del español (93,63 %), pero por eso mismo adquiere mayor

relieve dentro del espacio hispánico la coexistencia de lenguas; es decir, el bilingüismo o multilingüismo.

El bilingüismo es una realidad psico y sociolingüística a la que da carta de naturaleza la legislación de numerosos países hispanohablantes. A este respecto, el instrumento más decisivo en el proceso de oficialización de una o varias lenguas de un país es la constitución, junto a las leyes y normativas complementarias. Este corpus jurídico no suele crearse a partir de razones o evidencias puramente lingüísticas, sino más bien con fundamentos sociales y políticos. Esas razones políticas hacen posible que una lengua sin arraigo popular en un territorio pueda alcanzar el rango de oficial, como ocurre con el francés en Benín o Níger, o que una lengua bien asentada en un país no alcance el reconocimiento de la oficialidad nacional, como ocurre con el inglés en Estados Unidos o con el español en Argentina. En el caso de los territorios hispanohablantes, cada país ha resuelto la inclusión o no de la lengua española o castellana –o de otras lenguas– en su constitución dependiendo de su particular circunstancia sociopolítica. Esto significa que, en el marco de sus respectivas legislaciones, cada país presenta una dinámica de bilingüismo diferenciada. Ciertamente, existen paralelismos y correspondencias entre muchos de ellos, pero los diferentes entornos condicionan de forma específica a cada comunidad bilingüe de los países hispánicos.

| CUADRO 2   LENGUAS OFICIALES EN PAÍSES HISPÁNICOS |              |                                                                                                          |
|---------------------------------------------------|--------------|----------------------------------------------------------------------------------------------------------|
| País                                              | Constitución | Lenguas oficiales                                                                                        |
| Argentina                                         | 1994         | No se menciona lengua oficial.                                                                           |
| Bolivia                                           | 2009         | Castellano (y todos los idiomas de las naciones y de los pueblos indígenas originarios campesinos [36]). |
| Chile                                             | 2001         | No se menciona lengua oficial.                                                                           |
| Colombia                                          | 1991         | Castellano; lenguas y dialectos de los grupos étnicos en sus territorios.                                |
| Costa Rica                                        | 1949         | Español.                                                                                                 |
| Cuba                                              | 1976         | Español.                                                                                                 |
| Ecuador                                           | 1998         | Castellano (quichua, shuar y otros idiomas ancestrales de uso oficial para los pueblos indígenas).       |

| País                 | Constitución | Lenguas oficiales                                                                                                                                                                                            |
|----------------------|--------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| El Salvador          | 1983         | Castellano.                                                                                                                                                                                                  |
| España               | 1978         | Castellano. Las demás lenguas españolas en las respectivas comunidades autónomas según sus estatutos (catalán, gallego, valenciano, vasco o euskera). Protección a las «distintas modalidades lingüísticas». |
| Guatemala            | 1985         | Español.                                                                                                                                                                                                     |
| Guinea Ecuatorial    | 2011         | Español, francés, portugués y las que la ley determine (lenguas autóctonas como integrantes de la cultura nacional).                                                                                         |
| Honduras             | 1982         | Español.                                                                                                                                                                                                     |
| México               | 1917         | No se menciona lengua oficial.                                                                                                                                                                               |
| Nicaragua            | 1987         | Español (lenguas de las comunidades de la Costa Atlántica, uso oficial).                                                                                                                                     |
| Panamá               | 1972         | Español.                                                                                                                                                                                                     |
| Paraguay             | 1992         | Castellano y guaraní (lenguas indígenas minoritarias, patrimonio cultural de la nación).                                                                                                                     |
| Perú                 | 1993         | Castellano (quechua, aimara y demás lenguas aborígenes, también donde predominen).                                                                                                                           |
| Puerto Rico          | 1952; 1970   | No se menciona (español e inglés necesarios para ser miembros de la Asamblea).                                                                                                                               |
| República Dominicana | 1994         | No se menciona lengua oficial.                                                                                                                                                                               |
| Uruguay              | 1967         | No se menciona lengua oficial.                                                                                                                                                                               |
| Venezuela            | 1999         | Castellano (idiomas indígenas, de uso oficial para los pueblos indígenas en todo el territorio).                                                                                                             |

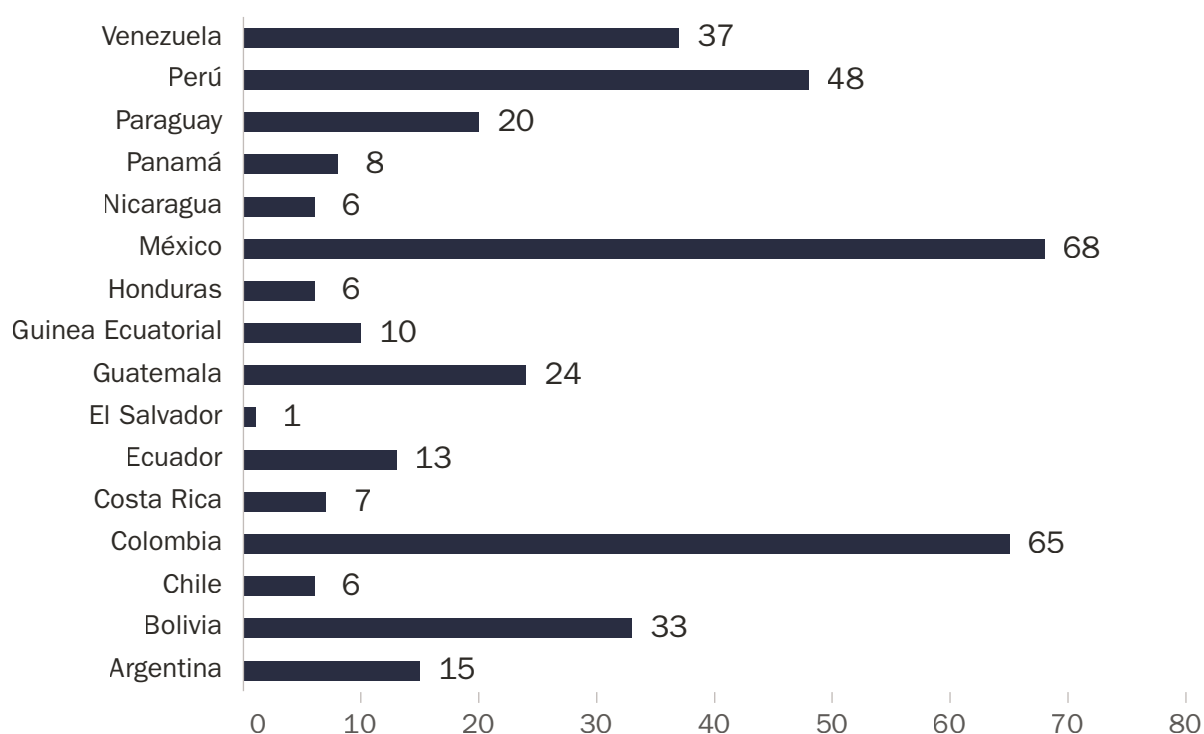
**FUENTE:** Moreno Fernández (2020).

Los países que integran la América hispana, salvo notables excepciones, conocen el uso de dos o más lenguas por parte de determinados grupos de su población y para determinados ámbitos comunitarios. Las lenguas que conviven con el español son de distintos orígenes y características: existen variedades criollas, como el palanquero colombiano, y del inglés, como el sanandresano o las que se conservan en Centroamérica; se hallan modalidades de mezcla bilingüe, como el fronterizo de Uruguay; y se conservan enclaves de otras lenguas europeas, como los del sur de Chile que aún hablan alemán o los del sur de Argentina que solo recientemente han perdido el galés. Entre todas esas variedades lingüísticas, las más importantes cuantitativamente

son las etiquetadas como lenguas indígenas u originarias (también nativas o amerindias), que a lo largo de la historia han mantenido con el español todo tipo de contactos, con muy diversas soluciones sociales y lingüísticas.

El panorama general de las lenguas indígenas en la América hispánica presenta algunas cifras reveladoras. En cuanto al número de lenguas indígenas vivas, es interesante recordar que, a finales del siglo <sup>xix</sup>, el conde de la Viñaza (1892) ofrecía referencias ciertas de más de 800 lenguas y que las cifras que se manejan para las últimas décadas del siglo <sup>xx</sup> hablan de alrededor de 725 lenguas vivas (Moreno Cabrera, 2003) y de 560 en la segunda década del siglo <sup>xxi</sup> (Banco Mundial, 2015). Las lenguas indígenas u originarias que históricamente han gozado de un mayor peso demográfico y cultural en la América hispana –importancia que mantienen hasta la actualidad– han sido el arahuaco, el náhuatl, el maya, el quechua, el aimara, el chibcha, el araucano o mapuche, y el guaraní.

GRÁFICO 1 | NÚMERO DE LENGUAS INDÍGENAS U ORIGINARIAS EN PAÍSES HISPANOHABLANTES



FUENTE: Observatorio Regional de Derechos de los Pueblos Indígenas (2020).

Uno de los factores de mayor importancia a la hora de describir las comunidades bilingües en territorios hispánicos es el de su vitalidad; esto es, el peso demográfico de los grupos originarios, indígenas o minoritarios. El análisis demolingüístico de estas comunidades etnolingüísticas encuentra enormes dificultades tanto conceptuales como metodológicas. En el plano conceptual, la dificultad básica reside en definir qué es ser indígena, por lo que suele recurrirse a la autoadscripción; en el plano metodológico, la dificultad está en identificar y posteriormente analizar la realidad étnica y lingüística específica de los grupos e individuos considerados indígenas. Estas dificultades metodológicas, se proyectan en la carencia de datos y, cuando los hay, en su falta de actualización y de comparabilidad.

**CUADRO 3 | POBLACIONES INDÍGENAS U ORIGINARIAS EN PAÍSES HISPANOHABLANTES<sup>4</sup>**

| País      | Población indígena | Población indígena que no habla español |         | Población indígena que sí habla español |          |
|-----------|--------------------|-----------------------------------------|---------|-----------------------------------------|----------|
| México    | 25.703.141         | 949.971                                 | 3,7 %   | 24.753.170                              | 96,30 %  |
| Guatemala | 7.850.978          | 1.638.867                               | 20,87 % | 6.212.111                               | 79,13 %  |
| Bolivia   | 5.841.360          | 267.729                                 | 4,58 %  | 5.573.631                               | 95,42 %  |
| Perú      | 5.659.459          | 335.576                                 | 5,93 %  | 5.323.883                               | 94,07 %  |
| Colombia  | 2.318.622          | 148.392                                 | 6,40 %  | 2.170.230                               | 93,60 %  |
| Chile     | 1.908.206          | 85.869                                  | 4,50 %  | 1.822.337                               | 95,50 %  |
| Ecuador   | 1.319.954          | 451.533                                 | 34,21 % | 868.421                                 | 65,79 %  |
| Argentina | 1.306.730          |                                         | 0,00 %  | 1.306.730                               | 100,00 % |
| Honduras  | 860.659            |                                         | 0,00 %  | 860.659                                 | 100,00 % |
| Venezuela | 813.089            | 219.534                                 | 27,00 % | 593.555                                 | 73,00 %  |

4. El informe de la OEI titulado *Lenguas indígenas de los países iberoamericanos. Datos y gráficos* (Prado, 2021) ofrece cifras diferentes a las aquí presentadas. Esto se debe, en gran parte, a que las fechas de las fuentes manejadas son dispares (en su mayoría, entre 2009 y 2015). Las fuentes allí citadas más cercanas en el tiempo (2021) son Ethnologue y Wikipedia, pero no son fuentes primarias. En nuestro análisis, se ha accedido a informaciones, proyecciones y estimaciones a partir de los censos nacionales más recientes.

| País              | Población indígena | Población indígena que no habla español |         | Población indígena que sí habla español |          |
|-------------------|--------------------|-----------------------------------------|---------|-----------------------------------------|----------|
| Panamá            | 698.114            | 119.205                                 | 17,08 % | 578.909                                 | 82,92 %  |
| Nicaragua         | 616.514            | 69.667                                  | 11,30 % | 546.847                                 | 88,70 %  |
| Guinea Ecuatorial | 415.000            |                                         | 0,00 %  | 415.000                                 | 100,00 % |
| Paraguay          | 162.445            | 19.253                                  | 11,85 % | 143.192                                 | 88,15 %  |
| Costa Rica        | 69.025             | 897                                     | 1,30 %  | 68.128                                  | 98,70 %  |
| El Salvador       | 12.799             |                                         | 0,00 %  | 12.799                                  | 100,00 % |
| Total             | 55.556.095         | 4.306.493                               | 7,75 %  | 51.249.602                              | 92,25 %  |

**FUENTE:** elaboración propia. Se incluyen todas aquellas poblaciones reconocidas oficial o formalmente, con indicación de hablantes monolingües y hablantes de español, ordenados por volumen de población indígena.

Fuentes e información adicional del cuadro 3

Argentina

Datos de *Censo Nacional de Población y Viviendas 2022* (INDEC, 2023). El censo pregunta en español qué lenguas se hablan o entienden. No hay constancia de que los miembros de pueblos originarios no hablen español, con la excepción tal vez de los niños en edad preescolar. La castellanización está generalizada principalmente a través de la escolarización, aunque existen iniciativas para potenciar la educación bilingüe (Acuña, 2009).

Bolivia

Datos del Instituto Nacional de Estadística de Bolivia (2020) y del Instituto Plurinacional de Estudios de Lenguas y Culturas (IPELC, 2024). El porcentaje de población indígena u originaria es del 48 %.

Chile

Datos del Instituto Nacional de Estadística de Chile (2017a), del Ministerio de Desarrollo Social y Familia (2013 y 2017) y del Centro de Estudios Públicos (CEP, 2016) para estimación de bilingües y monolingües. Del número total de personas censadas que se declaraban indígenas en 2017 (2.144.479), más de 146.000 no pertenecían a un pueblo originario reconocido por el Gobierno de Chile.

Colombia

Datos del Departamento Administrativo Nacional de Estadística (DANE, 2018). El porcentaje de población indígena en 2018 (4,4 %) se aplica a 2024. Se estima que no hablan español los indígenas no alfabetizados.

Costa Rica

Datos del Instituto Nacional de Estadística y Censos (INEC, 2024). El porcentaje de población indígena es del 1,3 %. El analfabetismo entre la población indígenas es del 4,8 %.

### Ecuador

Datos del Instituto Nacional de Estadísticas y Censos (INEC, 2022). Sobre nivel de alfabetización, se consideran los datos de UNICEF (2014). Acceso a datos mediante el sistema REDATAM.

### El Salvador

Datos de la Oficina Nacional de Estadística y Censos sobre el VI Censo de Población y Vivienda 2007 (ONEC, 2024). El porcentaje de población indígena u originaria es del 0,2 %. No hay constancia de población monolingüe en lengua indígena.

### Guatemala

Datos del censo de 2018 y estimaciones del Instituto Nacional de Estadística de Guatemala (2024). El porcentaje de población indígena u originaria es del 44 %. Acceso mediante el sistema REDATAM: Sabe hablar otro idioma; Otro idioma que habla.

### Guinea Ecuatorial

Datos del Instituto Nacional de Estadística de Guinea Ecuatorial (INEGE, 2023). El porcentaje de población alfabetizada, según cifras oficiales es del 94,8 %. La alfabetización se desarrolla en español. Comunicaciones personales estiman que no existe población que no hable español en absoluto.

### Honduras

Datos del Instituto Nacional de Estadística de Honduras (2013). Estimaciones sobre el censo de 2013. Acceso a base de datos en línea BASEINE del Instituto Nacional de Estadística de Honduras (2024). La población negra de habla inglesa es del 0,15 % aproximadamente. No hay constancia de monolingües en una lengua que no sea el español.

### México

Datos del Instituto Nacional de Estadística y Geografía (INEGI, 2022). Se ha proyectado a 2024 el 19,4 % de población indígena recontado en 2020. La proporción de personas que no hablan español está calculada sobre el total de personas autoadsritas como indígenas, no a partir de los hablantes de lengua indígena que no hablan español, cuya proporción sería del 3,7% en proyección de los datos de 2020. Esta información puede ampliarse en el artículo sobre el español en México incluido en esta publicación. De las lenguas indígenas que se hablan de forma predominante, destacan el náhuatl, hablado en 15 estados; seguido por el maya, que predomina en Campeche, Quintana Roo y Yucatán; el huichol, en Jalisco y Zacatecas; y el mayo, en Sinaloa y Sonora.

### Nicaragua

Datos del Instituto Nacional de Información de Desarrollo (INIDE, 2005) con proyección de un incremento poblacional del 3 % anual. El porcentaje de población indígena es del 8,63 %, de la que el 90,2 % habla lengua indígena, pero ello no supone que no hable español.

### Panamá

Datos del Instituto Nacional de Estadística y Censo a partir del censo de 2023 (INEC, 2024). El porcentaje de población que se declara indígena es del 16,6 %. El porcentaje de indígenas alfabetizados es del 71,6 %. Información del Consejo del Sector Privado para la Asistencia Educacional (COSPAE, 2002; 2024).

### Paraguay

Datos del Instituto Nacional de Estadística de Paraguay (2022; 2024). Acceso a datos mediante el sistema REDATAM. Debe distinguirse la población indígena hablante de guaraní de la población no indígena hablante de guaraní yopará.

### Perú

Datos del Ministerio de Cultura de Perú (2017) y del Instituto de Estudios Peruanos (IEP, 2023). El porcentaje de población indígena u originaria se aplica sobre la población de 2024. Se estima que no hablan español las personas analfabetas: 6,7 % de lengua materna indígena.

### Venezuela

Datos del Instituto Nacional de Estadística de la República Bolivariana de Venezuela (2015). El porcentaje de población indígena u originaria se estima en un 3 % en 2022 y se proyecta a 2024. Acceso a datos mediante sistema REDATAM. Se estima que hablan español las personas indígenas alfabetizadas (70,8 %).

Las cifras correspondientes a la población indígena que se expresa en español se han llevado al cuadro 1, bien en forma de hablantes nativos de español, bien como hablantes del grupo de competencia limitada, según los criterios explicados más arriba. A este respecto, se debe insistir en la idea de que el dominio nativo de una lengua no impide el dominio nativo de una segunda y hasta una tercera lengua. Del mismo modo, la competencia limitada en una o varias lenguas suele complementarse con el dominio nativo de otra u otras.

El cuadro 3 presenta cifras referidas al volumen de las poblaciones indígenas u originarias, de la población de ese origen que no habla español y de la población que sí es capaz de hablar español. Los criterios generales que se han seguido para la recolección y análisis de estos datos son los siguientes:

1. Se considera población indígena u originaria aquella que las autoridades de cada país consideran población indígena u originaria, generalmente basándose en la autoadscripción de los integrantes de estas poblaciones.
2. Se consideran hablantes de español aquellos miembros de estas comunidades que se reconocen o identifican como tales en los datos oficiales con que cuenta cada país; cuando no existen datos específicos sobre este conocimiento, se consideran hablantes de español aquellos individuos alfabetizados y escolarizados en español, tanto en la enseñanza secundaria como en la universitaria, o que han migrado a las ciudades, según los recuentos oficiales de cada país.

3. No se consideran hablantes de español aquellos individuos, miembros de una comunidad indígena con vitalidad lingüística, que no han sido alfabetizados o que no han alcanzado el nivel de la escuela secundaria.

Somos conscientes de que estos criterios pueden favorecer las cifras de los no hablantes de español, pero, de no haberlos seguido, estos grupos habrían resultado aún más minorizados en su representación cuantitativa, y se podría haber ocultado una realidad cultural y humanamente significativa, por pequeños que sean sus guarismos globales.

Los datos incluidos en el cuadro 3 proceden de los censos y las encuestas de hogares realizados en los últimos años en los países hispanohablantes y reflejan algunas cifras interesantes. Así, la población reconocida oficialmente como indígena es superior a los 55 millones de personas. Como es lógico, las cantidades referidas a cada país muestran una notable disparidad: desde la ausencia de población censada como indígena en Cuba, la República Dominicana, Puerto Rico o Uruguay, hasta la existencia de más de 5 millones de indígenas en Perú, Bolivia o Guatemala, y de más de 25 millones en México. La población indígena en su conjunto supone un 11,5 % de la población total.

Dos décadas atrás, la información derivada de los censos realizados entre 2000 y 2004 indicaba que, en los países hispanohablantes de América, la población indígena se situaba alrededor de los 35 millones, lo que suponía en torno a un 11 % de la población total. A comienzos de siglo XXI, el 65 % de los indígenas eran capaces de comunicarse en su lengua autóctona, de los que un 15 % eran monolingües, por lo que cerca de un 35 % de los indígenas eran monolingües en español. Los datos actuales indican, por un lado, un mantenimiento del porcentaje general de población indígena y, por otro, un crecimiento de la población capaz de comunicarse en español (del 85 % al 92,25 %), mientras que la proporción de monolingües en sus respectivas lenguas ha pasado del 15 % al 7,75 %.

La distribución geográfica de la población indígena o hablante de una lengua nativa que no sea el español permite apreciar aspectos muy significativos

desde el punto de vista lingüístico y social. Como se observa en el mapa 1, son muy pocos los países hispanohablantes en los que no existe algún contingente, de mayor o menor tamaño, de hablantes nativos de lenguas que no sean el español. En algunos países, como México, Perú o Bolivia, la presencia indígena u originaria no es solo intensa, sino también extensa, teniendo además en cuenta que el mapa 1 solo refleja las mayores concentraciones de hablantes de otras lenguas. Esto supone que, en la mayoría de los países hispánicos, existe población que vive cotidianamente el bilingüismo o el multilingüismo o que, como mínimo, cuenta con experiencia vital sobre lo que supone una realidad bilingüe o multilingüe, con toda su diversidad de posibles manifestaciones. En unos casos, las lenguas nativas pueden escucharse en los medios de comunicación social y encontrarse en las redes sociales; en otros casos, las lenguas nativas quedan reducidas a un espacio social de escasa trascendencia. Asimismo, el mapa 1 permite vislumbrar las consecuencias lingüísticas que en muchos lugares puede tener el habla bilingüe, bien en forma de transferencias, bien en forma de alternancia de lenguas, bien en forma de mezcla bilingüe.

MAPA 1

ÁREAS DE MAYOR PRESENCIA DE POBLACIÓN INDÍGENA O HABLANTE DE UNA LENGUA NATIVA QUE NO ES EL ESPAÑOL



**FUENTE:** elaboración propia.

En lo que se refiere a España, el Instituto Nacional de Estadística (INE, 2021)<sup>5</sup> ofrece directamente la cantidad de personas según su nivel de conocimiento de las lenguas más frecuentes, entre las que se incluyen, además del español o castellano, el catalán (mallorquín, menorquín, ibicenco, formenterense), el valenciano, el gallego, el euskera, el inglés, el francés,

5. Véase el artículo «El español en España», incluido en este mismo volumen.

el árabe, el rumano y el italiano<sup>6</sup>. La población monolingüe o bilingüe que en España no habla español o castellano se limita a una proporción del 0,6 %, extranjeros en su mayoría. De igual modo, resulta interesante destacar que el número de personas que hablaba bien dos lenguas en 2021 era de 14.721.371 y que eran 3.224.263 personas las que declaraban hablar bien tres lenguas. Así pues, un 31,8 % de la población española experimenta de primera mano la coexistencia de lenguas; a esta población podría sumarse la monolingüe en español que frecuenta o reside en los entornos bilingües de España.

### 2.3 | El español en los países no hispanohablantes

La historia de las comunidades hispanohablantes es larga en el tiempo y ancha en su geografía. Fuera del espacio hispánico propiamente dicho, los países o territorios en los que el español tiene una presencia suficientemente estimable supera la treintena. 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 casos, la presencia hispanohablante tiene su origen en movimientos migratorios más recientes, como en Canadá, Australia, el Reino Unido, Suiza o la propia Unión Europea. En otros casos, en fin, el conocimiento de español puede deberse a la contigüidad geográfica y a su efecto transfronterizo, a la intensidad de los intercambios socioeconómicos entre países limítrofes (Belice, Brasil, Andorra...) o a diverso tipo de razones históricas, como ocurre con el pueblo saharauí.

Cuadro 4 · Página siguiente >

---

**6.** Las denominaciones presentadas (glotónimos) son las que figuran en el INE para indicar el nivel de conocimiento de lenguas (INE, 2021). Véase: <https://www.ine.es/jaxi/Tabla.htm?tpx=55481&L=0>. Mantenemos esas denominaciones (p. e. catalán diferenciado de valenciano) para no modificar las cantidades aportadas por el INE.

| CUADRO 4   POBLACIÓN HISPANOHABLANTE EN PAÍSES Y TERRITORIOS NO HISPÁNICOS |                            |                                  |                                          |
|----------------------------------------------------------------------------|----------------------------|----------------------------------|------------------------------------------|
| País                                                                       | Grupo Dominio Nativo (GDN) | Grupo Competencia Limitada (GCL) | Total de usuarios potenciales de español |
| Estados Unidos                                                             | 41.899.801                 | 15.354.000                       | 57.253.801                               |
| UE-27                                                                      | 1.267.000                  | 28.047.000                       | 29.314.000                               |
| Reino Unido                                                                | 216.518                    | 4.882.741                        | 5.099.259                                |
| Marruecos                                                                  | 12.774                     | 1.741.711                        | 1.754.485                                |
| Canadá                                                                     | 707.400                    | 664.959                          | 1.372.359                                |
| Brasil                                                                     | 522.443                    | 170.444                          | 692.887                                  |
| Filipinas                                                                  | 4.584                      | 562.337                          | 566.921                                  |
| Suiza                                                                      | 212.970                    | 343.161                          | 556.131                                  |
| Australia                                                                  | 175.491                    | 384.000                          | 559.491                                  |
| Sáhara Occidental                                                          |                            | 423.739                          | 423.739                                  |
| Belice                                                                     | 58.893                     | 157.363                          | 216.256                                  |
| Antillas Neerlandesas                                                      | 46.621                     | 156.715                          | 203.339                                  |
| Argelia                                                                    | 1.149                      | 173.600                          | 174.749                                  |
| Japón                                                                      | 131.000                    | 29.000                           | 160.000                                  |
| Israel                                                                     | 104.000                    | 45.000                           | 149.000                                  |
| Aruba                                                                      | 14.737                     | 74.650                           | 89.387                                   |
| Andorra                                                                    | 34.132                     | 37.545                           | 71.677                                   |
| Trinidad y Tobago                                                          | 4.000                      | 66.401                           | 70.401                                   |
| Noruega                                                                    | 13.000                     | 24.000                           | 37.000                                   |
| Guam e Islas Marianas del Norte (Estados Unidos)                           | 1.309                      | 30.924                           | 32.233                                   |
| Rusia                                                                      | 28.924                     |                                  | 28.924                                   |
| Nueva Zelanda                                                              | 22.000                     |                                  | 22.000                                   |
| Turquía                                                                    | 5.460                      | 16.200                           | 21.660                                   |
| Islas Vírgenes (Estados Unidos)                                            | 16.788                     |                                  | 16.788                                   |
| China                                                                      | 15.130                     |                                  | 15.130                                   |

| País         | Grupo Dominio Nativo (GDN) | Grupo Competencia Limitada (GCL) | Total de usuarios potenciales de español |
|--------------|----------------------------|----------------------------------|------------------------------------------|
| Jamaica      | 8.500                      |                                  | 8.500                                    |
| Bulgaria     |                            | 5.000                            | 5.000                                    |
| India        | 4.855                      |                                  | 4.855                                    |
| Otros países | 160.323                    |                                  | 160.323                                  |
| Total        | 45.689.802                 | 53.390.493                       | 99.080.295                               |

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

Fuentes e información adicional del cuadro 4

Andorra

Para hablantes nativos, sobre la población total de Andorra en 2024 (Govern d’Andorra, 2022, p. 10), se ha aplicado un porcentaje del 40 %, que corresponde a los hablantes que declaran el español como su lengua materna. Para hablantes con competencia limitada, sobre la población total de Andorra en 2024 (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.

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

Los hablantes nativos de Curazao son lo que figuran en el censo como hablantes de español en el hogar (Central Bureau of Statistics Curaçao, 2023). Los hablantes con competencia limitada son los que el censo da como hablantes de papiamento. Los hablantes de español y de papiamento en los demás territorios se calculan sobre los porcentajes que ofrece Statista (2024), 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 %. Como se ha mencionado, los hablantes de papiamento se consideran hablantes con competencia limitada.

Argelia

En los hablantes nativos, incluimos a los españoles residentes en Argelia (INE, 2023). Entre los hablantes con competencia limitada, incluimos a los refugiados saharauis en territorio argelino (UNHCR), donde el español es cooficial y tiene una presencia notable en la enseñanza.

Aruba

Datos del Central Bureau of Statistics Aruba (2023).

Australia

A partir del censo de 2021 (Australian Bureau of Statistics, s. f.) con proyección de crecimiento del 2,4 % para 2024. Para los hablantes de competencia limitada, la estimación realizada por Moreno Fernández y Otero Roth (2006, p. 39) se ha actualizado. Agradecemos la ayuda de Coral Martínez Íscar, directora del Instituto Cervantes de Sídney.

### Belice

Los hablantes nativos se estiman a partir de población extranjera de origen hispanohablante. Según el censo de 2022 de Statistical Institute of Belize (SIB) (2022), el 52 % de la población habla español. Los hablantes de competencia limitada se calculan restando los nativos a este último porcentaje.

### Brasil

Los hablantes nativos se estiman a partir del Registro Nacional Migratorio (inmigrantes internacionales registrados por la Policía Federal) (UNICAMP, 2024). Los hablantes con competencia limitada se estiman a partir del registro de extranjeros residentes y fronterizos del Instituto Brasileiro de Geografia e Estatística (IBGE, 2024). Es posible un solapamiento parcial de la cifra de inmigrantes y de extranjeros residentes, pero no se ha encontrado información metodológica que lo confirme. Los hablantes de competencia limitada incluyen los miembros de la comunidad sefardí de São Paulo (alrededor de 35.000). Dado que podrían incluirse en el GALE, se excluyen los 5,5 millones de jóvenes que pueden mantener una conversación en español, según la Fundación Siglo para el Turismo y las Artes de Castilla y León (2013, p. 23).

### Canadá

A partir de datos de hablantes de español en casa: 1,76 % (Statistics Canada, 2022). Porcentaje proyectado a la población de 2024 (40.913.162). Para los hablantes de competencia limitada, al porcentaje de población de origen hispano (3,3 %) se le ha restado el porcentaje de hablantes de español en casa; asimismo, se han sumado los miembros de las comunidades sefardíes.

### China

En 2020, se tenían registrados 10.455 inmigrantes peruanos (ONU, 2020) y 4.675 españoles (INE, 2023). Los datos son escasos y de difícil acceso.

### Estados Unidos

A partir de datos de la Oficina del Censo de los Estados Unidos (2021a y 2021b). Población hispana y no hispana que hablaba español en el hogar en 2021, proyectada a 2024 (4,6 %), excluidos menores de 5 años. Según la Oficina del Censo de los Estados Unidos (2021a y 2021b), la población total del país en marzo de 2024 sería de 341.203.592 individuos de los que el 18,8 % se definían como hispanos en 2021. Entre los hispanos, el 67,6 % hablaba español en el hogar. El GCL ofrece una cifra redondeada por ser una estimación que incluye a los hispanos que potencialmente tendrían una competencia limitada en español y a personas que han estudiado español en distintos niveles. Se excluyen los 7,7 millones de inmigrantes no autorizados procedentes de México, Centroamérica, América del Sur y el Caribe, de acuerdo con estimaciones de 2017 (González Barrera, Krogstad y Noe Bustamante, 2020).

### Filipinas

Para los hablantes de dominio nativo, número de residentes españoles en 2023 (INE, 2023). Para los hablantes de competencia limitada, se proyecta crecimiento de población ofrecido por la Oficina Estadística de Filipinas (2010) a 2024. En este cálculo se incluyen los hablantes del criollo chabacano.

### Guam e Islas Marianas del Norte

Los hablantes nativos son extranjeros de origen hispanohablante. Dato proyectado a 2024. El grupo de competencia limitada en español incluye a los hablantes de chamorro, (Lewis, Simons y Fenning, 2016). Se estima que un 17,8 % de la población puede hablar chamorro (The Bureau of Statistics and Plans, 2020 y Oficina del Censo de los Estados Unidos, 2022b).

**India**

Cifra obtenida de la suma de 900 argentinos (según datos proporcionados por Pew Research Interactive Global Migrant); 2.055 mexicanos (según datos proporcionados por el Instituto de los Mexicanos en el Exterior); 670 colombianos (según datos proporcionados por Registro Electoral de la Embajada de Colombia en India); 20 chilenos (Instituto Nacional de Estadísticas de Chile, 2017b); 311 venezolanos (seguidores de «Venezolanos en India» en Facebook); y 899 españoles (INE, 2023).

**Islas Vírgenes**

Del total de la población (98.299 en 2024) habla español el 17,2 % (Oficina del Censo de los Estados Unidos, 2023).

**Israel**

Cifra estimada a partir de datos de la Oficina del Censo de Israel (2022). El grupo de competencia limitada incluye hablantes de judeoespañol. No se tienen en cuenta las pocas decenas de hispano-hablantes que pueda haber en territorio palestino.

**Jamaica**

El Joshua Project (2024) afirma que hay 8.500 hispanos en Jamaica.

**Japón**

Datos estimados a partir de Japanese Government Statistics (2023).

**Marruecos**

Hablantes nativos a partir de población española en Marruecos (INE, 2023). Hablantes de competencia limitada, a partir de Fernández Vítors (2019), proyectando a 2024 el porcentaje de hablantes de español.

**Reino Unido**

Proyección a 2024 del censo de 2021 (Office for National Statistics, 2022). Para los hablantes de competencia limitada, se ha hecho una estimación para 2024 a partir del porcentaje de personas que afirman poder mantener una conversación en español (Eurostat, 2012a).

**Rusia**

Estimación basada en seguidores de la página de Facebook «Latinos en Rusia / Latinos en Moscú (26.700), consultada en enero de 2024. Esta información sería más adecuada para un estudio de demografía digital, pero se aporta aquí en ausencia de otros datos. Hay 2.224 españoles en Rusia (INE, 2023).

**Sáhara Occidental**

Población adscrita al pueblo saharauí occidental, menos la residente en los campos de refugiados de Argelia. Véase: <https://saharaoccidental.es/>.

**Suiza**

A partir de datos de Kabatek (2022). Datos de población de 2020 (8.688.000) y de 2023 (9.006.664) de la Oficina Federal de Estadísticas de Suiza (<https://www.bfs.admin.ch/bfs/en/home.html>).

**Trinidad y Tobago**

Hablantes con competencia limitada a partir de estimación del 5 % de hablantes de español sobre la población total del país en 2023 (Grau Perejoan y Gea Monera, 2006, p. 209).

### Turquía

Hablantes nativos de origen español (INE, 2023). El grupo de competencia limitada 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

Cifras basadas en estimaciones. El 7 % de los ciudadanos de la UE tiene el español como lengua extranjera (Eurostat, 2024a), excluido el Reino Unido. Se excluyen también los más de seis millones de estudiantes de español que actualmente puede haber en la UE.

### Otros países

La cantidad anotada corresponde a migrantes procedentes de países hispanohablantes (ONU, 2020).

El cuadro 4 presenta las cifras de hablantes de español con dominio nativo y con competencia limitada en los países y territorios no hispanohablantes. La diversidad de procesos históricos y condiciones geográficas que han dado lugar a la formación de esas comunidades justifica que en unos países solo se documenten hablantes nativos (normalmente inmigrantes de origen hispanohablante) o hablantes con competencia limitada (normalmente vinculados a una anterior presencia hispánica o a la descendencia de hispanohablantes). En la mayoría de los países, sin embargo, es posible encontrar tanto hablantes nativos como hablantes con competencia limitada. Los criterios que se han seguido para la confección de este cuadro se resumen en los siguientes puntos:

1. La población está actualizada a 2024 de acuerdo con la información ofrecida por las respectivas oficinas del censo o institutos de estadísticas. Cuando no se cuenta con cifras oficiales para 2024, se adopta la proyección realizada por los mismos organismos oficiales o por organismos internacionales. En algunos casos, no se han obtenido datos suficientemente fiables, aunque se ha tenido en cuenta que en los años de la pandemia los movimientos poblacionales se redujeron de manera drástica.

2. La cantidad de hablantes integrantes del Grupo de Dominio Nativo se calcula a partir de la población inmigrante de primera generación de origen hispánico. También en este caso, el dominio nativo de la lengua española no es óbice para que se tenga un dominio nativo de otra u otras lenguas.
3. 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.

De los datos que nos ofrece el cuadro 4, merecen destacarse algunas realidades significativas. La primera es el volumen de hablantes de español fuera del espacio hispánico. Frente a la supuesta universalidad del inglés, es habitual limitar el alcance del español a los países hispanohablantes, pero lo cierto es que hay una significativa población fuera de esos países que, por distintos motivos, usan y mantienen el español. La cifra ronda los 100 millones de usuarios potenciales de español –aproximadamente la mitad de ellos hablantes nativos y la mitad, hablantes con competencia limitada– y debe valorarse en su justa medida. Junto al volumen total, el dato más llamativo que este cuadro revela es el peso demográfico de la población hispanohablante en Estados Unidos y en la Unión Europea. En el primer caso, se trata de una realidad bien conocida y analizada; en el segundo, se trata de una realidad que hasta fechas recientes no ha recibido la atención que merece. Si bien el peso del español en Europa recae mayormente en España, no puede ignorarse la dimensión hispana de países como Francia, Italia, Suiza o Alemania, principalmente por razones migratorias (véase el Observatorio del Español en Europa en <https://ose-hcias.de/es/>). La movilidad de los hispanohablantes es asimismo causa directa de la presencia del español en espacios tan distintos y distantes como China o Jamaica. Y merecen una mención particular situaciones como la del español en el Sáhara, donde es lengua cooficial; la de Guam y las Islas Marianas, donde aún se usa una mezcla bilingüe con un fuerte componente hispánico; o la de los sefardíes que mantienen vivo el español en sociedades como la

turca, aunque su conocimiento y uso está en franca decadencia entre los menores de 40 años<sup>7</sup>.

Finalmente, como novedad metodológica, este informe tiene en cuenta la población emigrante de origen hispanohablante registrada en países diferentes de los presentados en el cuadro 4, según datos procedentes de los registros migratorios de la ONU.

## 2.4 | Aprendices de español

El español es una de las lenguas extranjeras más estudiadas en el mundo, con más de 24 millones de alumnos en la estimación realizada para 2024. Esta cifra ha venido creciendo constante pero lentamente en los últimos diez años a un ritmo del 2 % anual. El cuadro 5 desglosa el número de aprendices en diferentes niveles educativos, desde la enseñanza primaria, secundaria y universitaria hasta otros espacios de enseñanza no reglada como los centros de idiomas o el Instituto Cervantes en sus distintas sedes.

Cuadro 5 · Página siguiente >

---

**7.** Además de las atendidas 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 usos lingüísticos. 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. No se identifican específicamente las comunidades sefardíes en países o territorios que cuentan con estadísticas de hablantes de español en casa.

**CUADRO 5 | NÚMERO APROXIMADO DE ESTUDIANTES DE ESPAÑOL EN EL MUNDO<sup>8</sup>**

| País            | Enseñanza primaria, secundaria y formación profesional | Enseñanza universitaria | Otros   | Centros del Instituto Cervantes y AVE 2022-2023 | Total     |
|-----------------|--------------------------------------------------------|-------------------------|---------|-------------------------------------------------|-----------|
| Estados Unidos  | 8.003.000                                              | 584.453                 |         | 3.697                                           | 8.591.150 |
| Brasil          | 4.001.530                                              | 43.517                  |         | 3.291                                           | 4.048.338 |
| Francia         | 3.339.793                                              | 244.080                 | 5.978   | 2.997                                           | 3.592.848 |
| Reino Unido     | 2.017.796                                              | 21.505                  |         | 2.915                                           | 2.042.216 |
| Italia          | 866.026                                                | 61.000                  | 41.685  | 5.830                                           | 974.541   |
| Costa de Marfil | 795.488                                                | 1.752                   | 855     |                                                 | 798.095   |
| Alemania        | 549.911                                                | 52.563                  | 178.255 | 5.765                                           | 786.494   |
| Benín           |                                                        |                         |         |                                                 | 412.515   |
| Camerún         | 400.000                                                | 3.000                   |         |                                                 | 403.000   |
| Senegal         |                                                        |                         |         | 66                                              | 356.066   |
| Suecia          | 255.690                                                | 7.612                   | 7.455   | 329                                             | 271.086   |
| Polonia         | 222.404                                                | 18.680                  | 3.928   | 2.101                                           | 247.113   |
| Gabón           | 200.000                                                |                         |         |                                                 | 200.000   |
| Canadá          | 152.501                                                | 11.221                  |         |                                                 | 163.722   |
| Rusia           | 100.446                                                | 28.347                  | 4.418   | 1.219                                           | 134.430   |
| Portugal        | 95.506                                                 | 5.600                   |         | 672                                             | 101.778   |
| Noruega         | 95.301                                                 | 587                     |         |                                                 | 95.888    |

8. 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. El caso de Guinea Ecuatorial es sumamente complicado. En informes anteriores se ha considerado la población estudiantil de Guinea como aprendiz de español (Serrano, 2014; Instituto Cervantes, 2023). Consideramos, sin embargo, más coherente con el tratamiento dado a los datos para otros países hispanohablantes no incluir esta población en el grupo de aprendices, teniendo en cuenta además la proporción de hablantes que consideramos de dominio nativo. Dada la peculiaridad de la situación idiomática de Guinea, aceptamos la cifra que proponen Bibang Oyee y Larre Muñoz en esta publicación y que, al no superar los 30.000 estudiantes, no se incluye en este cuadro.

| País            | Enseñanza primaria, secundaria y formación profesional | Enseñanza universitaria | Otros  | Centros del Instituto Cervantes y AVE 2022-2023 | Total      |
|-----------------|--------------------------------------------------------|-------------------------|--------|-------------------------------------------------|------------|
| Argelia         | 40.290                                                 | 45.409                  |        | 2.980                                           | 88.679     |
| Austria         | 56.060                                                 | 4.000                   | 9.131  | 724                                             | 69.915     |
| Australia       | 52.726                                                 | 9.300                   |        | 578                                             | 62.604     |
| Marruecos       | 43.261                                                 | 1.772                   | 4.839  | 12.386                                          | 62.258     |
| China           | 8.874                                                  | 34.823                  | 8.866  | 1.335                                           | 53.898     |
| Irlanda         | 45.589                                                 | 3.600                   |        | 1.115                                           | 50.304     |
| República Checa | 39.183                                                 | 8.317                   |        | 1.191                                           | 48.691     |
| Bélgica         | 14.159                                                 | 7.799                   | 24.338 | 726                                             | 47.022     |
| Países Bajos    | 32.014                                                 | 14.160                  |        | 459                                             | 46.633     |
| India           | 25.000                                                 | 7.681                   | 11.001 | 2.567                                           | 46.249     |
| Túnez           | 39.500                                                 |                         |        | 1.128                                           | 40.628     |
| Nueva Zelanda   | 34.275                                                 | 2.098                   |        |                                                 | 36.373     |
| Dinamarca       | 31.793                                                 | 248                     |        |                                                 | 32.041     |
| Otros países    |                                                        |                         |        |                                                 | 304.238    |
| Total           |                                                        |                         |        |                                                 | 24.208.813 |

**FUENTE:** elaboración propia. Si en alguna celda no aparecen cifras, se debe a la ausencia de datos específicos o a que los datos se ofrecen agregados.

Fuentes e información adicional del cuadro 5

Alemania

Estudiantes en enseñanza primaria, secundaria y formación profesional, datos de la Oficina Federal de Estadística correspondientes al año escolar 2022-23 (Statistisches Bundesamt, 2024). 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.*, 2022). Estudiantes en enseñanza no reglada inscritos en universidades populares (*Volkshochschulen*) en 2019 (Loureda Lamas *et al.*, 2022). De la cifra global proporcionada por Loureda Lamas *et al.* (2022), se ha restado 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). Los 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 las de 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, 2018, 2022). Para 2022, esta fuente reporta 35.306 estudiantes de español en primaria y 17.420 en secundaria durante el ciclo escolar 2021. 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 2021-2022 (Statistik Austria, 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 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 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 2021-2022. 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 en 2019, al que se ha aplicado un factor multiplicador uniforme calculado a la baja. Estudiantes en enseñanza no reglada, el MEFP (2022) reporta 5.270 alumnos de español en la comunidad francesa, 18.958 en la comunidad flamenca y 110 en la comunidad germanófona durante el ciclo escolar 2019-2020.

#### Benín

A partir de Serrano Avilés (2014).

#### Brasil

Estudiantes en enseñanza primaria, secundaria y formación profesional en 2021 según el Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP, 2020), por gentileza de la Consejería de Educación de España en Brasil. Estudiantes universitarios según INEP (2020).

#### 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. Agradecemos la ayuda de Gonzalo Leiva Fidalgo, segunda jefatura de la Embajada de España en Yaundé, Camerún.

#### Canadá

A falta de datos representativos a escala nacional (véase 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 2021-2022 (Statistics Canada, 2023a y 2023b). El número total de estudiantes será muy superior al reportado.

### China

Enseñanza primaria, secundaria y formación profesional, a partir de «El español: una lengua viva. Informe 2023» (Instituto Cervantes, 2023). Estudiantes universitarios, a partir de González Puy (2018). Estudiantes en enseñanza no reglada, número de alumnos chinos inscritos en universidades españolas en el año escolar 2016-2017.

### Costa de Marfil

A partir de datos suministrados en febrero de 2024 por la Consejería de Educación de la Embajada de España en Marruecos (acreditada en Costa de Marfil), consultados al Ministerio de Educación Secundaria y Alfabetización. Estudiantes en enseñanza primaria, secundaria y formación profesional, 528.248 en primaria y 266.840 en secundaria durante el ciclo escolar 2021-2022. El número de estudiantes en enseñanza no reglada corresponde a profesores que han recibido formación de español gracias al acuerdo entre la AECID, el Instituto Cervantes y la Embajada de España de 2015 a 2023. Agradecemos la ayuda de Alberto Herranz Martín, secretario general de la Consejería de Educación de la Embajada de España en Marruecos.

### 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, 2021 y 2023). Para los estudiantes universitarios, se toman datos de Lusin *et al.* (2023, p. 6).

### Francia

Estudiantes en enseñanza primaria, secundaria y formación profesional, datos de Eurostat (2024b) correspondientes al año escolar 2021-2022. Estudiantes universitarios, estimación de Tano (2023, p. 157). Estudiantes en enseñanza no reglada, datos a partir de «El español: una lengua viva. Informe 2023» (Instituto Cervantes, 2023).

### Gabón

Estimación para estudiantes de enseñanza secundaria suministrada en febrero de 2024 por la Embajada de España. Agradecemos la ayuda de Lucía Lecumberri Galofré, segunda jefatura de la Embajada de España en Gabón.

### India

Estimación a la baja a partir de cifras de Pujol Riembau (2020).

### Irlanda

Estudiantes en enseñanza primaria, secundaria y formación profesional, datos de Eurostat (2024b) correspondientes al año escolar 2021-2022. 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 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 2021-2022. 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 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 (en prensa). Para evitar una doble contabilización, se ha restado de la cifra global proporcionada por Bonomi, Calvi y Uberti-Bona (en prensa) el número de alumnos del Instituto Cervantes en Italia para el curso 2020-2021.

#### Marruecos

Estudiantes en enseñanza primaria, secundaria y formación profesional, a partir de «El español: una lengua viva. Informe 2023» (Instituto Cervantes, 2023). Estudiantes universitarios, datos del MEFP (2020). Estudiantes en enseñanza no reglada, datos suministrados en febrero de 2024 por la Consejería de Educación de la Embajada de España en Marruecos, correspondientes al alumnado en centros de titularidad del Estado español dependientes de la Consejería de Educación. Agradecemos la ayuda de Alberto Herranz Martín, secretario general de la Consejería de Educación de la Embajada de España en Marruecos.

#### Noruega

Estudiantes en enseñanza primaria, secundaria y formación profesional, datos de Eurostat (2024b) correspondientes al año escolar 2021-2022. 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 en 2019, al que se ha aplicado un factor multiplicador uniforme calculado a la baja.

#### Nueva Zelanda

A partir de datos del MEFP (2020).

#### Países Bajos

Estudiantes en enseñanza primaria, secundaria y formación profesional, datos de Eurostat (2024b) correspondientes al año escolar 2021-2022. 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 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 2021-2022. Estudiantes universitarios, datos a partir de Zielinski *et al.* (en prensa). Estudiantes en enseñanza no reglada, a partir de «El español: una lengua viva. Informe 2023» (Instituto Cervantes, 2023).

#### Portugal

Estudiantes en enseñanza primaria, secundaria y formación profesional, datos de Eurostat (2024b) correspondientes al año escolar 2021-2022. Estudiantes universitarios, dato a partir de Santos Róvira y Serrano Lucas (2022).

#### Reino Unido

Estudiantes en enseñanza primaria, secundaria y formación profesional, datos de Eurostat (2024b) correspondientes al año escolar 2018-2019. 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 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 2021-2022. 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 en 2019, al que se ha aplicado un factor multiplicador uniforme calculado a la baja.

### Rusia

A partir del MEFP (2020, p. 670; 2022, p. 325). Estudiantes en enseñanza no reglada, datos a partir de «El español: una lengua viva. Informe 2023» (Instituto Cervantes, 2023).

### Senegal

A partir de cifras oficiales de la Dirección de la Planificación de la Reforma de la Educación de Senegal, a través del Instituto Cervantes (noviembre de 2020).

### Suecia

Estudiantes en enseñanza primaria, secundaria y formación profesional, datos de Eurostat (2024b) correspondientes al año escolar 2021-2022. 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 en 2019, al que se ha aplicado un factor multiplicador uniforme calculado a la baja. Estudiantes en enseñanza no reglada, datos proporcionados por Statistics Sweden para el proyecto «El español en Europa».

### Túnez

Datos suministrados en febrero de 2024 por la Embajada de España en Túnez, facilitados por el Ministerio de Educación Tunecino para el ciclo escolar 2022-2023. Agradecemos la ayuda de Carmen Carnes Borrajo, segunda jefatura de la Embajada de España en Túnez.

### 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, Kazajistán, Kirguistán, 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,) así como de países otros países no desglosados de Europa (para Albania, Bosnia y Herzegovina, Macedonia del Norte, Montenegro y Serbia, véase Loureda et al. en prensa), y de la Unión Europea-27 (para 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 2021-2022. 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.

Los datos revelan que la geografía y el grado de aprendizaje de la lengua española dependen, sobre todo, de los gigantes demográficos de América, Estados Unidos y Brasil, y de la Unión Europea. Merece la pena, por tanto, detenerse a explicar brevemente las dinámicas de estos tres espacios.

Estados Unidos es el país con más estudiantes de español en el mundo (8,6 millones). En términos comparativos, la lengua española es, con diferencia, el idioma más aprendido en la educación reglada del país: lo estudia un 76 % del total del alumnado de lenguas extranjeras de la enseñanza primaria y secundaria, y un 50 % de los universitarios. Este hecho, sin embargo, no debe ocultar una realidad desfavorable para el español y el resto de las lenguas extranjeras, pues solo el 20 % del alumnado de primaria y secundaria, y el 6,5 % de los universitarios, aprenden una lengua distinta del inglés en Estados Unidos.

Brasil, por su parte, es uno de los países con más estudiantes de español: algo más de 4 millones. Como se observa en el cuadro 5, la estimación de estudiantes depende prácticamente de su estudio en los espacios de educación primaria y secundaria. En ellos, sin embargo, su presencia se ha visto debilitada tras la reforma de 2016 (Medida Provisória n.º 746) que derogó la *ley del español* de 2005 eliminando su obligatoriedad, reservada ahora exclusivamente al inglés. El efecto de la medida es evidente: en los últimos años, el español ha perdido más de 1 millón de alumnos, una dinámica cuya evolución habrá que observar en los próximos años.

En la Unión Europea se contabilizan alrededor de 6,5 millones de estudiantes de español. Buena parte del aprendizaje de la lengua española tiene lugar en las aulas de enseñanza secundaria, donde es la lengua extranjera más estudiada después del inglés, de implantación casi universal como primera o segunda lengua, y por delante del francés, el alemán y el italiano. En lo que llevamos de siglo XXI, el número de aprendices de español ha venido aumentando en un contexto de diversificación de la oferta de lenguas extranjeras promovido por políticas favorables a la incorporación de dos o más lenguas extranjeras a los currículos educativos de los estados miembros. Ejemplo de ello es la incorporación de una segunda lengua obligatoria en Italia en 2005,

que impulsó el número de alumnos de español, que pasaron de 144.000 a los más de 800.000 actuales. En Alemania, la apertura de espacios para más de una lengua extranjera obligatoria duplicó el número de aprendices y llevó al español a ser la segunda lengua extranjera en la etapa superior de la enseñanza secundaria. En términos cuantitativos, la presencia del español depende considerablemente de su alta implantación en Francia como segunda lengua extranjera, aprendida por cerca del 60 % del total del alumnado de secundaria. En el país galo, además, el español despierta un alto interés no solo en estudios de carácter filológico, sino también como lengua extranjera orientada a fines profesionales. El español es, asimismo, la segunda lengua extranjera en otros sistemas educativos, como es el caso de Suecia e Irlanda, la tercera en Polonia o Portugal, y la cuarta o quinta en otros países como Bélgica, los Países Bajos, o algunos de los estados miembros de incorporación más reciente como la República Checa, Eslovaquia o Hungría.

El cuadro 5 permite también mostrar otras dinámicas del aprendizaje de la lengua española. Una de ellas es la presencia del español en espacios educativos que no suelen ser escudriñados con la atención que merecen. Es el caso de países de África como Costa de Marfil, Benín, Camerún, Senegal o Gabón, con un número considerable de estudiantes, por encima de los 200.000. Otra dinámica importante es el peso relativo de las universidades como espacio de aprendizaje (tardío) del español, como sucede en China, donde el 68 % de los aprendices estimados se encuentran en espacios de educación superior, o de la oferta del Instituto Cervantes en lugares como Marruecos, donde el 20 % de los estudiantes de la estimación aprenden español en las aulas de sus seis centros en Casablanca, Fez, Marrakech, Rabat, Tánger y Tetuán.

Por último, es necesario advertir que la estimación propuesta, sin dejar de ser lo más exhaustiva posible, no puede ofrecer una imagen completa de la enseñanza de español en los espacios de enseñanza no reglada. La situación documental y estadística del sector del aprendizaje de idiomas es muy limitada en todo el mundo e invisibiliza un grupo de aprendices que en otras ediciones del informe anual del Instituto Cervantes se ha estimado en más de 5 millones y que no se han tenido en cuenta en los recuentos, criterio que

seguimos en este informe. El aumento del uso de espacios digitales para el aprendizaje de lenguas extranjeras, desde cursos en línea hasta aplicaciones para dispositivos móviles, podría apuntar a estimaciones aún mayores. Según los informes de Duolingo (Blanco, 2023), que actualmente cuenta con 88 millones de usuarios mensuales, el español es la segunda lengua con más aprendices, la primera en 34 países en 2023 entre los que se incluyen Estados Unidos, el Reino Unido, Irlanda, Australia, Sudáfrica, Suecia o Noruega.

## 2.5 | Hablantes potenciales de lengua española

La suma de los totales ofrecidos en los cuadros estadísticos anteriores queda recogida en el cuadro 6. El total corresponde al Grupo de Usuarios Potenciales de español dentro y fuera de los países mayoritariamente hispanohablantes. Como se ha explicado, 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 potencialmente limitaciones de naturaleza estilística o sociolingüística y aquellos que lo están aprendiendo dentro de un sistema determinado.

**CUADRO 6 | HABLANTES DE ESPAÑOL DE DOMINIO NATIVO, CON COMPETENCIA LIMITADA Y APRENDICES DE LENGUA, POR TIPOS DE PAÍSES Y CON TOTALES**

|                            | GDN         | GCL        | GALE       | Total       |
|----------------------------|-------------|------------|------------|-------------|
| Países hispanohablantes    | 452.807.955 | 24.510.743 |            | 477.318.698 |
| Países no hispanohablantes | 45.689.802  | 53.390.493 | 24.208.813 | 123.289.108 |
| Total                      | 498.497.757 | 77.901.236 | 24.208.813 | 600.607.806 |

**FUENTE:** elaboración propia.

El cuadro 6 revela algunas dinámicas que merecen resaltarse. En primer lugar, la comunidad hispanohablante con dominio nativo sigue rozando los 500 millones de hablantes, por mucho que otras fuentes menos minuciosas en su atención al español la suelen situar al menos un 10 % por debajo de ese umbral. En segundo lugar, se constata el crecimiento de los usuarios potenciales de español fuera del espacio mayoritariamente hispanohablante,

aumento que se debe en gran parte a la población de trasfondo migratorio. En cuanto a los aprendices de español, se constata el crecimiento sostenido de este grupo, que ha venido incrementándose a un ritmo del 2 % anual. En conjunto, los hablantes potenciales de lengua española en el mundo superan también por primera vez los 600 millones, lo que mantiene al español, junto al inglés y el chino mandarín, entre las lenguas con mayor potencial demográfico del planeta.

### 3 | El español entre las lenguas del mundo

Una vez analizada y actualizada la población hispanohablante, tanto en países hispánicos como en países que no tienen el español como lengua mayoritaria, resulta imprescindible para el análisis demográfico del español establecer algunas comparaciones con otras lenguas del mundo. De este modo es posible calibrar las dimensiones de la comunidad hispanohablante en relación con otras comunidades lingüísticas internacionales, con especial referencia, aunque no exclusivamente, a las lenguas que cuentan con el marchamo de oficial en la ONU: junto al español, el inglés, el francés, el chino mandarín, el árabe y el ruso.

Con el fin de obtener una caracterización actualizada del español entre las lenguas del mundo<sup>9</sup>, ofrecemos a continuación algunos datos comparativos referidos a cuatro tipos de factores: la dimensión demográfica e internacional de las lenguas; el peso internacional de las culturas que expresan; el nivel de desarrollo de las comunidades de hablantes; y finalmente, la presencia en internet de las lenguas. Dada la exigüidad y dispersión de los datos existentes, utilizaremos como base comparativa los datos que se ofrecen en el barómetro Calvet para 634 lenguas del mundo (Calvet y Calvet, 2022). Más allá de la fecha de recolección y publicación, los datos para la comunidad hispanohablante que este barómetro maneja no se corresponden exactamente con los aquí

---

<sup>9</sup> Una caracterización del español frente a las demás lenguas oficiales de la ONU puede hallarse en los diversos estudios que se han realizado sobre el peso internacional de las lenguas (Otero Roth, 1995; Chan, 2016; Moreno Fernández y Álvarez Mella, 2022). En esos trabajos se proponen y manejan diversos factores como base para la adjudicación de un mayor o menor peso en el concierto lingüístico internacional.

presentados para 2024, dado que no son fruto de un análisis tan exhaustivo para cada una de las lenguas como el que aquí se hace para el español. Con todo, la información aportada por el barómetro ciertamente permite una comparación más sensata que la que pueda ofrecer el manejo de fuentes dispares.

### 3.1 | Dimensión demográfica e internacional de las lenguas

Esta dimensión puede concretarse mediante el análisis de diversos factores, entre los que destacamos tres singularmente: el número de países donde las lenguas son oficiales, nacionales o mayoritarias; el volumen de la población hablante de cada lengua y la tasa de vehicularidad de las lenguas. Como se ha comentado, en este caso nos ceñiremos a la consideración de las lenguas que gozan de oficialidad en las agencias de la ONU.

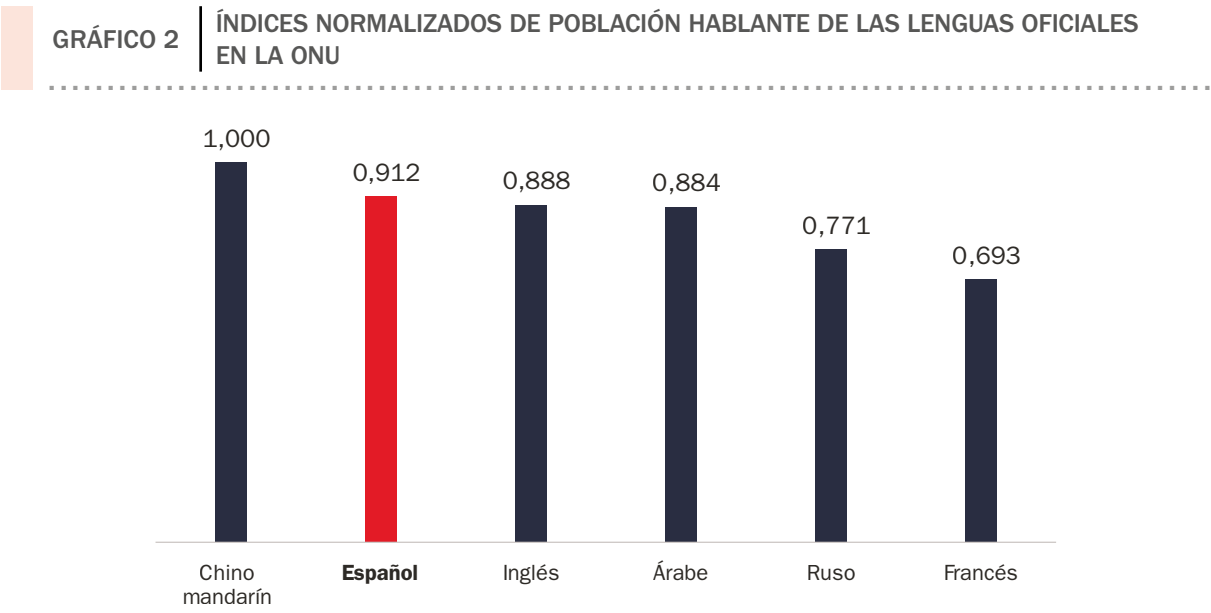
El número de países donde una lengua se habla debería ser un valor relativamente fácil de concretar, pero la realidad política y administrativa de cada territorio hace que la tarea presente ostensibles márgenes de variabilidad. Por poner unos ejemplos, según el criterio político o diplomático que se adopte, cabría la posibilidad de incluir o no incluir a Taiwán entre los países hablantes de chino mandarín o a Transnistria entre los territorios rusófonos, por más que nuestros argumentos se basen, no en la geopolítica, sino en los hablantes, los territorios y sus lenguas. Además, existen entidades no soberanas y territorios subnacionales susceptibles de ser considerados independientemente, así como diferentes niveles en la oficialidad de las lenguas, por no hablar del diferente tratamiento que puede darse a las variedades de una lengua, como en el caso del árabe. Estas dificultades pueden observarse en diversas regiones del mundo para un gran número de lenguas, incluidas el inglés, el francés, el ruso y el español.

En este informe se manejan los números de países por lengua que aparecen en el cuadro 7. En la determinación de esos números se ha otorgado mayor peso a la consideración sociopolítica internacional y a la realidad sociolingüística de los territorios que al mero y estricto reconocimiento político, generalmente asimétrico, entre países.

| CUADRO 7   NÚMERO DE PAÍSES O TERRITORIOS DONDE ES OFICIAL CADA LENGUA DE LA ONU |         |         |                |       |      |
|----------------------------------------------------------------------------------|---------|---------|----------------|-------|------|
| Inglés                                                                           | Francés | Español | Chino mandarín | Árabe | Ruso |
| 55                                                                               | 28      | 21      | 3              | 26    | 6    |

FUENTE: elaboración propia.

Los recuentos de población hablante de cada lengua tampoco están exentos de problemas, dada la diversidad de criterios con que cada lengua es tratada por cada fuente. Así, el árabe, por ejemplo, a menudo no se considera en su conjunto, sino en sus variedades nacionales, dada la distancia lingüística existente entre ellas, aunque también es posible un tratamiento unificado. Por ese motivo, consideramos interesante atender al índice poblacional que presenta el barómetro Calvet que, si bien no refleja unos recuentos unánimemente aceptados, sí incluye unos valores normalizados que ayudan a ponderar de modo relativo el peso demográfico de cada lengua. Estos valores se ajustan entre 0 y 1 mediante un logaritmo en el que la lengua de mayor número de hablantes, el chino mandarín, tiene el valor 1. Ese índice normalizado ofrece la ventaja añadida de que su variación interanual no es excesivamente grande (Calvet y Calvet, 2022). De acuerdo con este criterio, el índice poblacional del español es el segundo más alto del mundo, por encima del inglés.



FUENTE: Calvet y Calvet (2022).

Los porcentajes de hablantes de estas lenguas en todo el mundo, en relación con la población, es otro índice valioso para estimar la implantación internacional de una lengua. Dado que ni siquiera existe un único recuento absolutamente fiable de la población mundial y dado que el número de hablantes por lengua no se calcula con la misma metodología por parte de todas las fuentes, las proporciones que se ofrecen han de ser por fuerza aproximadas<sup>10</sup>.

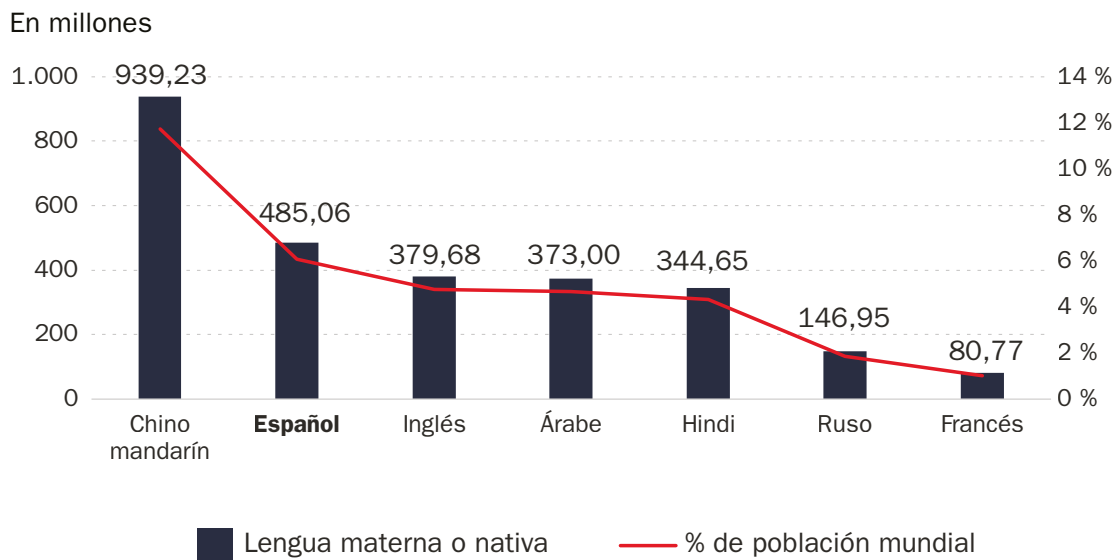
| CUADRO 8   NÚMERO APROXIMADO DE HABLANTES EN EL MUNDO POR LENGUA MATERNA O NATIVA, JUNTO A PORCENTAJE DE POBLACIÓN MUNDIAL Y HABLANTES DE LENGUA NATIVA Y SEGUNDA <sup>11</sup> |                         |                        |                         |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------|------------------------|-------------------------|
| Lengua                                                                                                                                                                          | Lengua materna o nativa | % de población mundial | Lengua nativa + segunda |
| Chino mandarín                                                                                                                                                                  | 939.237.350             | 11,74 %                | 1.138.222.350           |
| Español                                                                                                                                                                         | 485.063.960             | 6,06 %                 | 559.078.890             |
| Inglés                                                                                                                                                                          | 379.682.200             | 4,75 %                 | 1.456.448.320           |
| Árabe                                                                                                                                                                           | 373.000.000             | 4,66 %                 | 373.989.700             |
| Hindi                                                                                                                                                                           | 344.650.870             | 4,31 %                 | 609.454.770             |
| Ruso                                                                                                                                                                            | 146.954.150             | 1,84 %                 | 254.997.130             |
| Francés                                                                                                                                                                         | 80.770.190              | 1,01 %                 | 309.804.220             |

FUENTE: Eberhard, Simons y Fennig (2023) y Encyclopedia Britannica (s. f.).

10. Los datos del español se presentan de este modo en aras de la comparabilidad, si bien este informe ofrece datos que responden a análisis más completos y detallados.

11. Datos correspondientes a las lenguas oficiales de la ONU más el hindi. Aunque los datos se obtienen de fuentes no sincronizadas, en general son posteriores al año 2000. La población mundial de referencia para el cálculo es de 8.000 millones, cifra alcanzada en noviembre 2022 según la ONU (2024).

GRÁFICO 3 | NÚMERO APROXIMADO DE HABLANTES EN EL MUNDO POR LENGUA MATERNA O NATIVA, JUNTO A PORCENTAJE DE POBLACIÓN MUNDIAL



**FUENTE:** Eberhard, Simons y Fennig (2023) y Encyclopedia Britannica (s. f.).

Recordemos que, en aras de la comparabilidad, los datos sobre el español que ofrecen estas ilustraciones no reflejan el análisis que aquí se presenta, sino datos de estudios anteriores. El gráfico 3 muestra que los hablantes de español, como lengua nativa o materna, suponen una proporción de la población mundial tan solo por debajo del chino mandarín, si bien el crecimiento demográfico de la India y de los territorios donde se habla hindi hacen prever para esta lengua un crecimiento rápido e intenso que habrá de valorarse en los próximos años. El gráfico 4 muestra las proporciones aproximadas del volumen de hablantes nativos y totales en torno a 2022. Ahora bien, teniendo en cuenta los hablantes potenciales de español, en todos los perfiles analizados en este informe y las cuantificaciones actualizadas a 2024, el porcentaje de la población mundial hablante español sería del 7,5 %, proporción que se viene manteniendo estable desde 2020.

GRÁFICO 4 | VOLÚMENES RELATIVOS DE HABLANTES POR SU LENGUA MATERNA O NATIVA Y DE HABLANTES TOTALES DE CADA LENGUA



**FUENTE:** elaboración propia a partir de datos de Eberhard, Simons y Fennig (2023) y Encyclopedia Britannica (s. f.).

Por otro lado, la tasa de «vehicularidad» de las lenguas se define como la relación entre el número de hablantes que utilizan cada lengua como segunda lengua y su número total de hablantes. Este índice varía entre 0, para una lengua que solo tiene hablantes de L1, y 1, para una lengua cuyos hablantes la hablan todos como L2. La dificultad para la elaboración de esa tasa está en la existencia y, si es el caso, en la disponibilidad y precisión de las fuentes. Además, debe tenerse en cuenta que el nivel de conocimiento y uso de una segunda lengua no es equivalente en todos los casos, sino que existe una multiplicidad de posibles relaciones entre la lengua primera y segunda de una comunidad. Con todo, esta aproximación resulta sugestiva porque refleja el tipo de dinámica internacional de cada lengua y permite una interesante comparación.

CUADRO 9 | TASA DE VEHICULARIDAD DE LENGUAS OFICIALES EN LA ONU (EXCEPTO ÁRABE)

| Lengua         | Lengua materna o nativa | % de población mundial |
|----------------|-------------------------|------------------------|
| Francés        | 0,689                   | 0,693                  |
| Inglés         | 0,641                   | 0,888                  |
| Ruso           | 0,437                   | 0,912                  |
| Chino mandarín | 0,183                   | 1,000                  |
| Español        | 0,178                   | 0,912                  |

FUENTE: Calvet y Calvet (2022).

El cuadro 9 muestra la tasa de vehicularidad de cinco de las lenguas oficiales de la ONU. El barómetro Calvet incluye el idioma árabe, pero fragmentado en variedades dialectales, por lo que no ofrece una tasa unificada. El cuadro refleja un panorama en parte bien conocido: por una parte, el español es una lengua que muestra un muy alto coeficiente de hablantes de primera lengua, mientras que el inglés y el francés son lenguas frecuentemente habladas como segundas lenguas, especialmente el inglés; sin embargo, este mismo hecho es el que supone que la tasa de vehicularidad del español sea mucho más baja que la de estas otras lenguas, dado que la fuerza demográfica del español radica en su comunidad de hablantes nativos. La tasa de vehicularidad ligeramente más alta para el mandarín que para el español se debe en buena medida a la diáspora china por el mundo.

3.2 | Dimensión cultural-educativa

Para dar cuenta de la incidencia del factor cultural en la internacionalidad de las lenguas, recurrimos a dos indicadores cuantificados en el barómetro de Calvet. Estos factores serán las traducciones, de y hacia las respectivas lenguas, y la inclusión de las lenguas en los sistemas universitarios internacionales.

El barómetro Calvet incluye entre sus factores el de la traducción, distinguiendo los índices de las traducciones hacia las lenguas de referencia (lengua meta) y desde las lenguas de referencia (lengua fuente). El volumen de

traducciones es un indicador muy significativo de la importancia internacional de las lenguas como reflejo del peso internacional de las culturas y conocimientos que expresan, con la particularidad de que los valores anuales son susceptibles de acumularse a los valores de años anteriores, dada la prolongada vida de gran parte de las traducciones.

Los valores correspondientes a las traducciones están basados en datos procedentes del *Index Translationum* de la UNESCO. Este servicio publica el número de traducciones por idioma desde 1979 y los datos se presentan por el país en que se realizó la traducción, el año en que se realizó y el tema tratado. Las traducciones se clasifican en nueve categorías: general y bibliografía; filosofía y psicología; religión y teología; derecho, ciencias sociales y educación; ciencias exactas y naturales; ciencias aplicadas; artes, juegos y deportes; literatura y, por último, historia, geografía y biografía.

**CUADRO 10 | ÍNDICES DE TRADUCCIONES DE LAS LENGUAS OFICIALES DE LA ONU (LENGUAS FUENTE Y META)**

| Lengua         | Traducciones meta | Traducciones fuente | Promedio |
|----------------|-------------------|---------------------|----------|
| Inglés         | 0,949             | 1,000               | 0,974    |
| Francés        | 0,981             | 0,871               | 0,926    |
| Español        | 0,977             | 0,765               | 0,871    |
| Ruso           | 0,908             | 0,813               | 0,860    |
| Chino mandarín | 0,869             | 0,663               | 0,766    |
| Árabe          | 0,263             | 0,655               | 0,459    |

**FUENTE:** Index Translationum (como se citó en Calvet y Calvet, 2022).

La posición del español en cuanto a este factor es relativamente destacada en el panorama internacional y muestra un peso menor que el inglés y el francés. Según los datos del *Index Translationum*, el español ofrece como promedio un índice de 0,871, por encima del ruso, el chino mandarín y el árabe. En ese índice destacan las traducciones que se vierten hacia el español, con un valor más alto que las del inglés, mientras que las traducciones que tienen al

español como lengua fuente muestran un índice menor, no solo que las del inglés y el francés, sino también que las del ruso.

En lo que se refiere a la presencia de las lenguas en los programas universitarios internacionales, el barómetro Calvet ofrece unos índices que implican un complejo cálculo sobre algunas universidades que se consideran de especial relevancia, dado que solamente se escudriñan un 10 % de las aproximadamente 20.000 universidades que existen en el mundo. En este caso, se trata de cuantificar la importancia de las lenguas a través de su enseñanza en los primeros niveles de la educación superior y excluyendo los programas de posgrado.

| CUADRO 11   ÍNDICE DE PRESENCIA DE LAS LENGUAS OFICIALES DE LA ONU EN ENSEÑANZA UNIVERSITARIA |                         |
|-----------------------------------------------------------------------------------------------|-------------------------|
| Lengua                                                                                        | Enseñanza universitaria |
| Inglés                                                                                        | 1,000                   |
| Francés                                                                                       | 0,947                   |
| Español                                                                                       | 0,861                   |
| Chino mandarín                                                                                | 0,815                   |
| Ruso                                                                                          | 0,812                   |
| Árabe                                                                                         | 0,219                   |

FUENTE: Calvet y Calvet (2022).

Este factor revela con claridad la universalidad de la enseñanza del inglés en el nivel universitario, así como la importante oferta internacional de francés. El español, por su parte, tiene una presencia universitaria muy amplia que, en numerosas ocasiones, no se ajusta a la auténtica demanda de aprendizaje de español por parte de las sociedades y específicamente de los estudiantes, dado que la elección de segundos o terceros idiomas se ve limitada por unas ofertas docentes que suelen ser bastante rígidas y comodistas respecto a un previo estado de la cuestión.

### 3.3 | Dimensión de desarrollo humano

La certeza de la correlación entre el peso de las lenguas y el nivel de desarrollo de los pueblos y naciones que las hablan se ha evidenciado prácticamente en todos los análisis realizados sobre la importancia internacional de las lenguas. Verdaderamente, el nivel de desarrollo de los países es un factor decisivo en la percepción del peso de las lenguas, que puede traducirse en un mayor uso internacional de estas y en un mayor interés por aprenderlas.

Con el fin de apreciar cuál es la posición de las lenguas oficiales de la ONU en relación con el nivel de desarrollo de los respectivos países en que son mayoritarias, hemos dirigido nuestra mirada a dos conocidos índices internacionales: el Índice de Desarrollo Humano (IDH) y el Índice de Capital Humano (ICH)<sup>12</sup>.

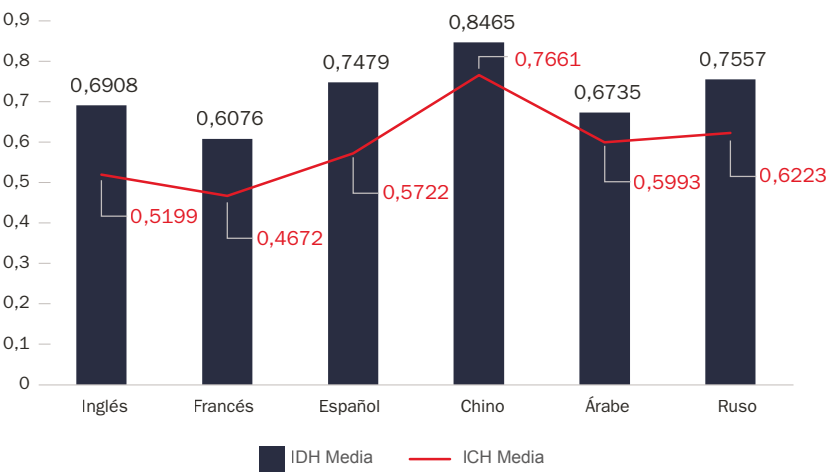
El IDH es un índice compuesto que tiene en cuenta el producto nacional bruto por persona, la esperanza de vida al nacer y el nivel educativo. En definitiva, este índice cuantifica el nivel de desarrollo de un país. Es probable que en muchos países el IDH no sea el mismo para todas las regiones, para todos los grupos étnicos que en ellas viven y, por tanto, para todas las lenguas que estos grupos étnicos utilizan. Sin embargo, no disponemos de datos desagregados y hemos de trabajar con un índice único para cada país, a pesar de su posible heterogeneidad. Para asignar un valor a cada lengua, tomamos una media ponderada del índice en cada uno de los países en los que se habla la lengua.

En cuanto al ICH, se trata de un valor anual, elaborado por el Banco Mundial, que mide la capacidad de los países para movilizar su capital humano; esto es, el potencial económico y profesional de su ciudadanía. El índice viene a cuantificar la contribución de la salud y la educación a la productividad en la siguiente generación. Ambos índices se expresan en valores comprendidos entre 0 y 1.

---

**12.** Sobre el producto interior bruto de cada comunidad lingüística y su evolución, puede encontrarse información en «El español: una lengua viva. Informe 2023» (Instituto Cervantes, 2023).

GRÁFICO 5 | PROMEDIOS DE IDH Y DE ICH EN LOS PAÍSES HABLANTES DE LAS LENGUAS OFICIALES DE LA ONU



|           | Inglés<br>(55 países) | Francés<br>(28 países) | Español<br>(21 países) | Chino<br>(3 países) | Árabe<br>(19 países) | Ruso<br>(7 países) |
|-----------|-----------------------|------------------------|------------------------|---------------------|----------------------|--------------------|
| IDH Media | 0,6908                | 0,6075                 | 0,7478                 | 0,8465              | 0,6735               | 0,7557             |
| ICH Media | 0,5199                | 0,4671                 | 0,5722                 | 0,7661              | 0,5993               | 0,6222             |

FUENTE: elaboración propia a partir de datos del Banco Mundial (2023 y 2024).

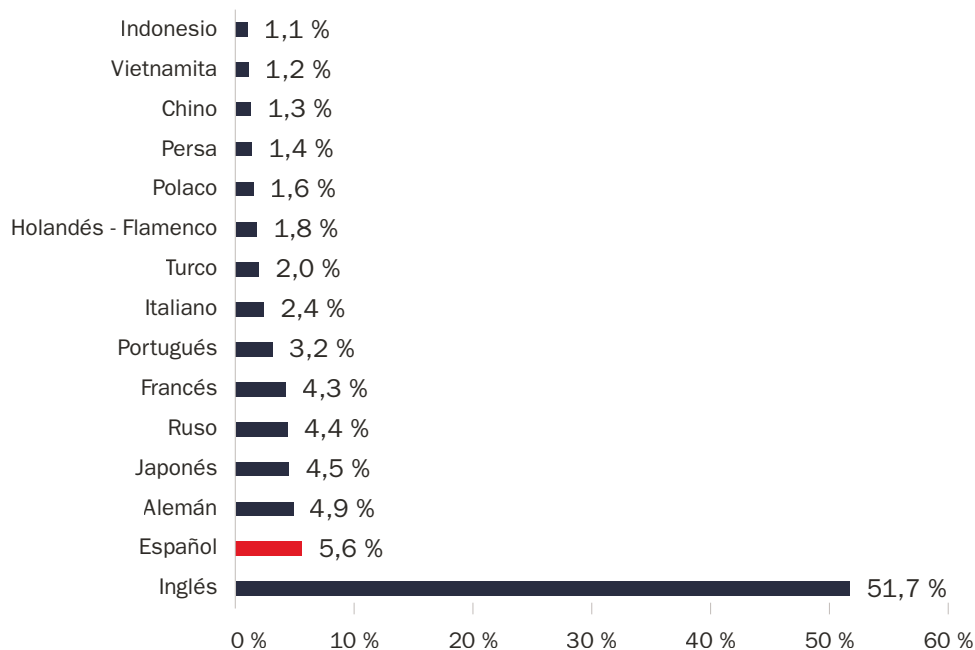
El gráfico 5 muestra la posición de las comunidades hispanohablantes en relación con las poblaciones de otras lenguas oficiales de la ONU. Aquí se aprecia que la posición del español, en su conjunto, se ubica por encima de los ámbitos anglófono y francófono. Para comprender adecuadamente el gráfico, debe tenerse en cuenta que, si bien Estados Unidos o Francia cuentan con un IDH e ICH más alto que el de China, por ejemplo, los datos se refieren al promedio de estos índices en todos los países hablantes de cada lengua. En este sentido, el inglés y el francés ven reducida su dimensión cuantitativa por el peso de los países en desarrollo donde se hablan, especialmente en el continente africano. Al mismo tiempo, el IDH del espacio hispanohablante sería más parejo con el de China, si de la comunidad sinohablante se excluyeran Taiwán y Singapur<sup>13</sup>.

13. El número de países adscritos a cada lengua puede variar dependiendo de si algunos de ellos cuentan o no con reconocimiento. Algunos países cuentan con más de una lengua oficial, por lo

### 3.4 | Dimensión digital

En la actualidad, las comunidades idiomáticas también se perfilan de un modo decisivo por su presencia en la esfera digital. Son muchos los indicadores que revelan la dimensión de tal presencia y la mayoría de ellos son convergentes en cuanto a la significación de las lenguas manejadas en las redes. Después del inglés, según Internet World Stats (2023) las lenguas más usadas en la red por número de usuarios son el español y el chino mandarín. Sin embargo, las posiciones varían si se miran los sitios web.

GRÁFICO 6 | PORCENTAJES DE USO DE IDIOMAS EN SITIOS WEB



FUENTE: W3Techs (2024).

El gráfico 6 muestra los porcentajes de sitios web según los idiomas que usan para su contenido. Los datos se reúnen atendiendo a los sitios web que se consideran relevantes. Un sitio web es relevante si tiene contenidos o funcionalidades significativos, para lo cual se analiza su contenido y su estructura de enlaces entre sitios web. De acuerdo con estos recuentos, basados en

que se adscriben a las comunidades idiomáticas correspondientes. Los índices correspondientes al chino mandarín incluyen los de China, Taiwán y Singapur. Estos últimos elevan significativamente el promedio.

millones de sitios web, el inglés es utilizado por el 51,7 % de todos los sitios web relevantes. El español, como las demás lenguas del mundo, muestra un porcentaje diez veces menor que el de la lengua inglesa (5,7 %), pero se sitúa a la cabeza del resto de las lenguas, incluidos el alemán o el francés.

#### **4 | Conclusiones parciales sobre demografía del español**

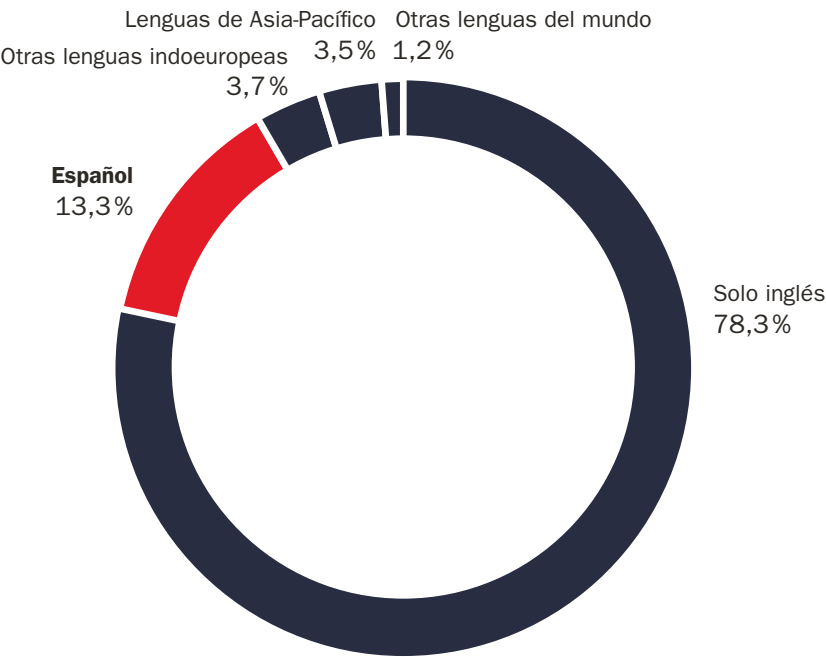
En resumen, la posición de la lengua española entre las lenguas del mundo aparece como relevante en relación con la mayoría de los índices que suelen manejarse para calibrar el peso internacional de las lenguas. El español muestra una demografía vigorosa, especialmente en cuanto a su comunidad nativa y un bajo índice de vehicularidad y, por lo tanto, revela un notable margen de crecimiento como segunda lengua en todo el mundo, en parte como consecuencia de las diferencias en los procesos de colonización y descolonización protagonizados por España y por el Reino Unido, Estados Unidos o Francia. Ese crecimiento podría llegar por medio de la educación y de los potenciales movimientos migratorios del futuro. En el ámbito educativo y cultural (traducciones, enseñanza universitaria), la comunidad hispanohablante se sitúa en posiciones internacionales punteras, próximas a las del francés especialmente. En cuanto al nivel de desarrollo de los países hispanohablantes, estos han experimentado una mejora progresiva que los sitúa en una mejor posición que la del conjunto de los países anglófonos y francófonos. La mejora también es apreciable en la esfera digital, en la que los indicadores muestran una posición destacada del español respecto a las demás lenguas, que no amenazan la hegemonía del inglés.

#### **5 | Notas de actualidad demolingüística: el español en Estados Unidos**

La situación del español en Estados Unidos afecta de manera importante a la trascendencia internacional de esta lengua. Estados Unidos es un foco de prestigio económico y cultural, cuya relevancia alcanza a las lenguas que vehiculan la comunicación dentro del país. A este respecto, la significación de la lengua inglesa es evidente, pero, entre las demás lenguas habladas en el

país norteamericano, ninguna otra se acerca a la penetración y la extensión de la lengua española. Los datos de la American Community Survey (Oficina del Censo de los Estados Unidos, 2022a) revelan que el español se habla en el hogar de un 13,3 % de los estadounidenses, lo que supone casi el doble que el resto de las lenguas juntas, a excepción naturalmente del inglés.

GRÁFICO 7 | LENGUA HABLADA EN EL HOGAR



|                            |             |
|----------------------------|-------------|
| Solo inglés                | 244.232.103 |
| Español                    | 41.434.050  |
| Otras lenguas indoeuropeas | 11.697.220  |
| Lenguas de Asia-Pacífico   | 11.013.422  |
| Otras lenguas del mundo    | 3.715.873   |

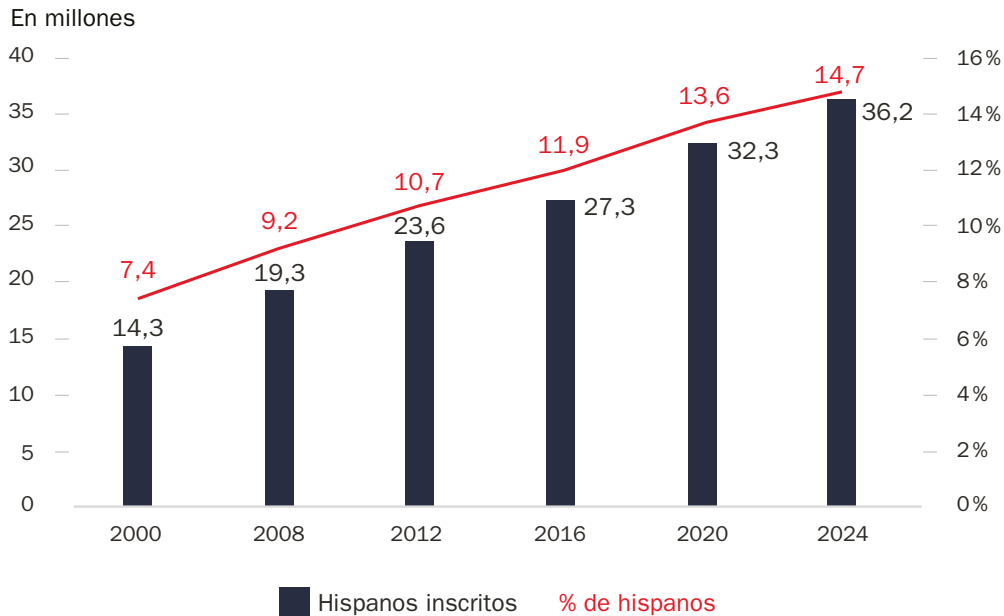
FUENTE: Oficina del Censo de los Estados Unidos (2022a).

Teniendo en cuenta que la población hispana de Estados Unidos en 2022 era de 63,7 millones (19,1 % de la población total), el porcentaje de esta población que mantenía el español en casa se situaría alrededor del 67 %, aunque son un 75 % los hispanos que afirman ser capaces de mantener una conversación en español (Mora y López, 2023). Como es sabido (Moreno

Fernández, 2023), el porcentaje de los que hablan español en casa se ha ido reduciendo paulatinamente desde los años noventa, si bien no a tanta velocidad como preveían quienes presagiaban una desaparición del español en tres generaciones, similar a la experimentada por otras lenguas. En este caso, la diferencia entre lenguas como el italiano o el alemán y el español reside en que el español en Estados Unidos es mucho más que una lengua migratoria. Con todo, la reducción comentada estaría vinculada a la idea de que la identidad hispana o latina ha ido debilitando su identificación con el hecho de hablar español. Un estudio del Pew Research Center revela que algunos latinos utilizan el hecho de hablar español para distinguir quién es hispano o latino y quién no lo es, pero que la mayoría de los adultos latinos (78%) afirma que no es necesario hablar español para ser considerado uno de ellos, mientras que el 21% afirma que sí lo es (Mora y López, 2023). Al mismo tiempo, crece el movimiento contra los llamados «No sabo kids»; esto es, la discriminación entre los hispanos hacia aquellos que no saben español (Brooks, 2023).

Los datos disponibles revelan el crecimiento continuo que la masa de electores hispanos ha experimentado en las dos últimas décadas. Debe a este respecto recordarse que, en Estados Unidos, no todos los potenciales votantes están inscritos para votar, por lo que el número de ciudadanos con derecho al voto efectivo es menor que el de potenciales votantes. En el caso de los hispanos, ese número ha sido tradicionalmente inferior al de otros grupos étnicos asentados en el país, si bien, cada año, alrededor de 1,4 millones más de hispanos cumplen los requisitos para votar en Estados Unidos.

GRÁFICO 8 | PORCENTAJE DE HISPANOS INSCRITOS PARA VOTAR EN 2024 EN ESTADOS UNIDOS



FUENTE: Krogstad et al. (2024).

En las elecciones presidenciales de 2024 se prevé que podrán votar 36,2 millones de personas hispanas, frente a los 32,3 millones de 2020. Esto representa el 50 % del crecimiento total de votantes con derecho a voto durante este periodo. A partir de datos de la Oficina del Censo y de proyecciones propias, el Pew Research Center señala cinco datos clave que han de tenerse en cuenta para este año (Krogstad et al., 2024). Los datos de mayor significación son los siguientes:

En noviembre de 2024, los hispanos con derecho a voto alcanzarán un nuevo máximo, al representar el 14,7 % de todos los votantes. Esto significa que el peso de los votos hispanos será mayor que en 2020, aunque está por ver, como se ha comentado, cuál es la orientación final de esos votos. Esta proporción resulta especialmente significativa cuando se compara con el exiguo 7,4 % de votantes hispanos que gozaron de este derecho en el año 2000.

Apenas un 53 % de los hispanos de Estados Unidos tenía derecho a votar en 2022. Esto se debe fundamentalmente a dos razones; por un lado, se trata de una población joven que, en un porcentaje considerable (29 %) aún no tiene

edad para votar; por otro lado, existe un contingente de población hispana (19 %) que aún no dispone de la nacionalidad estadounidense.

Los hispanos con derecho a votar tienen generalmente menor edad que el conjunto de los votantes. La proporción de votantes hispanos mayores de 50 años es del 33 %, proporción que contrasta con el 48 % de todos los votantes estadounidenses con esa misma edad. Por otro lado, solamente el 21 % de los hispanos con derecho a voto tiene un título universitario, frente al 33 % del conjunto de los votantes.

California es el estado con más votantes hispanos, unos 8.5 millones, lo que supone la cuarta parte de todos los votantes hispanos del país. Siguen en peso a California los estados de Texas, Florida, Nueva York y Arizona, y entre todos ellos suman el 65 % de todos los votantes hispanos con derecho a voto.

El estado con mayor proporción de votantes hispanos es Nuevo México, donde el 45 % de ellos tienen derecho a voto. Este dato se corresponde con la alta proporción de población hispana o latina, que supera a todos los demás grupos étnicos y raciales, incluidos los estadounidenses blancos no hispanos.

Estos cinco puntos dibujan un perfil cuantitativo y cualitativo de los hispanos con derecho a voto con vistas a las elecciones presidenciales de 2024. Asunto distinto es cuál será la orientación final del voto hispano. De igual modo que se establece un vínculo entre hablar español y la identidad hispana, también suele creerse que existe una relación directa entre ser hispano y votar por el partido demócrata, dado que sus propuestas en materia migratoria podrían beneficiar supuestamente a la comunidad hispana. La realidad, sin embargo, es más compleja, como revela el hecho de que, en las últimas elecciones presidenciales, los votantes hispanos que optaron por Joe Biden fueron el 59 %. Las informaciones disponibles señalan una mayor indecisión por parte de los hispanos que votaron por los demócratas en las últimas elecciones (Ramón, 2024).

## Referencias bibliográficas

- ACUÑA, L. (2009). De la castellanización a la educación intercultural bilingüe: Sobre la atención de la diversidad lingüística en la Argentina. *Revista Nebrija de Lingüística Aplicada a la Enseñanza de las Lenguas*, 6. <https://revistas.nebrija.com/revista-linguistica/article/view/122>
- 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](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>
- AUSTRALIAN BUREAU OF STATISTICS (S. F.). CENSUS OF POPULATION AND HOUSING 2021. *Language spoken at home*. <https://profile.id.com.au/australia/language>
- BANCO MUNDIAL (2015). *Latinoamérica indígena en el siglo XXI*. Banco Mundial. <https://documents1.worldbank.org/curated/en/541651467999959129/pdf/Latinoam%C3%A9rica-ind%C3%ADgena-en-el-siglo-xxi-primera-d%C3%A9cada.pdf>
- BANCO MUNDIAL (2023). *Informe anual 2023*. <https://www.bancomundial.org/es/about/annual-report>
- BANCO MUNDIAL (2024). *Datos de libre acceso del Banco Mundial*. <https://datos.bancomundial.org/>
- BLANCO, C. (4 de diciembre de 2023). 2023 Duolingo Language Report. *Duolingo blog*. <https://blog.duolingo.com/2023-duolingo-language-report/>
- BONOMI, M., CALVI, M. V. y UBERTI-BONA, N. (en prensa). *Demolingüística del español en Italia. Con un anexo sobre el español en San Marino, Malta y Ciudad del Vaticano*. Instituto Cervantes, Universidad de Heidelberg y Universidad de Zúrich.
- BROOKS, D. (27 de noviembre de 2023). «No sabo kids», los latinos discriminados en Estados Unidos por no dominar el español. *BBC News Mundo*. <https://www.bbc.com/mundo/articles/cx01l7yx2jno>
- CALVET, A. Y CALVET, L. J. (2022). *Baromètre des langues dans le monde*. Ministère de la Culture. <https://www.culture.gouv.fr/Thematiques/langue-francaise-et-langues-de-france/Agir-pour-les-langues/Innover-dans-le-domaine-des-langues-et-du-numerique/Soutenir-et-encourager-la-diversite-linguistique-dans-le-domaine-numerique/Barometre-des-langues-dans-le-monde-2022>
- CENTRAL BUREAU OF STATISTICS ARUBA (2023). *Population projections*. <https://cbs.aw/wp/index.php/2023/05/04/population-projections/>

- CENTRAL BUREAU OF STATISTICS CURAÇAO (2023). *Population by most spoken language (in private households)*. <https://senso.cbs.cw/education-characteristics>
- CENTRO DE ESTUDIOS PÚBLICOS DE CHILE (CEP) (2016). *Estudio Nacional de Opinión Pública n.º 77*. <https://www.cepchile.cl/encuesta/estudio-nacional-de-opinion-publica-n77-julio-agosto-2016/>
- CHAN, K. (2016). *Power language index: Which are the world's most influential languages?* Institute Européen d'Administration des Affaires. [http://www.kailchan.ca/wp-content/uploads/2016/12/Kai-Chan\\_Power-Language-Index-full-report\\_2016\\_v2.pdf](http://www.kailchan.ca/wp-content/uploads/2016/12/Kai-Chan_Power-Language-Index-full-report_2016_v2.pdf)
- CONDE DE LA VIÑAZA (1892). *Lenguas indígenas de América*. Estab. Tip. Suc. de Rivadeneyra.
- CONSEJO DEL SECTOR PRIVADO PARA LA ASISTENCIA EDUCACIONAL DE PANAMÁ (COSPAE) (2002). *El reto es avanzar. Informe de Progreso Educativo de Panamá*. <https://www.thedialogue.org/wp-content/uploads/2015/12/2002-El-reto-es-avanzar-Panam%C3%A11.pdf>
- CONSEJO DEL SECTOR PRIVADO PARA LA ASISTENCIA EDUCACIONAL DE PANAMÁ (COSPAE) (2024). *Cospae Digital*. <https://cospae.org/>
- CONSEJO NACIONAL DE POBLACIÓN DE MÉXICO (CONAPO) (2024). *Indicadores demográficos de la República Mexicana, en el año 2024*.
- DEPARTAMENTO ADMINISTRATIVO NACIONAL DE ESTADÍSTICA DE COLOMBIA (DANE) (2018). *Censo Nacional de Población y Vivienda*. <https://www.dane.gov.co/index.php/estadisticas-por-tema/demografia-y-poblacion/censo-nacional-de-poblacion-y-vivenda-2018>
- DOMINGO, J. (1968). Entrevistas. Gabriel García Márquez. *Ínsula*, 259, xxiii, 6-11.
- EBERHARD, D., SIMONS, G. Y FENNIG, C. (2023). *Ethnologue: Languages of the World*. (26.ª ed.). SIL International.
- ENCYCLOPAEDIA BRITANNICA (S. F.). BRITANNICA [HTTPS://WWW.BRITANNICA.COM/TOPIC/LANGUAGES-BY-NUMBER-OF-NATIVE-SPEAKERS-2228882](https://www.britannica.com/topic/languages-by-number-of-native-speakers-2228882)
- EUROSTAT (2012). *Europeans and their Languages. Special Eurobarometer 386*. [https://data.europa.eu/data/datasets/s1049\\_77\\_1\\_ebs386?locale=en](https://data.europa.eu/data/datasets/s1049_77_1_ebs386?locale=en)
- EUROSTAT (2024a). *Europeans and their Languages. Special Eurobarometer 386*. <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\\_uoe\\_lang01](https://ec.europa.eu/eurostat/web/products-datasets/-/educ_uoe_lang01)
- FERNÁNDEZ VÍTORES, D. (2019). El Magreb hispano: de la frontera al espacio de intercambio. *Archiletras científica*, 2, 45-63.
- FUNDACIÓN SIGLO PARA EL TURISMO Y LAS ARTES DE CASTILLA Y LEÓN (2013). *II Plan del Español como Lengua Extranjera de la Comunidad de Castilla y León 2013-2016*. Junta de Castilla y León.

- GAY, M. DE 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.educacionyfp.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>
- GONZÁLEZ BARRERA, A., KROGSTAD, J. M. Y NOE BUSTAMANTE, L. (11 de febrero de 2020). *Path to legal status for the unauthorized is top immigration policy goal for Hispanics in U.S.* Pew Research Center. <https://www.pewresearch.org/fact-tank/2020/02/11/path-to-legal-status-for-the-unauthorized-is-top-immigration-policy-goal-for-hispanics-in-u-s/>
- 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](https://cvc.cervantes.es/lengua/anuario/anuario_18/gonzalez/p01.htm)
- GOVERN D'ANDORRA (2022). *Coneixements i usos lingüístics de la població d'Andorra. Situació actual i evolució 1995-2022*. Govern d'Andorra y Andorra Recerca + Innovació. [https://www.cultura.ad/images/stories/Llengua/Documents/Presentaci%C3%B3\\_Indicador2022\\_10-11-2022\\_compressed.pdf](https://www.cultura.ad/images/stories/Llengua/Documents/Presentaci%C3%B3_Indicador2022_10-11-2022_compressed.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. 209-211). Instituto Cervantes, Plaza & Janés y Círculo de Lectores. [http://cvc.cervantes.es/lengua/anuario/anuario\\_06-07/pdf/paises\\_41.pdf](http://cvc.cervantes.es/lengua/anuario/anuario_06-07/pdf/paises_41.pdf)
- INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE) (2024). *Sistema de Registro Nacional Migratório. SisMigra*. <https://portaldeimigracao.mj.gov.br/pt/dados/microdados/1733-obmigra/dados/microdados/401205-sismigra>
- INSTITUTO CERVANTES (2012). El español: una lengua viva. Informe 2012. En VV. AA., *El español en el mundo. Anuario del Instituto Cervantes 2012* (pp. 17-52). Instituto Cervantes y AEBOE. [https://cvc.cervantes.es/lengua/anuario/anuario\\_12/el\\_espanol\\_en\\_el\\_mundo\\_anuario\\_instituto\\_cervantes\\_2012.pdf](https://cvc.cervantes.es/lengua/anuario/anuario_12/el_espanol_en_el_mundo_anuario_instituto_cervantes_2012.pdf)
- 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](https://cvc.cervantes.es/Lengua/anuario/anuario_23/el_espanol_en_el_mundo_anuario_instituto_cervantes_2023.pdf)
- INSTITUTO DE ESTADÍSTICAS DE PUERTO RICO (2023). *Annual Estimates of the Resident Population for Puerto Rico Municipios: April 1, 2020 to July 1, 2022 - (PRM-EST2022-POP), Population Division - U.S. Census Bureau*. <https://www.census.gov/data/tables/time-series/demo/popest/2020s-total-puerto-rico-municipios.html>
- INSTITUTO DE ESTUDIOS PERUANOS (IEP) (2023). *Percepciones hacia las poblaciones indígenas u originarias*. CooperAcción y Coordinadora Nacional de Derechos Humanos (CNDDHH).

INSTITUTO NACIONAL DE ESTADÍSTICA DE BOLIVIA (2020). *Encuesta de hogares 2016-2018*. <https://www.ine.gob.bo/index.php/publicaciones/encuesta-de-hogares-2016-2018/>

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

INSTITUTO NACIONAL DE ESTADÍSTICA DE ESPAÑA (INE) (2023). *Población española residente en el extranjero a 1 de enero de 2023*. <https://www.ine.es/jaxi/Datos.htm?tpx=57845>

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

INSTITUTO NACIONAL DE ESTADÍSTICA DE GUATEMALA (2024). *Estimaciones y proyecciones nacionales de población 1950-2020*. <https://www.ine.gob.gt/proyecciones/>

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 HONDURAS (2013). *Proyecciones de población 2013-2050. XVII Censo Nacional de Población y VI de Vivienda*. [https://www.ine.gob.hn/publicaciones/Censos/Censo\\_2013/09Tomo-IX-Proyecciones-de-Poblacion/index%20Censo.html](https://www.ine.gob.hn/publicaciones/Censos/Censo_2013/09Tomo-IX-Proyecciones-de-Poblacion/index%20Censo.html)

INSTITUTO NACIONAL DE ESTADÍSTICA DE HONDURAS (2024). *BASEINE. Base de datos en línea*. <https://ine.gob.hn/v4/baseine/>

INSTITUTO NACIONAL DE ESTADÍSTICA DE LA REPÚBLICA BOLIVARIANA DE VENEZUELA (2015). *Censo nacional de población y vivienda 2011. Empadronamiento de la población indígena*. República Bolivariana de Venezuela.

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

INSTITUTO NACIONAL DE ESTADÍSTICA DE 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) (2022). *Población y vivienda*. <https://www.inei.gob.pe/estadisticas/indice-tematico/poblacion-y-vivienda/>

INSTITUTO NACIONAL DE ESTADÍSTICA Y CENSO DE PANAMÁ (INEC) (2024). *XII Censo Nacional de Población y VIII de Vivienda 2023*. [https://www.inec.gob.pa/publicaciones/Default3.aspx?ID\\_PUBLICACION=1199&ID\\_CATEGORIA=19&ID\\_SUBCATEGORIA=71](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) (2024). *Estadísticas y fuentes*. <https://inec.cr/estadisticas-fuentes>
- INSTITUTO NACIONAL DE ESTADÍSTICA Y CENSOS DE ECUADOR (INEC) (2022). *Censo Ecuador. Estadísticas*. <https://www.ecuadorencifras.gob.ec/estadisticas/>
- INSTITUTO NACIONAL DE ESTADÍSTICA Y CENSOS DE LA REPÚBLICA ARGENTINA (INDEC) (2023). *Censo Nacional de Población y Viviendas 2022. Indicadores demográficos, por sexo y edad*. INDEC. [https://www.indec.gob.ar/ftp/cuadros/poblacion/censo2022\\_indicadores\\_demograficos.pdf](https://www.indec.gob.ar/ftp/cuadros/poblacion/censo2022_indicadores_demograficos.pdf)
- 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*. Disponible en: [https://inegi.org.mx/app/saladeprensa/noticia.html?id=7519#:~:text=De%20los%2023.2%20millones%20de,millones%20\(69.2%20%25\)%2C%20no.&text=En%202020%2C%20se%20contabilizaron%202,35%20219%20141\)%20en%20M%C3%A9xico](https://inegi.org.mx/app/saladeprensa/noticia.html?id=7519#:~:text=De%20los%2023.2%20millones%20de,millones%20(69.2%20%25)%2C%20no.&text=En%202020%2C%20se%20contabilizaron%202,35%20219%20141)%20en%20M%C3%A9xico)
- INSTITUTO NACIONAL DE ESTADÍSTICA Y GEOGRAFÍA DE MÉXICO (INEGI) (2024). *Indicadores sociodemográficos*. <https://www.inegi.org.mx/>
- INSTITUTO NACIONAL DE ESTADÍSTICAS DE CHILE (2017a). *Censo de Población y Vivienda. 2017*. <https://www.inec.cl/estadisticas/sociales/censos-de-poblacion-y-vivienda/censo-de-poblacion-y-vivienda>
- INSTITUTO NACIONAL DE ESTADÍSTICAS DE CHILE (2017b). *Segundo Registro de Chilenos en el Exterior*. Instituto Nacional de Estadísticas y Ministerio de Relaciones Exteriores. <https://biblioteca.digital.gob.cl/items/8bbc271d-c64d-45f7-a233-44b06582f0bd>
- 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) (2005). *VIII Censo de Población y IV de Vivienda 2005*. [https://www.inide.gob.ni/docu/censos2005/ResumenCensal/Res\\_Anexos3.pdf](https://www.inide.gob.ni/docu/censos2005/ResumenCensal/Res_Anexos3.pdf)
- INSTITUTO NACIONAL DE INFORMACIÓN DE DESARROLLO DE NICARAGUA (INIDE) (2024). *Anuarios estadísticos*. <https://www.inide.gob.ni/>
- INSTITUTO PLURINACIONAL DE ESTUDIO DE LENGUAS Y CULTURAS DE BOLIVIA (IPELC) (2024). *Boletín informativo*. Ministerio de Educación. <https://prensa.ipelec.gob.bo/ipelec>
- INTERNET WORLD STATS (2023). *Internet World Stats. Usage and Population Statistics*. <http://www.internetworldstats.com>

- JAPANESE GOVERNMENT STATISTICS (2023). *Statistics of Japan*. [https://www.e-stat.go.jp/en/stat-search/database?page=1&layout=datalist&stat\\_infid=000032142402](https://www.e-stat.go.jp/en/stat-search/database?page=1&layout=datalist&stat_infid=000032142402)
- 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](https://cvc.cervantes.es/lengua/anuario/anuario_21/africa/argelia.htm)
- 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* (pp. 245-266). Instituto Cervantes y McGraw Hill. [https://cvc.cervantes.es/lengua/anuario/anuario\\_22/kabatek/p01.htm](https://cvc.cervantes.es/lengua/anuario/anuario_22/kabatek/p01.htm)
- KROGSTAD, J., PASSEL, J., 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/>
- LATINOS EN MOSCÚ (s. f.). *Inicio* [Página de Facebook]. Facebook. Recuperado en enero de 2024. [https://www.facebook.com/groups/LATINOSenRusia/?locale=es\\_LA](https://www.facebook.com/groups/LATINOSenRusia/?locale=es_LA)
- LEWIS, M. P., SIMONS, G. F. y FENNING, C. D. (Eds.) (2016). *Ethnologue. Languages of the World* (19.<sup>a</sup> edición). SIL International.
- 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](https://cvc.cervantes.es/lengua/espanol_europa/espanol_alemania_precovid/default.htm)
- LOUREDA LAMAS, Ó., 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., y PIHLER CIGLIČ, B. (EN PRENSA): DEMOLINGÜÍSTICA DEL ESPAÑOL EN EL SURESTE DE EUROPA.
- LUSIN, N., PETERSON, T., SULEWSKI, C. y ZAFER, R. (2023). *Enrollments in Languages Other Than English in United States Institutions of Higher Education, Fall 2021*. MLA. <https://www.mla.org/content/download/191324/file/Enrollments-in-Languages-Other-Than-English-in-US-Institutions-of-Higher-Education-Fall-2021.pdf>
- MEDIDA PROVISÓRIA 746 (23 de septiembre de 2016). *Diário Oficial da União* n.º 184-A. <https://pesquisa.in.gov.br/imprensa/jsp/visualiza/index.jsp?jornal=1000&pagina=1&data=23/09/2016>
- MINISTERIO DE CULTURA DE PERÚ (2017). *Política nacional de lenguas originarias, tradición oral e interculturalidad al 2040*. Ministerio de Cultura.

- MINISTERIO DE DESARROLLO SOCIAL Y FAMILIA DE CHILE (2013). *Encuesta de Caracterización socioeconómica nacional (CASEN)*. Ministerio de Desarrollo Social y Familia. <https://observatorio.ministeriodesarrollosocial.gob.cl/encuesta-casen-2013>
- MINISTERIO DE DESARROLLO SOCIAL Y FAMILIA DE CHILE (2017). *Pueblos indígenas. Síntesis de resultados. CASEN*. Disponible en: [https://observatorio.ministeriodesarrollosocial.gob.cl/storage/docs/casen/2017/Casen\\_2017\\_Pueblos\\_Indigenas.pdf](https://observatorio.ministeriodesarrollosocial.gob.cl/storage/docs/casen/2017/Casen_2017_Pueblos_Indigenas.pdf)
- 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. <http://www.educacionyfp.gob.es/mc/redele/el-mundo-estudiaespa-ol/2020.html>
- MINISTERIO DE EDUCACIÓN Y FORMACIÓN PROFESIONAL (MEFP) (2022). *El mundo estudia español 2022*. Secretaría General Técnica del Ministerio de Educación. <http://www.educacionyfp.gob.es/mc/redele/el-mundo-estudiaespa-ol/2022.html>
- 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-ethnicity/2023/09/20/latinos-views-of-and-experiences-with-the-spanish-language/#:~:text=Latino%20identity%20in%20the%20U.S.,while%2021%25%20say%20it%20is>
- MORENO CABRERA, J. C. (2003). *El universo de las lenguas*. Castalia.
- MORENO FERNÁNDEZ, F. (2020). La internacionalización del español y su análisis. En L. Antero Reto y R. Gutiérrez Rivilla (Coords.), *La proyección internacional del español y el portugués: el potencial de la proximidad lingüística / Projeção mundial do espanhol e do português: o potencial da proximidade lingüística* (pp. 59-86). INCM, Instituto Cervantes e Instituto Camões.
- MORENO FERNÁNDEZ, F. (2023). *La lengua de los hispanos unidos de América*. Catarata.
- MORENO FERNÁNDEZ, F. y ÁLVAREZ MELLA, H. (2022). Reexamining the international importance of languages. *HCIAS Working Papers on Ibero-America*, 1, 1-23. <https://doi.org/10.48629/hcias.2021.1.84517>
- 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*. Instituto Complutense de Estudios Internacionales y Fundación Telefónica. <https://eprints.ucm.es/8936/1/DT03-06.pdf>

MORENO FERNÁNDEZ, F. y OTERO ROTH, J. (2007). *Atlas de la lengua española en el mundo* (3.<sup>a</sup> ed. en 2016). Ariel y Fundación Telefónica.

NATIONAL CENTER FOR EDUCATION STATISTICS U.S. (2021). *Enrollment and percentage distribution of students enrolled in private elementary and secondary schools, by school orientation and grade level: Selected years, fall 2009 through fall 2019*. [https://nces.ed.gov/programs/digest/d21/tables/dt21\\_205.20.asp](https://nces.ed.gov/programs/digest/d21/tables/dt21_205.20.asp)

NATIONAL CENTER FOR EDUCATION STATISTICS U.S. (2023). *Fast Facts. Back to school statistics*. <https://nces.ed.gov/fastfacts/display.asp?id=372#PK12-enrollment>

OBSERVATORIO REGIONAL DE DERECHOS DE LOS PUEBLOS INDÍGENAS (2020). *Revitalización de lenguas indígenas*. FILAC. <https://www.ordpi.org/wp-content/uploads/2021/03/Informe-Regional-Lenguas-Indigenas-Derechos-Intercult-2.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>

OFICINA DEL CENSO DE ISRAEL (2022). *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>

OFICINA DEL CENSO DE LOS ESTADOS UNIDOS (2021a). *American Community Survey. S1601. Language spoken at home*. <https://data.census.gov/table?q=S1601:+LANGUAGE+SPOKEN+AT+HOME&t=Language+Spoken+at+Home&g=0100000US>

OFICINA DEL CENSO DE LOS ESTADOS UNIDOS (2021b). *American Community Survey. B16006. Language spoken at home by ability to speak English for the population 5 years and over (hispanic or latino)*. <https://data.census.gov/table?q=Language+Spoken+at+Home+latino&t=Hispanic+or+Latino:Language+Spoken+at+Home&tid=ACSDT1Y2021.B16006>

OFICINA DEL CENSO DE LOS ESTADOS UNIDOS (2022a). *American Community Survey. DP02. Selected Social Characteristics in the United States*. <https://data.census.gov/table/ACSDP5Y2022.DP02>

OFICINA DEL CENSO DE LOS ESTADOS UNIDOS (2022b). *2020 Island Areas Censuses Data on Demographic, Social, Economic and Housing Characteristics Now Available for Guam*. <https://www.census.gov/newsroom/press-releases/2022/2020-island-areas-guam.html#:~:text=The%20Chamorro%20population%20was%20the,Islander%20group%20or%20race%20group>

OFICINA DEL CENSO DE LOS ESTADOS UNIDOS (2023). *USVI 2010 Census Detailed Crosstabulations Part1 v3.xlsx Archived July 8, 2022, at the Wayback Machine (Tablas 2-9)*. [https://www2.census.gov/census\\_2010/10-Island\\_Areas\\_Detailed\\_Cross\\_Tabulations/Virgin\\_Islands/](https://www2.census.gov/census_2010/10-Island_Areas_Detailed_Cross_Tabulations/Virgin_Islands/)

- OFICINA ESTADÍSTICA DE FILIPINAS (2010). *2010 Census of Population and Housing* [https://psa.gov.ph/system/files/main-publication/NCR\\_Report%2520No.1.pdf](https://psa.gov.ph/system/files/main-publication/NCR_Report%2520No.1.pdf)
- OFICINA NACIONAL DE ESTADÍSTICA DE LA REPÚBLICA DOMINICANA (ONE) (2023). *Indicadores principales*. <https://www.one.gob.do/>
- OFICINA NACIONAL DE ESTADÍSTICA E INFORMACIÓN DE CUBA (ONEI) (2023). *Anuario Estadístico de Cuba 2022*. Oficina Nacional de Estadística e Información.
- OFICINA NACIONAL DE ESTADÍSTICA Y CENSOS DE EL SALVADOR (ONEC) (2024). *Población y estadísticas demográficas*. <https://onec.bcr.gob.sv/poblacion-y-estadisticas-demograficas/>
- ORGANIZACIÓN DE NACIONES UNIDAS (ONU) (2020). *International Migrant Stock 2020*. Departamento de Asuntos Económicos y Sociales. Division de Población. <https://www.un.org/development/desa/pd/content/international-migrant-stock>
- ORGANIZACIÓN DE NACIONES UNIDAS (ONU) (2022). *World Population Prospects 2022*. <https://population.un.org/wpp/>
- ORGANIZACIÓN DE NACIONES UNIDAS (ONU) (2024). *Peace, dignity and equality on a healthy planet*. <https://www.un.org/en/global-issues/population#:~:text=The%20global%20human%20population%20reached,and%202%20billion%20since%201998>.
- OTERO ROTH, J. (1995). Una nueva mirada al índice de importancia internacional de las lenguas. En Marqués de Tamarón (Dir.), *El peso de la lengua española en el mundo* (pp. 235-282). Fundación Duques de Soria, INCIPE y Universidad de Valladolid. [https://cvc.cervantes.es/lengua/peso\\_lengua/otero.htm](https://cvc.cervantes.es/lengua/peso_lengua/otero.htm)
- PRADO, D. (2021). *Lenguas indígenas de los países iberoamericanos. Datos y gráficos*. Organización de Estados Iberoamericanos. <https://oei.int/oficinas/secretaria-general/publicaciones/lenguas-indigenas-de-los-paises-iberoamericanos-datos-y-graficos-2>
- PUJOL RIEMBAU, Ó. (2020). El español en la India: una lengua favorecida. En VV. AA., *El español en el mundo 2020. Anuario del Instituto Cervantes* (pp. 389-424). Instituto Cervantes y Bala Perdida S. L. [https://cvc.cervantes.es/lengua/anuario/anuario\\_20/pujol/p01.htm](https://cvc.cervantes.es/lengua/anuario/anuario_20/pujol/p01.htm)
- RAMÓN, P. (5 de febrero de 2024). *Latinos in the US: A shifting Voting Bloc*. Politiko. <https://politiko.com.ph/2024/02/05/latinos-in-the-us-a-shifting-voting-bloc/politiko-global/>
- REDATAM-CENTRO LATINOAMERICANO Y CARIBEÑO DE DEMOGRAFÍA (2023). *Censos Población y Vivienda*. <https://redatam.org/es/procesar-en-linea/censos/pobviv>
- SANTOS ROVIRA, J. M. Y SERRANO LUCAS, C. (2022). *Demolingüística del español en Portugal*. Instituto Cervantes, Universidad de Heidelberg y Universidad de Zúrich. [https://cvc.cervantes.es/lengua/espanol\\_europa/espanol\\_portugal/default.htm](https://cvc.cervantes.es/lengua/espanol_europa/espanol_portugal/default.htm)

- SERRANO AVILÉS, J. (ED.) (2014). La enseñanza del español en el 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/capitulo1.htm>
- STATISTA (2024). *Spoken languages of Bonaire, Saint-Eustatius and Saba 2021*. <https://www.statista.com/statistics/1377799/caribbean-netherlands-spoken-languages-2021/#:~:text=Papiamentu%2C%20a%20Portuguese%2Dbased%20language,88%20percent%20of%20the%20population>
- STATISTICAL INSTITUTE OF BELIZE (SIB) (2022). *Belize Population and Housing Census 2022*. <https://sib.org.bz/census-data/>
- STATISTICS CANADA (2022). *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 CANADA (2023a). *Table 37-10-0007-01 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 (2023b). *Table 37-10-0018-01 Postsecondary enrolments, by registration status, institution type, status of student in Canada and gender*. <https://doi.org/10.25318/3710001801-eng>
- STATISTIK AUSTRIA (2023). *Fremdsprachenunterricht der Schüler: innen im Schuljahr 2021/22*. <https://www.statistik.at/statistiken/bevoelkerung-und-soziales/bildung/schulbesuch/schuelerinnen>
- STATISTISCHES BUNDESAMT DEUTSCHLAND (2024). *Schüler mit Fremdsprachen-Unterricht*. <https://www-genesis.destatis.de/genesis//online?operation=table&code=21111-0006&bypass=true&levelindex=1&levelid=1709031181542#abreadcrumb>
- TANO, M. (2023). Dinámicas de la lengua española en los espacios de la educación superior francesa. *Archiletras científica*, 10, 147-163.
- THE BUREAU OF STATISTICS AND PLANS. GOVERNMENT OF GUAM (2020). *The 2020 Census of Guam*. <https://bsp.guam.gov/2020-census/>
- UNICAMP (2024). *Inmigrantes internacionais registrados no Brasil*. Observatório das Migrações em São Paulo. <https://www.nepo.unicamp.br/observatorio/bancointerativo/numeros-imigracao-internacional/sincretismo-migra/>
- UNICEF (2014). *Panorama de la situación de la niñez y adolescencia indígena en América Latina*. Oficina Regional para América Latina y el Caribe. <https://www.unicef.org/lac/informes/panorama-de-la-situacion-de-la-ninez-y-adolescencia-indigena-en-america-latina>

- UNIVERSIDAD CATÓLICA ANDRÉS BELLO (2022). *Condiciones de vida de los venezolanos. ENCOVI 2022*. Universidad Católica Andrés Bello. [https://assets.website-files.com/5d14c6a5c4ad42a4e794d0f7/636d0009b0c59ebfd2f24acd\\_Presentacion%20ENCOVI%202022%20completa.pdf](https://assets.website-files.com/5d14c6a5c4ad42a4e794d0f7/636d0009b0c59ebfd2f24acd_Presentacion%20ENCOVI%202022%20completa.pdf)
- VENEZOLANOS EN INDIA (S. F.). INICIO [PÁGINA DE FACEBOOK]. FACEBOOK. RECUPERADO EN ENERO DE 2024. <https://www.facebook.com/venezolanosenindia/>
- VV. AA. (1998). *El español en el mundo. Anuario del Instituto Cervantes 1998*. Instituto Cervantes y Arco Libros. [https://cvc.cervantes.es/lengua/anuario/anuario\\_98](https://cvc.cervantes.es/lengua/anuario/anuario_98)
- VV. AA. (2006). *Enciclopedia del español en el mundo. Anuario del Instituto Cervantes 2006-2007*. Instituto Cervantes, Plaza & Janés y Círculo de Lectores. [https://cvc.cervantes.es/lengua/anuario/anuario\\_06-07](https://cvc.cervantes.es/lengua/anuario/anuario_06-07)
- VV. AA. (2008). *Enciclopedia del español en los Estados Unidos. Anuario del Instituto Cervantes 2008*. Instituto Cervantes y Santillana. [https://cvc.cervantes.es/lengua/anuario/anuario\\_08](https://cvc.cervantes.es/lengua/anuario/anuario_08)
- VV. AA. (2009). *El español en el mundo. Anuario del Instituto Cervantes 2009*. Instituto Cervantes y BOE. [https://cvc.cervantes.es/lengua/anuario/anuario\\_09/default.htm](https://cvc.cervantes.es/lengua/anuario/anuario_09/default.htm)
- VV. AA. (2011). *El español en el mundo. Anuario del Instituto Cervantes 2010-2011*. Madrid: Instituto Cervantes y BOE. [https://cvc.cervantes.es/lengua/anuario/anuario\\_10-11/default.htm](https://cvc.cervantes.es/lengua/anuario/anuario_10-11/default.htm)
- W3TECHS (2024). *Usage statistics of content languages for websites*. [https://w3techs.com/technologies/overview/content\\_language](https://w3techs.com/technologies/overview/content_language)
- ZIELINSKI, A., BERTA, T., MEDVECZKÁ, M., ŠTRBÁKOVÁ, R., TATOJ, C., PESKOVA, J. y ÁLVAREZ MELLA, H. (en prensa). *Demolingüística del español en Europa CENTRAL*. Instituto Cervantes, Universidad de Heidelberg y Universidad de Zúrich.

## II. ANÁLISIS COMPLEMENTARIOS

La transversalidad de la lengua española hace posible que el análisis de su situación en el mundo trascienda la demografía para interesarse por su presencia y uso en muy diversos sectores de la actividad social y de la vida comunitaria. En este informe, se prestará atención a cuatro de estos sectores o ámbitos, todos ellos de gran importancia para la sociedad y para la lengua misma.

El primero de ellos es el de los medios de comunicación social, cuya presencia y capacidad de influencia sobre la sociedad se han ido incrementando conforme los soportes se han ido actualizando y diversificando. En el año del centenario de la radio en España, no deja de sorprender la vitalidad de los medios que se transmiten en formatos de audio.

El segundo de los informes sectoriales está dedicado a la traducción. Este viejo proceso de traslación sigue siendo fundamental para la expresión y transmisión de la cultura y el conocimiento, si bien los desarrollos tecnológicos están obligando a su adaptación a nuevas formas de practicarlo.

El tercer informe complementario que aquí se ofrece se dedica a la situación actual de las industrias culturales. Estas industrias incluyen muy distintos campos de la información, la formación y el entretenimiento, y demuestran su naturaleza esencial en todo tipo de circunstancias.

Finalmente, estos análisis complementarios se cierran con un texto dedicado a presentar un panorama de la situación actual de la inteligencia artificial en español. Para el futuro de esta lengua en la esfera digital, resulta clave conocer qué lugar ocupa la lengua española en el desarrollo de las tecnologías del lenguaje y cuáles son sus principales carencias.

# EL ESPAÑOL EN LOS MEDIOS DE COMUNICACIÓN

Luis Miguel Pedrero Esteban  
UNIVERSIDAD FRANCISCO DE VITORIA

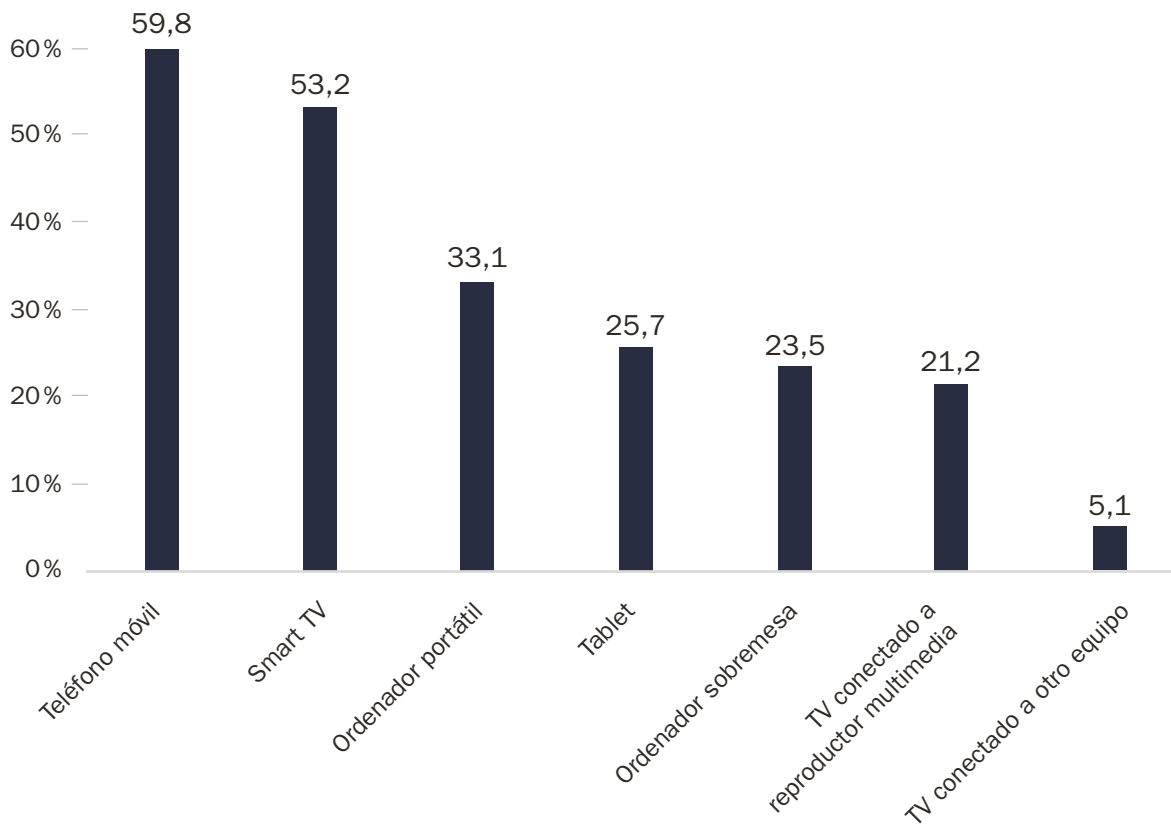
## 1 | Acotaciones, contextos y reglas básicas del español en los medios

Como precisa con rigor la académica M.<sup>a</sup> Victoria Romero Gualda, la referencia al español en los medios de comunicación debe considerarse inscrita en una denominación más amplia, la del lenguaje periodístico: «el propio, usado o creado por los periodistas, aparecido en los periódicos y en otros medios de comunicación, que contribuye a que el periodismo sea una actividad profesional caracterizable por su actuación lingüística diferenciada» (Romero Gualda, 1993, p. 9). Ello no implica que el español en los medios deba ser considerado como un lenguaje sectorial o marginal, pues debido a su alcance es el que más influye en el hablante común, incluso con mayor influjo en el contemporáneo ecosistema digital.

De hecho, la normalización de la conectividad en línea y la convergencia en los dispositivos móviles de todo tipo de contenidos (Asociación para la Investigación de Medios de Comunicación [AIMC], 2024) ha rebajado la penetración y el alcance de los medios tradicionales –prensa, radio y televisión– como fuentes de acceso preferentes a la información y el entretenimiento (gráfico 1). En su lugar, ganan presencia las redes sociales, las plataformas de distribución de vídeos y, de forma cada vez más extendida, los formatos de audio hablado (gráfico 2).

Gráfico 1 · Página siguiente >

GRÁFICO 1 | EQUIPO DE VISIONADO DE ACTIVIDADES AUDIOVISUALES POR INTERNET EN LOS ÚLTIMOS 30 DÍAS (%)

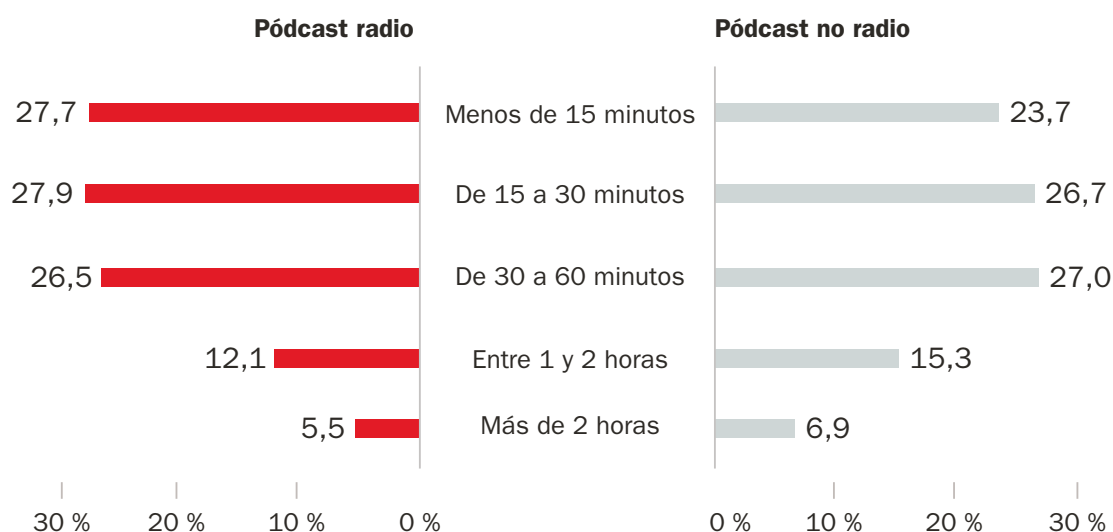
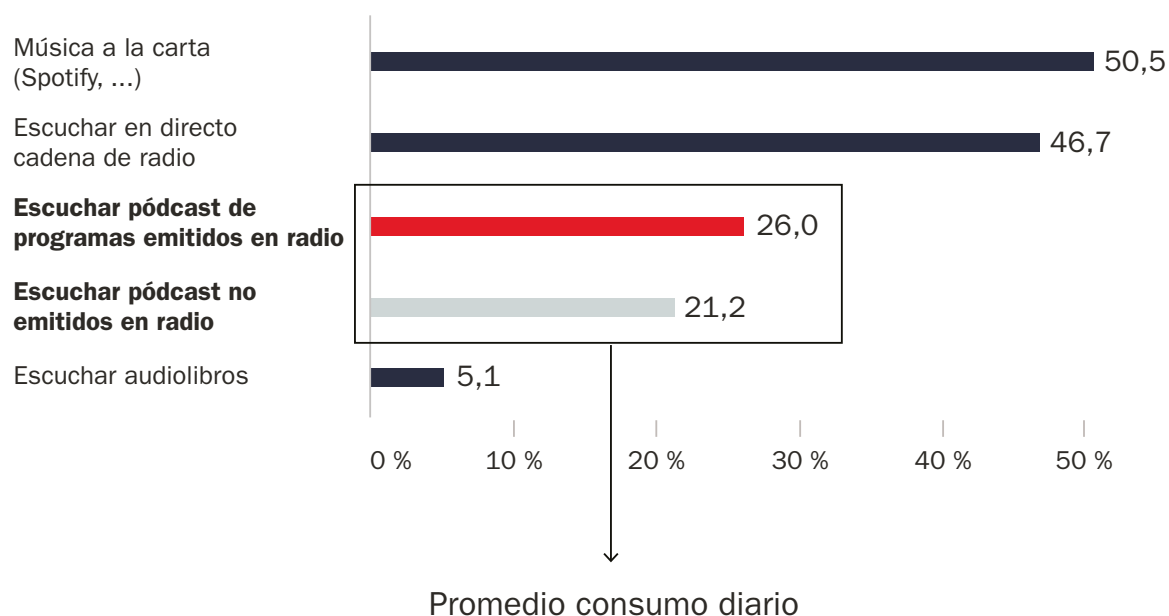


FUENTE: elaboración propia a partir de la encuesta «Navegantes en la Red». 26.ª edición (AIMC, 2024).

Coincidiendo con la celebración en todo el mundo de los primeros cien años de la radio, el medio sonoro por excelencia, asistimos a la eclosión de la que se ha denominado «era de la audificación»: frente al crecimiento exponencial de propuestas formativas y lúdicas asociadas a pantallas fijas y móviles, la popularización del pódcast, la revitalización de los audiolibros y el creciente peso de la voz como interfaz de conexión a internet a través de los asistentes virtuales han devuelto al lenguaje hablado su relevancia como primigenio sistema de comunicación y transmisión del conocimiento (Pedrero Esteban, 2023).

Gráfico 2 · Página siguiente >

GRÁFICO 2 | AUDIO POR INTERNET EN LOS ÚLTIMOS 30 DÍAS (%)



**FUENTE:** elaboración propia a partir de la encuesta «Navegantes en la Red». 26.ª edición (AIMC, 2024).

De ahí que, sin restar significación a la trascendencia del español en los medios impresos, este capítulo aborde con especial profundidad el horizonte que los canales y tecnologías de distribución del audio –hablado y audiovisual– brindan a la expansión de un idioma que en 2023 dominan de forma nativa casi 500 millones de personas en todo el mundo, y que pueden utilizar potencialmente más de 600 millones. Un idioma que, tanto en

su forma escrita como oral, asume reglas muy concretas en la plasmación de los datos, hechos o situaciones relacionadas con fines informativos o de entretenimiento:

- a. *Claridad y precisión.* A los periodistas se les exige el uso de un lenguaje preciso y claro para garantizar que los contenidos resulten comprensibles para el público en general. En el caso de los medios impresos, ese requisito atiende fundamentalmente a la gramática y a la sintaxis, y en los medios orales se extiende, además, a una locución pulcra y una entonación natural, que marque las inflexiones y pausas para una comprensión inmediata de los mensajes que el espectador u oyente escuchan en las programaciones de TV y radio: así como un lector puede releer un texto de forma autónoma, la linealidad de una emisión audiovisual condiciona la recepción y asimilación del contenido, de ahí la conveniencia de su inteligibilidad.
- b. *Formalidad.* Aunque el tono empleado en la redacción o locución de un contenido periodístico depende del tipo de medio, su línea editorial, su especialización, su género (noticia, reportaje, crónica, entrevista, artículo de opinión...) o el público al que se dirige, el lenguaje que se lee o se escucha en los medios eleva los estándares de expresiones o conversaciones coloquiales. Si bien la proliferación y el impacto de las redes sociales y las plataformas de distribución de contenidos audiovisuales generados por usuarios han rebajado la formalidad de ciertos periódicos y programas de radio y TV, esta cualidad es intrínseca a la calidad y credibilidad de un medio.
- c. *Objetividad.* Una de las premisas consustanciales al ejercicio periodístico remite a la obligatoriedad de un conocimiento exacto de los hechos y datos relatados, lo cual se traduce en una presentación imparcial y neutra de la información, sin la inclusión de valoraciones por parte del redactor. A pesar de que la objetividad periodística es un concepto cuya esencia filosófica convierte su aplicación en una utopía (Restrepo, 2001), sí puede –debe– concretarse en la ausencia de adjetivos, adverbios u

otras expresiones que connoten, maticen o introduzcan cualquier sesgo que oriente de un modo u otro la interpretación de la noticia.

- d. *Neologismos y anglicismos.* La influencia de otros idiomas en la comunicación, en especial el inglés, ha normalizado la inclusión de neologismos y anglicismos en el lenguaje habitual de los medios, sobre todo en áreas como la tecnología, la moda o el entretenimiento; esa tendencia se ha visto reforzada con la popularización de las redes sociales y los canales digitales. Si bien los términos neológicos enriquecen la lengua (Grijelmo García, 2022), no siempre mejoran el estilo o facilitan el entendimiento: de ahí la conveniencia de aclarar su significado para que los lectores, espectadores u oyentes comprendan cada palabra y no abandonen ante esa barrera léxica.
- e. *Respeto a la diversidad lingüística y cultural.* Sin perjuicio de lo indicado sobre un uso moderado de neologismos, los periodistas han de ser conscientes de las distintas variedades del español y de las particularidades lingüísticas de cada región; por ello, han de evitar los estereotipos o discriminar los términos propios de los entornos en los que tienen lugar los hechos relatados, aclarando, una vez más, el significado y sentido con el que dichos términos son utilizados. En todo caso, y como práctica transversal y aplicable a todas las reglas enunciadas, los comunicadores siempre han de adaptar el lenguaje y el estilo al medio específico y al público objetivo; lo ideal es utilizar un tono y un registro apropiados al formato y audiencia de cada medio.

## 2 | Factores que condicionan la evolución del español en los medios

Además de la irrupción de internet, son muchos y muy heterogéneos los factores que han incidido en la evolución del uso del idioma español en los medios de comunicación; por sus efectos sobre las rutinas de creación, comercialización y distribución del contenido periodístico, cabe destacar las siguientes dinámicas:

- a. *Expansión de los medios.* El desarrollo de la imprenta, más tarde el de la radio y la televisión y, con el nuevo siglo, el de internet, han multiplicado de manera casi exponencial el alcance y la accesibilidad de los medios de comunicación en todos los idiomas. El 66,2 % de la población mundial (5.350 millones de personas) tienen acceso a la red en 2024 y más de 5.000 millones (62,3 %) disponen de, al menos, un perfil en redes sociales (Kemp, 2024). Son indicadores que no dejan de crecer, y que afianzan la influencia de los medios a través de cada vez más tecnologías y canales de distribución.
- b. *Globalización y diversificación.* La globalización impulsada por la normalización de la conectividad y la exigencia de inmediatez en las demandas de información y entretenimiento cada vez más adaptadas al perfil de cada usuario han favorecido la creación de medios analógicos y digitales cuya creciente y cada vez más variada especialización permite atender a audiencias dispersas por todo el mundo. Solo en España se contabilizaban en 2023 casi 3.000 medios digitales activos: ediciones de diarios en internet, sitios web de revistas, emisoras de radio y canales de TV, y un creciente número de medios nativos, es decir, nacidos en internet (Salaverría Aliaga, 2023). Como se muestra en el cuadro 1, la mayoría se concentran en Madrid y Cataluña.

| CUADRO 1   DISTRIBUCIÓN DE MEDIOS DIGITALES POR COMUNIDADES AUTÓNOMAS EN ESPAÑA (2023) |                                 |                               |                      |                                |
|----------------------------------------------------------------------------------------|---------------------------------|-------------------------------|----------------------|--------------------------------|
|                                                                                        | medios digitales activos (2023) | % medios digitales nacionales | % población nacional | medios digitales vs. población |
| Andalucía                                                                              | 294                             | 10,0 %                        | 18,0 %               | - 8,0 %                        |
| Aragón                                                                                 | 73                              | 2,5 %                         | 2,8 %                | - 0,3 %                        |
| Asturias                                                                               | 45                              | 1,5 %                         | 2,1 %                | - 0,3 %                        |
| Baleares                                                                               | 67                              | 2,3 %                         | 2,6 %                | - 0,3 %                        |
| Canarias                                                                               | 80                              | 2,7 %                         | 4,7 %                | - 2,0 %                        |
| Cantabria                                                                              | 41                              | 1,4 %                         | 1,2 %                | - 0,2 %                        |
| Castilla-La Mancha                                                                     | 117                             | 4,0 %                         | 4,3 %                | - 0,3 %                        |

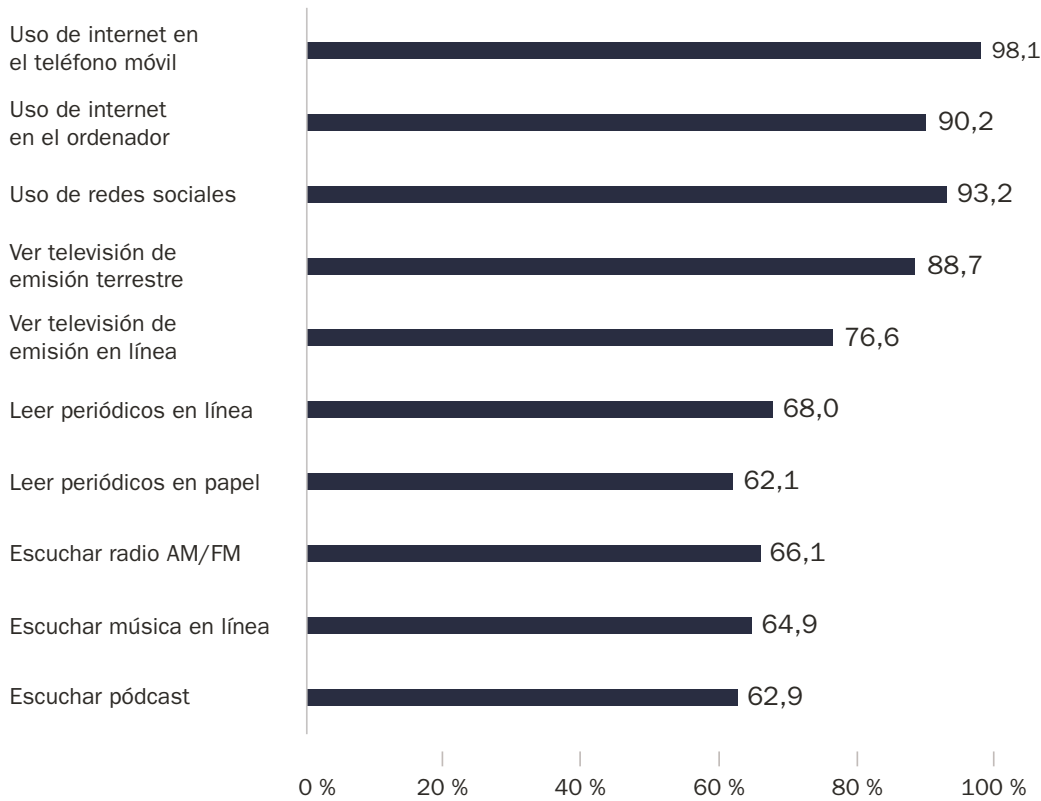
|                         | medios digitales<br>activos (2023) | % medios digitales<br>nacionales | % población<br>nacional | medios digitales<br>vs. población |
|-------------------------|------------------------------------|----------------------------------|-------------------------|-----------------------------------|
| Castilla y León         | 138                                | 4,7 %                            | 5,0 %                   | - 0,3 %                           |
| Cataluña                | 585                                | 20,0 %                           | 16,2 %                  | + 3,8 %                           |
| Comunidad<br>Valenciana | 203                                | 6,9 %                            | 10,7 %                  | - 3,8 %                           |
| Extremadura             | 56                                 | 1,9 %                            | 2,2 %                   | - 0,2 %                           |
| Galicia                 | 148                                | 5,1 %                            | 5,7 %                   | - 0,6 %                           |
| La Rioja                | 27                                 | 0,9 %                            | 0,7 %                   | - 0,2 %                           |
| Madrid                  | 738                                | 25,2 %                           | 14,2 %                  | + 11,0 %                          |
| Murcia                  | 57                                 | 1,9 %                            | 3,2 %                   | - 1,3 %                           |
| Navarra                 | 58                                 | 2,0 %                            | 1,4 %                   | + 0,6 %                           |
| País Vasco              | 133                                | 4,5 %                            | 4,6 %                   | - 0,1 %                           |
| Ceuta                   | 10                                 | 0,3 %                            | 0,2 %                   | + 0,1 %                           |
| Melilla                 | 3                                  | 0,1 %                            | 0,2 %                   | - 0,1 %                           |

**FUENTE:** elaboración propia a partir de datos de Salaverría Aliaga (2023).

- c. *Innovación tecnológica.* La continua renovación de los sistemas de producción, los canales de distribución y los soportes de acceso afecta directamente al modo en que se producen, distribuyen y consumen los medios. El dispositivo dominante es el *smartphone*, que utiliza un 98,1 % de los usuarios para acceder a los contenidos digitales; un 90,2 % emplea el ordenador fijo o portátil; el 93,2 % lo hace a través de las redes sociales; ven la TV tradicional el 88,7 % y la conectada el 76,6 %; los periódicos digitales son leídos por el 68 % de la población mundial, mientras que los diarios impresos por el 62,1 %; escuchan radio en FM dos de cada tres usuarios (66,1 %), música en plataformas de *streaming* un 64,9 %, y pódcast un 62 % (gráfico 3).

Gráfico 3 · Página siguiente >

GRÁFICO 3 | CONSUMO DE CONTENIDOS DIGITALES EN INTERNET POR USUARIOS DE 16 A 64 AÑOS (%)



FUENTE: elaboración propia a partir de datos de Digital 2024. Global Overview Report (Kemp, 2024).

d. *Profesionalización y estándares de calidad.* A medida que la industria de los medios ha crecido y se ha internacionalizado, los editores y operadores han puesto un énfasis creciente en la profesionalización y el cumplimiento de estándares de calidad en la producción de contenidos impresos y audiovisuales, de modo que las diferencias en otro tiempo tan acusadas entre las compañías estadounidenses o de países europeos como el Reino Unido y Alemania, y las del resto de territorios en los mismos continentes, se han diluido tanto en los procesos y rutinas como en su estética y puesta en escena. La proliferación de plataformas digitales ha permitido, además, el desarrollo de nuevos formatos y estilos de narración, especialmente en las plataformas de vídeo y las redes sociales, cuyos códigos expresivos se han ido extendiendo a los canales tradicionales hasta impulsar la renovación de sus lógicas productivas y la incorporación de nuevos perfiles creativos y técnicos.

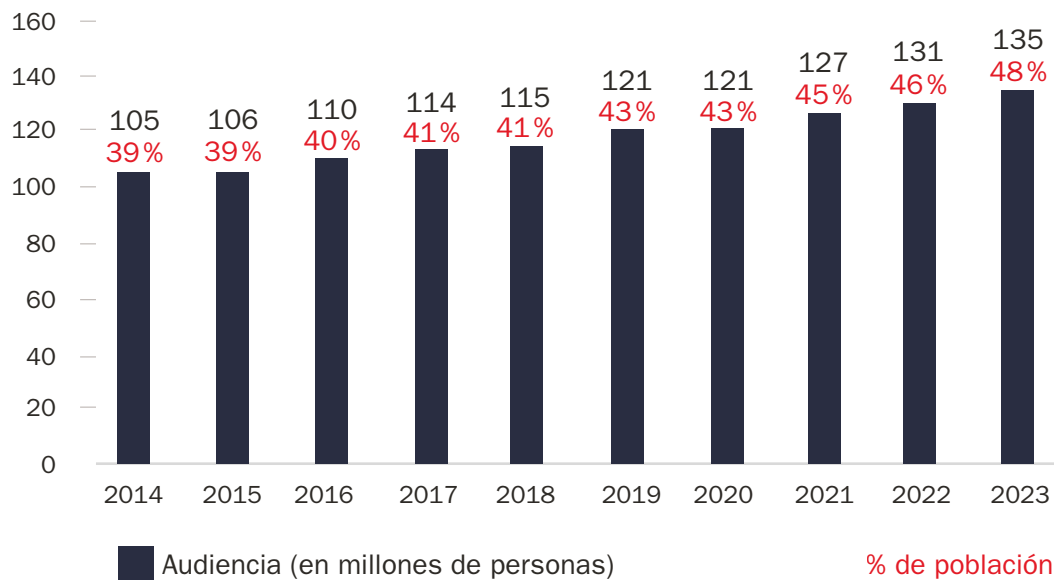
### 3 | La revalorización del audio hablado en los medios

El imparable caudal de productos digitales que circulan por la red y están disponibles en cualquier instante o lugar y a través de cualquier dispositivo ha redefinido la relación de los lectores, oyentes y espectadores con los medios de comunicación; se ha difuminado el impacto de las parrillas de radio y televisión, que durante muchos años impusieron los horarios y moldearon los hábitos de consumo. Aquella función propia y casi exclusiva de los medios analógicos la asumen en la actualidad las plataformas, espacios virtuales que configuran inmensos repertorios de servicios interactivos donde, según sus respectivas motivaciones, cada usuario selecciona el contenido, su apariencia, su frecuencia y hasta su velocidad de reproducción (Zuboff, 2019).

Más allá de percepciones o cálculos prospectivos, los indicadores en torno a la creciente presencia del audio en la dieta cotidiana de consumo de contenidos digitales resultan cada vez más contundentes (Edison Research, 2024). En un entorno de propuestas informativas y de entretenimiento indisolublemente ligadas a las pantallas fijas y móviles, los formatos sonoros están arraigando con fuerza en gran parte de la población por su ductilidad y por su compatibilidad con la realización de otras actividades, un atributo muy valorado en la obsesiva y generalizada búsqueda por maximizar el tiempo dedicado al conocimiento, la formación y la evasión.

Apoyada en la multiplicación de plataformas y en la proliferación de variantes, la escucha de audio hablado (radio, pódcast y audiolibros) se multiplica y alcanza niveles que no se habían registrado antes de la aparición de internet. En Estados Unidos, casi la mitad de la población mayor de 13 años (48 %, 135 millones de personas) dedica atención diaria hacia algún contenido sonoro basado en la palabra (gráfico 4); de hecho, cerca de la tercera parte del tiempo cotidiano de escucha (31 %) se dirige a la radio generalista, los pódcast o los audiolibros, lo que supuso en 2023 un aumento del 55 % con respecto a 2014. La demanda de este contenido se realiza sobre todo en casa (60 %), y en menor medida en el automóvil (14 %) o el trabajo (13 %), según destaca *The Spoken Word Audio Report* (National Public Media y Edison Research, 2023).

GRÁFICO 4 | EVOLUCIÓN DE LA AUDIENCIA DE AUDIO HABLADO EN ESTADOS UNIDOS  
(RADIO GENERALISTA, PÓDCAST, AUDIOLIBROS)



**FUENTE:** elaboración propia a partir de datos de The Spoken Word Audio Report (National Public Media y Edison Research, 2023).

España es uno de los países occidentales donde los internautas consumen más pódcast: lidera la escucha en Europa con una penetración del 45 %, superior en nueve puntos a la media de 24 países en el mundo consignados en ese estudio (Amoedo Casals, 2023). El perfil del consumidor es el de un varón (casi el 58 %, frente al 40 % de mujeres) universitario (casi el 40 % de los usuarios), que dedica al día una hora y veinte minutos. Un 42,5 % los escuchan a diario y un 28,1 % con periodicidad semanal. La mayoría (58 %) oye los episodios de principio a fin, mientras un 41,27 % lo hace de forma fragmentada, es decir, interrumpe la escucha en función de su rutina.

El 40 % de los oyentes españoles están suscritos a entre 4 y 10 pódcast, frente a un 20,7 % que sigue con asiduidad hasta 20 títulos. Por géneros, los pódcast que más interesan son los conversacionales (casi 60 %) y los documentales (46,78 %). En cuanto a los temas, la mayoría se decanta por la historia, la comedia y el humor, la investigación periodística, el análisis político y la salud o el bienestar. Más del 64 % de los usuarios asegura escuchar con auriculares y el 48 % lo hace a través de su *smartphone* (iVoox, 2022).

Uno de los hallazgos que mejor ilustran el cambio de paradigma en la relación de los ciudadanos con los medios de comunicación y con el lenguaje para estar informados y/o entretenidos es que, por primera vez, el *smartphone* se ha convertido en el receptor más utilizado para escuchar la palabra hablada: el 39 % de ese contenido en Estados Unidos ya se recibe a través de ese dispositivo; ocurre en el hogar y también en el trabajo, y el único espacio donde aún reina el transistor de AM/FM es el vehículo. No se trata solo de una cuestión técnica, ni tampoco económica, pese a que sintonizar la radio desde el móvil exige de una conexión a internet (con el consiguiente coste de contratación); lo realmente disruptivo es que, en los terminales de pantalla, la naturaleza del lenguaje radiofónico trasciende su dimensión sonora y se asocia a la visualización de contenidos con una significación que va más allá de lo auditivo.

Se altera así la centenaria esencia de un medio que nació basado en atraer la atención del oyente a través del oído y se centra ahora en la vista: en el entorno digital el audio primero se ve..., y luego se escucha, de forma que lo sonoro deviene en subsidiario de la apariencia con la que se presenta. Semejante evidencia justifica cuantas iniciativas intenta explotar la radio para ramificarse por entre los múltiples canales de distribución en los que puede seguir sonando: las cámaras en unos estudios que se asemejan hoy a platós de televisión, las fotografías, vídeos, gráficos o tarjetas sonoras (audiogramas) incrustados en las redes sociales, o las transmisiones en vivo a través de plataformas como YouTube o Twitch, en las que la programación irrumpe como imagen e interpela desde la narrativa visual.

Esa tendencia se ha extendido también a una de las tres grandes variantes de género del pódcast, el denominado «conversacional», basado en entrevistas, charlas, tertulias y hasta monólogos en torno a temáticas de enorme amplitud y variedad (Pedrero Esteban *et al.*, 2022): la fórmula de vídeo y pódcast se está demostrando ideal para atraer a la escucha a la generaciones más jóvenes, habituadas a consumir contenidos audiovisuales precisamente por la preponderancia del *smartphone* como soporte de conexión a internet.

Frente a esa categoría cobran creciente reconocimiento las otras dos variantes del pódcast: las denominadas «narrativos de ficción» y «narrativos de no ficción». Ambas remiten a la idea del relato, de las historias que, mucho antes de la aparición de la imprenta, constituían el vehículo básico de transmisión del conocimiento. «La narración, tal como brota lentamente en el círculo del artesano –el campesino, el marítimo y, posteriormente, también el urbano– es, de por sí, la forma similarmente artesanal de comunicación», defendía Walter Benjamin (1991, s. p.) en un ensayo en el que cuestionaba la capacidad de la información periodística para reflejar la actualidad de forma eficiente dada su obstinación por la novedad y la verificabilidad.

#### ILUSTRACIÓN 1 | EJEMPLOS DE PÓDCAST NARRATIVOS EN ESPAÑOL (2016-2024)



FUENTE: elaboración propia.

Sobre tales presupuestos se asientan los pódcast narrativos de no ficción –documentales, reportajes, ensayos, propuestas formativas y de desarrollo personal...– y los narrativos de ficción –series, películas sonoras, teatro, cuentos...–, de naturaleza estrictamente sonora y sobre bases expresivas orientadas hacia el poso y la serenidad. Estos formatos permiten contextualizar vidas y situaciones, abordar casos irresueltos, ofrecer nuevos enfoques a cuestiones ya tratadas, profundizar en hechos pasados o explicar las consecuencias de lo sucedido incluso a través de historias imaginadas que, en muchos casos, y del mismo modo que en la literatura, el cómic, el cine o la televisión, sirven como «anticipación» a lo que está por venir.

Somos seres narrativos, sin historias nos morimos: tomamos conciencia de quiénes somos y de dónde estamos a partir de lo que vemos, leemos y escuchamos. Lo hacemos a través de un lenguaje que también evoluciona para adaptarse a los nuevos canales, sistemas de distribución y soportes de acceso que vehiculan su difusión y recepción, que materializan el sentido y la naturaleza de los medios de comunicación. Como enfatiza el sociólogo Nik Couldry, los medios impresos, sonoros y audiovisuales constituyen, en primera instancia, tecnologías que sirven para transmitir y preservar lo que ellos denominan «contenido», es decir, símbolos, imágenes y sonidos con significado: «El significado es una dimensión básica de la vida humana: un mundo sin significado, que no tuviera absolutamente ningún sentido para nosotros, sería un mundo terrorífico. Poniendo en circulación significados, los medios desempeñan un papel básico para hacer nuestras vidas llevaderas» (Couldry, 2021, p. 26).

## Referencias bibliográficas

- ASOCIACIÓN PARA LA INVESTIGACIÓN DE MEDIOS DE COMUNICACIÓN (AIMC) (2024). *Navegantes en la Red. Encuesta a usuarios de Internet* (26.<sup>a</sup> edición). <https://www.aimc.es/otros-estudios-trabajos/navegantes-la-red/>
- AMOEDO CASALS, A. (2023). La escucha de pódcast se consolida en España. *Digital News Report España 2023*. <https://www.digitalnewsreport.es/2023/la-escucha-de-podcast-se-consolida-en-espana/>
- BENJAMIN, W. (1991). *El narrador*. Traducción de Roberto Blatt. Taurus. (Original publicado en 1936).
- COULDRY, N. (2021). *Medios de Comunicación*. Alianza Editorial.
- EDISON RESEARCH (2024). *The Infinite Dial 2024*. <https://www.edisonresearch.com/the-infinite-dial-2024/>
- GRIJELMO GARCÍA, A. (2022). *El estilo del periodista. Consejos lingüísticos, profesionales y éticos para escribir en los medios*. Taurus.
- KEMP, S. (31 de enero de 2024). *Digital 2024. Global Overview Report*. DataReportal, Meltwater y We Are Social. <https://datareportal.com/reports/digital-2024-global-overview-report>
- IVOXX (2022). *Estadísticas del consumo de podcast en español 2022. II Informe anual del Observatorio Ivoox*. [https://www.ivoox.com/\\_ik\\_observatorio2022\\_1.html](https://www.ivoox.com/_ik_observatorio2022_1.html)
- NATIONAL PUBLIC MEDIA Y EDISON RESEARCH (2023). *The Spoken Word Audio Report 2023*. <https://www.edisonresearch.com/the-spoken-word-audio-report-2023-from-npr-and-edison-research/>
- PEDRERO ESTEBAN, L. M. (2023). La consolidación del pódcast en España: parámetros para el reconocimiento de una nueva industria cultural. *Cuadernos de Periodistas*, 46, 39-50.
- PEDRERO ESTEBAN, L. M., MARTÍNEZ OTÓN, L., CASTILLO LOZANO, E. y MARTÍN NIETO, R. (2022). *Cómo suenan los podcasts en España. Radiografía de la producción original en las plataformas de audio digital en 2022*. Observatorio Nebrija del Español. <https://www.nebrija.com/catedras/observatorio-nebrija-espanol/pdf/tendencias-produccion-mercado-del-podcasting.pdf>
- RESTREPO, J. D. (2001). La objetividad periodística: utopía y realidad. *Chasqui. Revista Latinoamericana de Comunicación*, 74, 10-13.
- ROMERO GUALDA, M. V. (1993). *El español en los medios de comunicación*. Arco Libros.
- SALAVERRÍA ALIAGA, R. (2023). Miles de medios «online» en un mapa: cartografía ibérica del periodismo digital. *Cuadernos de Periodistas*, 46, 59-66.
- ZUBOFF, S. (2019). *The Age of Surveillance Capitalism*. Public Affairs

# EL ESPAÑOL Y LA TRADUCCIÓN EN EL MUNDO

Ignacio Infante Fernández

WASHINGTON UNIVERSITY IN ST. LOUIS

## 1 | Introducción

Como lengua materna de más de 498 millones de personas en el mundo, y con un número de hablantes que en el año 2024 ya supera los 600 millones<sup>1</sup>, el español desempeña un papel esencial dentro del campo de la traducción a escala global, tanto desde una dimensión de intercambio lingüístico internacional (en diferentes niveles profesionales, legales, políticos y diplomáticos en diferentes partes del mundo), como dentro de la producción cultural global (incluyendo el campo editorial, audiovisual y de medios digitales). Como punto de partida, dentro del mundo de la traducción literaria a escala global, el español constituye históricamente la tercera lengua receptora o, usando el término técnico dentro del campo de la traductología, lengua terminal (LT), después del alemán y el francés, mientras que constituye la sexta lengua de partida o lengua original (LO) de traducciones globalmente (tras inglés, francés, alemán, ruso e italiano) en el campo de la literatura mundial –a partir de los datos incluidos en el *Index Translationum* (UNESCO, s. f.), datos que serán analizados con más detalle en este artículo<sup>2</sup>.

---

1. Véase el primer apartado de este mismo artículo: I. DEMOGRAFÍA DEL ESPAÑOL EN EL MUNDO 2024.

2. El proyecto *Index Translationum* de la UNESCO constituye una bibliografía internacional de traducciones y una base de datos única para poder comparar y analizar las traducciones de libros (dentro de diferentes campos y géneros) de los estados miembros de la UNESCO. Desarrollado a partir del registro de traducciones creado por la Liga de las Naciones en 1932, y continuado por la ONU tras su creación en 1946, los registros del *Index Translationum* fueron computarizados a partir de 1979 y están disponibles de forma gratuita a través de la red. Las actualizaciones de los datos varían por país, con la mayoría de los datos actualizados en la última década (previo al año 2019).

Al mismo tiempo, dentro del creciente campo de la traducción automática, es importante resaltar que el español fue una de las primeras lenguas incluidas en Google Translate (cuando fue lanzado en 2006), una herramienta de traducción global que actualmente es usada por más de 600 millones de personas al día y en la que el español tiene un papel esencial como una de las seis lenguas con mayor demanda (junto a inglés, árabe, ruso, portugués e indonesio) (Turovsky, 2016). La traducción desde y al español es esencial para el funcionamiento de diferentes instituciones políticas internacionales, al ser una de las lenguas oficiales tanto de la ONU (junto a árabe, chino, francés, inglés y ruso)<sup>3</sup> como de la Unión Europea, constituyendo de esta forma un aspecto de la traducción en español (tanto como LO y LT) merecedor de su propio estudio y análisis. No obstante, el resto de este artículo se centrará de forma más específica en el papel del español en el campo de traducción literaria y, en particular, en el de la literatura y en el del mundo editorial más reciente.

## 2 | El español como lengua de traducción dentro de la literatura mundial: perspectiva general

Desde una perspectiva general, dentro del campo de la literatura mundial el español constituye históricamente la tercera lengua terminal (LT), a partir de los datos recogidos por el *Index Translationum*. Este hecho denota por una parte la considerable importancia e impacto del mundo literario y editorial en español a escala global en diferentes partes del mundo, al mismo tiempo que demuestra la demanda cultural de traducciones al español desde otras lenguas requerida por los sectores editoriales, culturales y educativos no solo en España y en América Latina, sino también en Estados Unidos, en África y en diferentes partes del resto del mundo hispanohablante. Considerando el español como LT dentro del campo de la traducción desde una perspectiva histórica, las cinco principales lenguas de origen (LO) han sido, en orden decreciente, el inglés, el francés, el alemán, el italiano y el

---

3. Para un análisis detallado del papel del español como lengua oficial dentro del sistema de la ONU, véase Fernández Vítóres (2014).

ruso, con una hegemonía considerable del inglés (128.384 obras recogidas por el *Index Translationum*) con respecto al resto, ya que supera en casi más de 90.000 obras traducidas al francés (36.249), lengua en segundo puesto.

En cuanto a su papel como lengua original (LO) traducida a diferentes lenguas dentro del campo de la literatura mundial, el español, como ya ha sido mencionado, se encuentra en el sexto lugar mundial a partir de los datos proporcionados por el *Index Translationum*. Parte de esta discrepancia entre el lugar del español como LO y como LT está directamente conectada a la prevalencia del inglés como lengua global durante los siglos xx y xxi, y en menor medida a la relación histórica entre el inglés y el español como lenguas y mercados editoriales, ambos con una dimensión global desde finales del siglo xx. A partir de los datos proporcionados por *Index Translationum*, podemos confirmar que mientras que las traducciones del inglés al español como LT alcanzan 128.384 obras, las traducciones del español al inglés como LT suman 10.062 obras en total, aproximadamente casi un 8 % del total de obras traducidas del inglés al español.

Por otra parte, desde una perspectiva comparativa, esta considerable diferencia es similar a los casos de las lenguas alemana y francesa, en los que el inglés constituye una LO que supera al número de obras procedentes de ambas lenguas: en el primer caso, las traducciones del inglés al francés alcanzan 151.508 obras, mientras que las del francés al inglés son 32.548; en el segundo caso, las traducciones del inglés al alemán llegan 190.931 obras, mientras que las del alemán al inglés son 32.339. Esta similitud demuestra que este aspecto diferencial del inglés como lengua de partida (LO) para francés, alemán y español constituye un elemento estructural dentro del campo de la traducción (y de la industria editorial) en el mundo global contemporáneo. Al mismo tiempo, estos datos pueden ayudar a contextualizar el contraste del impacto global entre el español y el inglés como lenguas de traducción desde una perspectiva internacional, pueden demostrar el valor incalculable de la traducción literaria como modo de fomentar una mayor circulación de obras originalmente escritas en español como lengua global más allá del

mundo hispanohablante y, consecuentemente, pueden generar un mayor impacto mundial de las culturas hispanohablantes en el siglo <sup>XXI</sup>.

### 3 | El español como lengua de origen dentro de la literatura mundial: la traducción de autores y naciones hispanohablantes

Es importante analizar con más detalle el tipo de obras más traducidas, al igual que los países hispanohablantes representados dentro del índice de traducciones creado y mantenido por la UNESCO. Por ejemplo, entre los 50 autores más traducidos a escala mundial recogidos por el *Index Translationum*, tan solo aparece un escritor hispanohablante, Gabriel García Márquez, en el puesto 49 con 1.382 traducciones (la autora en el primer lugar de la lista es Agatha Christie con 7.236 traducciones). En cuanto a países de origen dentro de los 50 primeros, España aparece como el segundo país con mayor número total de obras traducidas (como LO) a otras lenguas dentro de la literatura mundial (después de Alemania) y constituye, junto a México (en el puesto 43) y Argentina (en el puesto 49), uno de los tres países hispanohablantes incluidos dentro de los cincuenta primeros en la base de datos de la UNESCO. Al mismo tiempo, es importante destacar los diez escritores y escritoras hispanohablantes más traducidos a otras lenguas, incluido el número total de traducciones publicadas por cada autor en todas las lenguas recogidas en el *Index Translationum*.

CUADRO 1 | LOS 10 AUTORES HISPANOHABLANTES MÁS TRADUCIDOS A OTRAS LENGUAS

| Puesto | Autor                        | N.º de traducciones |
|--------|------------------------------|---------------------|
| 1      | Gabriel García Márquez       | 1.382               |
| 2      | Isabel Allende               | 832                 |
| 3      | Mario Vargas Llosa           | 659                 |
| 4      | Miguel de Cervantes Saavedra | 638                 |
| 5      | Jorge Luis Borges            | 580                 |

| Puesto | Autor                        | N.º de traducciones |
|--------|------------------------------|---------------------|
| 6      | José María Parramón Vilasaló | 468                 |
| 7      | Federico García Lorca        | 433                 |
| 8      | Pablo Neruda                 | 395                 |
| 9      | Julio Cortázar               | 353                 |
| 10     | Manuel Vázquez Montalbán     | 341                 |

**FUENTE:** elaboración propia a partir de datos del (UNESCO, s. f.).

La lista de los 10 escritores y escritoras en español proporcionada por el *Index Translationum* nos muestra una variedad geográfica, histórica y de géneros que caracteriza la producción literaria y cultural en español globalmente: la lista incluye autores originarios de Colombia, Chile, Perú, Argentina y España; al mismo tiempo ofrece una dimensión histórica al incluir a autores y autoras de diferentes periodos (desde Cervantes a Isabel Allende, por ejemplo), mientras que demuestra, en cuanto a géneros literarios, una preponderancia de la narrativa sobre la poesía (Neruda y García Lorca) o el teatro (García Lorca), e incluye también en el sexto puesto la obra divulgativa sobre la pintura de José María Parramón Vilasaló.

#### 4 | La traducción en el mercado editorial global: estudio comparativo del papel de la traducción en los mercados editoriales de España y Estados Unidos

En el caso concreto del mundo editorial en España, el número de traducciones anuales en la última década (2013–2022) suelen corresponder a un cuarto (entre un 20 % y un 25 %) de la producción editorial anual con unos 13.000 libros traducidos (13.440 en el año 2022) (Ministerio de Cultura, 2024). Es importante resaltar que este número de traducciones en 2022 es considerablemente menor que el número de traducciones publicadas una década antes (2003–2012), cuando el número de traducciones publicadas anualmente en España era de unas 25.000 obras, con un máximo de 28.851 obras traducidas en el año 2008. Parte de esta reducción se puede atribuir parcialmente al impacto

global de la pandemia del COVID-19 y a una disminución general del número de ejemplares totales editados en España durante el mismo periodo (2019-2021).

Estos datos mencionados (entre un 20 % y un 25 % de obras traducidas dentro del mercado editorial en España) contrastan de forma marcada con el número de traducciones en el mercado editorial en Estados Unidos, en el que las cifras suelen rondar el 3 % de libros publicados anualmente (incluidas todas las lenguas)<sup>4</sup>. Dentro del mundo editorial de Estados Unidos, esta cifra del 3 % se ha convertido en un número paradigmático que denota, por una parte, la baja cantidad de traducciones que caracterizan al mercado editorial estadounidense y, por otra parte, el monolingüismo preponderante dentro del mismo mercado, que conlleva una marginalización considerable de la producción literaria y cultural no anglosajona (como lenguas de origen) dentro de Estados Unidos. Para analizar las causas de esta cifra del 3 %, y para documentar e intentar aumentar el número de traducciones publicadas en Estados Unidos, desde la Universidad de Rochester (estado de Nueva York) se creó en el año 2007 la iniciativa Three Percent (Tres por ciento), que busca promover la literatura internacional (y consecuentemente la traducción literaria) dentro de Estados Unidos (University of Rochester (s. f.). Como parte de esta iniciativa, Three Percent y la editorial Open Letter Books (también parte de la misma universidad) crearon en ese mismo año una base de datos de traducciones aparecidas en Estados Unidos que, desde 2018, es publicada por la revista *Publishers Weekly* bajo el título de *Translation Database*. Aunque no constituye un registro oficial, esta base de datos es una herramienta esencial y única para cuantificar el número de traducciones publicadas anualmente en Estados Unidos, e incluye un gran número de lenguas de origen, las editoriales que publican estas obras, al igual que los nombres de los traductores y traductoras de las obras publicadas.

A pesar del bajo número de traducciones publicadas en Estados Unidos, ya mencionado, particularmente en comparación con el número de traducciones desde el inglés publicadas en el mercado editorial en español, los datos

---

4. Para un análisis reciente y detallado del porcentaje de traducciones de lenguas al inglés publicadas en Estados Unidos en los últimos años (y un estudio pormenorizado del porcentaje del 3 % mencionado dentro del contexto editorial estadounidense), véase Rutherford, Levitt y Zhang (2024).

proporcionados por la base de datos de traducción de *Publishers Weekly* son extremadamente importantes dentro del contexto del presente estudio sobre el español y la traducción en el mundo. A partir de ellos podemos establecer las siguientes observaciones sobre el impacto de la traducción de obras en español (como LO) dentro del mercado editorial de Estados Unidos en las últimas dos décadas:

- El número de obras en español (como LO) traducidas dentro del mercado editorial de Estados Unidos ha crecido considerablemente en los últimos 10 años. Por ejemplo, mientras que el año 2010 el número de obras publicadas fue de 54 (con 38 obras de ficción), en el año 2022 fue de 79 obras, lo que supone un 46 % más (con 55 obras de ficción). Pese al bajo número de traducciones correspondiente al 3 % del total editorial en Estados Unidos (en comparación con la media de entre el 20 % y 25 % de traducciones dentro del mercado editorial de España), este aumento en la última década en Estados Unidos sugiere un mayor interés por las literaturas escritas en español y una mayor demanda por parte del público y los medios culturales estadounidenses. Para aportar un dato comparativo y de contraste en este contexto, el número de obras en inglés (como LO) traducidas dentro del mercado editorial de España fue de 9.211 en el año 2011 y de 6.530 (un 7 % del total de obras publicadas) en el año 2022 (Ministerio de Cultura, 2024).
- En comparación con las obras traducidas desde el español recogidas en Translation Database, el número de obras traducidas en Estados Unidos originalmente escritas en alemán (LO) fue de 38 en 2010 y 32 en 2022; por otra parte, el número de obras traducidas en Estados Unidos escritas en francés (LO) fue de 65 en 2010 y 76 en 2022.
- La preferencia por las obras de ficción parece ser una constante que no solo refleja los intereses editoriales y de lectura del siglo XXI en Estados Unidos, sino que también es un hecho análogo a los datos históricos proporcionados por el *Index Translationum* de la UNESCO (en el que 6 de los 10 primeros escritores en español eran novelistas, o cuentistas en los

casos de Borges y Cortázar). Tanto en los datos proporcionados anteriormente para 2010 y 2022, el porcentaje de obras de ficción corresponde aproximadamente a un 70 % del total de traducciones del español (LO) publicadas en Estados Unidos desde el año 2008.

- Este aumento del número de traducciones publicadas en Estados Unidos también conlleva una mayor representación de autores originarios de diferentes países hispanohablantes dentro del mercado editorial estadounidense y, consecuentemente, del mercado anglosajón global.
- Pese a ser un género de menor recepción que la ficción, la poesía en español ha conseguido mantener un papel relevante dentro del mercado editorial estadounidense, sobre todo a través de sellos editoriales independientes, aunque con un menor número de publicaciones en los últimos diez años. Según los datos de *Translation Database*, mientras que en el año 2010 se publicaron 16 colecciones de poesía en español como LO, con obras que incluían a autores como Juan Gelman (Argentina), Marosa di Giorgio (Uruguay), Ida Vitale (Uruguay) y Raúl Zurita (Chile), por mencionar algunos ejemplos, en el año 2022 se publicaron 9 colecciones, con obras de autores como Antonio Gamoneda (España), Coral Bracho (México) y María Negroni (Argentina) (*Publishers Weekly*, s. f.).

## 5 | Conclusiones

En el año 2024 resulta esencial llevar a cabo un análisis detallado del papel del español en relación con la traducción en el mundo. Como ha sido brevemente descrito en este ensayo, el español constituye a través de la traducción una lengua totalmente central en tanto en cuanto a la articulación del campo de la literatura mundial históricamente, así como de la producción editorial global desde una perspectiva contemporánea. Al mismo tiempo, es importante resaltar que la infraestructura de la industria editorial en español incorpora la traducción de otras lenguas como una parte considerable de su producción total y proyección internacional, lo que constituye en el caso particular de España un 25 % de la producción editorial anual, con el inglés (como LO) con aproximadamente un 7% anual

del total de obras publicadas. Por otra parte, es evidente que la creciente demanda de traducciones desde el español (como LO) dentro del mercado editorial de Estados Unidos contrasta con la falta de una infraestructura de traducción adecuada y necesaria para aumentar de forma considerable el volumen de traducciones a un nivel más allá del paradigmático porcentaje del 3 % que ha definido el número de traducciones anuales en el mercado estadounidense en las primeras dos décadas del siglo XXI. Dentro de este contexto, poder desarrollar programas o iniciativas que continúen mostrando la riqueza literaria y cultural de las literaturas en español a través de la traducción, y apoyando tanto a traductores y traductoras como a editoriales que posibiliten la articulación de estas iniciativas internacionalmente, constituye una oportunidad única para la futura expansión del español en el siglo XXI como lengua global (y lengua de traducción) en diferentes partes del mundo contemporáneo.

## Referencias bibliográficas

- FERNÁNDEZ VÍTORES, D. (2014). *Spanish in the United Nations System* (serie Informes del Observatorio/Observatory Reports, n.º 4). Instituto Cervantes at the Faculty of Arts and Sciences of Harvard University. [https://cervantesobservatorio.fas.harvard.edu/sites/default/files/004\\_informes\\_dfv\\_spanish\\_un\\_system.pdf](https://cervantesobservatorio.fas.harvard.edu/sites/default/files/004_informes_dfv_spanish_un_system.pdf)
- MINISTERIO DE CULTURA (2024). *Estadística de la edición española de libros con ISBN. Resultados*. <https://www.cultura.gob.es/servicios-al-ciudadano/estadisticas/cultura/mc/culturabase/libro/resultados-libro.html>
- RUTHERFORD, M. B., LEVITT, P. Y ZHANG, E. (2024). *Whence the 3 Percent?: How Far Have We Come toward Decentering America's Literary Preference? Global Perspectives*, 5, 1, 93034. <https://doi.org/10.1525/gp.2024.93034>
- PUBLISHERS WEEKLY (s. f.). *Translation Database*. <https://www.publishersweekly.com/pw/translation/home/index.html>
- UNESCO (s. f.). *Index Translationum*. <https://www.unesco.org/xtrans/bsform.aspx>
- UNIVERSITY OF ROCHESTER (s. f.). *Three Percent*. <https://www.rochester.edu/college/translation/threepcent/>
- TUROVSKY, B. (28 de abril de 2016). *Ten Years of Google Translate*. Google blog. <https://blog.google/products/translate/ten-years-of-google-translate/>

# EL ESPAÑOL Y LAS INDUSTRIAS CULTURALES

Héctor Álvarez Mella

OBSERVATORIO DEL ESPAÑOL EN EUROPA

CENTRO DE ESTUDIOS IBEROAMERICANOS (HCIAS)

DE LA UNIVERSIDAD DE HEIDELBERG

## 1 | Introducción

Las industrias culturales son espacios que relacionan a individuos y grupos humanos a través de la cultura. Desde un punto de vista comunicacional, los bienes simbólicos a los que dan acceso editoriales, plataformas de *streaming* o televisiones digitales conectan a los ciudadanos en torno a experiencias compartidas, lugares comunes, valores, patrones de comportamiento o ideologías.

La relación entre las industrias culturales y las lenguas dista de ser banal o anecdótica. Por un lado, la lengua es materia de creación de muchas de las manifestaciones culturales más relevantes de nuestras sociedades, desde los libros a las canciones, las televisiones y las pantallas de cine. Por otro, las lenguas conforman espacios geolingüísticos de consumo cultural, como el hispanohablante, que definen patrones de circulación regional, nacional e internacional de bienes culturales.

A lo largo de las siguientes páginas se ofrece una panorámica de tres industrias culturales, la editorial, la musical y la audiovisual, desde el punto de vista de la lengua española. Se presentarán diferentes indicadores de producción, distribución y consumo cultural tanto a escala nacional como internacional. La relevancia de esta doble perspectiva radica en el impacto de la transformación digital que está reconfigurando la geografía de la circulación de creaciones culturales.

## 2 | La industria editorial

Desde la aparición de las primeras imprentas en el incipiente mundo hispano-hablante (1472 en Segovia, 1540 en el Virreinato de Nueva España) hasta la era digital del presente, los libros en español han llegado a millones de lectores de todo el mundo a través del papel y las pantallas. En 2022, los países hispanohablantes registraron el 7 % de la producción mundial de libros, con casi 220.000 ISBN solicitados. Como se observa en el cuadro 1, España, Argentina y México se encuentran entre los mercados editoriales más activos, aunque bastante por debajo de otros espacios como Estados Unidos, mercados asiáticos como los de Japón, Corea del Sur, la India e Indonesia, mercados europeos como los de Alemania, el Reino Unido, Italia y Polonia, o el de Brasil en América del Sur. Dado que la mayoría de los libros publicados en un país lo son en sus lenguas mayoritarias, el español es la sexta lengua editorial en la actualidad, después del inglés, el japonés, el coreano, el hindi y el alemán.

| CUADRO 1   LOS 20 MERCADOS EDITORIALES CON MÁS REGISTROS DE ISBN <sup>1</sup> |                          |
|-------------------------------------------------------------------------------|--------------------------|
| País                                                                          | ISBN registrados en 2022 |
| Estados Unidos                                                                | 3.279.217                |
| Japón                                                                         | 902.311                  |
| Corea del Sur                                                                 | 338.237                  |
| India                                                                         | 281.091                  |
| Alemania                                                                      | 277.000                  |
| Brasil                                                                        | 179.042                  |
| Reino Unido                                                                   | 153.167                  |
| Italia                                                                        | 139.970                  |
| Indonesia                                                                     | 107.856                  |
| Polonia                                                                       | 99.995                   |
| España                                                                        | 95.811                   |

1. Los datos para Francia se refieren a libros registrados, no a ISBN, por lo que podrían variar ligeramente.

| País                    | ISBN registrados en 2022 |
|-------------------------|--------------------------|
| Francia                 | 81.909                   |
| Turquía                 | 83.653                   |
| Holanda                 | 59.283                   |
| Suecia                  | 37.338                   |
| Argentina               | 35.500                   |
| República Checa         | 34.985                   |
| Australia               | 31.708                   |
| México                  | 27.534                   |
| Países hispanohablantes | 218.969                  |

FUENTE: World Intellectual Property Organisation (WIPO, 2023).

El mercado del libro en español ha venido creciendo en las últimas décadas. Según los datos más recientes del Centro Regional para el Fomento del Libro en América Latina y el Caribe (CERLALC, 2021), entre 2001 y 2020 la actividad editorial en español duplicó su producción pasando de 92.921 libros a 185.217. El crecimiento en la producción de libros vino impulsado tanto por el aumento de la edición en España, sobre todo a comienzos del siglo, como por la actividad de editoriales latinoamericanas (que suponen el 56 % de la edición de la región), y más concretamente por los mercados de México, que multiplicó por cuatro su producción de libros, Colombia y Argentina.

| CUADRO 2   EVOLUCIÓN DE LA PRODUCCIÓN EDITORIAL EN PAÍSES HISPANOHABLANTES <sup>2</sup> |        |        |
|-----------------------------------------------------------------------------------------|--------|--------|
| País                                                                                    | 2001   | 2020   |
| España                                                                                  | 53.528 | 82.282 |
| Argentina                                                                               | 12.368 | 27.694 |
| México                                                                                  | 5.214  | 20.925 |
| Colombia                                                                                | 7.600  | 20.378 |

2. No se incluye Nicaragua por falta de datos actualizados. Tampoco se dispone de información para Puerto Rico y Guinea Ecuatorial.

| País                 | 2001  | 2020  |
|----------------------|-------|-------|
| Chile                | 2.580 | 8.353 |
| Perú                 | 1.392 | 5.698 |
| Ecuador              | 1.250 | 4.264 |
| Cuba                 | 1.433 | 2.950 |
| Uruguay              | 500   | 2.220 |
| Venezuela            | 3.100 | 2.186 |
| República Dominicana | 621   | 1.765 |
| Costa Rica           | 998   | 1.724 |
| Bolivia              | 384   | 1.103 |
| Guatemala            | 517   | 1.073 |
| Panamá               | 430   | 884   |
| Paraguay             | 353   | 823   |
| El Salvador          | 144   | 675   |
| Honduras             | 209   | 220   |

**FUENTE:** CERLALC (2021).

Aunque el sector editorial continúe siendo una de las industrias culturales más analógicas, fruto en gran parte del fetichismo del libro en papel, la digitalización de la lectura es una importante tendencia que se desprende de los datos estadísticos disponibles. Según el informe de la plataforma Bookwired (2023), la venta de libros digitales en español no ha dejado de aumentar durante los últimos años, llegando a nuevos lectores, sobre todo en Estados Unidos y en la Unión Europea, donde los ingresos crecieron un 30,6 % y un 14,5 % respectivamente. Tales cifras son reflejo del impacto de la digitalización en la circulación de libros en español en espacios donde la distribución en papel se encuentra con limitaciones. El audiolibro, por su parte, ha venido ganando terreno a un ritmo mayor impulsado por la audificación del consumo de libros en plataformas digitales de suscripción o descarga (un 53 % entre 2021 y 2022). La mayor parte de oyentes de libros se encuentra en España (38 %), México (28 %) y en Estados Unidos (23 %).

### 3 | La industria audiovisual

La industria audiovisual es un importante espacio de creación y consumo cultural en español. Desde el reconocimiento mundial de cineastas como Buñuel y Almodóvar hasta el éxito internacional de telenovelas y, más recientemente, series en español como *La casa de papel*, numerosas producciones se proyectan en grandes y pequeñas pantallas de todo el mundo.

#### 3.1 | El cine en español

El español es la segunda lengua en la que se producen más películas (véase cuadro 3). Entre 2007 y 2017, la base de datos de estadísticas culturales de la UNESCO (2019b) recoge 4.371 producciones en español, muy por detrás del inglés (11.365), y seguido por lenguas asiáticas como el hindi, el telugu y el tamil –todas ellas lenguas de la importante industria fílmica de Bollywood–, el japonés, lenguas europeas como el francés, el italiano y el alemán, y otras lenguas como el yoruba o el árabe.

| CUADRO 3   PRINCIPALES LENGUAS DE PRODUCCIÓN DE PELÍCULAS (2007-2017) |                      |
|-----------------------------------------------------------------------|----------------------|
| Lengua                                                                | Películas producidas |
| inglés                                                                | 11.365               |
| español                                                               | 4.371                |
| hindi                                                                 | 2.689                |
| telugu                                                                | 2.508                |
| francés                                                               | 2.494                |
| tamil                                                                 | 2.376                |
| yoruba                                                                | 1.768                |
| japonés                                                               | 1.201                |
| árabe                                                                 | 808                  |
| italiano                                                              | 708                  |
| alemán                                                                | 490                  |

FUENTE: UNESCO (2019b).

El mundo hispanohablante es uno de los espacios culturales de mayor producción filmica, cuyos principales centros se encuentran en España, México, Colombia y Argentina. Las industrias cinematográficas de estos países no solo producen la mayoría de las películas en español sino también más éxitos. Entre las quince películas de mayor éxito mundial producidas o coproducidas en países hispanohablantes (Letterboxd, 2024a y 2024b), solo una se encuentra entre los cien films con mayor taquilla a escala global: *La sociedad de la nieve*, nominada al Óscar a la mejor película internacional. Entre las quince películas, predominan las producciones españolas (7) y las mexicanas (4), seguidas de las argentinas (2) y las chilenas (2).

La mayor parte de las películas producidas en países hispanohablantes están orientadas a mercados cinematográficos y televisivos nacionales, aunque un número no desdeñable de producciones circula más allá de las fronteras en las que se produjo. Como se observa en el cuadro 4, los mayores éxitos de las películas en español se dan en las salas de cine de países hispanohablantes (en el mercado doméstico o en otros países). En los diez años que abarca la base de datos de la UNESCO, las películas en español estuvieron al menos en 54 ocasiones entre las cinco películas más vistas de algún país o, en otras palabras, cada año al menos cinco países contabilizaron una película en español entre las cinco películas más taquilleras. Usando este dato como indicador y desglosando el interés de las audiencias por países, se pueden sacar dos conclusiones. Primera, que las películas en español tienen más éxito en países hispanohablantes, ya sea en sus propios países de producción o bien en otros países de habla hispana. Segunda, que las producciones cinematográficas en español también son capaces de conquistar mercados no hispanohablantes, sobre todo en países europeos y en Brasil.

Cuadro 4 · Página siguiente >

CUADRO 4 | ESPACIOS DE ÉXITO DE LAS PELÍCULAS EN ESPAÑOL

| Espacios                          | N.º de veces en que una película hispanohablante se sitúa entre las 5 primeras |
|-----------------------------------|--------------------------------------------------------------------------------|
| <b>Países hispanohablantes</b>    | <b>32 (59 %)</b>                                                               |
| <b>Países no hispanohablantes</b> | <b>22 (41 %)</b>                                                               |
| Brasil                            | 3                                                                              |
| Portugal                          | 3                                                                              |
| Andorra                           | 2                                                                              |
| Rumanía                           | 2                                                                              |
| Francia                           | 1                                                                              |
| Holanda                           | 1                                                                              |
| Azerbaiyán                        | 1                                                                              |
| Macedonia del Norte               | 1                                                                              |
| Egipto                            | 1                                                                              |
| Serbia                            | 1                                                                              |
| Bélgica                           | 1                                                                              |
| República Checa                   | 1                                                                              |
| Islandia                          | 1                                                                              |
| Italia                            | 1                                                                              |
| Malta                             | 1                                                                              |
| Polonia                           | 1                                                                              |

FUENTE: UNESCO (2019a).

3.2 | Netflix y las series en español

La proyección internacional de series en español depende hoy de plataformas de suscripción entre las que destaca Netflix. Desde sus comienzos, la plataforma ha orientado sus estrategias de expansión internacional más allá del espacio anglófono hacia los países hispanohablantes entendidos como

un mercado común. Desde el punto de vista de los mercados televisivos, el mundo hispanohablante constituye una «región geolingüística» interconectada. En 2011, América Latina se convirtió en la segunda región en la que se expandió la empresa, después de Canadá, y pronto se expandiría a España en 2015. Netflix se encuentra en América Latina con un espacio audiovisual receptivo con una gran presencia de contenidos televisivos estadounidenses, no solo en las parrillas y la oferta de televisión por satélite, sino también en las preferencias de las audiencias. Sin embargo, la plataforma pronto orienta su estrategia de expansión hacia la producción de contenidos en español. Como se observa en el cuadro 5, el español es con mucha diferencia la lengua distinta del inglés en la que Netflix produce o coproduce más series.

**CUADRO 5 | SERIES PRODUCIDAS POR NETFLIX SEGÚN LA LENGUA DE PRODUCCIÓN**

| Idioma         | Series terminadas | Series activas | Series programadas | Total     |
|----------------|-------------------|----------------|--------------------|-----------|
| <b>Español</b> | <b>75</b>         | <b>17</b>      | <b>1</b>           | <b>93</b> |
| Coreano        | 32                | 11             | 2                  | 45        |
| Hindi          | 25                | 10             | 1                  | 36        |
| Portugués      | 22                | 7              | 0                  | 29        |
| Francés        | 21                | 4              | 1                  | 26        |
| Alemán         | 18                | 4              | 0                  | 22        |
| Japonés        | 18                | 4              | 0                  | 22        |
| Turco          | 13                | 8              | 1                  | 22        |
| Italiano       | 12                | 6              | 1                  | 19        |
| Polaco         | 12                | 6              | 0                  | 18        |
| Otras lenguas  | 56                | 23             | 2                  | 81        |

**FUENTE:** Wikipedia<sup>3</sup>.

3. Véase: [https://en.wikipedia.org/w/index.php?title=List\\_of\\_Netflix\\_original\\_programming&oldid=1219027516](https://en.wikipedia.org/w/index.php?title=List_of_Netflix_original_programming&oldid=1219027516) y [https://en.wikipedia.org/w/index.php?title=List\\_of\\_ended\\_Netflix\\_original\\_programming&oldid=1218379469](https://en.wikipedia.org/w/index.php?title=List_of_ended_Netflix_original_programming&oldid=1218379469).

La evidente intención de incorporar un prisma local a la estrategia de Netflix se refuerza además por la creciente demanda por parte de las audiencias y las empresas productoras de la difusión de series y películas producidas en sus propios países de origen. Siguiendo esta misma estrategia de «glocalización», la intención de los productores de contenidos originales para Netflix en los mercados hispanohablantes es ofrecer productos audiovisuales que reflejen la realidad de la cultura de habla hispana y satisfagan esta demanda de contenidos locales, y que, al mismo tiempo, puedan resultar atractivos para el público global. Como se observa en el cuadro 6, tanto las listas de series que se encuentran activas actualmente en la plataforma, como las que están terminadas tienden a una presencia mayor de contenidos originales de países diferentes a España y México, los principales países productores de la industria audiovisual hispanohablante, aunque todavía muy reducida en comparación.

| CUADRO 6   SERIES EN ESPAÑOL PRODUCIDAS POR NETFLIX<br>SEGÚN EL PAÍS HISPANOABLANTE DE PRODUCCIÓN |                   |                |                    |       |
|---------------------------------------------------------------------------------------------------|-------------------|----------------|--------------------|-------|
| País                                                                                              | Series terminadas | Series activas | Series programadas | Total |
| España                                                                                            | 32                | 6              | 1                  | 39    |
| México                                                                                            | 27                | 6              | 0                  | 33    |
| Colombia                                                                                          | 8                 | 3              | 0                  | 11    |
| Argentina                                                                                         | 6                 | 1              | 0                  | 7     |
| Chile                                                                                             | 1                 | 1              | 0                  | 2     |
| Perú                                                                                              | 1                 | 0              | 0                  | 1     |

FUENTE: Wikipedia<sup>4</sup>.

Más allá de la producción, interesa analizar los patrones de distribución y consumo de las series en español. Entre 2020 y 2023, 18 series en español se encuentran entre las 100 más vistas en Netflix (cuadro 7). Al dominio de las producciones anglosajonas se contrapone una importante presencia de

4. Véase: [https://en.wikipedia.org/w/index.php?title=List\\_of\\_Netflix\\_original\\_programming&oldid=1219027516](https://en.wikipedia.org/w/index.php?title=List_of_Netflix_original_programming&oldid=1219027516) y [https://en.wikipedia.org/w/index.php?title=List\\_of\\_ended\\_Netflix\\_original\\_programming&oldid=1218379469](https://en.wikipedia.org/w/index.php?title=List_of_ended_Netflix_original_programming&oldid=1218379469)

series en coreano y en español, que confirman una orientación internacional de la plataforma estadounidense.

| CUADRO 7   PAÍSES DE PRODUCCIÓN DE LAS 100 SERIES MÁS VISTAS EN NETFLIX (2020-2023) |                                        |
|-------------------------------------------------------------------------------------|----------------------------------------|
| País de producción                                                                  | N.º de series entre las 100 más vistas |
| Estados Unidos                                                                      | 51                                     |
| Corea del Sur                                                                       | 14                                     |
| Reino Unido                                                                         | 10                                     |
| Colombia                                                                            | 8                                      |
| España                                                                              | 6                                      |
| México                                                                              | 4                                      |
| Alemania                                                                            | 2                                      |
| Francia                                                                             | 1                                      |
| Japón                                                                               | 1                                      |
| Canadá                                                                              | 1                                      |
| Italia                                                                              | 1                                      |
| Brasil                                                                              | 1                                      |

FUENTE: Flixpatrol (2024).

4 | La industria musical

En las últimas décadas, la música se ha convertido en un altavoz global del español. Impulsada sobre todo desde Estados Unidos, la música en español supone por primera vez una alternativa a la hegemonía del inglés. Al menos así se pronunciaba la revista *The Economist* (2022) en un artículo en el que analizaba los datos de escuchas de la plataforma Spotify. Aunque la música en inglés siga dominando las listas de éxitos, las plataformas digitales globales son escenario de una inusitada popularidad de nuevos (y no tan nuevos) géneros en otros idiomas, como es la música urbana latina en español. Tras el éxito global de «Despacito» en 2017, la música latina en español, liderada por artistas como Bad Bunny, está experimentando un crecimiento sin precedentes. El año 2022 fue para *Billboard*, importante revista musical

estadounidense, el «año del conejo» en alusión a Bad Bunny, después de haber sido el artista más reproducido en Spotify en 2020 y 2021; más recientemente, en marzo de 2024, la cantante colombiana Karol G fue portada de la revista tras haber sido nombrada mujer del año, la primera artista latina. Detrás de estos dos «hitos» de la música en español está un fenómeno basado en los cambios producidos por la digitalización de la producción, la distribución y el consumo musical y ciertos cambios de paradigma.

En la actualidad, el consumo de música ha diversificado sus espacios de escucha más allá de las listas de éxitos y las cifras de ventas de sencillos y LP. Las plataformas de *streaming*, como Spotify y YouTube, se han convertido en el principal medio de consumo musical. Prueba de ello es el impacto del *streaming* en la industria: según el informe *Global Music Report* (International Federation of the Phonographic Industry [IFPI], 2024), en 2023 el 67 % de los ingresos globales del sector es a través de la suscripción o los anuncios en plataformas. Ese mismo año, el español es la lengua del 21 % de las 500 canciones más reproducidas en Spotify, por detrás del inglés (75 %) y muy por delante de otras lenguas demográficamente importantes como el coreano, el hindi, el árabe o el portugués, que también colocan alguna canción en el *ranking* (Elgiryewithana, 2023). En YouTube, el mismo porcentaje de canciones, 21 %, de las 100 más escuchadas a escala global son música urbana latina, mayoritariamente cantada en español (Moreno Cazalla, 2023); y un 7 % de los 30 videoclips más vistos en la historia de la plataforma son en español, con un total de 13.000 millones de visualizaciones (YouTube, s. f.). Las cifras mareantes de estas dos plataformas contrastan con cuotas más humildes en otros espacios de consumo. Entre enero de 2020 y julio de 2023, solo un 3 % de las 200 canciones más descargadas en iTunes eran música urbana latina (en español); en el *ranking* Billboard de las 100 canciones más emitidas por la radio solo lo eran el 2 % (Moreno Cazalla, 2023).

Las estadísticas comentadas parecen indicar que el éxito global de la música en español es un fenómeno de *streaming*. Para entender esta dinámica es necesario comprender la lógica de este nuevo modelo de consumo en la industria cultural. El éxito de la música urbana latina es fácilmente

atribuible a la combinación entre demografía y algoritmos. La alta concentración de hispanohablantes que consumen música en español a través de internet potencia su popularidad global pues posiciona fácilmente las canciones en las listas de éxitos globales. Efectivamente, las listas de éxitos en países hispanohablantes están dominadas por música en español, especialmente la música urbana latina, y la popularidad de los artistas hispanohablantes en el espacio geolingüístico del español se globaliza en las plataformas. Además, este tipo de música se ha posicionado como uno de los géneros más prominentes en la industria musical estadounidense y ha adoptado el español como marca de identidad. Este fenómeno surge en un contexto de resistencia a las políticas lingüísticas restrictivas y representa una dimensión transnacional que fortalece la identidad hispanohablante en Estados Unidos.

## 5 | Apuntes finales

Desde el comienzo del siglo **xxi**, las industrias culturales suponen una reorganización de la hegemonía cultural. Ante un evidente dominio cultural del inglés labrado durante la segunda mitad del siglo **xx**, el impacto global de productos culturales en otras lenguas como el español constituye una compleja dinámica de «contraflujos» que debe ser tomada en cuenta si se quiere comprender el mapamundi policéntrico de las lenguas en la actualidad.

Las industrias culturales son relevantes para la *geopolítica del español* en tanto en cuanto forman parte de la diplomacia cultural. Tradicionalmente, la lengua se ha comprendido y usado como uno de los pilares de las estrategias de *poder blando* de los países. Sin embargo, desde el punto de vista de las lenguas internacionales, se puede hablar de una diplomacia cultural de las propias lenguas que tendría que ver con la capacidad de seducción del idioma, manifestada en el interés que despierta, y con sus dinámicas de proyección internacional. Es en este horizonte en el que artistas como Bad Bunny y Rosalía, Pedro Almodóvar y Penélope Cruz, Carlos Ruiz Zafón o Isabel Allende son, voluntaria o involuntariamente, embajadores del español, pues contribuyen a construir su imagen como lengua.

## Referencias bibliográficas

- BOOKWIRE (2023). *Informe Bookwire. Evolución del mercado digital (ebooks y audiolibros) en España y América Latina*. [https://www.bookwire.es/fileadmin/media/\\_footer/unternehmen/knowledge\\_hub/BOOKWIRE\\_INFORME\\_2023.pdf](https://www.bookwire.es/fileadmin/media/_footer/unternehmen/knowledge_hub/BOOKWIRE_INFORME_2023.pdf)
- CENTRO REGIONAL PARA EL FOMENTO DEL LIBRO EN AMÉRICA LATINA Y EL CARIBE (CERLALC) (2021). *El espacio iberoamericano del libro 2020*. Centro Regional para el Fomento del Libro en América Latina y el Caribe UNESCO. [https://cerlalc.org/wp-content/uploads/2021/10/Cerlalc\\_Publicaciones\\_El\\_espacio\\_iberoamericano\\_2020\\_010921.pdf](https://cerlalc.org/wp-content/uploads/2021/10/Cerlalc_Publicaciones_El_espacio_iberoamericano_2020_010921.pdf)
- ELGIRIYEWITHANA, N. (2023). *Most Streamed Spotify Songs 2023* [conjunto de datos]. Kaggle. <https://www.kaggle.com/datasets/nelgiriyeewithana/top-spotify-songs-2023>
- FLIXPATROL (2024). *The Most Watched Netflix Movies and TV Shows by Total Hours Viewed*. <https://flixpatrol.com/streaming-services/most-hours-total/netflix/>
- INTERNATIONAL FEDERATION OF THE PHONOGRAPHIC INDUSTRY (IFPI) (2024). *Global Music Report 2024. State of the Industry*. IFPI. <https://globalmusicreport.ifpi.org/>
- LETTERBOXD (2024a). *Films Released in the 2020s. Sorted by popularity*. <https://letterboxd.com/films/popular/decade/2020s/>
- LETTERBOXD (2024b). *Films Released in the 2020s in Spanish. Sorted by popularity*. <https://letterboxd.com/films/popular/decade/2020s/language/spanish/>
- MORENO CAZALLA, L. (2023). *El boom de la música urbana latina y su influencia sobre la expansión mundial del español* [Documento de trabajo del Observatorio Nebrija del Español]. [https://www.nebrija.com/catedras/observatorio-nebrija-espanol/pdf/16\\_musica\\_urbana\\_latina.pdf](https://www.nebrija.com/catedras/observatorio-nebrija-espanol/pdf/16_musica_urbana_latina.pdf)
- THE ECONOMIST (29 de enero de 2022). *What Spotify data show about the decline of English*. <https://www.economist.com/interactive/graphic-detail/2022/01/29/what-spotify-data-show-about-the-decline-of-english>
- UNESCO INSTITUTE FOR STATISTICS (2019a). *Exhibition of Feature Films. Origin of Top 5 countries of All Feature Films Exhibited*. [http://data.uis.unesco.org/Index.aspx?DataSetCode=CUL\\_DS](http://data.uis.unesco.org/Index.aspx?DataSetCode=CUL_DS)
- UNESCO INSTITUTE FOR STATISTICS (2019b). *Feature Film Production. Language of production*. [http://data.uis.unesco.org/Index.aspx?DataSetCode=CUL\\_DS](http://data.uis.unesco.org/Index.aspx?DataSetCode=CUL_DS)
- WORLD INTELLECTUAL PROPERTY ORGANISATION (WIPO) (2023). *World Intellectual Property Indicators 2023*. WIPO.
- YOUTUBE (S. F.). *MOST VIEWED VIDEOS OF ALL TIME (OVER 700M views)* (Consultado el 18 de marzo de 2024). [https://www.youtube.com/playlist?list=PLirAqAtl\\_h2r5g8xGajEwdXd3x1sZh8hC](https://www.youtube.com/playlist?list=PLirAqAtl_h2r5g8xGajEwdXd3x1sZh8hC)

# EL ESPAÑOL ARTIFICIAL

Antonio Moreno Sandoval

UNIVERSIDAD AUTÓNOMA DE MADRID

## 1 | Introducción

Arthur C. Clarke y Stanley Kubrick imaginaron que en 2001 la inteligencia artificial (IA) llegaría a dominar a la del ser humano. No parece, por el momento, que las oscuras predicciones se vayan a cumplir. No lo han hecho, desde luego, en cuanto a la fecha, pues ha habido que esperar hasta finales de 2022 para que la inteligencia artificial irrumpa de una forma palpable en nuestra vida cotidiana y lo ha hecho precisamente a través de aplicaciones que imitan una de las capacidades más intrínsecamente humanas: la habilidad lingüística. Hablamos de los grandes modelos generativos de lenguaje (LLM por su sigla en inglés) como ChatGPT y otras aplicaciones similares. Entender cómo funcionan estas herramientas es la mejor forma de disipar los temores de la sociedad, en general, y los de los profesionales que trabajan con la lengua, en particular. ¿Son realmente una amenaza para el futuro de profesores, escritores, traductores y periodistas o constituyen más bien una oportunidad? ¿Pueden afectar a la forma en la que usamos nuestra propia lengua, a nuestra capacidad para expresarnos o a la evolución del español? Son cuestiones difíciles y no disponemos de una bola de cristal para responderlas, pero bucear en las interioridades de estos modelos ayuda a comprender sus utilidades y limitaciones.

## 2 | La IA y el procesamiento automático del lenguaje

Desde sus orígenes en la década de los cincuenta del siglo pasado, la IA tiene el objetivo de desarrollar programas que actúen como «agentes inteligentes» (robots) que puedan comportarse como seres humanos. Además de contar con visión, razonamiento y aprendizaje del entorno, estos robots

tienen que comunicarse en una lengua. De esta manera, el tratamiento computacional del lenguaje es el componente esencial de la interacción humano-máquina.

Este proceso comunicativo implica que ambos actores (humano y ordenador) procesan determinada información en función de un conocimiento lingüístico y un conocimiento del mundo compartidos. La cuestión es cómo codificar en un programa de ordenador ambos tipos de conocimiento. Necesariamente hay que crear un modelo matemático que se pueda implementar en algoritmos informáticos.

En la primera época de la IA, aproximadamente entre los años sesenta y los años noventa, la tecnología predominante era la de los modelos simbólicos. Inspirados en la lógica y la lingüística teórica, los programas se alimentaban de gramáticas y diccionarios escritos por lingüistas. Las palabras se trataban como símbolos almacenados en lexicones, que se analizaban mediante reglas que describían los patrones de combinación gramaticales. Estas reglas creaban estructuras sintácticas que luego se traducían a expresiones lógicas para interpretar el significado de las oraciones. Con el significado codificado en una fórmula lógica se podía buscar información, realizar inferencias o generar una oración en otra lengua, siguiendo el proceso inverso. Esta tecnología fracasó por su ineficacia en la desambiguación entre varios análisis posibles. Efectivamente, la actuación lingüística humana produce mensajes ambiguos de manera continua. «Me encantan mis pensamientos» puede ser traducido al inglés como *I love my thoughts* o *I love my pansies*. Para escoger la traducción correcta hay que entender el contexto y para ello se necesita un método computable que decida cuál es la interpretación correcta en un contexto comunicativo concreto. Los modelos simbólicos presentan además otras limitaciones: no existe un método general (o formalismo) aceptado para escribir gramáticas computacionales y no son robustos, en el sentido de que no es posible describir todas las reglas de una lengua. Además, tienen poca tolerancia con las oraciones aceptables que no sean claramente gramaticales. En resumen, los modelos simbólicos no eran prácticos para emular la actuación humana en contextos reales. Por

tanto, no servían para crear aplicaciones que satisficieran las necesidades de los usuarios.

Este fracaso impulsó un cambio de paradigma y se buscó la solución en la aproximación estadística. Los modelos probabilísticos son robustos, porque siempre dan una solución en función del análisis más probable a partir de las frecuencias observadas en un corpus de datos. Si fallan, es básicamente por alguna de estas dos razones: o el corpus de entrenamiento no es apropiado o faltan datos. A principios del nuevo milenio estos modelos ya superaban claramente a los simbólicos, gracias a la disponibilidad creciente de datos en internet. Sobre esas grandes colecciones de datos se empezó a aplicar la tecnología conocida como «aprendizaje automático» (*machine learning*), que se basa en el reconocimiento de patrones a partir de datos en bruto («aprendizaje no supervisado») o anotados por especialistas humanos («aprendizaje supervisado»). Entre 2000 y 2018, compitieron diferentes escuelas de aprendizaje automático –clasificadores bayesianos, algoritmos genéticos, redes neuronales (Domingos, 2015)– hasta que finalmente apareció el modelo *Transformer* (Vaswani et al., 2017) basado en redes neuronales, que superó en rendimiento y eficacia a los demás y es la base de la actual IA neuronal.

Hay que destacar un avance previo esencial: la aparición de la «semántica vectorial», los *embeddings* o «representaciones vectoriales» (Mikolov, Yih y Zweig, 2013). La técnica consiste en convertir las palabras en vectores (*word2vec*): cada palabra se representa no por un signo, sino por una matriz de números que expresan sus relaciones con otras palabras en su entorno. Esta conversión se realiza de manera automática y no supervisada mediante un algoritmo de aprendizaje automático, por eso ningún humano entiende qué propiedades ha codificado el algoritmo, cómo y por qué. Las representaciones vectoriales (*embeddings*) son opacas a la interpretación humana, pero tienen una cualidad muy importante: se pueden disponer en un espacio. Dos representaciones vectoriales son similares si se encuentran muy cercanas en ese espacio. Las que son más parecidas en significado aparecen agrupadas: el grado de relación entre dos unidades lingüísticas

se puede conocer calculando la distancia entre ellas, el coseno del ángulo que las separa. De esta manera, se ha conseguido formalizar matemáticamente el concepto de semántica distribucional de la lingüística estructuralista de Harris y Firth<sup>1</sup>: el significado de una palabra concreta depende del contexto y de las palabras que la acompañen. Es un método eficiente y seguro para reconocer sinónimos, oraciones con significado similar y traducciones equivalentes en otras lenguas: sus representaciones vectoriales serán muy parecidas.

Durante una década, las diferentes aproximaciones al aprendizaje automático compitieron para determinar cuál era el modelo más efectivo para reconocer patrones en grandes corpus textuales. Con la aparición de los *embeddings*, las redes neuronales ganaron la competición al resto de modelos probabilísticos, ya que consiguen generar los valores de los vectores más eficientemente a partir de enormes cantidades de texto. Así comenzó la era de los grandes modelos del lenguaje o «modelos fundacionales».

Recapitulemos: para capturar el significado de las emisiones lingüísticas (ya sean oraciones, diálogo o textos) los ordenadores necesitan un modelo matemático que formalice las relaciones entre las unidades lingüísticas que las componen. La primera opción fueron los modelos simbólicos, con gramáticas formales escritas explícitamente por lingüistas. Estas gramáticas producen resultados aplicando reglas. La segunda opción es que los algoritmos de aprendizaje automático, al reconocer patrones que se repiten y empleando técnicas estadísticas, generen un modelo probabilístico de las relaciones entre las palabras, asignando a cada palabra una probabilidad mayor o menor de aparecer en función del contexto previo. Un paso más allá es la conversión de los elementos lingüísticos en matrices de vectores numéricos que capturan las relaciones semánticas. Por ejemplo, una palabra como «perro» se puede representar con un vector de múltiples dimensiones como [0.49, 0.14, -0.04, 0.01...]. Estas representaciones

---

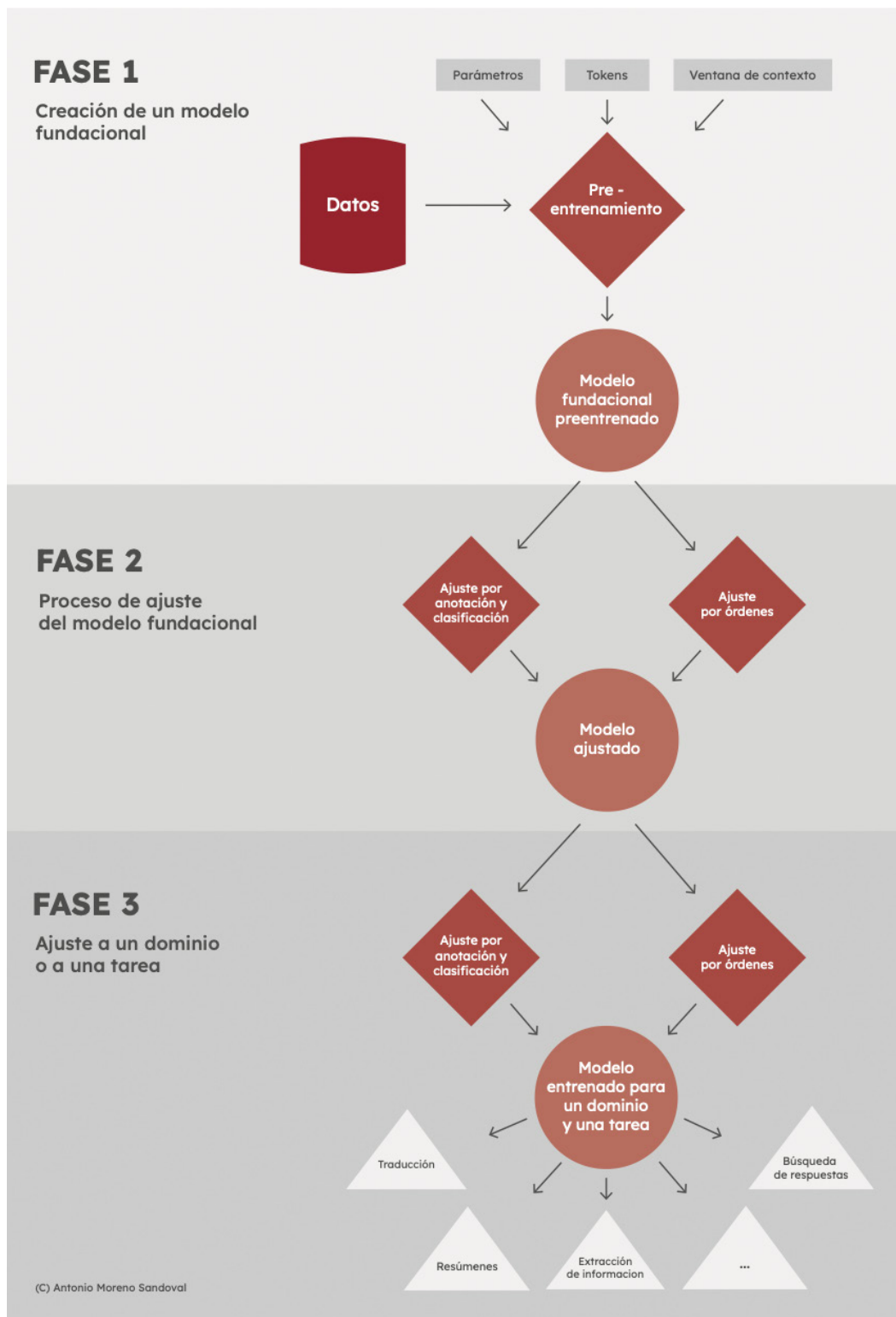
1. Consúltense Harris (1954) y Firth (1957), donde se establecen los fundamentos de la semántica basada en la distribución de unidades en un contexto de uso real, recogido en un corpus.

numéricas solo las pueden interpretar los propios algoritmos. A diferencia de los modelos simbólicos, que son explicativos, los modelos probabilísticos solo son predictivos: predicen cuál es la palabra más probable que va a continuación. Por tanto, la actual IA es predictiva, no explicativa. Construye oraciones gramaticales y con significado, pero no es posible rastrear cómo las ha generado. Por eso, es esencial que los resultados se revisen.

### 3 | Los modelos fundacionales y la arquitectura *Transformer*

Los modelos fundacionales del lenguaje son ficheros que contienen representaciones vectoriales de palabras y oraciones. Para crearlos es necesario contar con: 1) enormes cantidades de textos en formato electrónico; 2) una gran infraestructura informática de GPU; 3) algoritmos de aprendizaje automático, entre los que se ha impuesto la arquitectura denominada *Transformer* (Vaswani et al., 2017). Se les denomina «modelos fundacionales» porque son la base de conocimiento para aplicarlo a dominios y tareas concretas. Veamos las fases de creación de los actuales modelos de IA (gráfico 1).

GRÁFICO 1 | FASES DE UN MODELO FUNDACIONAL DE IA



**FUENTE:** elaboración propia.

### 3.1 | Fase 1: creación de un modelo fundacional

Para crear un modelo hay que contar con la mayor cantidad posible de textos en formato electrónico. La procedencia de esos datos es uno de los secretos mejor guardados de las empresas que generan modelos IA (véase el epígrafe 4 de este artículo). ¿Por qué? Porque el tamaño y calidad de ese corpus de datos es lo que marca la diferencia entre distintos modelos y hay que tener en cuenta que muchos datos solo están disponibles para una empresa. Veamos, por ejemplo, las fuentes que ha utilizado Google para generar su modelo PaLM (cuadro 1).

**CUADRO 1 | FUENTES UTILIZADAS POR GOOGLE PARA EL MODELO PaLM**

| Fuente                                               | Inglés          | No inglés       | Bilingüe      | Traducción    |
|------------------------------------------------------|-----------------|-----------------|---------------|---------------|
| Recuento de tokens (palabras y signos de puntuación) |                 |                 |               |               |
| Conversaciones en redes sociales (multilingües) 50 % | 756.378.913.006 | 169.908.649.039 | 6.404.486.427 | 1.448.443.476 |
| Páginas web (multilingües) 27 %                      | 459.437.466.428 | 38.653.502.458  | 7.387.577.398 | 4.260.754.907 |
| Wikipedia (multilingües) 4 %                         | 12.851.315.601  | 42.010.300.146  | 2.514.892.098 | 1.403.598.754 |
| Libros (en inglés) 13 %                              | 258.396.969.011 | 597.753.715     | 1.605.687.335 | 2.323.744.561 |
| Noticias (en inglés) 1 %                             | 26.244.234.449  | 26.455.407      | 45.117.552    | 5.554.488     |

**FUENTE:** Briakou, Cherry y Foster (2023).

La base de datos de PaLM está formada por una gran variedad de casos de usos lingüísticos en más de 100 lenguas, entre las que destaca el inglés con el 78 %. El español ocupa la cuarta posición, detrás del alemán y el francés, con solo un 2,1 % (Chowdhery *et al.*, 2023). Además de la distribución de las lenguas, nos interesa especialmente conocer la distribución de las fuentes: la mitad corresponde a conversaciones multilingües en las redes sociales, seguida por páginas web (27 %). Toda la Wikipedia multilingüe supone el 4

% del total. El enorme corpus de Google Books en inglés solo representa el 13 %. Dependiendo de la empresa propietaria de los datos o de las redes sociales, se puede inferir su procedencia.

Además de la cantidad y la procedencia de los datos, es crucial su calidad. Para mejorarlos se eliminan los duplicados y el texto ilegible, así como las temáticas o lenguas que no interesan. Este proceso de limpieza es costosísimo y se realiza de manera semiautomática. En cualquier caso, los productores de modelos fundacionales no proporcionan muchos detalles sobre el proceso de compilación, limpieza y preparación de los datos de entrenamiento. Todo ello está en manos de muy pocas organizaciones que cuentan con los medios materiales y humanos para generar esas colecciones de datos.

El proceso de «preentrenamiento» consiste en producir las representaciones vectoriales (*embeddings*) sin supervisión humana. Hay tres factores que se pueden ajustar en función de los recursos y el resultado deseado: el número de parámetros que se van a considerar, la cantidad de *tokens* (palabras y signos de puntuación) que van a componer el corpus de entrenamiento y la longitud de la ventana de contexto. El primero está en función de la infraestructura informática disponible y del tiempo dedicado a producir el modelo. Normalmente, se generan modelos de diferentes tamaños según el número de parámetros considerados. Los *tokens* con los que se alimenta el modelo están en función de los datos disponibles. Algunos equipos han optado por abastecer una gran cantidad de datos, mientras otros han preferido dedicar más tiempo a mejorar su calidad (o simplemente no disponían de suficientes medios informáticos para manejar cantidades ingentes). La ventana de contexto es el número de palabras en secuencia que tiene en cuenta el algoritmo para generar la representación vectorial. Obviamente, cuando se amplía la ventana, se complica el proceso. Cada equipo puede jugar con estos tres factores para producir un modelo distinto, aunque el algoritmo utilizado es básicamente el mismo: *Transformer* o «transformador». Consta de dos partes –un codificador (*encoder*) y un decodificador (*decoder*)– y utiliza la técnica de «autoatención» para enfocarse en los elementos más importantes de la secuencia textual procesada. De esta manera, se generan representaciones

vectoriales contextualizadas que capturan dependencias estructurales y semánticas relevantes entre las palabras de una oración.

El rasgo más sobresaliente de los modelos fundacionales es que son muy generales y a partir de ellos se pueden desarrollar aplicaciones para realizar tareas tan diferentes como responder preguntas, traducir o resumir un texto. Sin embargo, esa flexibilidad para tareas muy diversas limita su especificidad. Por ello, se recurre al «ajuste» (*fine-tuning*) de un modelo de lenguaje adaptando los parámetros a un dominio o tarea específicos.

### 3.2 | Fase 2: proceso de ajuste de un modelo fundacional

El ajuste consiste en volver a entrenar el modelo fundacional con nuevos datos especializados para actualizar los parámetros. Por ejemplo, se puede usar un modelo general del español y adaptarlo a una variante dialectal o a un dominio como la jerga médica o judicial. O se puede adaptar un modelo fundacional para traducir un par de lenguas. Este proceso, aunque requiere menos recursos que preentrenar un modelo fundacional, suele ser también muy costoso para la mayoría de las empresas y equipos. Se necesitan equipos técnicos bien formados en los métodos de ajuste y optimización, además de grandes volúmenes de datos nuevos.

Para el ajuste se emplean dos tipos de técnicas, ambas supervisadas por especialistas:

1. El ajuste por anotación y clasificación, que consiste en anotar unidades y conceptos relevantes para que el modelo se fije en ellos y aprenda las relaciones. Otra opción es clasificar manualmente documentos o dar ejemplos de resúmenes para que el modelo «aprenda» a realizar la tarea.
2. El ajuste por órdenes (*prompting*), que consiste en dar instrucciones al modelo para que aprenda a partir de ellas. Los especialistas lo entrenan de la siguiente manera: a) se le proporciona la instrucción (por ejemplo: «traduce la siguiente oración del inglés al español»); b) se introducen los datos de entrada (en este caso, una oración en inglés); c) se le muestra el

resultado de salida correcto (la oración equivalente en español). De esta manera, proporcionando ejemplos, se consigue ajustar el modelo fundacional para traducir con mayor precisión. La cantidad y calidad de los datos supervisados que se aporten influirán en los resultados que pueda producir el nuevo modelo.

El modelo así entrenado se puede seguir especializando y refinando todavía más. En este punto, aparecen empresas o grupos de investigación que se concentran en tareas o dominios aún más particulares. Es un campo en el que las grandes multinacionales no están interesadas en participar y se abre una buena oportunidad para muchos actores en la industria y el mundo académico. Lo tratamos en el siguiente apartado.

### 3.3 | Fase 3: proceso de ajuste a una tarea o un dominio

La especialización en la comunicación lingüística puede manifestarse de muchas maneras: una traducción en una variante del español, un buscador de información dentro de la legislación de un país, un generador automático de informes de inspección técnica, un extractor de términos y conceptos relevantes en un documento, etc. Para estas tareas específicas se puede volver a ajustar un modelo existente, por ejemplo, un traductor automático o un agente conversacional.

Se emplean las mismas técnicas que en la fase 2. Supongamos que queremos desarrollar un buscador de relaciones de causa y efecto en noticias financieras. Los especialistas lo entrenan de la siguiente manera: a) proporcionan la instrucción (por ejemplo: «muestra la causa y el efecto en la siguiente oración»); b) dan la entrada: «las previsiones de mejora de empleo han subido porque hay más reservas hoteleras y compra de billetes de avión»; c) se muestra la salida, en el formato deseado: «CAUSA: hay más reservas hoteleras y compra de billetes de avión; EFECTO: las previsiones de mejora de empleo han subido».

La evolución de los modelos de IA ha sido tan dinámica en los últimos cinco años que es difícil predecir cómo serán los próximos, aunque se observa una

tendencia hacia la «democratización» de la IA y la aparición de pequeñas empresas que apuestan por diferenciarse en la calidad y el valor de los datos con los que realizan el ajuste a la tarea.

## 4 | Dos modelos de negocio: código abierto o cerrado

Los sistemas que se han hecho populares en estos dos últimos años los han desarrollado grandes multinacionales como Google o Meta en Estados Unidos, o bien compañías de innovación como Open AI o Anthropic, más pequeñas pero apoyadas por inversores muy fuertes económicamente (Microsoft). Los desarrollos europeos están apareciendo con cierto retraso, pero en los últimos meses la compañía francesa Mistral AI o la alemana Aleph Alfa han lanzado sus propios modelos y también hay empresas españolas que se están incorporando al panorama, como veremos luego más en detalle.

Algunas de estas compañías han preferido dejar su código abierto para que los usuarios puedan acceder a él, mientras que otras han optado por modelos propietarios con un código cerrado. Esta decisión tiene repercusiones en sus modelos de negocio.

### 4.1 | Sistemas abiertos

Meta, la compañía propietaria de Facebook, WhatsApp e Instagram, optó por un sistema abierto cuando lanzó su modelo LLaMa en 2023. También la europea Mistral ha adoptado esta aproximación. Google, por su parte, dispone de modelos de ambos tipos: abierto (Gemma 7B) y cerrado (Gemini), con un modelo de negocio mixto. Son estos modelos de código abierto los que ofrecen más oportunidades para que empresas de menor tamaño o entidades académicas puedan desarrollar sus propias aplicaciones. El mejor ejemplo de sistema abierto es BERT, modelo original de *Transformers* y que ha servido de base para la mayoría de los desarrollos de modelos fundacionales y ajustados en español. Su popularidad se debe no solo a que fue innovador y altamente efectivo, sino sobre todo a que Google ha dado acceso público a su código y a modelos preentrenados con él. BERT ha sido adaptado a diferentes

lenguas y tareas específicas, lo que permite a los investigadores y desarrolladores personalizarlo para sus necesidades. Esto ha dado lugar a muchos «descendientes» de BERT que han mejorado y optimizado el modelo original.

## 4.2 | Sistemas cerrados

Open AI, con sus conocidos modelos GPT4 y GPT4 Turbo, o Anthropic<sup>3</sup>, con sus modelos Claude 1 y Claude2, han optado por ofrecerlos como servicios con un coste por *token*. El modelo de negocio es más clásico, en el sentido de que se cobra por la utilización de los datos y sus herramientas asociadas.

La competencia y actualización de modelos es tan rápida que es difícil estar al día. Con esa necesidad en mente, LMSYS Org, un grupo de investigadores de las universidades de Berkeley, San Diego y Carnegie Mellon, ha creado una plataforma de evaluación de modelos, denominada LMSYS Chatbot Arena<sup>2</sup>, donde se pueden probar de manera anónima modelos y votar los resultados ofrecidos. El cuadro 2 muestra los 10 mejores modelos a fecha de 26 de marzo de 2024. También existe una plataforma similar para evaluar modelos abiertos<sup>3</sup>.

El LMSYS Chatbot Arena funciona de la siguiente manera: los usuarios entran en la aplicación y pueden realizar la misma pregunta a dos modelos anónimos. Cuando analizan los resultados de cada *chatbot*, pueden valorar su impresión sobre las respuestas (mejor una que otra, las dos correctas, las dos equivocadas). Después de la votación se descubre el nombre de los dos sistemas usados. Con la suma de los votos de los usuarios se confecciona la clasificación de la tabla, que está funcionando dinámicamente. Sugerencia: pregunten por algún tema o cuestión que conozcan bien.

Ilustración 1 · Página siguiente >

---

2. Véase LMSYS en: <https://arena.lmsys.org/>.

3. Véase: [https://huggingface.co/spaces/HuggingFaceH4/open\\_llm\\_leaderboard](https://huggingface.co/spaces/HuggingFaceH4/open_llm_leaderboard).

ILUSTRACIÓN 1 | LISTA DE LOS 10 MODELOS MEJOR VALORADOS EN LMSYS CHATBOT ARENA

LMSYS Chatbot Arena Leaderboard

[Vote](#) | [Blog](#) | [GitHub](#) | [Paper](#) | [Dataset](#) | [Twitter](#) | [Discord](#)

LMSYS Chatbot Arena is a crowdsourced open platform for LLM evals. We've collected over 400,000 human preference votes to rank LLMs with the Elo ranking system.

Arena Elo

Full Leaderboard

Total #models: 75. Total #votes: 477471. Last updated: March 26, 2024.

Contribute your vote  at [chat.lmsys.org](https://chat.lmsys.org)! Find more analysis in the [notebook](#).

| Rank |  Model |  Arena Elo |  95% CI |  Votes | Organization | License         | Knowledge Cutoff |
|------|-----------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------|--------------|-----------------|------------------|
| 1    | <a href="#">Claude 3 Opus</a>                                                           | 1253                                                                                        | +5/-5                                                                                    | 33250                                                                                   | Anthropic    | Proprietary     | 2023/8           |
| 1    | <a href="#">GPT-4-1106-preview</a>                                                      | 1251                                                                                        | +4/-4                                                                                    | 54141                                                                                   | OpenAI       | Proprietary     | 2023/4           |
| 1    | <a href="#">GPT-4-0125-preview</a>                                                      | 1248                                                                                        | +4/-4                                                                                    | 34825                                                                                   | OpenAI       | Proprietary     | 2023/12          |
| 4    | <a href="#">Bard (Gemini Pro)</a>                                                       | 1203                                                                                        | +5/-7                                                                                    | 12476                                                                                   | Google       | Proprietary     | Online           |
| 4    | <a href="#">Claude 3 Sonnet</a>                                                         | 1198                                                                                        | +5/-5                                                                                    | 32761                                                                                   | Anthropic    | Proprietary     | 2023/8           |
| 6    | <a href="#">GPT-4-0314</a>                                                              | 1185                                                                                        | +5/-4                                                                                    | 33499                                                                                   | OpenAI       | Proprietary     | 2021/9           |
| 6    | <a href="#">Claude 3 Haiku</a>                                                          | 1179                                                                                        | +5/-5                                                                                    | 18776                                                                                   | Anthropic    | Proprietary     | 2023/8           |
| 8    | <a href="#">GPT-4-0613</a>                                                              | 1158                                                                                        | +4/-5                                                                                    | 51860                                                                                   | OpenAI       | Proprietary     | 2021/9           |
| 8    | <a href="#">Mistral-Large-2402</a>                                                      | 1157                                                                                        | +5/-4                                                                                    | 26734                                                                                   | Mistral      | Proprietary     | Unknown          |
| 9    | <a href="#">Qwen1.5-72B-Chat</a>                                                        | 1148                                                                                        | +5/-5                                                                                    | 20211                                                                                   | Alibaba      | Qianwen LICENSE | 2024/2           |
| 10   | <a href="#">Claude-1</a>                                                                | 1146                                                                                        | +6/-6                                                                                    | 21908                                                                                   | Anthropic    | Proprietary     | Unknown          |
| 10   | <a href="#">Mistral Medium</a>                                                          | 1145                                                                                        | +5/-4                                                                                    | 26196                                                                                   | Mistral      | Proprietary     | Unknown          |

FUENTE: LMSYS Org (consultado el 26 de marzo de 2024).

5 | Modelos fundacionales en español

Las primeras aplicaciones en español usaron como base el mBERT (BERT multilingüe). Es una versión de la primera arquitectura de transformadores que integraba fuentes en más de 100 lenguas. A partir de 2020, empiezan a crearse modelos fundacionales exclusivamente en español. A continuación, hacemos un breve repaso histórico de estas aportaciones.

5.1 | BETO

En 2020, apareció el primer modelo español basado en la arquitectura BERT creada por Google. Fue desarrollado por la Universidad de Chile. El corpus empleado para el entrenamiento contenía unos 3 billones de palabras, tomadas de la Wikipedia en español y del proyecto OPUS (Tiedeman, 2012), que incluye documentos de la ONU, de diferentes gobiernos, noticias, charlas TED, etc. El modelo fue entrenado con 110 millones de parámetros en unas TPU cedidas por Google. Los resultados ofrecidos en su momento mostraron

que BETO mejoraba a m-BERT en la mayoría de las tareas, salvo en respuestas a preguntas. El artículo de referencia es Cañete *et al.* (2023).

## 5.2 | BERTIN

En julio de 2022 se publicó BERTIN, un nuevo modelo español basado también en la arquitectura BERT. En este caso, se empleó el corpus mC4, la variante multilingüe del corpus C4 (*Colossal, Cleaned version of Common Crawl*), que es una versión de Common Crawl, un corpus en el que Amazon recopila periódicamente desde 2008 el contenido de internet y lo almacena en sus servidores. El acceso a esta información es gratuito y los investigadores pueden usar libremente conjuntos de esos datos e, incluso, ejecutar sus propias tareas de investigación en los servidores de Amazon. La parte en español de este corpus contiene unos 235 billones de palabras. La principal innovación de este modelo es que sus desarrolladores utilizaron solo una muestra de unos 50 millones de documentos (en lugar de los 400 millones originales) para rebajar el nivel de ruido contenido en los datos en bruto. Demostraron que cuidando la calidad de los datos de entrenamiento se pueden conseguir resultados tan buenos o mejores que con mayores cifras de datos sin limpiar, lo que permite emplear recursos informáticos reducidos al alcance de grupos de investigación con presupuestos limitados (de la Rosa *et al.*, 2022).

## 5.3 | MarIA

MarIA es el primer modelo español basado en la arquitectura RoBERTa, evolucionada de BERT, y en la arquitectura GPT2. Los datos de entrenamiento proceden del Archivo Web del Español recopilado por la Biblioteca Nacional de España entre 2009 y 2019, con más de 135 billones de palabras. Ha sido desarrollado por el Barcelona Supercomputing Center y financiado por el Gobierno español con el objetivo central de crear una serie de modelos públicos para la industria y la comunidad científica. Además del volumen de datos, destaca por el proceso de limpieza y de eliminación de duplicados. Los dos modelos liberados públicamente, MarIA base y MarIA large son actualmente los más grandes en español y se pueden usar para ajustarlos a diferentes dominios y tareas (Gutiérrez Fandiño *et al.*, 2022).

## 5.4 | RigoBERTa

A diferencia de los tres modelos anteriores, que han sido promovidos y financiados por instituciones públicas, RigoBERTa es una iniciativa del Instituto de Ingeniería del Conocimiento, institución de investigación privada sin ánimo de lucro localizada en el campus de la Universidad Autónoma de Madrid. El modelo fue publicado en julio de 2022 (Vaca *et al.*, 2022) pero es de acceso restringido. El corpus de entrenamiento está formado por el corpus OSCAR (constituido por textos de internet), por la parte española del mC4, así como por fuentes internas. Uno de los rasgos destacados es que fue sometido a un meticuloso proceso de limpieza y eliminación de duplicados, como en BERTIN y MarIA. Se empleó la arquitectura DeBERTa y se utilizaron unos recursos informáticos similares a los usados por BERTIN, que eran menos de la mitad de los empleados en BETO y MarIA. Los resultados de RigoBERTa superaron a los modelos más antiguos y también a los de MarIA en algunas tareas.

Como resumen, los modelos fundacionales en español han seguido los diferentes avances en arquitecturas basadas en transformadores y se han centrado en mejorar la calidad de los datos de entrenamiento en español, dadas las limitaciones presupuestarias. A este respecto, en la inauguración del último Mobile World Congress en Barcelona, a principios de 2024, el Gobierno español expresó su intención de impulsar los modelos del lenguaje en castellano y lenguas cooficiales anunciando la construcción de grandes modelos del lenguaje y la creación de una Sociedad Española para la Transformación Tecnológica (SETT) que aglutine las inversiones.

## 6 | La recogida de datos: sus implicaciones éticas, legales y sociales

Como hemos visto, para entrenar estos modelos fundacionales es imprescindible contar con grandes volúmenes de datos. La procedencia de estos datos y la forma en la que se obtienen tiene diversas implicaciones, tanto éticas como legales.

## 6.1 | Internet y otras fuentes abiertas

Muchos datos se obtienen descargando información que se encuentra disponible en internet. Una de las fuentes más comunes es el corpus gigante recogido por Amazon Common Crawl, muy valioso por su amplitud, multilingüismo y porque el acceso está abierto a quienes lo necesitan, pero que plantea también algunos inconvenientes. No existen garantías de que los autores de esos contenidos hayan autorizado su uso, de forma que al utilizarlos se pueden estar vulnerando los derechos de autor o la privacidad de los datos personales.

Algunos proyectos de conocimiento colaborativo, como Wikipedia, constituyen otra fuente de datos utilizable y, en este caso, con garantías de que sus autores los han puesto a disposición de los internautas para su consulta y utilización.

## 6.2 | Datos de acceso restringido

Las grandes empresas tecnológicas pueden restringir el acceso a las grandes colecciones de datos que recogen a través de sus redes. Así, Meta puede utilizar las conversaciones a través de Facebook o WhatsApp, o Google puede emplear los contenidos que se almacenan en Google Books o YouTube. Estos datos pueden marcar la diferencia en diferentes tareas.

## 6.3 | Sesgos, alucinaciones y otros riesgos conocidos

Esta recogida masiva de datos sin filtrar presenta implicaciones sociales, legales y hasta medioambientales, que conviene valorar:

- **Sesgos.** Si nadie controla el origen y equilibrio de esos datos, pueden estar sesgados en una dirección o en otra. Un sesgo es que el idioma mejor representado en estos corpus es el inglés, por lo que los modelos resultan mejor entrenados para esta lengua que para otras. Otro sesgo posible es que representen mejor un registro de la lengua que otro. Por ejemplo, un corpus con mucho peso de redes sociales y aplicaciones de mensajería sobrerrepresentará el lenguaje coloquial en tanto que otro con abundancia

de textos académicos tenderá a generar un modelo de la variante más formal de la lengua. Para mejorar la capacidad en una lengua concreta, hay que aportar al modelo más y mejores datos en español. Lo mismo ocurre si se requiere entender o generar lenguaje en un dominio específico, ya sea médico, jurídico, financiero, etc. Por supuesto, puede haber también sesgos sociológicos relativos al nivel económico, la raza o el género, mejor o peor representados en los datos.

- *Alucinaciones.* Tampoco hay filtros que vigilen la gramaticalidad de los enunciados ni la veracidad de los contenidos. En consecuencia, se pueden generar expresiones incorrectas formalmente y se proporcionan informaciones falsas, por lo que hay que mantenerse vigilante al usar sus resultados.
- *Implicaciones legales y éticas.* Si no se conocen las fuentes con las que se ha entrenado un modelo, hay riesgo de vulnerar los derechos a la privacidad de los datos personales o los de propiedad intelectual de los autores originales. Tampoco estaría claro de quién es la propiedad intelectual de los contenidos generados por la IA: ¿de los autores de la información en la que se ha basado para producirlos, de la empresa o entidad que controla el modelo de lenguaje o del usuario que, con sus instrucciones, ha obtenido el contenido? Del mismo modo, también es complicado determinar quién sería el responsable de los posibles daños que se deriven del uso de estos modelos.
- *Implicaciones sociales.* Este desarrollo tecnológico solo es posible para quienes disponen de unas infraestructuras de computación muy potentes, lo que implica que solo pueden participar quienes disponen de un gran respaldo económico (grandes empresas o gobiernos de algunas naciones, principalmente).
- *Implicaciones medioambientales.* El entrenamiento de uno de estos modelos implica un procesamiento computacional masivo que consume muchos recursos de energía y agua. Según un estudio reciente (de Vries, 2023), el preentrenamiento de un modelo como GPT3 consume tanta electricidad y agua como una casa estadounidense durante 121 años.

## 7 | Repercusiones de los modelos en el español y sus usuarios

No se puede obviar que la IA influye e influirá en la difusión y en la normalización del español. El uso intensivo de sus aplicaciones afectará a la comunicación, tanto escrita como oral. Los peligros descritos en la sección anterior, especialmente la difusión de sesgos y noticias falsas, así como la violación de derechos de autor son aspectos sobre los que habrá que legislar y mantenerse vigilantes.

En un plano distinto, la IA afecta a diferentes profesiones que trabajan con el idioma:

- En la educación, tanto los profesores como los alumnos usan las herramientas de la IA para crear materiales y resolver ejercicios, respectivamente. Si se usan de manera apropiada, sirven de apoyo, pero no sustituyen a los productos de la inteligencia humana. Muchos centros educativos se preocupan por la originalidad de los trabajos presentados por sus alumnos. Es esencial transmitir a los alumnos que los textos generados artificialmente no están controlados por ninguna inteligencia y, por tanto, habitualmente caen en falsedades o en inexactitudes.
- Traductores e intérpretes los usan y al mismo tiempo temen su competencia. Sin embargo, la traducción de calidad solo se puede conseguir con una revisión y posterior edición de los textos traducidos automáticamente. Como hemos visto, los sistemas no tienen conciencia, sentimiento ni coherencia. Por tanto, no es extraño que produzcan traducciones que no se ajustan al contenido del texto original (porque no lo entienden) o que la expresión en la lengua meta no sea apropiada (porque cometen errores gramaticales, de estilo o en las referencias culturales).
- Los escritores y periodistas pueden caer en la tentación de usar estas herramientas, pero tienen que ser conscientes de su falta de veracidad y de originalidad: lo que predomina en los datos de entrenamiento son textos de baja calidad expresiva. Además, algunos sistemas están filtrados para omitir opiniones (se percibe en el escaso uso de adjetivos valorativos) o

no hablar de determinados temas (sexo, política, etc.), lo que produce «creaciones literarias» de muy baja creatividad.

En conclusión, la IA generativa es una herramienta útil que, de momento, no sustituye a la intervención humana ni mucho menos. De lo que no cabe duda es de que hay que avanzar en los desarrollos legislativos que establezcan los límites a su uso y delimiten las responsabilidades. La Ley de IA promulgada por la Unión Europea es un avance en ese sentido, pero todavía incompleta.

Por otro lado, se abre un apasionante campo de investigación, pero es necesario impulsar la intervención de pequeñas empresas y entidades académicas para que su desarrollo sea justo y mantener la independencia, pues ya hemos visto que la ingente cantidad de recursos que requiere podría dejar esta tecnología en manos de unas pocas grandes corporaciones. Finalmente, hay que adaptar la pedagogía en los centros educativos para enseñar el uso responsable de estas herramientas.

## Referencias bibliográficas

- BRIAKOU, E., CHERRY, C. y FOSTER, G. (2023). Searching for Needles in a Haystack: On the Role of Incidental Bilingualism in PaLM's Translation Capability. arXiv:2305.10266
- CAÑETE, J., CHAPERON, G., FUENTES, R., HO, J. H., KANG, H. y PÉREZ, J. (2023). Spanish Pre-trained BERT Model and Evaluation Data. arXiv:2308.02976
- CHOWDHURY, A., NARANG, S., DEVLIN, J., BOSMA, M., MISHRA, G., ROBERTS, A., BARHAM, P., CHUNG, H. W., SUTTON, C., GEHRMANN, S., SCHUH, P., SHI, K., TSVYASHCHENKO, S., MAYNEZ, J., RAO, A., BARNES, P., TAY, Y., SHAZEER, N., PRABHAKARAN, V. ... FIEDEL, N. (2023). PaLM: Scaling Language Modeling with Pathways. *Journal of Machine Learning Research*, 24, 240, 1-113.
- DOMINGOS, P. (2015). *The master algorithm: How the quest for the ultimate learning machine will remake our world*. Basic Books.
- FIRTH, J. R. (1957). *Papers in Linguistics 1934-1951*. Oxford University Press.
- GUTIÉRREZ FANDIÑO, A., ARMENGOL ESTAPÉ, J., PÀMIES, M., LLOP PALAO, J., SILVEIRA OCAMPO, J., CARRINO, C. P., GONZÁLEZ AGIRRE, A., ARMENTANO OLLER, C., RODRÍGUEZ PENAGOS, C. y VILLEGAS, M. (2022). MarIA: Spanish Language Models. arXiv:2107.07253
- HARRIS, Z. S. (1954). Distributional Structure. *Word*, 10, 2-3, 146-162.

- MIKOLOV, T., YIH, W. T. y ZWEIG, G. (2013). Linguistic Regularities in Continuous Space Word Representations. En L. Vanderwende, H. Daumé III y K. Kirchhoff (Eds.), *Proceedings of the 2013 Conference on the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (pp. 746-751). Association for Computational Linguistics.
- DE LA ROSA, J., PONFERRADA, E. G., VILLEGAS, P., GONZÁLEZ DE PRADO-SALAS, P., ROMERO, M., y GRANDURY, M. (2022). BERTIN: Efficient Pre-Training of a Spanish Language Model using Perplexity Sampling. arXiv:2207.06814
- TIEDEMANN, J. (2012). Parallel Data, Tools and Interfaces in OPUS. En *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)* (pp. 2.214-2.218). European Language Resources Association.
- VACA SERRANO, A., GARCÍA SUBIES, G., MONTORO ZAMORANO, H., ALDAMA GARCÍA, N., SAMY, D., BETANCUR SÁNCHEZ, D., MORENO SANDOVAL, A., GUERRERO NIETO, M. y BARBERO JIMÉNEZ, A. (2022). RigoBERTa: A State-of-the-Art Language Model for Spanish. arXiv:2205.10233
- VASWANI, A., SHAZEER, N., PARMAR, N., USZKOREIT, J., JONES, L., GÓMEZ, A. N., KAISER, L. y POLOSUKHIN, I. (2017). Attention is all you need. *Advances in neural information processing systems*, 30.
- DE VRIES, A. (2023). The growing energy footprint of artificial intelligence. *Joule*, 7, 10, 2.191-2.194.

## Conclusiones a partir de los análisis complementarios

De los datos ofrecidos en los cuatro análisis complementarios a la demo-  
grafía que acaban de presentarse, podemos destacar algunas informaciones  
de interés para el conocimiento de la actual situación del español.

### Medios de comunicación

En el acceso a los medios de comunicación, ganan presencia las redes so-  
ciales, las plataformas de distribución de vídeos y, de forma cada vez más  
extendida, los formatos de audio hablado. Puede decirse que asistimos a la  
eclosión de una era de la audificación.

Solo en España hay casi 3.000 medios digitales activos: ediciones de diarios  
en internet, sitios web de revistas, emisoras de radio y canales de TV, y un

creciente número de medios nacidos en internet, la mayoría concentrados en Madrid y Cataluña.

El dispositivo dominante es el *smartphone*, que utiliza un 98,1 % de los usuarios para acceder a los contenidos digitales; un 90,2 % emplea también el ordenador fijo o portátil; el 93,2 % lo hacen a través de las redes sociales; ven la TV tradicional el 88,7 % y la conectada el 76,6 %.

Los periódicos digitales son leídos por el 68 % de la población mundial, mientras que los diarios impresos por el 62,1 %. Escuchan radio en FM dos de cada tres usuarios (66,1 %), música en plataformas de *streaming* un 64,9 %, y podcast un 62,9 %.

España es uno de los países occidentales donde los internautas consumen más podcast: lidera la escucha en Europa con una penetración del 45 %. El 40 % de los oyentes españoles están suscritos a entre 4 y 10 podcast, frente a un 20,7% que sigue con asiduidad hasta 20 títulos. Los podcast que más interesan son los conversacionales (casi 60 %) y los documentales (46,78 %).

## Traducción

Dentro del campo de la literatura, a escala global, el español constituye históricamente la tercera lengua receptora de traducciones, después del alemán y el francés, mientras que constituye la sexta lengua de partida, tras el inglés, francés, alemán, ruso e italiano.

El español fue una de las primeras lenguas incluidas en Google Translate en 2006. Esta herramienta de traducción global es usada actualmente por más de 600 millones de personas al día en todo el mundo. Este dato evidencia la necesidad de una mejora constante en las traducciones que se ofrezcan por este medio.

La preferencia por la traducción de obras de ficción escritas originalmente en español parece ser una constante según reflejan los datos históricos proporcionados por el *Index Translationum* de la UNESCO.

En términos comparativos, las traducciones del español al inglés serían aproximadamente un 8 % respecto al total de obras traducidas desde el inglés al español.

España es el segundo país con mayor número total de obras traducidas (como lengua de partida) a otras lenguas dentro de la literatura mundial (después de Alemania); España, México y Argentina son los tres países hispanohablantes incluidos dentro de los cincuenta primeros en la base de datos de traducción de la UNESCO.

En España, el número de traducciones anuales en la última década (2013-2022) suele corresponder a un cuarto (entre un 20 % y un 25 %) de la producción editorial anual, con unos 13.000 libros traducidos. En Estados Unidos el volumen de traducciones publicadas no supera el 3 % de su oferta editorial. La creciente demanda de traducciones desde el español dentro del mercado editorial estadounidense contrasta con la falta de una adecuada infraestructura de traducción.

## Industrias culturales

El español es la sexta lengua con mayor producción editorial del mundo, con un 7 % de la edición mundial de libros.

El español es la segunda lengua más importante de producción cinematográfica en el mundo, con más de 4.300 películas producidas entre 2007 y 2017. La mayor parte de las películas producidas en países hispanohablantes están orientadas a mercados cinematográficos y televisivos nacionales. El 59 % de las películas en español se sitúan entre las cinco más vistas en los países hispanohablantes; y el 41 % entre las cinco más vistas en los países no hispanohablantes.

El español es la primera lengua distinta del inglés en plataformas de *streaming* como Netflix o Spotify. España y México son los países hispanohablantes en los que Netflix produce más series en español (terminadas, activas y programadas). España es el quinto país del mundo productor de series de Netflix situadas entre las 100 más vistas del mundo.

En 2023, el español fue la lengua del 21 % de las 500 canciones más reproducidas en Spotify, por detrás del inglés (75 %) y muy por delante de otras lenguas demográficamente importantes como el coreano, el hindi, el árabe o el portugués.

## Inteligencia artificial

La creación de modelos de IA requiere enormes cantidades de textos en formato electrónico, una gran infraestructura informática y algoritmos de aprendizaje automático, entre los que se ha impuesto la arquitectura denominada *Transformer*.

La procedencia de los datos es uno de los secretos mejor guardados de las empresas que generan modelos de IA. Asimismo, los productores de modelos fundacionales no proporcionan muchos detalles sobre el proceso de compilación, limpieza y preparación de los datos de entrenamiento.

La base de datos utilizada por Google está formada por una gran variedad de casos de usos lingüísticos en más de 100 lenguas, entre las que destaca el inglés con el 78 %. El español ocupa la cuarta posición, detrás del alemán y el francés, con solo un 2,1 %.

El corpus empleado para el entrenamiento del modelo BETO contenía unos 3 billones de palabras, tomadas de la Wikipedia en español y del proyecto OPUS. La parte en español del corpus BERTIN contiene unos 235 billones de palabras. Los datos de entrenamiento proceden del Archivo Web del Español recopilado por la Biblioteca Nacional de España entre 2009 y 2019, con más de 135 billones de palabras.

En los modelos de IA existe la posibilidad de un sesgo en la representación de registros de la lengua y en la proporcionalidad de datos de las distintas lenguas. También existen sesgos entre lenguajes de especialidad recogidos en el modelo y el origen social, racial, económico y de género de los datos representados, así como en los lenguajes de especialidad recogidos en el modelo. En general, se hace necesario aportar a los modelos más y mejores datos en español.

El desarrollo de la IA tiene implicaciones medioambientales, ya que el entrenamiento de uno de sus modelos implica un procesamiento computacional masivo que consume muchos recursos de energía y agua. El preentrenamiento de un modelo como GPT3 consume tanta electricidad y agua como una casa estadounidense durante 121 años.

### III. CONCLUSIONES FINALES Y RECOMENDACIONES

---

La realidad internacional de la lengua española en 2024 revela hechos significativos que son reflejo de su peso en los terrenos de la demografía, el desarrollo, la cultura o el mundo digital. El hecho de superar los 498 millones de hablantes como lengua nativa y los 600 millones de usuarios potenciales es revelador de una comunidad idiomática vigorosa y pujante. Este hecho resulta significativo en un panorama de diversidad y multilingüismo.

Los datos aquí reunidos, asimismo, son unos índices excelentes de los espacios de crecimiento y mejora de la lengua y de su comunidad, así como de los ámbitos en los que existen mayores deficiencias. Por ese motivo, nuestras conclusiones apuntan algunas líneas de acción que convendría valorar por parte de las autoridades y entes responsables:

- a. La evolución de la comunidad hispanohablante en el espacio hispánico revela un crecimiento de la población bilingüe. Este hecho muestra claramente la necesidad de crear programas de educación auténticamente bilingüe allí donde los sistemas educativos públicos no la contemplan. Los programas de educación bilingüe para indígenas, monolingües y bilingües, han de ir ligados a la lucha contra el analfabetismo y a la creación de oportunidades.
- b. La evolución del grupo de hablantes de español con competencia limitada revela la necesidad de crear, flexibilizar y desarrollar programas que atiendan las demandas formativas de la población migrante, allí donde la inmigración hispana es más cuantiosa, de modo que se posibilite el mantenimiento de la lengua española más allá de la segunda generación.
- c. En el campo de la traducción, es importante mejorar la aportación de los países hispanohablantes en especialidades como las ciencias naturales y

aplicadas, o la tecnología, para hacer más atractiva la traducción de conocimientos a otras lenguas.

- d. 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 creación, formación y actualización del profesorado de español.
- e. Para el desarrollo del español en las redes, es importante posibilitar técnicamente su uso en todo tipo de programas y aplicaciones, en todo tipo de tecnologías del lenguaje, desarrollando herramientas pensadas específicamente para el español, sin menoscabo de su desarrollo para otras lenguas del espacio hispánico.
- f. De nuestros análisis complementarios se obtienen conclusiones de gran interés para los sectores de referencia:
  - 1. En cuanto a los medios de comunicación, la evolución social de los medios basados en el audio aconseja prestar atención y apoyar tanto el desarrollo tecnológico de la audificación, como la elaboración de contenidos de calidad en español para este soporte.
  - 2. El ámbito de la traducción evidencia la necesidad de mejorar el conocimiento expresado en español más allá de las humanidades y de la creación literaria, especialmente en los ámbitos de la ciencia, la tecnología y, en general, la no ficción. La relativa debilidad del sector de la traducción en Estados Unidos ofrece la posibilidad de introducir desde fuera traducciones en lenguas distintas del inglés.
  - 3. En el ámbito de las industrias culturales, la producción en español presenta un amplio margen de mejora en la penetración y el consumo de productos televisivos. La fuerte posición de la industria de la música en español supone una oportunidad para la introducción y difusión de todo tipo de estilos y modalidades musicales.

4. En el ámbito de la inteligencia artificial, es evidente la necesidad de aumentar la base de corpus en español, tanto en volumen, como en diversidad y refinamiento, evitando o paliando los sesgos existentes de registros y estilos textuales, así como los sesgos socioculturales, económicos, raciales o de género.
- g. En el campo de la política, sigue siendo necesario contribuir al prestigio de la lengua mediante acciones diplomáticas para el uso institucional y público del español, así como mediante programas de difusión cultural de calidad, de apoyo a la difusión de la ciencia en español y de apoyo a los medios de comunicación en español en cada uno de los entornos comunitarios, hispanohablantes y no hispanohablantes.

Estas acciones deberían ser debidamente contextualizadas y perfiladas de acuerdo con el repertorio lingüístico de cada país o territorio. Asimismo, tales acciones deben proponerse desde la conciencia de que el peso internacional de la lengua española trasciende la suma de las aportaciones particulares de cada uno de los países en que el español tiene presencia. En este caso, de nuevo, el todo es mucho más que la suma de las partes.





