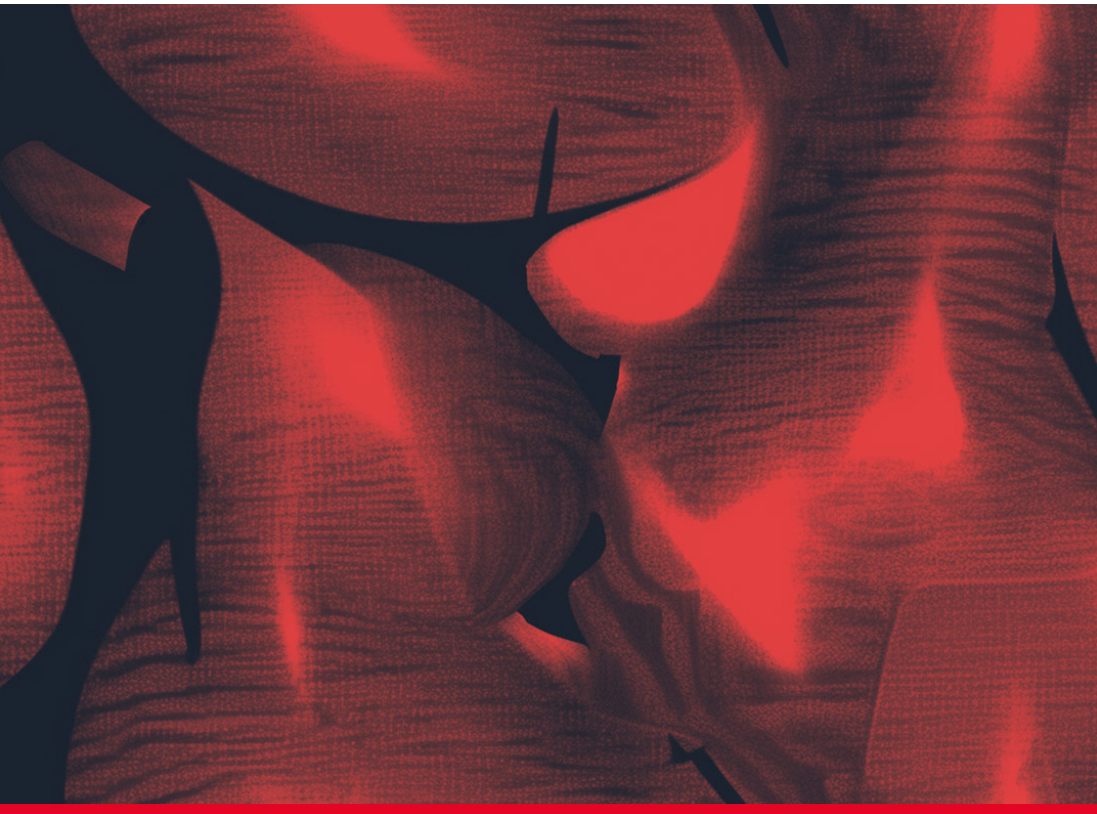
2025 Report

SPANISH: A LANGUAGE TO THE WORLD



Instituto Cervantes

Libreros, 23. 28801 Alcalá de Henares (Madrid)
Alcalá, 49. 28014 Madrid
E-mail: informa@cervantes.es
https://www.cervantes.org

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

HEAD EDITOR: Álvaro García Santa-Cecilia
EDITOR IN CHIEF: Rebeca Gutiérrez Rivilla
COPYEDITORS: Germán Hita Barrenechea, Miguel Ángel Carrón Sánchez, Almudena Labrador Bravo and María Dolores Pérez Vidal

SPANISH: A LANGUAGE TO THE WORLD 2025:

AUTHOR: Observatorio Global del Español
MANAGING EDITOR: Francisco Moreno Fernández
LAYOUT ADAPTING: Departamento de Cultura Digital

INSTITUTO CERVANTES

NIPO (PDF): 110-25-057-6

Cover design and layout: Departamento de Cultura Digital

2025 Report

SPANISH: A LANGUAGE TO THE WORLD

Instituto Cervantes

Contents

SPANISH:
A LIVING LANGUAGE

+635

million people (635.743.644) are potential Spanish speakers worldwide.

519.115.258
people with **native language proficiency**

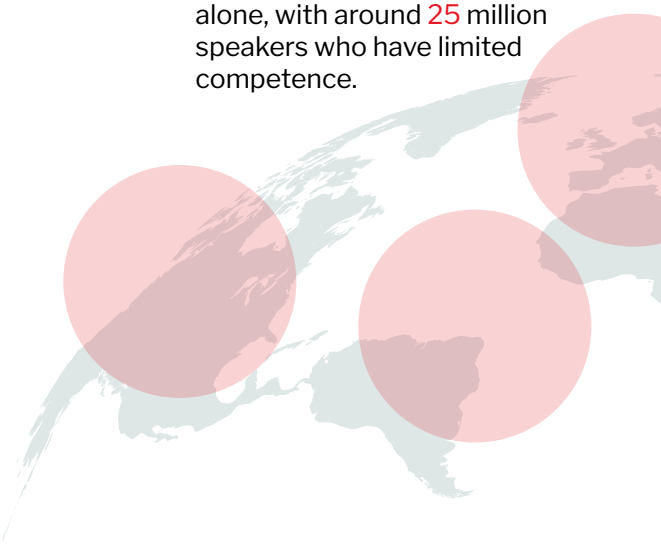
limited competence users
92.068.243

24.560.143
Spanish **learners**

There are around

460

million native Spanish speakers in Spanish-speaking countries alone, with around 25 million speakers who have limited competence.



Spanish is the

3rd **mother tongue** in the world, after Mandarin Chinese and Hindi and the 2nd among the official languages of the UN.

For the first time, **the number of native Spanish speakers** is set to exceed

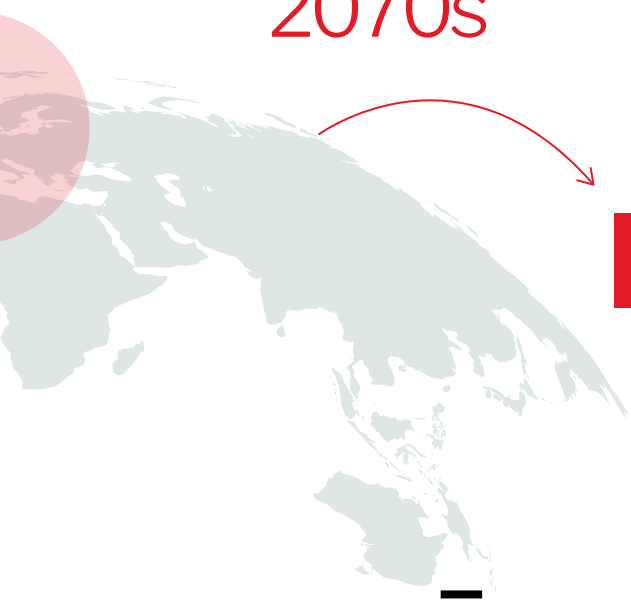
500 million
(+22 % since 2012)

The **Spanish-speaking world** is facing intense **urbanisation,** which will affect

86 %
of the population by 2050.

The estimated native Spanish-speaking population will peak in the

2050s and 2070s



Since 2012, the limited competence speaker group has grown by

↑ 79 %

This population could exceed **200 million by the end of the 21st century,** according to the most conservative estimate.

78 %

of **Spanish learners** are concentrated in the **United States, the European Union and Brazil.**

One in ten native Spanish speakers lives in a non-Spanish-speaking country, highlighting its significance as a language of migration.

More than 127 million people outside Spanish-speaking countries are potential Spanish speakers.

Since 2012, Spanish language learners **have increased** by

↑ 36 %

due to the institutionalisation of Spanish as a foreign language and its growing presence in educational and social setting.

SPANISH: A LANGUAGE TO THE WORLD 2025

OBSERVATORIO GLOBAL DEL ESPAÑOL

Spanish around the world: 20 key facts

1. Spanish is spoken by well over 630 million people worldwide.

2. By 2025, the number of native Spanish speakers is set to exceed 500 million for the first time, reaching 520 million.

3. The native Spanish-speaking community has become the third largest in the world, behind Mandarin Chinese and Hindi. The Hindi-speaking community is the world's second largest native community due to India's rapid population growth.

4. There are around 460 million native Spanish speakers in Spanish-speaking countries alone, with around 25 million speakers who have limited competence.

5. One in ten native Spanish speakers lives in a non-Spanish-speaking country, highlighting its significance as a language of migration.

6. Over 120 million people outside Spanish-speaking countries are potential Spanish speakers.

7. The European Union is home to more than 45 million Spanish speakers, excluding those living in Spain.

8. Since 2012, native speakers have increased by 22% overall, although natural demographic growth has slowed while emigration to Spanish-speaking countries has increased.

9. The limited competence speaker group has grown by 79% since 2012, driven by Spanish language teaching in school systems in several European countries, the United States and Brazil.

10. Since 2012, Spanish language learners have increased by 36% overall, due to the institutionalisation of Spanish as a foreign language and its growing presence in educational and social settings.

11. The number of Spanish learners between 2024 and 2025 has grown by 1.5%. The current growth rate could lead to 100 million students by the end of the century, assuming that the teaching of Spanish is sufficiently institutionalised.

12. The United States, Brazil and numerous European Union countries account for the largest share of Spanish learners.

13. The Spanish-speaking world is facing intense urbanisation, which will affect 86% of the population by 2050. This will negatively impact the use of minority languages, while benefiting Spanish acquisition in groups of speakers who are unfamiliar with the language.

14. The estimated native Spanish-speaking population will peak in the 2050s and 2060s.

15. The limited competence Spanish-speaking population could exceed 200 million by the end of the 21st century, according to the most conservative estimate. One in three people could be a limited competence Spanish speaker (second language or foreign language) in non-Spanish-speaking countries.

16. The combined use of Spanish and English holds the most promise for international communication, given the collective size of their native communities (around 1 billion speakers), the number of countries in which one or both are official languages (over 75), and their prestigious cultures and knowledge systems.

17. Overall, Spanish-speaking countries are less unequal in terms of Human Development Index (HDI) than countries where other languages are spoken, such as English or French.

18. The declaration that English is the official language of the United States does not necessarily affect the day-to-day life of the American people, but it may harm the social and linguistic rights achieved, inherited or acquired by the Hispanic population.

19. The percentage of Spanish-language web pages is the second highest in the world, although it is far behind the number of English-language pages.

20. Language analysis of user comments on social media reveals a trend toward shorter statements and lexical creation.

I. SPANISH AROUND THE WORLD 2025: DEMOGRAPHY[1]

1. Introduction

This report analyses the structure of the potential Spanish-speaking community in 2025, i.e. the group of people who, under different conditions and circumstances, can interact in Spanish. Spanish speakers can be found practically everywhere in the world, albeit with varying profiles and in widely differing proportions.

At present, Spanish language demographics reveal two quantitatively significant facts. First, there is the consolidation of a community of potential speakers over 630 million strong. Second, the number of native speakers has exceeded 500 million for the first time in history. These figures come from sources and language demographics methods that are constantly being reviewed and refined, but which are currently the most reliable.

Demography constitutes a major factor in describing and analysing any linguistic community. Demography underpins a language's vitality when it is thriving, ensuring intergenerational transmission while highlighting the different profiles of its speakers. Therefore, to understand the sociolinguistic composition and evolution of Spanish or any other language, it is essential to focus on the demographics of its speaker groups. The linguistic demography of Spanish, together with other quantitative and qualitative factors, provides an overview of the current situation of the language.

Similar to the 2024 report and previous editions, data is presented on Spanish speakers who constitute a *Native Proficiency Group* (NPG) in

1. Report by Francisco Moreno Fernández and Héctor Álvarez Mella, in collaboration with Mariel Elizondo Romo and Jana Weckesser, from the Global Observatory of Spanish.

countries where these speakers have a significant presence. These countries include both those that are historically and socio-politically considered Spanish-speaking, irrespective of whether the language is official or not, and those in which there are varying contingents of native Spanish speakers.

This report also covers speakers who, while not being native speakers of the language, have sufficient knowledge to communicate in Spanish with native and non-native speakers. This group of speakers is referred to here as the *Group with Limited Competence* (GLC) due to the limitations or restrictions in sociolinguistic or stylistic appropriateness, as well as fluency. These limitations may be evident in their use of the language for communicative interactions. Language skills and their development are difficult to measure because they are blurred by the complexity of cognitive abilities and the multiplicity of social and interpersonal functions that language can fulfil in each context. Different speaker profiles can be included in this group, from heritage speakers (descendants of Spanish-speaking immigrants) who have not benefited from adequate conditions for language maintenance, to graduates of education systems who did not achieve very high levels of linguistic proficiency.

Lastly, this report examines the linguistic demography of those who are engaged in institutionalised or regulated language learning processes in some way. The analysis of the *Group of Spanish Language Learners* (or in its Spanish acronym, GALE) is also faced with complex obstacles to overcome. Language teaching has countless variations across the Spanish-speaking world and in international settings, and no complete, comparable, or up-to-date data exists for each of these contexts. The methodological challenges mentioned above have been overcome by using linguistic demography resources and tools, along with epistemological decisions, which will be explained in due course.

The comparison of Spanish and its communities with other speaker groups, both *international languages* and other languages of great

vitality, is becoming increasingly relevant at a time when the number of potential Spanish speakers has exceeded 630 million. International languages are defined as those that are used as a means of communication between speakers from different countries as native, second or foreign languages, typically across transcontinental geographical areas. This means that a language of great vitality is not necessarily considered international. This report uses these assumptions to analyse how Spanish compares with other languages around the world, especially the official languages of the United Nations (UN) (English, French, Mandarin Chinese, Russian and Arabic), but also other languages of great vitality.

The proposed comparison is not intended to encompass all external dimensions of a language, such as economics, culture, education, publishing, media, technology, etc. This objective would far exceed the scope of this report, not to mention the practical difficulties involved in obtaining truly comparable data for a significant number of languages. Therefore, this comparative analysis will focus on certain relevant aspects of languages, particularly demography. The development indices of linguistic communities, their most recent evolution and their future prospects will also figure prominently. To this end, the annual reports issued by the Instituto Cervantes, which have been produced using a common or similar methodology for over a decade, are a valuable resource. They contain information that remains largely unchanged lately, either because the current data is similar or because no new data is available.

As in the report "El español: lengua para el mundo 2024" [Spanish around the world 2024] (Global Observatory of Spanish, 2024), the 2025 edition includes a dossier with several monographic analyses dedicated to different aspects of the international situation, use and dissemination of the Spanish language.

The first analysis focuses on issues pertaining to the economy of the Spanish language and examines forecasts for Spanish in terms of global

GDP, trade flows and the digital economy. The second provides an up-to-date assessment of artificial intelligence applied to the Spanish language, focusing on public strategies, resources, the business ecosystem, research groups and the impact of technologies on Spanish in all its geographical diversity. The third analyses the influence of Spanish in the linguistic resources contained in European digital infrastructures and explores the relevance and usefulness of Spanish terminology as a key linguistic resource for the development of different language technologies. Finally, an analysis and interpretation of the Hispanic or Latino vote in the latest US presidential elections is presented.

These subject-specific reports written by experts in their fields, together with the demographic report produced by a team of university academics, reflect a desire to ensure that this research is collegial rather than personal, seeking to draw on the diverse knowledge and perspectives of those involved in the task. To achieve this, the Instituto Cervantes' Hispanic Studies Department also carries out a thorough review, verification and editing process. As such, the Global Observatory of Spanish seeks to fulfil two of the aims set out in its founding regulations, namely to act as a permanent body for gathering and analysing information on the status of Spanish in different areas at an international level, and to promote studies and technical reports diagnosing the status of Spanish to foster knowledge and prestige of the Spanish language and Hispanic cultures.

2. Key linguistic demography of Spanish

2.1 Spanish in Spanish-speaking countries

Table 1 shows the population figures and linguistic demographic breakdown of the Spanish-speaking community in countries where Spanish is the official, national, common or, alternatively, majority language. The table includes data on native and limited competence Spanish speakers, updated to 2025. The key source for these figures is census data on the population of each of these countries, together with various other

pieces of information provided by their respective statistical institutes. The following guidelines were used to prepare Table 1::

1. Population data are updated to 2025, according to information provided by each country's census offices or statistical institutes. When official figures for 2025 are not available, projections made by other official or international bodies are used. In general, priority is given to national sources over international sources, except when national sources have obvious shortcomings.

2. The number of speakers in the Native Proficiency Group (NPG) is calculated based on the total population, minus potentially monolingual speakers of other languages and speakers in the Group with Limited Competence. The assumption is that native proficiency in Spanish does not preclude native proficiency in one or more other languages.

3. The number of speakers belonging to the Group with Limited Competence (GLC) is generally obtained by adding together the number of individuals belonging to indigenous or native peoples who have been educated in Spanish up to primary school level. This excludes monolingual speakers of indigenous languages who have been identified as such. This figure is supplemented by the number of foreigners from non-Spanish-speaking countries registered in each country. This is especially true for non-naturalised foreigners, for whom no evidence exists that they have achieved native proficiency in the language, although some may have done so. Other native speakers of languages other than Spanish who demonstrate or acknowledge limited use of Spanish are also included here.

4. Figures in this table and other quantitative references are provided with the accuracy resulting from the calculations. This is not to say that these are exact numbers of speakers, given the very nature of the object being quantified. The figures could be rounded according to standard mathematical criteria, but readers of these

reports can perform this exercise easily without losing the specific data resulting from the corresponding operations, which could then be subject to further verification.

Table 1 maintains the proportions of NPG and GLC speakers calculated for 2024. The percentages in the previous report were calculated *ex novo* according to updated data provided by the relevant demographic sources. These proportions are certainly not static but rather dynamic, as they are subject to conditions arising from very different social and migratory processes. However, it is unlikely that these percentages will be radically altered within a year. Furthermore, the primary sources have not been updated in such a way as to require a detailed review of these statistics. Unless the political, legislative, educational or migratory conditions of a country's population change drastically, these speaker ratios will be maintained for a reasonable period of time, pending a consistent update of the primary source data.

In the linguistic demographic estimates for Spanish-speaking territories, Cuba, Venezuela and Equatorial Guinea are the countries for which the most unreliable data are available. In the case of Caribbean countries, strong migratory movements, motivated by political and economic causes, have led to significant changes in the number of expatriates, migrants and residents, which makes it difficult to estimate the number of Spanish speakers. Official sources in these countries, if available, provide data that is outdated or somewhat unreliable, as can be seen from the sharp contrast with international sources. For Equatorial Guinea, in addition to the problem of sources, the interpretation given to the concept of proficiency in Spanish is also an issue. As explained in this report, this matter is resolved by using the concept of "proficiency in the language" as equivalent to that of a native speaker, rather than the concept of "native language", which is always delicate and open to different interpretations (Moreno Fernández, 2024).

Together with the challenges mentioned above, others have arisen, which have been dealt with in the most appropriate manner in each

case. On the one hand, residents abroad may or may not be included in the national censuses of Spanish-speaking countries, which could lead to double counting, as may also be the case with second-generation immigrants or the school-age population. On the other hand, speakers of Hispanic origin, when they become naturalised in a non-Spanish-speaking country, often see information about their origin disappear from statistics. This makes it impossible to identify them as native Spanish speakers, unless the census includes information on the language spoken at home. In such cases, the methodology applied always seeks to avoid duplication and misinterpretation, especially when dealing with large population groups.

TABLE 1 POPULATION, NATIVE SPEAKERS AND LIMITED COMPETENCE SPEAKERS IN SPANISH-SPEAKING COUNTRIES

Country	2025 Population Estimate (Calculation)	Native Proficiency Group (NPG)	Native speakers (%)	Group with Limited Competence (GLC)
Mexico	138,070,271	130,034,675	94.18%	7,952,848
Colombia	53,110,609	52,090,983	98.08%	1,019,724
Spain	49,270,000	47,250,026	95.90%	1,675,180
Argentina	47,473,760	46,866,195	98.72%	607,664
Peru	34,412,393	29,601,626	86.02%	977,312
Venezuela	29,395,334	28,631,153	97.40%	764,279
Chile	20,206,953	19,315,922	95.59%	891,127
Ecuador	18,103,660	16,963,223	93.70%	568,455

Country	2025 Population Estimate (Calculation)	Native Proficiency Group (NPG)	Native speakers (%)	Group with Limited Competence (GLC)
Guatemala	18,079,810	13,722,652	75.90%	3,283,293
Bolivia	11,425,746	9,197,806	80.50%	2,079,486
Dominican Republic	10,878,267	10,323,570	94.90%	554,792
Cuba	10,055,968	9,970,591	99.15%	1,006
Honduras	10,038,054	9,548,292	95.12%	489,857
Paraguay	7,758,263	5,222,154	67.31%	2,488,851
Nicaragua	7,218,143	7,001,696	97.00%	216,544
El Salvador	6,366,086	6,358,547	99.88%	7,639
Costa Rica	5,567,811	5,510,005	98.96%	57,905
Panama	4,544,816	3,925,898	86.38%	606,278
Uruguay	3,588,937	3,557,812	99.13%	31,224
Puerto Rico	2,980,532	2,834,581	95.10%	146,046
Equatorial Guinea	1,682,611	1,245,206	74.00%	437,479
TOTAL	490,228,024	459,172,613	93.67%	24,856,989

SOURCE: compiled by the author. Data in Table 1 is sorted by total population size in each country.

TABLE 1 SOURCES AND ADDITIONAL INFORMATION1

Argentina

Data from the National Institute of Statistics and Census of Argentina (INDEC, 2023) projected to 2025 (INDEC, 2024). The GLC includes literate speakers of indigenous languages with incomplete primary education, as well as foreigners of non-Spanish-speaking origin.

Bolivia

Data from the National Institute of Statistics of Bolivia based on the 2024 census, which amends the forecasts that existed for 2025 (INEBolivia, 2024). The 2025 population forecast is reached by applying the 1% annual population growth forecast by *The World Factbook* (Central Intelligence Agency, n.d.). The GLC includes non-monolingual speakers of indigenous languages and those with low levels of schooling, as well as foreigners of non-Spanish-speaking origin.

Chile

Based on data from the National Institute of Statistics of Chile (INE Chile, 2017a). Data from the 2024 Population and Housing Census have not yet been included. According to official data, the first set of indicators from the 2024 census will be released in June 2025. The population forecast for 2025 comes from INE Chile (2024). The GLC includes low-education, non-monolingual speakers of indigenous languages, as well as foreigners of non-Spanish-speaking origin. Data accessed through REDATAM (ECLAC, n.d.-b).

Colombia

Data from the National Administrative Department of Statistics of Colombia (DANE, 2025). The percentage of the population of indigenous people in 2018 (4.4%) is applied to 2025. Non-literate indigenous peoples are assumed to be non-Spanish speakers.

Costa Rica

Information from the National Institute of Statistics and Census of Costa Rica based on the 2022 census (INEC Costa Rica, n.d.). INEC forecast for 2025 (INEC Costa Rica, 2024). The population of indigenous peoples is 1.3%. Illiteracy among the population of indigenous peoples is 4.8%, who may not speak Spanish.

Cuba

The population figure includes registered residents (Cuban News Agency, 2024; Figueredo Reinaldo *et al.*, 2024), although they may have partial residence in other countries (Cuban News Agency, 2024). The number of permanent residents in Cuba is estimated to be 8,620,000 (Álvarez Quiñones, 2023). These figures coincide with those provided by the UN (2024b). Its forecasts point to the possibility that Cuba's population could fall to around 6 million by 2100.

Ecuador

Based on data from the National Institute of Statistics and Census of Ecuador (INEC Ecuador, 2022) and the institute's own population forecast for 2025 (INEC Ecuador, 2024). According to INEC, the population had already exceeded 8 million in January 2025. The

GLC includes low-education, non-monolingual speakers of indigenous languages, as well as foreigners of non-Spanish-speaking origin.

El Salvador

Data from the Central Reserve Bank of El Salvador (BCR, 2022) and the National Office of Statistics and Census of El Salvador (ONEC, n.d.). The GLC includes non-Spanish-speaking foreigners.

Spain

Data from the National Institute of Statistics of Spain (INE España, 2024a, 2024c). For the NPG, the proportion of the total population that does not speak any Spanish or Castilian and those who speak it poorly is calculated based on data from 2021 (INE España, 2021).

Guatemala

Data from the National Survey on Quality of Life and Well-being in Households (ENCABIH) (INE Guatemala, 2024). The population forecast for 2025 was provided by the National Institute of Statistics of Guatemala (INE Guatemala, 2020; INE Guatemala, National Statistics Office of the Dominican Republic and Foundation for Strategic Analysis and Development of Small and Medium-sized Enterprises, 2024). The GLC includes non-monolingual speakers of indigenous languages and those with low levels of schooling, as well as foreigners of non-Spanish-speaking origin.

Equatorial Guinea

Data from the 2015 census (Republic of Equatorial Guinea, 2016; National Institute of Statistics of Equatorial Guinea INEGE, 2023). A new census is scheduled for 2025. The population growth rate forecast by the UN Department of Economic and Social Affairs (2022a) has been applied. The Government of Equatorial Guinea forecast a population of 1,612,677 for 2023 (INEGE, 2024a). The GLC includes indigenous, Creole and European language speakers. The level of linguistic competence is a criterion that causes discrepancies when estimating the number and proportion of Spanish speakers in Equatorial Guinea. This report continues to use the criteria employed in previous analyses (Moreno Fernández and Otero Roth, 2006) to assess how the community of speakers has evolved, pending a thorough review of this estimate. The Native Proficiency Group includes those who are native speakers, speak the language with native-like proficiency, and can be considered as such in interactions with other speakers. This criterion is based on non-sequential grounds. It should be noted that: 1) no other African language (Swahili, Wolof, Hausa, etc.) is used as a national common language in Equatorial Guinea (Eberhard, Simons and Fennig, 2025), unlike in other countries; 2) that the percentage of the population that is literate in 2024, according to official figures, is 90.1% (INEGE, 2024b) and that literacy is taught in Spanish.

Honduras

Data from the Continuous Multi-Purpose Household Survey (EPHPM) of the National Institute of Statistics of Honduras (INE Honduras, 2017, 2024a), as well as from BASEINE (INE Honduras, 2024b). The population growth rate was taken from the UN Department of Economic and Social Affairs (2022a). The GLC includes speakers of indigenous languages and the English-speaking black population, as well as foreigners of non-Spanish-speaking origin.

Mexico

Population data from the National Institute of Statistics and Geography of Mexico (INEGI, 2020, n.d.), the National Population Council of Mexico (CONAPO, 2024) and the Government of Mexico. Forecast for 2025 by CONAPO and the Mexican Ministry of the Environment and Natural Resources (CONAPO-SEMARNAT, 2020). The GLC includes the corresponding proportion of indigenous speakers with poor Spanish proficiency and foreigners of non-Spanish-speaking origin (INEGI, 2022).

Nicaragua

Data from the National Institute of Development Information of Nicaragua (INIDE, n.d., 2025), as the last census was conducted in 2005 (INIDE, 2005). The country profile from the Diplomatic Information Office of the Nicaraguan Ministry of Foreign Affairs (OID, 2025) has been incorporated.

Panama

Data from the National Institute of Statistics and Census of Panama (INEC Panama, n.d.). The 2025 forecast is based on information from the Economic Commission for Latin America and the Caribbean (ECLAC, n.d.-a). The GLC includes low-education, non-monolingual speakers of indigenous languages and foreigners from non-Spanish-speaking countries. Additional information provided by the Private Sector Council for Educational Assistance (COSPAE, 2025).

Paraguay

Data from the 2022 census (INE Paraguay 2022; 2024a). National population forecast for 2025 by the National Institute of Statistics of Paraguay (INE Paraguay, 2024b). The NPG includes monolingual Spanish speakers and bilingual Spanish and Guaraní speakers, using data collected through the REDATAM system (ECLAC, n.d.-b).

Peru

The latest Peruvian census was in 2017, which has led to inconsistent forecasts and estimates. The most recent data from the National Institute of Statistics and Informatics of Peru (INEI Peru, 2024) reduces the forecast population in previous reports. However, the information from INEI and the Peruvian Ministry of Health (MINSA, 2024) does not match completely. The NPG includes indigenous peoples with secondary and university education in Spanish. Data accessed through the REDATAM system (ECLAC, n.d.-b). The GLC includes low-education, non-monolingual speakers of indigenous languages, as well as foreigners of non-Spanish-speaking origin.

Puerto Rico

Data from the census taken by the Institute of Statistics of Puerto Rico (State Data Center-Puerto Rico, 2023). The 2025 population forecast is from the same institute (State Data Center-Puerto Rico, 2025). To calculate the GLC, speakers of English and other languages spoken at home are deducted from the NPG.

Dominican Republic

Data from the National Statistics Office of the Dominican Republic (ONE, 2023). Estimated by ONE for 2025 (ONE, 2025). The population of the Dominican Republic is estimated to

grow by around 1% in 2025 (Domínguez, 2024). The GLC includes non-Spanish-speaking foreigners, mainly Haitians.

Uruguay

Forecast by the National Institute of Statistics of Uruguay, based on the 2023 census (INE Uruguay, 2023). The GLC includes non-Spanish-speaking foreigners.

Venezuela

Given the lack of clarity surrounding the figures provided by various sources, this report has kept the same population estimate as in the 2024 report. There are discrepancies between the information provided by the National Institute of Statistics of Venezuela, whose website is often down, and other sources, which estimate the population to be between 26 and 30 million. Some sources quote a population of 26.68 million, based on data from the International Monetary Fund (Statista, 2024b); other sources report a population of 30.55 million, supposedly based on data from the UN Department of Economic and Social Affairs (Countrymeters, n.d.). The Inter-Agency Coordination Platform for Refugees and Migrants from Venezuela puts the number of Venezuelan refugees or migrants worldwide at 7.89 million (R4V, 2024). No official population forecast or growth projection for 2025 is known to exist. The population forecasts of the Latin American Demographic Center (CELADE) are from 1983 and cannot be considered (ECLAC, 1983). The GLC includes non-monolingual speakers of indigenous languages and those with low levels of schooling, as well as foreigners of non-Spanish-speaking origin.

Besides Cuba and Venezuela, especially the latter, a comparison of the linguistic demographic data for 2024 and 2025 reveals some interesting facts. Firstly, the demography of Spanish-speaking countries has grown at an average rate of 1.3% per year, which confirms or consolidates the significance of the native community within the total number of potential Spanish speakers. This growth in total population implies a proportional growth in the NPG and the group of speakers with limited competence (GLC): the former stands at around 460 million native speakers and the latter at around 25 million.

The percentage of native speakers in Spanish-speaking countries as a whole is over 90%, as in previous years. If the figures for native speakers remain the same in 2025, Guatemala, Bolivia, Paraguay and Equatorial Guinea will remain below 85% within this group, which is the lowest in the entire Spanish-speaking world. As for GLC, Cuba continues to record low figures due to its small foreign population. Meanwhile, Mexico, Guatemala, Bolivia and Paraguay continue to

record high percentages due to the size of their bilingual indigenous peoples, as does Spain because of its non-Spanish-speaking immigrant population.

However, a deeper understanding of the data presented in Table 1 requires an overview of the countries where Spanish is the majority language and their immediate surroundings. In 2025, the Latin American and Caribbean region is expected to have a population of over 600 million, which naturally includes Spanish, Portuguese and speakers of other languages (World Bank, n.d.-a, 2024), although the former have a majority share.

For linguistic demography purposes, the annual percentage of population growth has been gradually declining. Latin America and the Caribbean have seen their growth rate fall from 1.5% in 2000 to 1.1% in 2012 and 0.7% today. In Spanish-speaking countries, this rate stands at around 1%, which clearly indicates that the Hispanic community is gradually seeing a decline in its demographic growth.

Furthermore, the future consequences of deportations from the United States will also have to be considered. These forced return migrations may influence the demography of Spanish-speaking countries in the near future. They could also have a knock-on effect on migration from Latin America to Europe, especially to Spain, a phenomenon that occurs every time the United States tightens its immigration policies.

Other demographic indicators point to a gradual improvement in living conditions across this vast region of the Americas. For example, the infant mortality rate (under 5 years old) has fallen from 34.4% in 2000 to 15.7% in 2024, and the percentage of the population with access to the internet has risen from 42.2% in 2012 to 75.6% in 2024. However, other indicators show a decline or stagnation in quantitative terms, despite improvements revealed by these figures. GDP per capita growth, as an annual percentage, fell from 2.1% in 2000 to 1.4% in 2024, improving by just 0.2% since 2012. At the same time, life expectancy at birth has stagnated since 2012, falling from 77 years for women and 76.7 years for

men to 76.7 years for women and 70.5 years for men. As for education, the primary school completion rate, as a percentage, has fallen since 2012 by 2.3%, from 96.9% to 94.6% (ECLAC, 2022).

This brief overview depicts a global region that, over the last quarter of a century, has seen a general improvement in its progress and well-being indicators. However, over the last twelve years, several areas have deteriorated. All this is set against a backdrop of highly diverse political environments that alienate Latin America from the image of a homogenised region that it has projected in other periods.

Finally, looking at it from a sociolinguistic angle, the development of trends towards bringing more regional or local dialects closer to national and international linguistic standards is noticeable in both the western and eastern Atlantic. This is undoubtedly influenced by factors stemming from easier mobility and contact between peoples of different origins through interregional media and social media, where contact tends to lead to standardisation and the search for common usage. In addition, the rise in the percentage of the population enrolled in education throughout the Hispanic world, and especially the number of university students, must also be considered. This leads to a convergence towards the most prestigious linguistic standards (Bravo García, 2008; Masuda, 2020).

2.2 Spanish in non-Spanish-speaking countries

It is extremely difficult to know the number of Spanish speakers living or residing in countries around the world, based on any of the linguistic profiles covered in this report. There may well be Spanish speakers in every country and territory in the world. A simple statistic illustrates this widespread presence: between 2024 and 2025, candidates from 200 countries took the test of constitutional and sociocultural knowledge of Spain (known as the CCSE test), organised by the Instituto Cervantes and designed to enable candidates to obtain Spanish citizenship based on residence. However, collecting up-to-date and

reliable information on Spanish speakers in all countries of the world is impossible in some cases, difficult in others and, ultimately, futile or inconsequential in many. Therefore, this report focuses individually on around thirty countries outside the Hispanic world where Spanish has a significant presence or could play an important role in the future expansion of the language. As explained in previous reports, certain cases involve territories where Spain or other Spanish-speaking countries have maintained Spanish-speaking populations for a long period of time, such as the United States and the Philippines. In other countries or territories, the Spanish-speaking presence has its origins in more recent migratory movements, such as in Canada, Australia, the United Kingdom, Switzerland and the European Union itself. Finally, knowledge of Spanish may be motivated by geographical proximity, cross-border dynamism, the importance of socio-economic exchanges between neighbouring countries (Belize, Brazil, Andorra, etc.) or by various historical or symbolic reasons that have emerged at different times. The remaining countries or territories, apart from the thirty mentioned, have been assessed as a whole based on the most reliable sources available.

TABLE 2 SPANISH-SPEAKING POPULATION IN NON-HISPANIC COUNTRIES AND TERRITORIES

Country	Native Proficiency Group (NPG)	Group with Limited Competence (GLC)	Total potential Spanish speakers
United States	45,481,557	20,000,000	65,481,557
EU-27	11,433,891	34,301,673	45,735,564
United Kingdom	217,974	5,564,106	5,782,080
Brazil	1,111,502	1,931,480	3,042,982

Country	Native Proficiency Group (NPG)	Group with Limited Competence (GLC)	Total potential Spanish speakers
Morocco	136,892	1,751,733	1,888,625
Canada	321,407	1,043,892	1,365,299
Australia	185,209	386,573	571,782
Switzerland	213,776	351,259	565,035
Philippines	6,834	547,696	554,530
Israel	79,827	360,150	439,977
Belize	38,908	199,256	238,164
Russia, Belarus, Kazakhstan and Kyrgyzstan	55,540	153,776	209,316
Netherlands Antilles	36,651	140,693	177,344
Algeria	1,348	173,600	176,948
Trinidad and Tobago	48,800	68,417	117,217
Aruba	17,975	78,140	96,115
Japan	74,413	17,004	91,417
Andorra	35,232	38,756	73,988
Guam and Northern Mariana Islands (United States)	-	42,300	42,300
China	19,849	12,510	32,359

Country	Native Proficiency Group (NPG)	Group with Limited Competence (GLC)	Total potential Spanish speakers
Norway	27,367	4,090	31,457
New Zealand	27,523	92	27,615
Iceland	2,985	19,554	22,539
South Korea	4,395	10,268	14,663
Moldova	10,632	5,203	15,835
Türkiye	6,858	5,115	11,973
India	7,466	3,446	10,912
Jamaica	8,514	330	8,844
Egypt	938	47	985
Palestine	657	95	752
Other countries	327,725	-	327,725
TOTAL	59,942,645	67,211,254	127,153,899

SOURCE: compiled by the author. Countries ranked by number of potential Spanish speakers.

TABLE 2 SOURCES AND ADDITIONAL INFORMATION

Andorra

The percentage of native speakers has been calculated at 40% of the population of Andorra in 2025 (88,082), corresponding to speakers who declare Spanish as their first language (Government of Andorra, 2022, p. 10). For speakers with limited competence, the percentage applied to the total population of Andorra (Government of Andorra, 2022, p. 10) is 44%, corresponding to speakers who declare Catalan as their first language, although they may also be native speakers of Spanish. It should be noted that research published in 2021 by Jiménez Salcedo found Spanish to be the "language that everyone knows how to

speak", the *lingua franca,* the "common language" or the "default language" of the country among the Andorran population (Jiménez Salcedo, 2021). The new Law on the Native and Official Language of Andorra, approved on 25 April 2024 at the ordinary session of the General Council, establishes that, as of 2026, 30 hours of Catalan language training will be required to renew residence and work permits. The population of Andorra as of 31 December 2024 was 87,097, of which 16.7% were of foreign origin, which has doubled recently (*La Ciutat,* 2025).

Netherlands Antilles (Bonaire, Curaçao, Sint Eustatius, Saba)

Native speakers from Curaçao are those listed in the census as Spanish speakers at home (Central Bureau of Statistics Curaçao, 2023), forecast to 2025. Limited competence speakers are those listed in the census as Papiamento speakers, forecast to 2025. Spanish and Papiamento speakers in the other territories are estimated based on percentages provided by Statista (2024a), which are calculated based on the total population. The percentages are as follows: Bonaire (Spanish speakers: 76%; Papiamento speakers: 88.4%); Sint Eustatius (Spanish speakers: 31.5%; Papiamento speakers: 31.5%); Saba (Spanish speakers: 31.8%; Papiamento speakers: 97.6%). These figures indicate that a large part of the population is bilingual or trilingual. General population data are taken from the Central Bureau of Statistics Curaçao (n.d.).

Algeria

Native speakers include Spanish nationals residing in Algeria (INE España, 2024b) and Mexican nationals residing in Algeria (Institute for Mexicans Abroad, 2020). Limited competence speakers include Sahrawi refugees in Algerian territory, where Spanish is co-official and has a significant educational presence. The number of Spanish speakers is estimated based on the population over 50 years of age: the total population is around 173,000, and those over 50 are around 30,000, according to data from the UN Refugee Agency (2018). Spanish speakers in Western Sahara are included in the figures for Morocco.

Aruba

Data from the Central Bureau of Statistics Aruba (2021, 2023). This includes residents from Spanish-speaking countries. Limited competence speakers include those who know Papiamento, which is approximately 70% of the population (Sato Center, 2012).

Australia

Based on the 2021 census (Australian Bureau of Statistics, 2021), with forecast figures for 2025. The number of Spanish speakers at home would be 171,370. For Group with Limited Competence speakers, the estimated figure for 2024 has been updated to 2025.

Belize

According to the 2022 census, 56.55% of Belizeans speak Spanish (Statistical Institute of Belize, 2022). A forecast has been made for 2025. According to the International Organisation for Migration (IOM, 2022), 80% of immigrants born abroad come from Spanish-speaking countries. Limited competence speakers are taken from the total number of Spanish speakers, minus those born in Spanish-speaking countries. Many Belizeans who speak Spanish may do so with native proficiency, but this is not assumed here.

Brazil

Native speakers from Spanish-speaking countries are measured based on the National Migration Registry (international immigrants registered by the Federal Police) (UNICAMP, 2024). In 2025, Venezuelans clearly outnumber those from other countries. According to information from the Brazilian school association Cultura Inglesa and Babbel, Juliana Pazetti puts the number of Brazilians who speak Spanish at less than 1% (Nuadda, n.d.). This figure would currently represent a total of more than 2 million native Spanish speakers. Limited competence speakers are estimated based on the border register of the Brazilian Institute of Geography and Statistics (IBGE, 2024) and a joint study of Brazil's land borders (Moura, 2024). Given that the border population is around 10 million, a conservative estimate of the number of speakers with linguistic skills that could be described as border Portuñol is put forward. This population group with limited Spanish competence should be joined by students who have already graduated from the education system (both formal and informal) and heritage speakers who have been able to preserve the language and who, based on IBGE data, could be estimated at between 500,000 and 1 million people. This group could also include Sephardic speakers who, especially in the older generations, have been able to preserve a Judaeo-Spanish variety of Spanish..

Canada

Based on data from Spanish speakers at home: 1.76% (Statistics Canada, 2025). Percentage projected for the population in 2025 (41,206,022). For limited competence speakers, the percentage of Hispanic origin (3.3%), which has increased significantly in recent years, has been subtracted from the percentage of Spanish speakers at home. The Canadian estimate of Judaeo-Spanish speakers, at different levels of competence, is around 5,000.

China

Data on Spanish speakers in China is limited and difficult to access. Records for 2020 are available for Peruvians (10,455), based on UN data (Datosmacro, n.d.-a), and Mexicans (3,261) (Institute for Mexicans Abroad, 2020). For Spanish residents in China, we have records from the INE (4,675) (INE España 2024b). The downside to these figures is that they cannot be projected with any degree of reliability and are based on consular records, which not all expatriates register with when they move to a country. For the NPG, those who obtained a DELE level C certificate before 2019 are included. For the GLC, they include those who obtained a DELE level B certificate before 2019.

South Korea

Data on Spanish speakers in South Korea comes from the South Korean government. The information published by the National Geographic Information Institute of South Korea (NGII, 2021) for Hispanic countries is not exact, but is reported in ranges. Average values from these ranges have been used for each country. The largest Hispanic foreign residents come from Mexico, Peru and Spain. The NPG includes people with a DELE level C certificate before 2019; the GLC includes people with a DELE level B certificate before 2019.

Egypt

Data on native speakers corresponds to Spanish residents (INE España, 2024b) and Mexican residents (ben el Exterior, 2020), in addition to those who obtained a DELE level C certificate before 2019. The GLC corresponds to people who obtained a DELE level B certificate before 2019.

United States

The US population on 1 January 2025 is 341,178,551 (US Census Bureau, 2025). The forecast for the Hispanic population in 2025 is 19.72%. Spanish is spoken at home by 67.6% of Hispanics (US Census Bureau, 2023). However, the Pew Research Center reports the results of a National Latino Survey which found that 75% of Hispanic adults say they can hold a conversation, understanding and speaking very well, at least (Mora and López, 2023). However, this percentage has been gradually declining throughout the 21st century due to migration patterns and the large share of Hispanics who were born in the United States. For limited competence speakers, an estimate is provided that includes Hispanics who would retain communication skills because they acquired Spanish as a heritage language, as well as those who learned Spanish through the education system and retain the Spanish they learned, and those who may have retained Judaeo-Spanish, especially older Sephardic Jews. As the identification between Hispanic and Spanish speaker is waning, a rounded figure for the GLC is used here. The unauthorised or undocumented Hispanic or Latino population, which currently stands at around 11 million (Passel and Krogstad, 2024), should be included in the census. The current political situation does not favour the assumption that this situation will remain stable.

Philippines

Data updated from the 2020 Population and Household Census and forecast to 2025 (Philippine Statistics Authority, 2024). For native speakers in the Philippines, figures for Spanish nationals (INE España, 2024b) and Mexican nationals living abroad (Institute for Mexicans Abroad, 2020) have been considered, to which residents from Argentina, Bolivia, Chile, Colombia, Costa Rica, Ecuador, Peru and Puerto Rico (UN, 2020) have been added. The members of the 127 households in which Spanish is spoken, according to the 2020 census (Badillo Matos, 2024), are included: 4.1 members per household. However, residents of Spanish origin (4,952 in 2020) are not included, as they may coincide with the Spanish population residing in the Philippines, which has already been counted. For the Group with Limited Competence, those registered in the 2020 census as Chabacanos Caviteños, Cotabateños, Davaeños and Zamboangueños (Philippine Statistics Authority, 2023) have been considered, with projections to 2025. Chabacano speakers from these ethnic groups are included in the Group with Limited Competence because of their foreseeable limitations in communicating with Spanish speakers from other backgrounds, not because their varieties show any internal linguistic deficiencies. The 1,730,000 speakers who graduated from higher education pre-1986 are not included in this group because the quantitative size of this group after 40 years is unknown (Madrid Álvarez-Piñer, 2018).

Guam and Northern Mariana Islands

The Hispanic population in Guam stands at 4,522 (US Census Bureau, 2020). They are not included in the tally, as they are included as Latinos in the total for the United States. The group of speakers with limited competence is mainly made up of Chamorro speakers. It is estimated that 17.8% of the population can speak Chamorro (The Bureau of Statistics and Plans, 2020). The data is taken from *Ethnologue* (Eberhard, Simons and Fening, 2025). The inclusion of Chamorro speakers in the Group with Limited Competence for Spanish may be questioned. However, we have chosen to do so because of Spanish's strong influence on this Creole language, especially in its lexicon, and because of this population's contact with speakers of Hispanic origin.

India

Data on Spanish speakers in India is limited and difficult to access. This estimate is based on the sum of 910 Spanish nationals (INE España, 2024b), 3,300 Argentinians (followers of "Argentinos en India" on Facebook); 77 Mexicans (Institute for Mexicans Abroad, 2020); 670 Colombians (according to data provided by the Electoral Register of the Colombian Embassy in India); 20 Chileans (INE Chile, 2017b); 329 Venezuelans (followers of "Venezolanos en India" on Facebook); 1,700 Peruvians (followers of "Peruanos en India" on Facebook). For the NPG, those who obtained a DELE level C certificate before 2019 are included. For the GLC, they include those who obtained a DELE level B certificate before 2019. While the Indian census provides detailed information on the number of speakers of L1, L2 and L3, it does not offer any information on Spanish, which is recorded under the category "other languages".

Iceland

Provisional data from the ongoing study on the linguistic demography of Spanish in the Nordic countries, included in the project "Spanish in Europe" by the University of Heidelberg and the University of Zurich.

Israel

Data on native Spanish speakers corresponds to foreign residents (Datosmacro, n.d.-b). The data is based on UN sources. 73.5% of the Israeli population is Jewish (Central Bureau of Statistics Israel, 2011, 2022a, 2022 b), of which approximately 50% are Sephardic. When calculating the Group with Limited Competence, 10% of the Sephardic population is estimated to be Judaeo-Spanish speakers (Central Bureau of Statistics Israel, 2023). In 2011, the Central Bureau of Statistics Israel published figures on speakers by first language in relation to their knowledge of Hebrew. At that time, 81,300 people over the age of 20 were reported to have Spanish as their first language. Spanish-speaking groups may exist, including descendants of immigrants who arrived after 1948, mainly from Argentina and Mexico. Since many of them are naturalised Israeli citizens, information about their origin is not readily available.

Jamaica

The Joshua Project (2024) reports 8,500 Hispanics in Jamaica, mainly Cubans. The NPG includes people with a DELE level C certificate before 2019; the GLC includes people with a DELE level B certificate before 2019.

Japan

Records from governments of Spanish-speaking countries. The number is in line with that reported by Badillo Matos (2021). No up-to-date and reliable information on speakers with limited competence has been published. The NPG includes people with a DELE level C certificate before 2019; the GLC includes people with a DELE level B certificate before 2019.

Morocco

Native speakers from the Spanish-speaking population in Morocco, which includes Spanish nationals (INE España, 2024b) and an estimate of the foreign population of non-US American origin (Haut-Commisariat au Plan, 2014). Since Western Sahara was excluded from Morocco's demographic counts in 2014, an estimate of Spanish speakers in this

region has now been added. Estimated population for Western Sahara in 2025: 673,266, of which those over 50 years of age would be Spanish speakers (18% of the population) (Gómez and Garrido, 2020). Given the lack of updated sources, speakers with limited Spanish competence are calculated based on Fernández Vítores and Benlabbah (2014). The proportion of Moroccans who speak Spanish has been kept at 4.6% to avoid disrupting the time series in these reports, although the figures are out of date pending a more comprehensive and recent study. Other sources suggest significantly different figures: the Colombian ambassador to Morocco, Farida Loudaya, estimated 7 million in 2017 (Rojas, 2017). In our estimates, Western Sahara is not considered an autonomous political entity, although the corresponding figures are included in Morocco and Algeria.

Moldova

Data from the publication *Demolingüística del español en Rumanía, Bulgaria y Moldavia* [Linguistic demography of Spanish in Romania, Bulgaria and Moldova] (Bleorțu *et al.*, 2024).

Norway

Estimated data for 2025 based on provisional information from an ongoing study on the linguistic demography of Spanish in the Nordic countries, included in the project "Spanish in Europe" by the University of Heidelberg and the University of Zurich. The population forecast for 2025 has been calculated using data from the Norwegian government (Statistics Norway, 2021).

New Zealand

Official data on the Latin American population (Stats New Zealand, 2018), projected to 2025 and excluding Brazilians, was used as a starting point for native speakers. Spanish residents in New Zealand (INE España, 2024b) were added. The NPG includes people with a DELE level C certificate before 2019; the GLC includes people with a DELE level B certificate before 2019.

Palestine

Data on Spanish nationals abroad (INE España, 2024b). As is well known, not everyone who moves to another country registers with the relevant consulates. There are no records from other Spanish-speaking countries on people who have moved to Palestine. The NPG includes people with a DELE level C certificate before 2019; the GLC includes people with a DELE level B certificate before 2019. The current state of affairs in Palestine makes it impossible to provide estimates or verified data beyond what is considered here.

United Kingdom

Information on Spanish speakers is obtained from the census (Office for National Statistics, 2022), with percentages forecast for 2025. As for the Group with Limited Competence, *Special Eurobarometer 386* (Eurostat, 2012) reported that 3% of British people could hold a conversation in Spanish, while the British Council put the figure at 4% for 2014, rising to 8% in 2015 for people aged 18 to 34 (British Council, 2017). Given the growth of Spanish in the United Kingdom, despite Brexit, due to migration, tourism and education, estimates suggest that by 2025 the percentage could be 12% of the population, especially among young people. *Special Eurobarometer 540* (Eurostat, 2024a) no longer includes data on the United Kingdom.

Russia, Belarus, Kazakhstan and Kyrgyzstan

Provisional data from the ongoing study on the linguistic demography of Spanish in Russian-speaking countries, included in the project "Spanish in Europe" by the University of Heidelberg and the University of Zurich. Cubans account for the largest group in Russia, along with Ecuadorians and Spanish.

Switzerland

Based on data from Kabatek (2022). Growth forecast for total population: 9,058,300 (Federal Statistical Office, 2020).

Trinidad and Tobago

In 2023, around 44,800 Venezuelan immigrants arrived in Trinidad and Tobago, joining the resident population from Spanish-speaking countries. Limited competence speakers are calculated based on an estimate of 5% of Spanish speakers in the total population of the country (Grau Perejoan and Gea Monera, 2006, p. 209).

Türkiye

Native speakers include Spanish nationals (INE España, 2024b), along with approximately one thousand Colombians and Mexicans. The NPG includes individuals with a DELE level C certificate obtained before 2019; the Group with Limited Competence includes individuals with a DELE level B certificate obtained before 2019. The GLC also includes speakers of Judaeo-Spanish. The most significant communities are in Istanbul and Izmir. According to Sephardic authorities, most Sephardic Jews in Istanbul can read Spanish, but only about 20% can understand it and slightly fewer can speak it. In Izmir, two-thirds of the 1,200 members of the community are speakers. The information provided by Karen Sharhon and Jak Sigura, leaders of the Sephardic communities in Istanbul and Izmir, is greatly appreciated. Special thanks also go to Fernando Vara de Rey, director of the Instituto Cervantes in Istanbul, for his assistance.

European Union–27

Information from *Special Eurobarometer 540* (Eurostat, 2024a), which includes data on the population over 15 years of age in 2023. Data refers to this sector of the population. According to Eurobarometer, 2,207,036 people (excluding Spain) have Spanish as their first language and 3% (11,433,891) speak it very well. Those with a good or basic knowledge of Spanish are included in the Group with Limited Competence: 9% (34,301,672). The population under 15 years of age, which is in school and counted separately, is not considered.

Other countries

Data obtained from official records of Spanish-speaking citizens abroad. Countries with the largest number of residents abroad have been included: Spanish (in all countries not included in this table); Mexican (in Brazil and Asia, excluding the countries in this table); Cuban (in Haiti). Most Spanish-speaking countries have residents in other Spanish-speaking countries, so they have not been considered in this table. Some of the world's most populous countries have a small native Spanish-speaking population, which is largely included in the overall figure. Pakistan (with an estimated 3,000 native speakers) and Thailand (with an estimated 1,750 native speakers) deserve special mention. Indonesia, the fourth most populous country in the world, has an estimated native Spanish-speaking

population of 900. The floating population, mainly due to tourism (e.g. in Thailand or Indonesia), is not included in our estimates.

Table 2 shows the number of native speakers and limited competence speakers of Spanish in non-Spanish-speaking countries and territories. The notes accompanying the figures for each country give an explanation for certain interpretations and methodological decisions related to linguistic demography. The criteria explained in the following points have been considered to count the numbers for both groups and to compile the table:

1. The list of countries has been reviewed, corrected and updated in accordance with the current geopolitical situation (e.g. regarding the Western Sahara) and in accordance with new data as it becomes available (e.g. thanks to the "Spanish in Europe" project by the University of Heidelberg and the University of Zurich).

2. The population figures are updated to 2025 in accordance with the information provided by the respective census offices or statistical institutes. Where no official figures are available for 2025, the forecast made by the same official bodies or by international organisations is used. In some cases, sufficiently reliable data could not be obtained. Furthermore, the forecast has not been made when the demographic or socio-political situation may have seriously altered trends, so estimates have been made only. Depending on the date of the data, the potential impact of COVID-19 has been assumed.

3. The number of speakers in the Native Proficiency Group is calculated primarily from the first-generation immigrant population of Hispanic origin. As explained, from a methodological angle, native proficiency in Spanish does not preclude native proficiency in one or more other languages. For countries where there is a very small Spanish-speaking population, people who have obtained the DELE certificate at level C up to 2019 have been included. The risk of double counting with other pools of speakers, especially foreign residents, is low in this case.

4. The number of speakers in the Group with Limited Competence is generally obtained by adding together the number of speakers recognised as descendants of Spanish speakers or continuers of a Hispanic population, culture or tradition. As such, the most difficult groups to assess are speakers of Judaeo-Spanish and Creole languages such as Chabacano and Chamorro. In countries with a small Spanish-speaking population, people who obtained the DELE certificate at level B before 2019 have been included. The risk of double counting with other pools of speakers, especially with potential students in secondary education, where applicable, or university students, is low in this case.

The data set reported here calls for a general comment and several specific considerations. The general comment concerns the total number of potential Spanish speakers outside the Hispanic domain, which exceeds 127 million, 28% above the data provided for 2024. In addition to the significance of the figure itself, the difference is also due to methodological reasons: some censuses and data have been updated or provide more complete information; more solid and detailed studies are available, such as those stemming from the project "Spanish in Europe"; population movements may have intensified, as in the case of Venezuelans moving to Brazil; in addition, migrant populations worldwide have been considered.

As for the figure itself, these 122 million potential Spanish speakers refer directly to one of the areas in which Spanish is most likely to grow in the coming years: its knowledge and use as a second and foreign language (generally with limited competence). Added to this is the potential increase in the number of students worldwide, which will be discussed later in this report. Consequently, these 122 million people make it possible to estimate the progressive growth of Spanish as an international and common language, beyond the Hispanic community's internal exchanges.

The specific considerations in Table 2 require us to look at the demographic significance of the Spanish-speaking population in the United States and the European Union. A slowdown in the increase in Spanish speakers can be observed in the United States, caused by the number of Hispanic or Latino births in the country itself. This leads to a more widespread knowledge and use of English among younger generations and, often, to the non-acquisition of Spanish. Furthermore, the country's migration and socio-economic policies will shape the future impact of both migration flows and the linguistic trends among Hispanic groups already living there. As for Europe, besides the impact of using a more refined methodology, *Special Eurobarometer* 540 (Eurostat, 2024a) clearly indicates that, while only 7% of Europeans over the age of 15 can currently hold a conversation in Spanish, this figure rises to 13% among young people aged 15 to 24, with an upward trend. This latest figure places Spanish 1 percentage point behind French in the same age group and 1 point above German, with both languages showing a downward trend among young people.

The 2025 report also reveals that Spanish-speaker mobility continues to be a direct cause behind the presence of Spanish in areas of the world that are linguistically and culturally distant from the Hispanic world. Data is available in these cases on foreign residents of Hispanic origin as native speakers of the language. However, not enough information is available on their limited competence, due to either the absence of established groups of Hispanic origin[2] or a lack of data on school-leavers with knowledge of Spanish. In Guam, the Mariana

2. In addition to those included in the table, other Sephardic communities can be found in various countries around the world (e.g. Bulgaria, Serbia, Singapore, Egypt), but they are very small and have not been included due to the difficulty of obtaining information about their linguistic repertoires and habits. The use of Judaeo-Spanish is mostly confined to liturgy, reflecting limited competence in the language. In other cases, it is not foreseeable that Judaeo-Spanish will be maintained because the Sephardic community lives alongside Jewish groups of non-Hispanic origin, in addition to coexisting with the non-Jewish population of their countries of residence. Sephardic communities are not specifically identified in countries or territories that have statistics on Spanish speakers at home.

Islands and Morocco, Hispanic communities have existed in the past, which may explain limited knowledge and use of Spanish.

2.3 Spanish language learners

In 2025, approximately 24.6 million people learn Spanish as a foreign language in different educational contexts. This figure (24,560,143 people) represents a 1.5% increase compared to 2024 and reflects the trend in recent years. Since the first comprehensive estimate was provided by the Instituto Cervantes report in 2012, Spanish learners worldwide have increased by more than a third, although learning is conditioned by the vastly different realities of the educational landscape.

Table 3 shows the number and global breakdown of Spanish learners in formal education (primary, secondary and university), as well as in other non-formal education settings such as language centres and Instituto Cervantes classrooms around the world. The data indicate two key points. On the one hand, Spanish language teaching is highly dependent on its degree of institutionalisation as a foreign language in formal non-university educational contexts. On the other hand, a teaching map can be drawn around three territories that mark its main areas of influence: the United States, Brazil and the European Union.

The quantitative relevance of these three areas as vectors for Spanish language learning worldwide is plain to see, as they account for four out of every five Spanish learners. Spanish is the number one foreign language in the United States, where 76% of primary and secondary school students learning a foreign language study it, and 50% of university students. In Brazil, 90% of students studying a second foreign language study Spanish, around 4 million students, although newly approved legislation is reducing this figure. Lastly, Spanish is the second or third most studied foreign language in the European Union, although its institutionalisation varies considerably depending on the country, educational stage or type of educational pathway considered.

The map of Spanish learners would not be complete without considering other regions, particularly the African continent, where one in ten learners can be found. The figures for learners in these countries are based primarily on the institutionalisation of Spanish as a second foreign language within the respective education systems at secondary level in five territories in West and Central Africa: Côte d'Ivoire (798,095), Benin (412,515), Cameroon (403,000), Senegal (356,224) and Gabon (200,000). For example, in Côte d'Ivoire, around two-thirds of students studying a second foreign language study Spanish (Djandué, 2021). In other North African countries such as Algeria, Morocco and Tunisia, Spanish is also studied, although it is not widely taught in formal education. Fortunately, this data will be updated throughout 2025, so next year's report will include it and provide updated information on this important world region.

In addition to secondary education, Spanish language teaching in universities and higher education institutions is worth mentioning. Tertiary education in some countries compensates to some extent for the lack of opportunities to study Spanish at earlier stages of education. For example, in China and Algeria, more than half of Spanish learners study the language in higher education centres. However, the documentary evidence on Spanish language learning in universities is considerably limited. Not only is there no systematic statistical data on students of Spanish and Spanish-related subjects, but countries generally do not have official figures on the number of university students learning Spanish as a foreign language or for specialists in other disciplines.

Non-formal education deserves a separate mention. Language teaching is not a very transparent sector in terms of information about its activity and coverage. As far as Spanish linguistic demography is concerned, Spanish is spoken by an estimated 5 million people who learn it in informal settings, which could be a hidden population in other editions of the report. The figure could be even higher if the interest in learning Spanish on platforms such as Duolingo is considered, as mentioned later in this report. Another important indicator is its employability value, based on official certification data. Over the last few years since the

COVID-19 health crisis, the number of candidates for Diplomas in Spanish as a Foreign Language (DELE) has almost doubled, reaching more than 78,000 examinees. The majority of these candidates are speakers of one of the courses offered by the Instituto Cervantes at its international centres, home to around 68,000 students in 2025.

TABLE 3 APPROXIMATE NUMBER OF SPANISH LANGUAGE LEARNERS AROUND THE WORLD

Country	Primary, secondary and vocational education	University education	Other	Instituto Cervantes and AVE centres 2023-2024	Total
United States	8,358,465	584,453		3,738	8,946,656
Brazil	4,001,530	43,517		3,371	4,048,418
France	3,357,221	244,080	5,978	2,037	3,609,316
United Kingdom	2,017,796	21,505		3,378	2,042,679
Italy	862,243	61,000	41,685	5,813	970,741
Côte d'Ivoire	795,488	1,752	855		798,095
Germany	518,021	52,563	178,255	5,442	754,281
Benin					412,515
Cameroon	400,000	3,000			403,000
Senegal				224	356,224
Sweden	261,598	7,612	7,455	283	276,948
Poland	245,983	18,680	3,928	2,510	271,101
Gabon	200,000				200,000
Canada	156,126	10,992			167,118

Country	Primary, secondary and vocational education	University education	Other	Instituto Cervantes and AVE centres 2023-2024	Total
Portugal	101,694	5,600		594	107,888
Norway	95,863	587			96,450
Algeria	40,290	45,409		3,024	88,723
Austria	56,507	4,000	9,131	774	70,412
Australia	52,726	9,300		803	62,829
Morocco	43,261	1,772	4,839	11,031	60,903
Japan				498	60,498
Russia	42,022	8,398	4,418	1,277	56,115
Ireland	49,451	3,600		1,013	54,064
China	8,874	34,823	8,866	1,477	54,040
Czech Republic	44,219	8,317		1,264	53,800
Belgium	14,320	7,799	24,338	735	47,192
India	25,000	7,681	11,001	2,670	46,352
Tunisia	39,500			899	40,399
New Zealand	34,275	2,098			36,373
Netherlands	19,097	14,160		457	33,714
Denmark	28,813	248			29,061
Other countries					304,238
TOTAL					24.560.143

SOURCE: compiled by the author. Maintaining the approach used in previous reports, and so as not to radically alter the statistical projection, the table only includes countries with more than 30,000 students, although the total figure does include data from countries with fewer students. The number of students given corresponds to 112 countries, so the total number worldwide must be higher than that shown in this table. If no figures appear in any cell, this is due to the absence of specific data or because aggregate data is provided.

TABLE 3 SOURCES AND ADDITIONAL INFORMATION

Germany

Students in primary, secondary and vocational education, data provided by the Federal Statistical Office for the 2023–2024 school year (Statistisches Bundesamt, n.d.). For university students, there were 5,393 students enrolled in Hispanic Studies courses in the 2019-2020 academic year, plus an estimated 47,170 students studying Spanish as a foreign language at 106 German universities (Loureda Lamas *et al.*, 2020, 2021, 2022). Non-formal education students enrolled in adult education centres (*Volkshochschulen*) in 2019 (Loureda Lamas *et al.*, 2022). The number of students at the Instituto Cervantes in Germany for the 2022-2023 academic year has been subtracted from the overall figure provided to avoid double counting.

Algeria

Based on data from Jiménez Caballero, Aous and Salhi (2021). University students include 3,409 students enrolled in university Spanish departments in the 2020-2021 academic year and 42,000 university students in other faculties such as translation and interpreting, as well as those studying Spanish as a foreign language, based on figures for 2016.

Australia

Data from the Ministry of Education and Vocational Training (MEFP). The ministry reports 35,306 students of Spanish in primary education and 17,420 in secondary education during the 2021 school year (MEFP, 2022). The figures are for public schools only. The number of university students corresponds to 2017 (MEFP, 2018).

Austria

Students in primary, secondary and vocational education, data provided by the Federal Statistical Office for the 2022–2023 school year (Statistik Austria, 2024). University students estimated based on the number of Erasmus students from each country enrolled in Spanish universities in 2019, to which a uniform downward multiplier has been applied. Students in non-formal education, data provided by the Association of Adult Education Centres (*Volkshochschulen*) on students enrolled in Spanish courses during 2019.

Belgium

Students in primary, secondary and vocational education, data provided by Eurostat (2024b) for the 2022-2023 school year. University students estimated based on the number of Erasmus students from each country enrolled in Spanish universities in 2019, to which a uniform downward multiplier has been applied. Students in non-formal education: the Ministry of Education and Vocational Training reports 5,270 students of Spanish in

the French-speaking community, 18,958 in the Flemish community and 110 in the German-speaking community during the 2019-2020 school year (MEFP, 2022).

Benin

Based on data from Serrano Avilés (2014).

Brazil

Students in primary, secondary and vocational education, data for 2021 provided by the National Institute for Educational Studies and Research Anísio Teixeira (INEP, 2020) through the Spanish Ministry of Education in Brazil. For university students, INEP (2020) was used. However, these figures should be viewed with caution, as the situation may be changing following the enactment of an educational reform in secondary education that makes Spanish language teaching "optional", whereas it had been compulsory between 2005 and 2017. The current situation is being studied, which could halve the number of Spanish students, as it is optional. For now, the latest official estimates are being used.

Cameroon

Estimates provided in February 2024 by the Spanish Embassy in Cameroon. The number of university students includes both Spanish language and translation studies.

Canada

Due to the lack of representative data at the national level (MEFP, 2018, 2020, 2022), an estimate of the school population in the province of Quebec has been made based on data from Gay (2014) for the 2022-2023 academic year (Statistics Canada 2024a, 2024b). The total number of students is expected to be much higher than reported. As this report is being written, a study on Spanish in Canadian universities is being completed by Lauren Beck and Yolanda Iglesias. This study could put the number of Spanish students at this educational level at over 15,000. The report "Spanish around the World. Report 2026" will include the final figures.

China

Primary, secondary and vocational education, based on Instituto Cervantes (2023). University students, based on González Puy (2018).

Côte d'Ivoirel

Based on data provided in February 2024 by the Instituto Cervantes in Côte d'Ivoire, consulted at the Ministry of Secondary Education and Literacy, and the Spanish Embassy. Students in primary, secondary and vocational education: 528,648 in primary and 266,840 in secondary during the 2021-2022 school year. Students in non-formal education correspond to teachers who have received Spanish -language training from the Embassy from 2015 to 2023.

United States

Students in primary, secondary and vocational education estimated based on the percentage of K-12 students studying a foreign language (20%) and the proportion of these choosing Spanish (76%) according to the report by American Councils for International Education (2017, p. 7) and the school population (National Center for Education Statistics US, 2023a, 2023b, 2024). For university students, the data is taken from Lusin *et al.* (2023, p. 6).

France

Students in primary, secondary and vocational education, data provided by Eurostat (2024b) for the 2022-2023 school year. University students estimated by Tano (2023, p. 157). Students in non-formal education, data from Instituto Cervantes (2023).

Gabon

Estimate for secondary school students provided in February 2024 by the Spanish Embassy.

India

Lower estimate based on figures from Pujol Riembau (2020). The actual figure could be much higher.

Ireland

Students in primary, secondary and vocational education, data provided by Eurostat (2024b) for the 2022-2023 school year. University students estimated based on the number of Erasmus students from each country enrolled in Spanish universities in 2019, to which a uniform downward multiplier has been applied.

Italy

Students in primary, secondary and vocational education, data provided by Eurostat (2024b) for the 2022-2023 school year. University students estimated based on the number of Erasmus students from each country enrolled in Spanish universities in 2019, to which a uniform downward multiplier has been applied. Students in non-formal education, based on Bonomi, Calvi and Uberti-Bona (2024). The number of students at Instituto Cervantes in Italy for the 2020-2021 academic year has been subtracted from the overall figure provided by the authors to avoid double counting.

Japan

Based on Ugarte Farrerons (2012) and the Japanese Ministry of Internal Affairs and Communications (2023). An exhaustive study on Spanish in Japanese universities is currently being completed by Juan Carlos Moyano, Yoshimi Hiroyasu and Ángela Yamaura. This study could estimate the number of Spanish students at this educational level to be over 60,000. The report "Spanish around the World. Report 2026" will include the final figures.

Morocco

Students in primary, secondary and vocational education, according to Instituto Cervantes (2023). University students, data from the Ministry of Education and Vocational Training (2020). Students in non-formal education, data provided in February 2024 by the Spanish Embassy, corresponding to students in centres owned by the Spanish State and attached to the Regional Ministry of Education.

Norway

Students in primary, secondary and vocational education, data provided by Eurostat (2024b) for the 2022-2023 school year. University students estimated based on the number of Erasmus students from each country enrolled in Spanish universities in 2019, to which a uniform downward multiplier has been applied.

New Zealand

Data from the Ministry of Education and Vocational Training (2020).

Netherlands

Students in primary, secondary and vocational education, data provided by Eurostat (2024b) for the 2022-2023 school year. University students estimated based on the number of Erasmus students from each country enrolled in Spanish universities in 2019, to which a uniform downward multiplier has been applied.

Poland

Students in primary, secondary and vocational education, data provided by Eurostat (2024b) for the 2022-2023 school year. University students, data from Zieliński *et al.* (forthcoming). Students in non-formal education, from Instituto Cervantes (2023).

Portugal

Students in primary, secondary and vocational education, data provided by Eurostat (2024b) for the 2022-2023 school year. University students, data from Santos Rovira and Serrano Lucas (2022).

United Kingdom

Students in primary, secondary and vocational education, data provided by Eurostat (2024b) for the 2018-2019 school year. University students estimated based on the number of Erasmus students from each country enrolled in Spanish universities in 2019, to which a uniform downward multiplier has been applied.

Czech Republic

Students in primary, secondary and vocational education, data provided by Eurostat (2024b) for the 2022-2023 school year. University students estimated based on the number of Erasmus students from each country enrolled in Spanish universities in 2019, to which a uniform downward multiplier has been applied.

Russia

Provisional data from the ongoing study on the linguistic demography of Spanish in Russian-speaking countries, included in the project "Spanish in Europe" by the University of Heidelberg and the University of Zurich. Students in non-formal education, data from Instituto Cervantes (2023).

Senegal

Based on official figures from the Senegalese Directorate for Education Reform Planning, provided to Instituto Cervantes in November 2020.

Sweden

Students in primary, secondary and vocational education, data provided by Eurostat (2024b) for the 2022-2023 school year. University students estimated based on the

number of Erasmus students from each country enrolled in Spanish universities in 2019, to which a uniform downward multiplier has been applied. For students in non-formal education, data from Statistics Sweden provided in the ongoing study on the linguistic demography of Spanish in the Nordic countries, included in the project "Spanish in Europe" by the University of Heidelberg and the University of Zurich.

Tunisia

Data provided in February 2024 by the Spanish Embassy, supplied by the Tunisian Ministry of Education for the 2022-2023 school year.

Other countries

This includes figures for countries with fewer than 30,000 students in Africa and Asia (Qatar, United Arab Emirates, Philippines, Ghana, Guinea-Bissau, Iran, Iraq, Israel, Libya, Malawi, Mali, Togo, Vietnam and Zimbabwe), whose data was provided by the Spanish embassies in the respective countries in February 2024; data from Brunei, Indonesia, Malaysia and Thailand from Badillo Matos (2023); data for Kazakhstan and Kyrgyzstan based on provisional data from the ongoing study on the linguistic demography of Spanish in Russian-speaking countries, included in the project "Spanish in Europe" by the University of Heidelberg and the University of Zurich; as well as data not broken down for other European countries: for Albania, Bosnia and Herzegovina, North Macedonia, Montenegro and Serbia, see Loureda Lamas *et al.* (2025); data for Belarus based on provisional data from the ongoing study on the linguistic demography of Spanish in Russian-speaking countries, included in the project "Spanish in Europe" by the University of Heidelberg and the University of Zurich; for the European Union-27 (Bulgaria, Cyprus, Croatia, Estonia, Finland, Greece, Hungary, Latvia, Lithuania, Luxembourg, Malta, Romania, Slovakia and Slovenia), see Eurostat (2024b).

The tally for the EU-27 countries includes students in primary, secondary and vocational education for the 2022-2023 school year. For university students, with a view to providing some continuity in the statistics, the criterion applied in "El español: una lengua viva. Informe 2023" [Spanish: a living language Report 2023] (Instituto Cervantes, 2023) has been followed: an estimate based on the number of Erasmus students from each country enrolled in Spanish universities, to which a uniform downward multiplier has been used. Regarding students of Spanish language courses in the Instituto Cervantes network of centres and on the AVE platform, data on students in Instituto Cervantes classrooms located in countries where there are no centres as such are excluded.

2.4 Potential Spanish language speakers

The sum of the totals provided in the previous tables is shown in Table 4. The total corresponds to the group of potential Spanish speakers inside

and outside the countries where Spanish is the majority language. The potential use of Spanish includes both native speakers and those whose proficiency or competence may be limited for stylistic or sociolinguistic reasons, as well as those who are learning the language within an educational system.

TABLE 4 NATIVE SPANISH SPEAKERS, LIMITED COMPETENCE SPEAKERS AND LANGUAGE LEARNERS, BY COUNTRY TYPE AND TOTALS

	NPG	GLC	GALE[3]	Total
Spanish-speaking countries	459,172,613	24,856,989	-	484,029,602
Non-Spanish-speaking countries	59,942,645	67,211,254	24,560,143	151,714,042
TOTAL	519,115,258	92,068,243	24,560,143	635,743,644

SOURCE: compiled by the author.

Table 4 consolidates several important features. Firstly, the Spanish-speaking community continues to grow and now has over 500 million native speakers. Secondly, the growth of potential Spanish speakers outside the mainly Spanish-speaking world can be observed. This is largely due to the population with a migrant background, but also to updated sources and data, as well as more refined analysis. As for Spanish learners, there has been sustained growth in this group, with an increase of 1.5% over the last year. Overall, the number of potential Spanish speakers worldwide exceeds 630 million, which makes Spanish one of the languages with the greatest demographic potential on the planet.

3. In GALE (Group of Spanish Language Learners) for Spanish-speaking countries, learners who are already counted as such in their respective countries of origin could be included. Therefore, they are not included in this table to minimise the possibility of double counting.

3. Spanish among international languages

3.1 Spanish as a native language, first language or mother tongue

This report reveals that the Spanish-speaking community with native proficiency in the language has surpassed 500 million speakers. This figure ranks it below the native Mandarin Chinese community, as has always been the case, but now also behind the native Hindi community, while remaining above the number of native English speakers. Table 5 contains the figures used to establish this hierarchy.

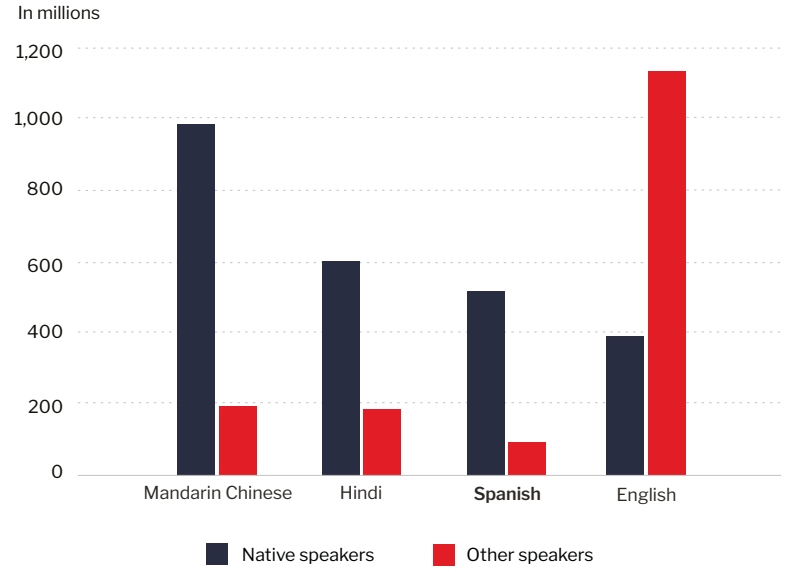
TABLE 5 MOST WIDELY SPOKEN LANGUAGES IN THE WORLD AS NATIVE, SECOND AND/OR FOREIGN LANGUAGES

Language	Native speakers	Other speakers
Mandarin Chinese	989,874,880	193,927,500
Hindi	603,372,494	186,393,633
Spanish	519,115,258	92,068,243
English	390,348,200	1,137,573,260

FUENTE: Mandarin Chinese and English data are from *Ethnologue* (Eberhard, Simons and Fenning, 2025); Hindi data is a forecast for 2025 based on the 2011 Census (Office of the Registrar General & Census Commissioner, 2021, 2022; National Commission on Population, 2019) for native speakers and *Ethnologue* for "other speakers"; Spanish data is as per this report.

To interpret this table, the characterisation of native language (and native speakers) varies depending on the source, alternating between mother tongue and first language (L1), which means that the reference is not the same for all counts (Moreno Fernández, 2024). Furthermore, an accurate interpretation of these numbers of speakers requires additional information to be handled and assessed regarding the social and linguistic nature of each of the languages considered.

GRAPH 1 MOST WIDELY SPOKEN LANGUAGES IN THE WORLD AS NATIVE (L1), SECOND AND/OR FOREIGN LANGUAGES (L2)



SOURCE: Mandarin Chinese and English data are from *Ethnologue* (Eberhard, Simons and Fenning, 2025); Hindi data is a forecast for 2025 based on the 2011 Census (Office of the Registrar General & Census Commissioner, 2021, 2022; National Commission on Population, 2019) for native speakers and *Ethnologue* for "other speakers"; Spanish data is as per this report.

Firstly, Mandarin Chinese, a dialect of the Han ethnic group in northern China, is a macrolanguage that comprises a dozen varieties that share a written script and tradition, and which are not intelligible to other varieties of Chinese (Min, Hakka), even though they share the same writing system. In geographical terms, Mandarin Chinese is most widely spoken in Beijing province and in northern, central and southwestern China (Eberhard, Simons and Fenning, 2025).

Standard Chinese or Putonghua is a form of Mandarin Chinese that is considered the official language of mainland China, although its rules do not apply in Macau or Hong Kong. Around 70% of Chinese people speak some version of Mandarin as their first or native language. Chinese is written in two different ways: traditional Chinese, which is used officially in Taiwan, Hong Kong and Macau, and simplified Chinese

(pinyin), which was created in the second half of the 20th century to make writing and, consequently, literacy more accessible. Simplified writing is the official form in mainland China and the most popular among younger generations of international Chinese speakers.

It is extremely difficult to know the exact number of speakers of standard Mandarin Chinese, as language censuses, if available, are not usually accessible internationally. As for the figure given for "Other speakers", this refers mainly to speakers in Taiwan, Macau and Singapore, as well as in the United States and Canada. At any rate, Mandarin Chinese is the language of the People's Republic of China (with its various regions), Singapore and Taiwan. Its international use therefore belongs primarily to migrant groups scattered throughout the world, who do not usually form a socially integrated language community in their respective host countries.

Secondly, Hindi is derived from Sanskrit, incorporating Arabic and Persian elements and words (Dipshikha, 2023). It is now an official language in India, although English functions as the *de facto* language of government due to the reluctance of several states to accept the use of Hindi (Department of Official Language, 2016). The Indian Constitution lists 22 languages that it officially recognises. This is set against a national backdrop in which 122 main languages are spoken and 1,369 other languages and varieties are classified, according to the 2011 Census (Office of the Registrar General & Census Commissioner, 2011).

In geographical terms, native Hindi speakers are primarily found in the north-central part of the country, namely in the states of Rajasthan, Uttar Pradesh, Madhya Pradesh, Bihar, Jharkhand, Chhattisgarh and Delhi. Other official languages are more widely spoken in the south (Marathi, Telugu, Kannada and Tamil)[4]. As mentioned above, Table 5

4. Likewise, references to Hindustani or Hindi-Urdu can be found in linguistic demography reports. This is due to the linguistic proximity of the two varieties (Urdu and Hindi), but significant differences exist between them: Urdu, the national language of Pakistan, although not the majority language, is written using an adaptation of the Persian alphabet and is used mainly by Muslims, while Hindi is written in the Devanagari alphabet and is used mainly by Hindus.

shows the number of native Hindi speakers in India forecast for 2025 based on population data from the 2011 Census (Office of the Registrar General & Census Commissioner, 2022).). In addition to L1 speakers, this source also includes figures for second (L2) and third (L3) language speakers in India (Office of the Registrar General & Census Commissioner, 2021). As a result, the Hindi-speaking community is much larger, with 158.6 million people speaking Hindi as their L2 and another 27.7 million as their L3. This is hardly surprising in one of the world's most populous countries, where Hindi is an official language. Furthermore, many Hindi L2 speakers are most likely fluent in Hindi as their native language, which would give even greater standing to its native speaker community.

This description of the Chinese and Hindi languages is required to understand and properly define the native Spanish-speaking community. The differences between these native-speaking communities are striking in many ways, as are the differences compared to the native English-speaking community. Table 6 shows the main differences and similarities between the four largest native-speaking communities in the world.

TABLE 6 A COMPARISON OF MANDARIN CHINESE, HINDI, SPANISH AND ENGLISH NATIVE-SPEAKING COMMUNITIES

	Family	Geography	UN Official	Writing system	L1 Community	L2 Community
Chinese	Sino-Tibetan	1 continent	Yes	China	Growing	Stable
Hindi	IE-Indo-Iranian	1 region	No	Devanagari	Growing	Stable
Spanish	IE-Romance	3 continents	Yes	Latin	Slow growth	Growing
English	IE-Germanic	5 continents	Yes	Latin	Slow growth	Growing

SOURCE: compiled by the author. IE = Indo-European

A comparison between these four speaker communities and their respective languages highlights substantial differences, regardless of the figures. Three of the languages belong to the great Indo-European family, although they are linked to different branches: Hindi to Indo-Iranian, Spanish to Romance and English to Germanic. The geography of these languages is also extremely diverse: while Hindi and Mandarin Chinese are mainly spoken as native languages in regions of their respective countries, Spanish is a geographically international native language (spoken in 21 countries in Europe, America and Africa), as is English to an even greater extent (spoken in 57 countries and 28 sovereign entities in Europe, America, Africa, Asia and Oceania). In addition, Hindi is the only native language that does not have official status in the UN system. As far as writing is concerned, English and Spanish share the use of the Latin alphabet, while Chinese and Hindi use alphabets that are less widely used internationally.

Lastly, the linguistic demography of Chinese and Hindi, on the one hand, and Spanish and English, on the other, are opposite dynamics: while native speaker demography is growing rapidly in China and India, especially in the latter case, it is slowing down in Spanish- and English-speaking countries, with the possible exception of African countries, where the percentage of native English speakers is relatively low. Conversely, Spanish and English L2 speakers are growing in number, particularly in the case of English, while Hindi L2 is only growing at the same rate as the Indian population. Mandarin Chinese as a second language, although increasingly popular, cannot match the figures for international languages, due to both the linguistic and written gap and the growing ability of Chinese speakers to communicate internationally in English and Spanish.

This comparison does not imply that some native (or mother tongue) languages are more important than others, as all are cognitively, linguistically, culturally and socially essential to their speakers. The comparison, however, provides a clear picture of the differences arising from factors external to the languages themselves, such as geography or official status. This comparison is highly valuable for understanding the special

characteristics of Spanish as a native language in contrast to other languages. Nevertheless, bilingual knowledge of Spanish and English is the most promising combination for international communication, given the combined weight of their native communities (around 1 billion speakers), the number of countries in which one or the other is an official language (over 75 on several continents), the widespread use of their shared alphabet, and the prestige of their cultures and knowledge systems.

3.2 Spanish and other UN languages

A comparison of the native Spanish-speaking community with Chinese, Hindi and English-speaking communities has great linguistic and cultural significance, while also situating Spanish in its demographic context. However, from an international socio-political standpoint, a demographic comparison between Spanish, English and Mandarin Chinese and the other official languages of the UN (French, Russian and Arabic) proves highly revealing.

A first stumbling block in this task lies in establishing reliable figures for the number of speakers of each of the six official languages of the UN. To achieve this, a distinction must be made between languages considered native or mother tongues (L1) and second and additional languages (L2; non-native proficiency). To this end, this will be done in the same way as for Hindi, by treating languages as native or L1. The most widely sought, cited and reproduced source for this purpose is Ethnologue. Other information sources, such as Wikipedia and Statista, do not shy away from using Ethnologue as a reference. However, the Global Observatory of Spanish has attempted to count the number of speakers of the six languages concerned, starting with the list of countries that have them as official or co-official languages and extending it to the rest of the countries, mainly using census sources. The figures sought refer to speakers of each language as L1 (native) in all countries of the world. As expected, the results obtained by the Observatory do not fully concur with the data provided by Ethnologue, mainly because neither the dates of the data used nor the sources usually match fully in all analyses.

TABLE 7 No. OF MEMBERS OF NATIVE COMMUNITIES (L1) OF UN OFFICIAL LANGUAGES WORLDWIDE, ACCORDING TO THE GLOBAL OBSERVATORY OF SPANISH AND ETHNOLOGUE

Language	Global Observatory of Spanish	Ethnologue
Mandarin Chinese	954,412,674	989,874,880
Spanish	519,115,258	483,979,360
English	443,268,546	390,348,200
Arabic	224,263,923	334,765,000
Russian	172,383,133	145,166,390
French	139,138,190	74,170,080

SOURCE: compiled by the author using data from the Global Observatory of Spanish and Eberhard, Simons and Fennig (2025).

As mentioned, the number of native Spanish speakers exceeds 515 million. Table 4 shows the total number of Spanish speakers in all countries where native speakers of each language have been counted. Again, these figures have not been rounded to respect the data used and raise many critical issues, such as the number of speakers in extremely multilingual countries, as is true of many African countries. French and English have a very low percentage of speakers at home or as a first language, often below 10% of the total population, and among the population that has them as a second language, difficulties arise in determining how many people are native speakers. These situations may be influenced not only by the multiplicity of languages, which makes each territory a linguistic mosaic, but also by the widespread knowledge and use of African *lingua francas*, such as Swahili, Wolof, Hausa or Lingala, or even pidgin languages (Nigerian Pidgin), which reduce the communicative scope of European languages. Furthermore, Table 7 does not attempt to provide figures for speakers with

limited competence or as second languages[5]. However, between the figures from the Observatory and those from Ethnologue, a total lack of correspondence in terms of the relative size of each community of speakers is not observed. This is likely due to the total population base used, which in the case of Ethnologue is not as up to date as that of the Global Observatory of Spanish.

Besides the exact numbers and the comments already made about Mandarin Chinese and English, there are notable differences between these language communities that should be highlighted. For example, the figures given for Arabic by Ethnologue refer to Standard Arabic, which does not have an equal presence in all Arabic-speaking countries. Furthermore, Arabic is spoken in numerous varieties that can present significant difficulties in terms of mutual intelligibility, especially between the Western (Maghreb) and Eastern (Sudanese, peninsular, etc.) varieties. The Egyptian variety is the most widely understood in the Arabic-speaking world due to its presence in the media and the arts. The Arabic world is at any rate characterised by the existence of diglossia, in which dialectal varieties occupy regional, colloquial and family spaces, while the standard or Egyptian variety is found in international, public and academic spheres. The differences compared to the Hispanic world are therefore obvious.

The Russian language covers a large geographical area, but with a proportionally smaller population. At present, Russian is only recognised as an official language in half a dozen countries (either as an official language or as a state language), as the former republics linked to the USSR have gradually dispensed with it as conflicts with the Russian Federation have occurred. This has not led to the radical disappearance

5. This is Instituto Cervantes' first attempt to quantify international languages. The research was carried out by the Global Observatory of Spanish, through independent research conducted by a research team based in San Millán de la Cogolla (Amalia Ovín, Alejandro Bocigas, Pol Pérez and Enrique Ramírez, supervised by Francisco Moreno Fernández). The methodology used is similar to that used for Spanish, although not as exhaustive, which will be remedied in future research. Nevertheless, the counts are based on more up-to-date data than that used by other sources, and the same criteria have been applied to all languages.

of Russian-speaking communities in these republics, but rather to the public relegation of Russian in national affairs. In Spanish-speaking countries, Spanish is not the official language in all of them (as is the case in Argentina, Mexico and Uruguay), but it does retain its value as a *de facto* national or general common language.

Finally, parallels could be drawn between French and Spanish, based on Spain and France's historical evolution. However, the reality is that the colonial styles in each case were rather dissimilar, as was the case with the British colonies, whose geopolitical outcome resulted in a group of countries, mainly in Africa, where French is the L1 of a relatively small proportion of the population, while Spanish is the L1 of the majority of the inhabitants of Hispanic countries.

3.3 Spanish and human development

When comparing linguistic contexts, the quantitative information provided by language demographics should not be the sole basis for assessment. Indeed, linguistic communities are defined by the characteristics of the languages themselves, as well as by the essential features of their demography. However, other factors related to politics, sociology, and economics also contribute to the overall image of these communities. Therefore, the profiles of the communities that speak the official languages of the UN should be compared in terms of their level of human development and other contributing factors, such as the level of poverty or gender equality.

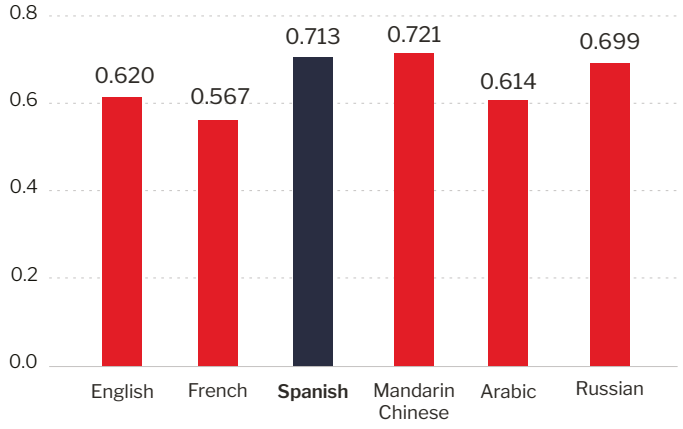
The latest *Human Development Report* 2023/2024 published by the United Nations Development Programme (UNDP) (UNDP, 2024a) provides a snapshot of the current situation in the world's nations. It begins with the state of affairs in 2021-2022, when uncertainty surrounding the pandemic seemed to slow human development and halt national evolution. At that time, the global Human Development Index (HDI) fell for the first time in recorded history. The HDI is a composite index that measures the average outcome of three basic dimensions:

longevity and health, knowledge, and standard of living (UNDP, 2024b). The 2023-2024 report indicates an upturn in the HDI, which reached an all-time high in 2023 when all aspects of the index improved.

The report also draws attention to two significant developments affecting transcontinental linguistic communities. On the one hand, inequality between countries with very high and low HDI is increasing, contrary to what would be expected in the long term. On the other hand, climate change may lead to a marked increase in inequalities, as evidenced by the number of deaths caused by phenomena related to such events. While American countries have seen this specific number of deaths stabilise in the Spanish-speaking world, developed countries, including the United Kingdom, have seen a decline.

For this analysis, the average HDI for all countries that speak each of the official languages of the UN was compared, with a proportional result that is not so dissimilar to that obtained by calculating these indices in previous reports. By comparing the six language communities, we see that the Sinophone and Spanish-speaking populations have an average HDI that leads the way in terms of language development (Mandarin Chinese: 0.721; Spanish: 0.713), while the French-speaking community (0.567) has the lowest average. Of course, these results are averages, as the values for French-speaking countries are influenced by the low HDI of many African nations, as is the case with English, while Mandarin Chinese is favoured by Hong Kong and Singapore's high indices. The Spanish-speaking world, which is less unequal in terms of HDI values, benefits from a high overall average, supported by Spain, Chile, Panama and Costa Rica's "very high" human development score.

GRAPH 2 2022 AVERAGES OF THE HUMAN DEVELOPMENT INDEX (HDI) IN COMMUNITIES THAT SPEAK OFFICIAL UN LANGUAGES (VALUES BETWEEN 0-1)



SOURCE: PNUD (2024a).

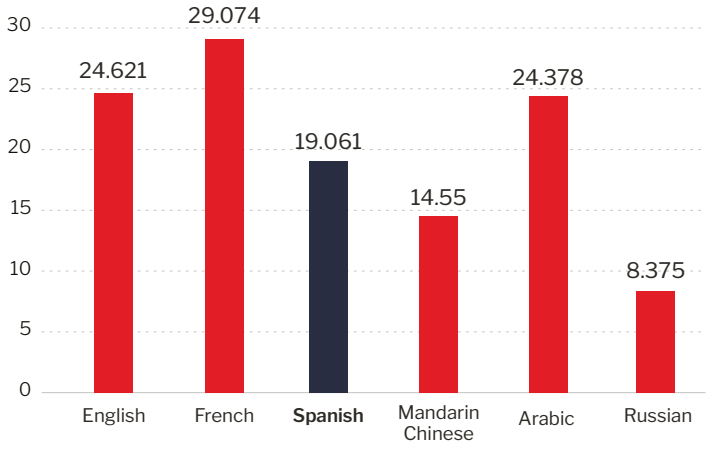
However, as a complex value, the HDI contains nuances that also merit discussion, beyond the obscurity inherent in dealing with averages. This can be demonstrated by focusing on two of these nuances: multidimensional poverty and gender inequality.

Multidimensional poverty is not defined solely by a lack of wealth, but also by deprivation or difficulty in accessing goods and services, such as adequate healthcare, quality housing, employment, or education, along with empowerment and fitness (e.g., body mass) (Multidimensional Poverty Peer Network (MPPN, n.d.)). As a result, poverty is affected by several dimensions, hence its name. Multidimensional poverty is quantified by an index that expresses the percentage of the population suffering from it, scaled according to the intensity of deprivation.

To compare the Multidimensional Poverty Index (MPI) in communities where the official UN languages are spoken natively, averages were

again used, with the results shown in Graph 3. Generally speaking, an inverse proportionality with the HDI of these communities can be observed, whereby Mandarin Chinese and Spanish-speaking communities have a lower MPI than French- and English-speaking communities. However, the proportionally high MPI of Arabic-speaking countries is noteworthy. The Russian-speaking community has the lowest poverty rate in this case.

GRAPH 3 2022 AVERAGES OF THE MULTIDIMENSIONAL POVERTY INDEX IN COMMUNITIES THAT SPEAK OFFICIAL UN LANGUAGES

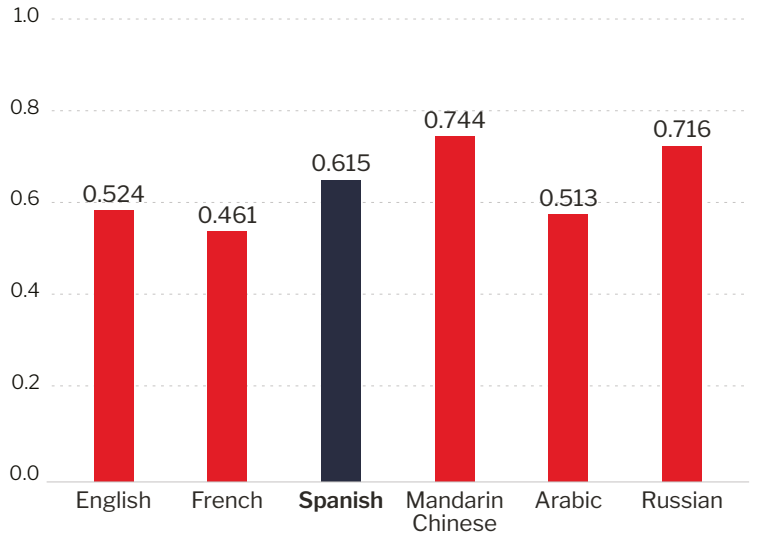


SOURCE: PNUD (2024a).

Finally, the Gender Inequality Index (GII) is a composite value that reflects inequality between women and men across three domains: reproductive health, empowerment, and the labour market (UNDP, 2024b). The lack of direct correlation between the above indices and gender inequality, which is such a significant aspect of community life, is particularly intriguing. The Mandarin Chinese-speaking (0.744) and Russian-speaking (0.716) communities show the highest average inequality, followed, albeit some way behind, by the Spanish-speaking community (0.615). The GII is less prominent for the English-speaking (0.524) and

French-speaking (0.461) communities, and contrary to what is generally thought in the West, Arab countries have a relatively low GII, lower than that of the English-speaking community. Nevertheless, all linguistic communities still have a long way to go.

GRAPH 4 2022 AVERAGES FOR THE GENDER INEQUALITY INDEX IN COMMUNITIES THAT SPEAK OFFICIAL UN LANGUAGES (VALUES BETWEEN 0-1)



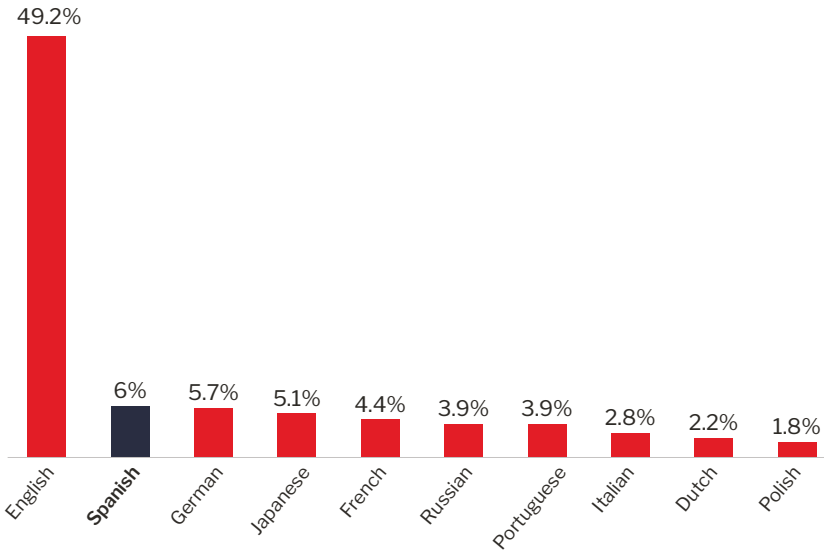
SOURCE: PNUD (2024a).

3.4 Spanish on the Internet

Spanish also plays an important role on the Internet. Spanish-language content currently accounts for 6% of all web pages worldwide, placing Spanish above German, Japanese and French, for example. Nevertheless, no language can rival English, which accounts for 49% of all web pages. Even so, Spanish-language websites are on a growth trend (1.5% since 2015, with 0.4% in 2024), while English-language websites have

declined by approximately 6.2% since 2014, Russian-language websites by 2.2%, and German-language websites by 0.3%[6].

GRAPH 5 PERCENTAGES OF WEBSITES BY CONTENT LANGUAGE (FEBRUARY 2025)



SOURCE: W3Techs (n.d.).

The growth of Spanish on the internet is unquestionably linked to user demand, which in turn correlates with the rise in mobile phone use. Mobile internet users in Latin America are increasing significantly,

6. The W3Tech statistics were chosen to be presented because the data is automatically obtained from websites that are considered relevant due to their content and features. This excludes duplicate sites, redirected sites, subdomains, and sites that only display a single page from the server, for example. Other sources use different criteria to quantify language prevalence on the internet. The Observatory of Linguistic and Cultural Diversity in the Internet (OBDILCI), based in Nice (France) and linked to the International Organisation of La Francophonie and previously to the Latin Union, states that 20% of web pages are in English, 19% in Chinese and 7.7% in Spanish. However, at the same time, OBDILCI has developed a somewhat ambiguous Cyberglobalisation Index (CGI) that includes factors such as the number of speakers of each language, the number of countries where it is spoken, and the number of speakers connected to the internet. The complex index quantifies English as accounting for 14.13% of global cyberglobalisation, French for 9.55%, and Spanish for 2.09% (OBDILCI, 2024). The speaker figures are taken from Ethnologue and other sources, with a 20% confidence range.

from 418 million in 2023, with a 65% penetration rate, to an estimated 485 million in 2030, with a 72% penetration rate. Likewise, the percentage of smartphone connections will increase from 80% in 2023 to 92% in 2030 (GSMA, 2024a). In Latin America, there are three types of content that account for the bulk of internet traffic: social media, web browsing and live-streaming. All three are video intensive, with a potential influence on spoken language (GSMA, 2024b).

Furthermore, the findings of a study on the linguistic complexity of user comments on multiple digital platforms (Facebook, X, YouTube, Voat, Reddit, Usenet, Gab, and Telegram), carried out on 300 million comments in English, deserve a closer look. The results indicate a consistent trend towards shorter texts and reduced lexical richness, even though new words are constantly being introduced. These English-language trends are likely to apply to Spanish and other languages around the world (Di Marco et al., 2024).

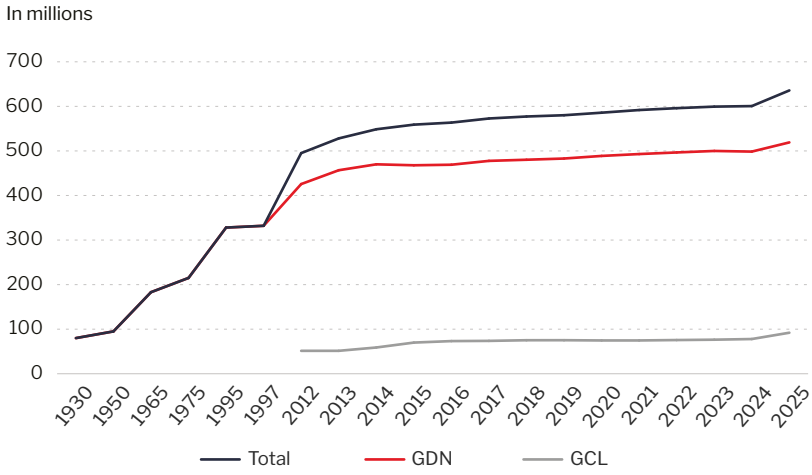
Lastly, regarding other aspects of the Spanish language on social media, Duolingo's language learning app is a worthwhile topic for brief consideration. The organisation's 2024 report reveals that Spanish was the second most popular language learned in 33 countries (compared to 135 for English and 16 for French) was most in demand in the United States, the Nordic countries, South Africa and Australia (Blanco, 2024).

4. The projection of the Spanish language

The current state of the Spanish language and its speaker communities around the world is not a contemporary phenomenon, but rather the result of an evolution that gained significant momentum in the 20th century. The scarce and imprecise linguistic demography references preceding the 1990s describe how Spanish evolved from the 1930s, when the Hispanic community had fewer than 100 million speakers, to the last decade of the 20th century, when it exceeded 300 million. The evolution describes an asymptotic curve that reaches its turning point

in the 1970s and 1980s, when the Hispanic population experienced a major boom as a result of the demographic transition that began in the 1960s (Agar Corbinos and Ferrer, 2001). This coincides with the time the United States began to give visibility to its Spanish-speaking population by using the label "Hispanic" in the census.

GRAPH 6 DEMOGRAPHIC TRENDS AMONG POTENTIAL SPANISH SPEAKERS, BY NATIVE PROFICIENCY AND LIMITED COMPETENCE (1930-2025)



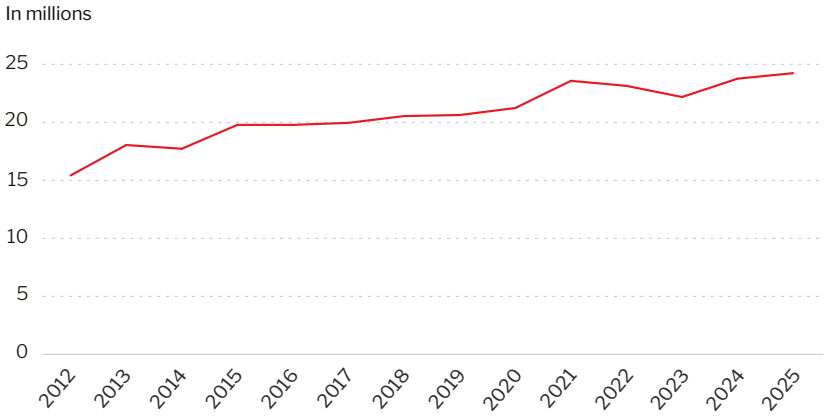
SOURCE: 1930 (Herrero Mayor, 1944); 1950 (Rosenblat, 1977); 1965 (Rosenblat, 1977: 183,000; Pottier, 1968: 130,000); 1996 (Gordon and Grimes, 2005: 266,000; Moreno Fernández and Otero Roth, 1998: 350,000; Hernández Alonso, 1994: 450,000); 2012-2023 (Instituto Cervantes, yearbooks from 2012 to 2023); 2024 (Global Observatory of Spanish, 2024).

Since 1992, the Instituto Cervantes has been tracking Spanish language development, providing detailed and valuable information on its use and learning[7]. This ongoing analysis reveals a steady increase in the number of Spanish speakers worldwide (in both Spanish-speaking and non-Spanish-speaking countries), with varying degrees of proficiency, to over 630 million potential speakers by 2024. The rise in the number of potential Spanish language learners also shows a persistent trend,

7. See the chapter "Observing the Spanish language" by Francisco Moreno Fernández in this edition.

from just over 18 million learners in 2012 to more than 24 million today. This trend was severely disrupted by the COVID-19 health crisis, which led to a three-year setback compared to the previous period. It appears that growth momentum has only recovered in 2025.

GRAPH 7 GLOBAL DEVELOPMENT OF THE GROUP OF SPANISH LANGUAGE LEARNERS (GALE) (2012-2025)



SOURCE: 2012-2023 (Instituto Cervantes, yearbooks from 2012 to 2023); 2024 (Global Observatory of Spanish (2024).

When comparing the different groups of Spanish speakers considered by the Instituto Cervantes series of studies, it is possible to see that the groups that have grown the most in terms of percentage are those with limited competence (79%) and Spanish learners (36%), while native speakers, who are more closely linked to pure demographics, have grown by 22%.

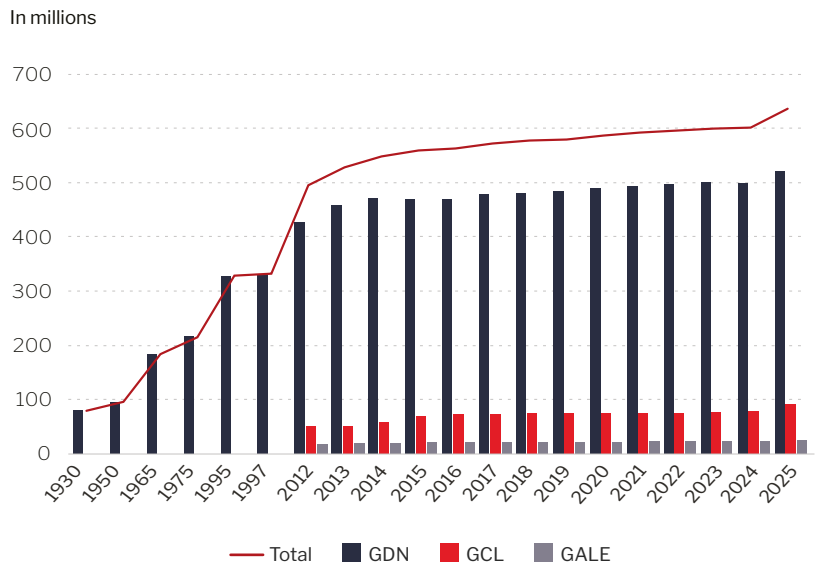
TABLE 8 CHANGE IN THE NUMBER OF POTENTIAL SPEAKERS WITH NATIVE PROFICIENCY (NPG), LIMITED COMPETENCE (GLC) AND SPANISH LEARNERS (GALE) BETWEEN 2012 AND 2025

Year	NPG	GLC	GALE	Total
2012	425,636,709	51,442,310	18,017,913	495,096,932
2013	456,676,339	51,442,310	19,962,500	528,081,149
2014	469,949,593	58,905,216	19,724,511	548,579,320
2015	467,801,452	69,922,994	21,252,789	558,977,235
2016	469,037,629	73,284,395	21,252,789	563,574,813
2017	477,659,770	73,749,162	21,382,247	572,791,179
2018	480,229,642	75,201,405	21,815,280	577,246,327
2019	482,949,086	75,267,790	21,882,448	580,099,324
2020	488,767,362	74,710,449	22,329,275	585,807,086
2021	492,990,519	74,716,239	24,069,206	591,775,964
2022	496,573,842	75,625,447	23,748,298	595,947,587
2023	499,947,796	76,422,128	23,035,198	599,405,122
2024	498,497,757	77,901,236	24,208,813	600,607,806
2025	519,115,258	92,068,243	24,560,143	635,743,644
% Δ	22%	79%	36%	28%

SOURCE: 2012-2023 (Instituto Cervantes, yearbooks from 2009 to 2023); (Global Observatory of Spanish (2024).

In this regard, the annual calculation of these percentages may have suffered from methodological issues, such as methodological refinement and the constant updating and revision of data, as well as the progressive discovery of new speaker groups, even though technical decisions have always leaned towards the most conservative approaches. This conservatism is essentially a matter of investigative caution, which is always necessary when dealing with social data, although the attempt to avoid institutional bias may also have played a role. Whatever the case, it is clear that the most significant group of potential Spanish speakers by volume is native speakers, who far outnumber speakers with limited socio-stylistic competence and Spanish learners. Similarly, the limited competence and learner groups are undeniably the ones that offer the greatest margin for growth, due to their nature and dynamics. Graph 8 shows the volume of speakers with different levels of proficiency over the last century, according to the available data.

GRAPH 8 DEMOGRAPHIC TRENDS AMONG POTENTIAL SPANISH SPEAKERS (1930-2025)



SOURCE: 1930 (Herrero Mayor, 1944); 1950 (Rosenblat, 1977); 1965 (Rosenblat, 1977: 183,000; Pottier, 1968: 130,000); 1996 (Gordon and Grimes, 2005: 266,000; Moreno Fernández and Otero Roth, 1998: 350,000; Hernández Alonso, 1994: 450,000); 2012-2023 (Instituto Cervantes, yearbooks from 2012 to 2023); 2024 (Global Observatory of Spanish, 2024.

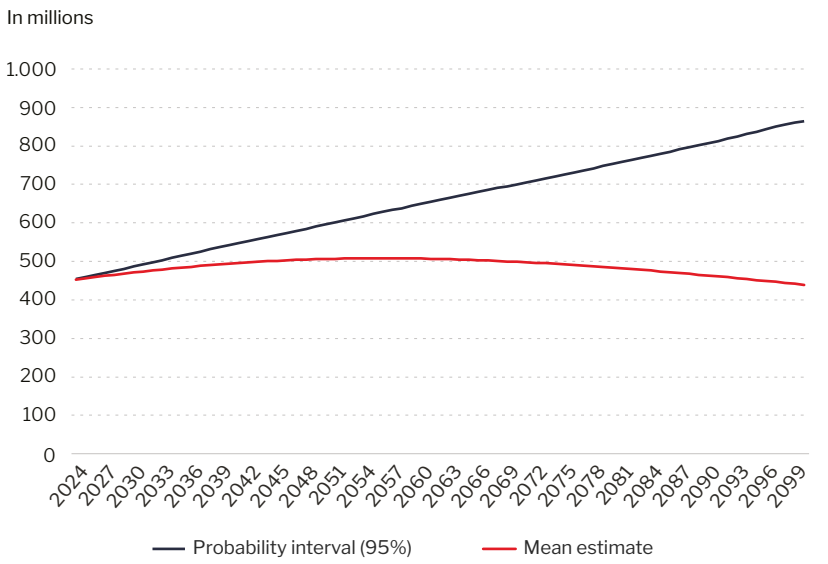
Together with the task of analysing the trend in the number of potential Spanish speakers, an attempt should be made to project the trend in the Spanish-speaking community over the coming decades. This exercise does not need to be carried out annually, given that demographic trends do not usually change from one year to the next, but it is worthwhile presenting it periodically, even if only as a reminder.

The global outlook between 2025 and 2100 depicts a growing population, which will begin to slow down in 2050 and peak in 2086, with 10.43 billion people (UN, 2022b; World Bank, 2025). Thereafter, the population curve will begin to decline towards 2100. By that time, 80% of the world's population will live in Africa and Asia. However, by 2050, Africa is expected to be the fastest-growing continent, and half of the world's population will live in the following nine countries: India, Nigeria, Pakistan, Democratic Republic of Congo, Ethiopia, United States, Indonesia, and Uganda. If we transpose this demography to the linguistic sphere, the outcome would be that, among the official UN languages, English would benefit most from this growth, both as L1 and L2, as it is the official language in five of these nine countries (India, Nigeria, Pakistan, the United States and Uganda), while French is the official language in only one of them (the Democratic Republic of Congo). Spanish would not benefit from prominent linguistic demography within this context.

To calculate the projected number of native Spanish speakers (L1) up to the year 2100, this report has used the population forecast for each Spanish-speaking country, as calculated by the UN, as shown in Graph 9. It plots two curves, the first of which represents an average estimate of the native Spanish-speaking community: this curve would peak in the middle of the century and then experience a gradual decline, in line with demographic trends, in the last third of the century. Secondly, the graph

shows a line of possible maximum growth which, although statistically possible, does not seem demographically plausible: an asymptote is perhaps more predictable than completely linear growth. Nevertheless, this projection should be viewed with caution as it may underestimate the impact of migration.

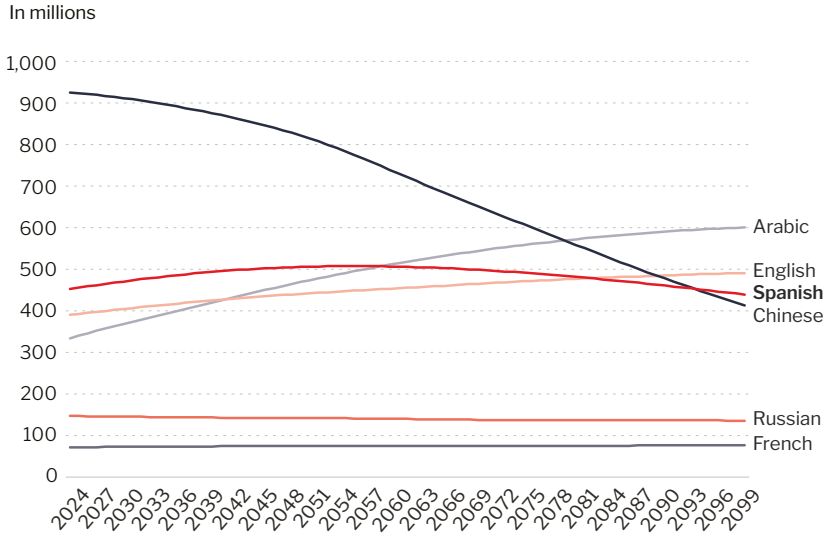
GRAPH 9 NATIVE SPANISH SPEAKER (L1) COMMUNITY PROJECTION UNTIL 2100



SOURCE: compiled by the author based on data from the UN (2022b) and Eberhard, Simons and Fennig (2025).

A comparison of the projections for the Spanish-speaking community with those for the other five official UN languages reveals highly disparate trends, which appear to be linked to the demographic trends in each linguistic community.

GRAPH 10 NATIVE SPEAKER COMMUNITIES (L1) OF THE OFFICIAL UN LANGUAGES PROJECTED UNTIL 2100



SOURCE: compiled by the author based on data from the UN (2022b) and Eberhard, Simons and Fennig (2025.

Graph 10 shows projections based on absolute estimated population figures. To estimate the number of native speakers, the percentage of speakers of each language as their first language (L1) is considered, according to Ethnologue data, in countries where it is the official language in 2025. As such, the current linguistic demography of international languages as native languages for each country is projected onto the UN's demographic forecasts for each country. For example, the projected scenario considers that the proportion of French speakers in Canada will remain constant (around 20%) throughout the timescale.

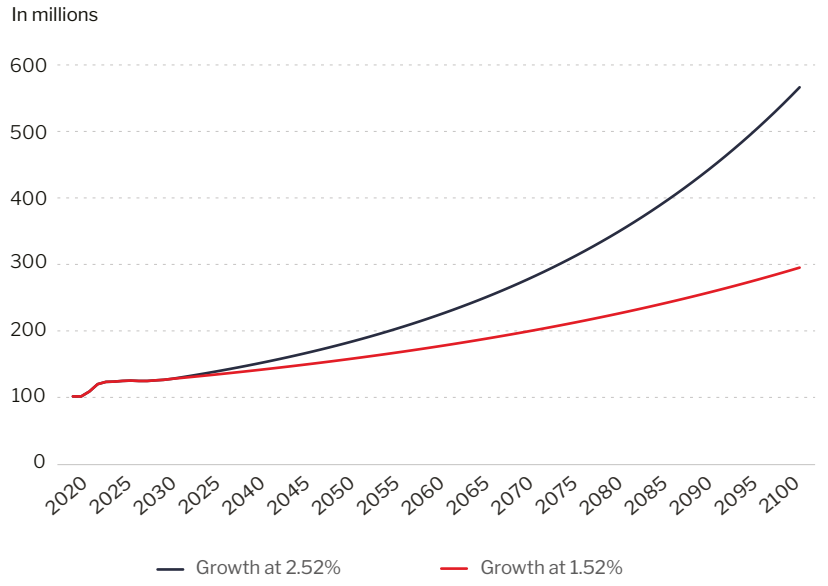
Regarding the results of the projections, both the decline in Mandarin Chinese speakers and the growth of Arabic-speaking communities stand out. Both cases are linked to the population percentages predicted for each community. Furthermore, the projection obtained for the French language could be called into question, given that Africa is

expected to have the highest demographic growth in the world and a large part of the continent is French-speaking. This paradox can be explained by the relatively low proportion of native French speakers (L1) in African countries. However, if literacy and schooling become widespread in these countries, it can be assumed that the percentage of native French speakers, even if they do not have it as their L1, would inevitably increase, which is not reflected in the graph. However, the 2022 French Language Worldwide (La langue française dans le monde) report noted that, since 2010, no country in Africa had experienced a significant change in the percentage of its French-speaking population. Most countries had less than 50% of their total population speaking French, which could be interpreted as a weak indicator of a less than optimistic future (International Organisation of La Francophonie, 2022).

As for projections specifically related to Spanish, the possible future of speakers considered to have limited competence (GLC) in terms of socio-stylistic skills and fluency is one of the most interesting aspects, given that this is one of the groups of speakers with the greatest potential for growth. It should be noted that over the last 13 years, the annual growth rate of the GLC has been 3.52%. If this growth rate continues, there will be 1 billion potential Spanish speakers by 2100, representing 11% of the world's population. Such a projection could be considered to exceed reasonable limits, given the constraints on its empirical basis. However, it seems reasonable to assume that, in the European Union, for example, according to Special Eurobarometer 540 (2024a), 7% of the population speaks Spanish as a foreign language, and this figure rises to over 12% among the under-15s, meaning that significant advances could be made in a single generation.

Nevertheless, given that the GLC depends partly on speakers who have graduated from the education system, as well as on the descendants of immigrants, and since both factors are subject to circumstantial elements, the projections have been adjusted downwards, with the GLC growing by one percentage point less in each of them; that is, 2.52% and 1.52% annually.

GRAPH 11 GROUP WITH LIMITED COMPETENCE GROWTH PROJECTION UNTIL 2100



SOURCE: compiled by the author based on UN data (2022b).

The projection shown in Graph 11 depicts an outlook with over 200 million speakers with limited competence, in its most conservative estimate, and over 500 million in its least conservative projection. When compared with demographic forecasts for native speakers, a future scenario emerges in which the proportion of speakers with limited competence will be much higher than at present (1 in 3 Spanish speakers). This has several implications. The most significant would be that the international dimension of Spanish is expected to increase, given that most of these speakers with limited competence are speakers of Spanish as a foreign language who reside in non-Spanish-speaking countries.

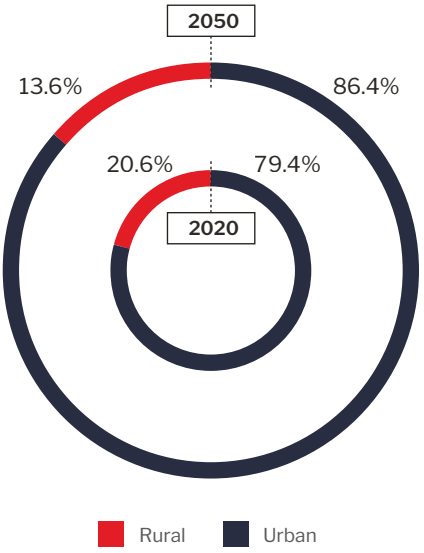
The caution in projecting the number of speakers with limited competence is reflected and even amplified in the possible projection of Spanish language learners. Therefore, no projection of the number of Spanish as a foreign language learners is presented here. In general,

conducting a demolinguistic projection of the number of language learners is a very uncertain task for several reasons. A major challenge is the difficulty of controlling the variables that influence interest in foreign languages; another is their institutionalisation in education systems. While the number of people learning a language is known to depend on several social factors associated with international relations, particularly economic ones (Ginsburgh, Melitz and Toubal, 2014), other decisive factors, such as educational policies relating to foreign language teaching, are difficult to predict. In the case of the Spanish language, there is an additional methodological difficulty, as the database available for projections, which has been built up since 2012 to compile reports for the Instituto Cervantes, has been expanded, adjusted and improved in successive editions. As such, it is impossible to distinguish between changes due to the evolution of learner numbers and those due to the incorporation of new learner pools or adjustments to quantification criteria.

A final projection task leads us to compare rural and urban population trends in Spanish-speaking countries between 2020 and 2050. This comparison is interesting from a linguistic perspective because cities constitute the primary habitat for the use of official, national, or de facto languages (in this case, Spanish), while indigenous, autochthonous, or minority languages, when they are preserved, tend to be found in rural areas.

The total populations of Spanish-speaking countries clearly show that the urban population predominates over the rural population, both in 2020 and in 2050. In the former, the proportion of the rural population barely exceeds 20%; in the latter, the rural population falls to 13.6%. This implies that urbanisation will remain an intense process in Spanish-speaking areas over the coming decades. As a result, indigenous and minority languages will be neglected, and the proportion of Spanish speakers will increase, particularly those with limited competence, such as first-generation internal migrants from rural areas to cities.

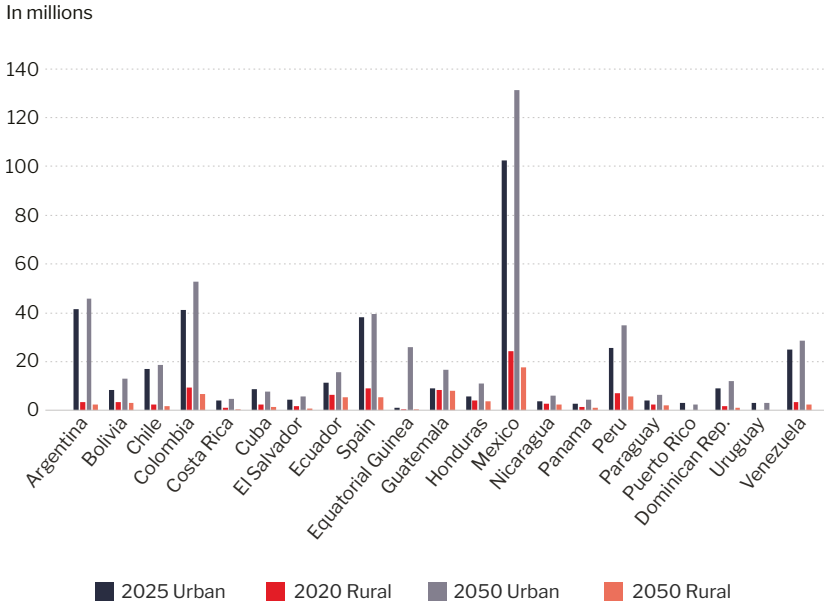
GRAPH 12 URBAN AND RURAL POPULATION OF SPANISH-SPEAKING COUNTRIES IN 2020 AND 2050



SOURCE: World Bank (n.d.-b).

Looking at the rural and urban population trends in each Spanish-speaking country, the same process of deruralisation or urbanisation can be observed as indicated by the overall figures. Between 2020 and 2050, urban populations will grow in absolute terms and rural populations will decline in virtually all countries. However, Mexico stands out as the country where urban growth and rural decline will be most apparent.

GRAPH 13 URBAN AND RURAL POPULATION OF SPANISH-SPEAKING COUNTRIES IN 2020 AND 2050 (BY COUNTRY)



SOURCE: World Bank (n.d.-b).

Therefore, these figures suggest that Spanish will grow in strength in Spanish-speaking countries while rural areas, a stronghold for traditional dialects, will experience demographic decline. For countries that have a small rural population, such as Puerto Rico and Uruguay, these trends will be of little consequence for the varieties of Spanish spoken there.

Referencias bibliográficas

Agar Corbinos, L. & Ferrer, M. (2001). Démographie et contexte socio-économique en Amérique Latine et dans les Caraïbes. In F. Gendreau (Dir.), *Les transitions démographiques des pays du sud* (pp. 25-40). Éditions Stem. https://horizon.documentation.ird.fr/exl-doc/pleins_textes/divers17-09/010027446.pdf

Álvarez Quiñones, R. (30 May 2023). Cuba, sin «revolución», tendría ahora 17 millones de habitantes. *Diario de Cuba*. https://diariodecuba.com/cuba/1685444611_47487.html#:~:text=Son

American Councils for International Education (2017). *The National K-12 Foreign Language Enrollment Survey Report*. https://www.americancouncils.org/sites/default/files/FLE-report-June17.pdf

Argentinos en India (n.d.). *Home* [Facebook page]. Facebook. Accessed on 30 March 2025. https://www.facebook.com/groups/277316635769716/

Australian Bureau of Statistics (2021). *Census of Population and Housing Language spoken at home*. https://www.abs.gov.au/ausstats/abs@.nsf/Lookup/2900.0main+features100622016

Badillo Matos, Á. (2021). *Lengua y cultura en español en el Japón de la era Reiwa*. Real Instituto Elcano. https://www.realinstitutoelcano.org/wp-content/uploads/2021/01/badillo-lengua-y-cultura-en-espanol-en-japon-era-reiwa-1.pdf

Badillo Matos, Á. (2024). *Español para tigres sudasiáticos. Lengua y cultura en español en Filipinas y el sudeste asiático*. Instituto Cervantes & Real Instituto Elcano.

Blanco, C. (2024). *The 2024 Duolingo Language Report*. Duolingo. https://blog.duolingo.com/2024-duolingo-language-report/

Bleorțu, C., Buzilă, P., Ilian, I., Țiței-Avădanei, A., Loureda Lamas, Ó., Álvarez Mella, H., Roșca, A. & Mancheva, D. (2024). *Demolingüística del español en Rumanía, Bulgaria y Moldavia*. Instituto Cervantes, University of Heidelberg & University of Zurich. https://cvc.cervantes.es/lengua/espanol_europa/espanol_rumania/

Bonomi, M., Calvi, M. V. & Uberti-Bona, N. (2024). *Demolingüística del español en Italia. Con un anexo sobre el español en Malta, San Marino y Ciudad del Vaticano*. Instituto Cervantes, University of Heidelberg & University of Zurich. https://cvc.cervantes.es/lengua/espanol_europa/espanol_italia/default.htm

Bravo García, E. (2008). *El español internacional*. Arco Libros.

British Council (2017). *Languages for the future*. https://www.britishcouncil.org/sites/default/files/languages_for_the_future_2017.pdf

Central Bureau of Statistics Aruba (2021). *Census 2020 Housing Tables*. https://cbs.aw/wp/index.php/2021/11/30/census-2020-housing-tables/

Central Bureau of Statistics Aruba (2023). *Population projections*. https://cbs.aw/wp/index.php/2023/05/04/population-projections/

Central Bureau of Statistics Curaçao (n.d.). *Population*. Accessed on 30 March 2025. https://www.cbs.cw/population

Central Bureau of Statistics Curaçao (2023). *Population by most spoken language (in private households)*. https://senso.cbs.cw/education-characteristics

Central Bureau of Statistics Israel (2011). *Jews and others aged 20 and over, by mother tongue and Hebrew usage*. https://www.cbs.gov.il/he/publications/doclib/2013/seker_hevrati11/pdf/t28.pdf

Central Bureau of Statistics Israel (2022a). *Population of Israel on the Eve of 2023*. https://www.cbs.gov.il/he/mediarelease/DocLib/2022/426/11_22_426e.pdf

Central Bureau of Statistics Israel (2022b). *Selected Data from the 2021 Social Survey on Languages*. https://www.cbs.gov.il/en/mediarelease/Pages/2022/Selected-Data-from-the-2021-Social-Survey%20on%20Languages.aspx

Central Bureau of Statistics Israel (2023). *Population Statistical Abstract of Israel 2023*. https://www.cbs.gov.il/en/publications/Pages/2023/Population-Statistical-Abstract-of-Israel-202-No-74.aspx

Central Intelligence Agency (CIA). (n.d.). *The World Factbook*. Accessed on 30 March 2025. https://www.cia.gov/the-world-factbook/

Central Reserve Bank of El Salvador (BCR) (2022). *Estimaciones y Proyecciones de Indicadores Demográficos*. https://www.bcr.gob.sv/documental/Inicio/busqueda/192

Cuban News Agency (2024). Actualizan cifras sobre la población efectiva en Cuba. *Tribuna de La Habana*. https://www.tribuna.cu/cuba/2024-07-19/actualizan-cifras-sobre-la-poblacion-efectiva-en-cuba

Countrymeters (n.d.). *Venezuela Population*. Accessed on 30 March 2025. https://countrymeters.info/en/Venezuela

Datosmacro (n.d.-a). *China Inmigración*. Accessed on 30 March 2025. https://datosmacro.expansion.com/demografia/migracion/inmigracion/china#:~:text=China%20tiene%2C%20seg%C3%BAn%20los%20%C3%BAltimos,%2C%20que%20son%20el%2038.63%25

Datosmacro (n.d.-b). *Israel Inmigración*. Accessed on 30 March 2025. https://datosmacro.expansion.com/demografia/poblacion/israel#:~:text=Solo%20el%2021%2C2%25%20de,442%20habitantes%20por%20Km2.

Department of Official Language (2016). *Constitutional Provisions*. Ministry of Home Affairs India. https://rajbhasha.gov.in/en/constitutional-provisions

Di Marco, N., Loru, E., Bonetti, A., Grazia Serra, A. O., Cinelli, M. & Quattrociocchi, W. (2024). Patterns of linguistic simplification on social media platforms over time. PNAS, 121 (50), 1-11. https://doi.org/10.1073/pnas.2412105121

Diplomatic Information Office (OID) (2025). *Nicaragua*. Spanish Ministry of Foreign Affairs, European Union and Cooperation. https://www.exteriores.gob.es/documents/fichaspais/nicaragua_ficha%20pais.pdf

Dipshikha, B. (2023). Hindi. In *Linguistic Survey of India*. Language Division. Office of the Registrar General. https://language.census.gov.in

Djandué, B. D. (2021). El español lengua extranjera en Costa de Marfil: ¿por qué y para qué se aprende? *redELE*, 33, 1-11. https://www.educacionfpydeportes.gob.es/dam/jcr:5fcce4ed-51a6-48b5-9e63-26f24b153766/ele-marfil.pdf

Domínguez, Y. (30 August 2024). Población dominicana crecerá casi un 1% para el 2025. *Listín Diario*. https://listindiario.com/la-republica/20240830/poblacion-dominicana-crecera-1-2025-one_823598.html

Eberhard, D., Simons, G. & Fennig, C. (2025). *Ethnologue: Languages of the World*. (28th ed.). SIL International. https://www.ethnologue.com/

Eurostat (2012). *Europeans and their Languages. Special Eurobarometer* 386. https://op.europa.eu/en/publication-detail/-/publication/f551bd64-8615-4781-9be1-c592217dad83

Eurostat (2024a). *Europeans and their Languages. Special Eurobarometer* 540. https://europa.eu/eurobarometer/surveys/detail/2979

Eurostat (2024b). *Pupils by education level and modern foreign language studied*. https://ec.europa.eu/eurostat/web/products-datasets/-/educ_uoe_lang01

Federal Statistical Office (2020). *Evolution of the permanent resident population according to the 3 basic scenarios, 1990-2050*. https://www.bfs.admin.ch/bfs/en/home/statistics/population/population-projections/national-projections.assetdetail.12947897.html

Fernández Vítores, D. & Benlabbah, F. (2014). *La lengua española en Marruecos*. Spanish Agency for International Development Cooperation (AECID) & Instituto de Estudios Hispano-Lusos.

Figueredo Reinaldo, Ó., Héctor Rodríguez, Y., Extremera Peregrín, D. A., Extremera San Martín, D., Carmona Tamayo, E. & Padrón Padilla, A. (19 July 2024). Hoy somos menos de 10 millones de cubanos residentes. *Cubadebate*. http://www.cubadebate.cu/noticias/2024/07/19/hoy-somos-menos-de-10-millones-de-cubanos-residentes/

Gay, M. L. (25-28 June 2014). *La situación del español en Canadá y la enseñanza de ELE en Quebec* [communication]. V Congreso internacional FIAPE: ¿Qué español enseñar y

cómo? Variedades del español y su enseñanza. Cuenca, Spain. https://www.educacionfpydeportes.gob.es/dam/jcr:e88c4c04-15ff-45ad-8eae-235703f0c114/13--la-situacion-del-espanol-en-canada-y-la-ensenanza-de-ele-en-quebec--gaylourdes-pdf.pdf

Ginsburgh, V., Melitz, J. & Toubal, F. (2014). Foreign Language Learning: An Econometric Analysis. *CESifo Working Papers*, 4923. https://www.cesifo.org/en/publications/2014/working-paper/foreign-language-learning-econometric-analysis

Global Observatory of Spanish (2024). El español: lengua para el mundo 2024. In Various Authors, *El español en el mundo 2024. Anuario del Instituto Cervantes* (pp. 21-164). Instituto Cervantes & McGraw-Hill. https://cvc.cervantes.es/lengua/anuario/anuario_24/el_espanol_en_el_mundo_anuario_instituto_cervantes_2024.pdf

Gómez, E. & Garrido, A. (2020). *Sáhara Occidental. Preguntas frecuentes*. Una mirada al Sáhara Occidental. https://Sáharaoccidental.es/faqs/Sáhara-occidental/

González Puy, I. (2018). El español, un valor en alza en China. In Various Authors, *El español en el mundo. Anuario del Instituto Cervantes 2018* (pp. 287-303). Instituto Cervantes & Official State Gazette National Agency (AEBOE). https://cvc.cervantes.es/lengua/anuario/anuario_18/gonzalez/p01.htm

Gordon, R. G. & Grimes, B. F. (2005). *Ethnologue: Languages of the World* (15th ed.). Summer Institute of Linguistics. https://www.sil.org/resources/archives/6143

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

Grau Perejoan, M. & Gea Monera, M. P. (2006). *El español en Trinidad y Tobago*. In Various Authors, *Enciclopedia del español en el mundo. Anuario del Instituto Cervantes 2006-2007* (pp. 209-2011). Instituto Cervantes, Plaza & Janés & Círculo de Lectores. https://cvc.cervantes.es/lengua/anuario/anuario_06-07/pdf/paises_41.pdf

GSMA (2024a). *The Mobile Economy in Latin America 2024*. https://www.gsma.com/solutions-and-impact/connectivity-for-good/mobile-economy/wp-content/uploads/2024/06/The-Mobile-Economy-Latin-America-2024.pdf

GSMA (2024b). *Mobile network usage in Latin America. Current data traffic and forecasts to 2030*. https://www.gsma.com/about-us/regions/latin-america/wp-content/uploads/2024/10/Uso-de-redes-moviles-en-America-Latina-ENG_-octubre-2.pdf

Haut-Commisariat au Plan (2014). *Les residentes étrangeres au Maroc*. https://www.hcp.ma/file/231891/

Hernández Alonso, C. (1994). Política lingüística española en América. In Various Authors, *Los derechos humanos en América. Una perspectiva de cinco siglos* (pp. 97-121). Cortes de Castilla y León.

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

Institute for Mexicans Abroad (2020). *Población mexicana en el exterior*. https://ime.gob.mx/estadisticas

Instituto Brasileiro de Geografia e Estatística (IBGE) (2024). *Sistema de Registro Nacional Migratório*. https://portaldeimigracao.mj.gov.br/pt/dados/microdados/1733-obmigra/dados/microdados/401205-sismigra

Instituto Cervantes (2023). El español: Una lengua viva. Informe 2023. In Various Authors, *El español en el mundo 2023. Anuario del Instituto Cervantes* (pp. 25-140). Instituto Cervantes & McGraw-Hill. https://cvc.cervantes.es/Lengua/anuario/anuario_23/el_espanol_en_el_mundo_anuario_instituto_cervantes_2023.pdf

Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) (2020). *Censo da Educação Superior 2019*. https://www.gov.br/inep/pt-br/acesso-a-informacao/dados-abertos/microdados/censo-da-educacao-superior

International Organization for Migration (IOM) (2022). *Migration and Belize. Results of the Migration Modules Attached to the 2021 Labor Force Survey*. https://sib.org.bz/wp-content/uploads/MigrationandBelize.pdf

International Organization of La Francophonie (2022). *The French Language Worldwide. Overview 2022*. Gallimard & International Organization of La Francophonie. https://observatoire.francophonie.org/wp-content/uploads/2022/07/Livret_OIF2_Anglais-VF.pdf

Japanese Ministry of Internal Affairs and Communications (2023). *Statistical Handbook of Japan*. Statistics Bureau of the Japanese Ministry of Internal Affairs and Communications. https://www.stat.go.jp/english/data/handbook/pdf/2023all.pdf

Jiménez Caballero, I., Aous, I. & Salhi, S. E. (2021). El español en Argelia. In Various Authors, *El español en el mundo 2021. Anuario del Instituto Cervantes* (pp. 527-538). Instituto Cervantes & Bala Perdida S. L. https://cvc.cervantes.es/lengua/anuario/anuario_21/africa/argelia.htm

Jiménez Salcedo, J. (2021). De lengua extranjera a lengua dominante: Representaciones epilingüísticas del castellano en el Principado de Andorra. *Revista de Investigación Lingüística*, 24, 119-145. https://doi.org/10.6018/ril.480081

Joshua Project (2024). *Country: Jamaica*. https://joshuaproject.net/countries/JM

Kabatek, J. (2022). El español en Suiza, con un apunte sobre su presencia en Liechtenstein. In Various Authors, *El español en el mundo 2022. Anuario del Instituto Cervantes*. Instituto Cervantes & McGraw Hill (pp. 245-266). https://cvc.cervantes.es/lengua/anuario/anuario_22/kabatek/p01.htm

La Ciutat (17 January 2025). *El 16,7 % de la población de Andorra es de otras nacionalidades*. https://laciutat.cat/es/laciutatandorra-es/el-167-de-la-poblacio-dandorra-es-daltres-nacionalitats-duplicant-se-en-els-darrers-anys

Loureda Lamas, Ó., Moreno Fernández, F., Álvarez Mella, H. & Scheffler, D. (2020). Alemania y sus hablantes de español. In Various Authors, *El español en el mundo 2020. Anuario del Instituto Cervantes* (pp. 321-354). Instituto Cervantes & Bala Perdida, S. L. https://cvc.cervantes.es/lengua/anuario/anuario_20/loureda_moreno/p01.htm

Loureda Lamas, Ó., Moreno Fernández, F., Álvarez Mella, H. & Scheffler, D. (2021). *Demolingüística del español en Alemania*. Instituto Cervantes, University of Heidelberg & University of Zurich. https://cvc.cervantes.es/lengua/espanol_europa/espanol_alemania/default.htm

Loureda Lamas, Ó., Moreno Fernández, F., Álvarez Mella, H. & Scheffler, D. (2022). *El español en la Alemania precovid-19*. Instituto Cervantes, University of Heidelberg & University of Zurich. https://cvc.cervantes.es/lengua/espanol_europa/espanol_alemania_precovid/default.htm

Loureda Lamas, Ó., Pihler Ciglič, B., Kapović, M., Bleorțu, C., Veselko, V., Palapanidi, K., Spahić, E., Ustamujić, I., Mihajlovikj-Kostadinovska, S., Jovanović, A., Šifrar Kalan, M., Popovski, I., Kovač Baretт, I., Vučina Simović, I., Kovač, J., Georgijev, I., Kaba, F. & Valero Fernández, P. (2025). *Demolingüística del español en el sureste europeo*. Instituto Cervantes, University of Heidelberg & University of Zurich. https://cvc.cervantes.es/lengua/espanol_europa/espanol_sureste/

Lusin, N., Peterson, T., Sulewski, C. & Zafer, R. (2023). *Enrollments in Languages Other Than English in US Institutions of Higher Education, Fall 2021*. Modern Language Association of America. https://www.mla.org/Resources/Guidelines-and-Data/Reports-and-Professional-Guidelines/Enrollments-in-Languages-Other-Than-English-in-United-States-Institutions-of-Higher-Education

Madrid Álvarez-Piñer, C. (2018). La evolución de la lengua y la cultura en español en Filipinas. In Various Authors, *El español en el mundo. Anuario del Instituto Cervantes 2018* (pp. 305-312). Instituto Cervantes & Official State Gazette National Agency (AEBOE). https://cvc.cervantes.es/lengua/anuario/anuario_18/madrid/p01.htm

Masuda, K. (2020). *Ideología del estándar y realidad plurinormativa de la lengua española: El caso de E/LE en Japón* [PhD diss., University of Barcelona]. https://www.tdx.cat/handle/10803/668728#page=1

Mora, L. & López, M. H. (20 September 2023). *Latinos' Views of and Experiences With the Spanish Language*. Pew Research Center. https://www.pewresearch.org/race-and-ethnicity/2023/09/20/latinos-views-of-and-experiences-with-the-spanish-language/

Moreno Fernández, F. (2024). *Demografía de las lenguas*. Iberoamericana Vervuert.

Moreno Fernández, F. & Otero Roth, J. (1998). Demografía de la lengua española. In Various Authors, *El español en el mundo. Anuario del Instituto Cervantes 1998* (pp. 59-86). Instituto Cervantes & Arco Libros.

Moreno Fernández, F. & Otero Roth, J. (2006). *Demografía de la lengua española*. Complutense University of Madrid. https://eprints.ucm.es/8936/1/DT03-06.pdf

Moura, R. (2024). Aspectos populacionais da fronteira terrestre segundo o censo 2022. *Boletim regional urbano e ambiental*, 31. http://dx.doi.org/10.38116/brua31art3

Multidimensional Poverty Peer Network (MPPN) (n.d.). ¿Qué es Pobreza Multidimensional? University of Oxford. Oxford Department of International Development (OPHI). Accessed on 30 March 2025. https://www.mppn.org/es/pobreza-multidimensional/que-es-el-ipm/

National Administrative Department of Statistics of Colombia (DANE) (2025). *Censos de población*. https://telencuestas.com/censos-de-poblacion/colombia/2025

National Center for Education Statistics US (2023a). *Private elementary and secondary school enrollment, percentage distribution of private school enrollment, and private enrollment as a percentage of combined enrollment in public and private schools, by school orientation and grade: Selected years, fall 1999 through fall 2021*. Digest of Education Statistics. https://nces.ed.gov/programs/digest/d23/tables/dt23_205.15.asp?current=yes

National Center for Education Statistics US (2023b). *Statistics U. S.* https://nces.ed.gov/fastfacts/display.asp?id=372#PK12-enrollment

National Center for Education Statistics US (2024). *Enrollment in public elementary and secondary schools, by region, state, and jurisdiction: Selected years, fall 1990 through fall 2023*. Digest of Education Statistics. https://nces.ed.gov/programs/digest/d24/tables/dt24_203.20.asp

National Comission on Population (2019). *Population Projections for India and States 2011-2036. Report of the Technical Group on Population Projections*. https://geographicinsights.iq.harvard.edu/files/geographicinsights2/files/report_population_projection_2019.pdf

National Geographic Information Institute (NGII) (2021). *National Atlas of Korea III*. (International Migration of Foreign Nationals, p. 2.683). http://nationalatlas.ngii.go.kr/pages/page_2683.php

National Institute of Development Information of Nicaragua (INIDE) (n.d.). Anuarios estadísticos. Accessed on 30 March 2025. https://www.inide.gob.ni/Home/Anuarios

National Institute of Development Information of Nicaragua (INIDE) (2005). *VIII Censo de Población y IV de Vivienda 2005*. https://www.inide.gob.ni/Estadisticas/censoCEPOV2005

National Institute of Development Information of Nicaragua (INIDE) (2025). *Estadísticas Sociales y Demográficos*. https://www.inide.gob.ni/

National Institute of Statistics and Census of Argentina (INDEC) (2023). *Censo Nacional de Población y Viviendas 2022 por sexo y edad. Indicadores demográficos*. INDEC. https://www.indec.gob.ar/ftp/cuadros/poblacion/censo2022_indicadores_demograficos.pdf

National Institute of Statistics and Census of Argentina (INDEC) (2024). *Proyecciones*. https://www.indec.gob.ar/

National Institute of Statistics and Census of Costa Rica (INEC Costa Rica) (n.d.). *Estadísticas demográficas*. Accessed on 30 March 2025. https://inec.cr/estadisticas-fuentes/estadisticas-demograficas

National Institute of Statistics and Census of Costa Rica (INEC Costa Rica) (2024). *Estimaciones y proyecciones nacionales de población 1950-2100*. https://admin.inec.cr/sites/default/files/2024-09/mepoblaceppnac%201950%20-%202100.pdf

National Institute of Statistics and Census of Ecuador (INEC Ecuador) (2022). *Censo Ecuador*. INEC. https://www.censoecuador.gob.ec/wp-content/uploads/2024/05/Presentacion_Nacional_2da_entrega.pdf

National Institute of Statistics and Census of Ecuador (INEC Ecuador) (2024). *Estimaciones y proyecciones de la población de Ecuador, Revisión 2024*. https://www.ecuadorencifras.gob.ec/documentos/web-inec/Poblacion_y_Demografia/Proyecciones_Poblacionales/censo_2022/revision_2024/Boletin_tecnico_proy_rev2024.pdf

National Institute of Statistics and Census of Panama (INEC Panamá) (n.d.). *XII Censo Nacional de Población y VIII de Vivienda 2023*. Accessed on 30 March 2025. https://www.inec.gob.pa/publicaciones/Default3.aspx?ID_PUBLICACION=1199&ID_CATEGORIA=19&ID_SUBCATEGORIA=71

National Institute of Statistics and Geography of Mexico (INEGI) (n.d.). Accessed on 30 March 2025. *Indicadores sociodemográficos*. https://www.inegi.org.mx/

National Institute of Statistics and Geography of Mexico (INEGI) (2020). *Censo General de Población y Vivienda*. https://www.inegi.org.mx/programas/ccpv/2020/

National Institute of Statistics and Geography of Mexico (INEGI) (2022). *Estadísticas a propósito del Día Internacional de los Pueblos Indígenas*. https://inegi.org.mx/app/saladeprensa/noticia.html?id=7519#:~:text=De

National Institute of Statistics and Informatics of Peru (INEI) (2024). *Situación de la Población Peruana 2024: Una mirada de la diversidad étnica*. https://cdn.www.gob.pe/uploads/document/file/6616587/5751291-situacion-de-la-poblacion-peruana-2024-una-mirada-de-la-diversidad-etnica.pdf

National Institute of Statistics of Bolivia (INE Bolivia) (2024). *Censo de Población y Vivienda 2024*. https://censo.ine.gob.bo/

National Institute of Statistics of Chile (INE Chile) (2017a). *Censo de Población y Vivienda*. https://www.ine.gob.cl/estadisticas/sociales/censos-de-poblacion-y-vivienda/censo-de-poblacion-y-vivienda

National Institute of Statistics of Chile (INE Chile) (2017b). *Segundo Registro de Chilenos en el Exterior*. National Institute of Statistics of Chile & Chilean Ministry of Foreign Affairs.

National Institute of Statistics of Chile (INE Chile) (13 August 2024). *Sala de Prensa*. https://www.ine.gob.cl/sala-de-prensa/prensa/general/noticia/2024/08/13/en-marzo-de-2025-se-entregan-los-primeros-resultados-del-censo#:~:text=Marzo%202025%3A%20entrega%20del%20n%C3%BAmero,de%20indicadores%20del%20Censo%202024

National Institute of Statistics of Equatorial Guinea (INEGE) (2023). *Demografía y población*. https://inege.org/wp-content/uploads/2024/01/Guinea-Ecuatorial-en-Cifras-2023.pdf

National Institute of Statistics of Equatorial Guinea (INEGE) (2024a). *Anuario Estadístico de Guinea Ecuatorial 2024*. INEGE. https://inege.org/wp-content/uploads/2024/11/ANUARIO-ESTADISTICO-DE-GUINEA-ECUATORIAL-2024.pdf

National Institute of Statistics of Equatorial Guinea (INEGE) (2024b). *II Encuesta Nacional de Hogares de Guinea Ecuatorial. Informe de resultados definitivos*. INEGE. https://inege.org/wp-content/uploads/2024/12/Informe-Definitivo-de-la-ENH2.pdf

National Institute of Statistics of Guatemala (INE Guatemala) (2020). *Estimaciones y proyecciones de población a nivel departamental y municipal. Resultados*. https://www.ine.gob.gt/ine/wp-content/uploads/2020/08/Presentacion_19_08-2020-resultados.pdf

National Institute of Statistics of Guatemala (INE Guatemala) (2024). *Encuesta Nacional de Calidad y Bienestar de los Hogares – ENCABIH*. INE. https://www.ine.gob.gt/wp-content/uploads/2024/03/INFORME_ENCABIH.pdf

National Institute of Statistics of Guatemala, National Statistics Office of the Dominican Republic (ONE) and Foundation for Strategic Analysis and Development of Small and Medium-sized Enterprises (FAEDPYME) (2024). *Informe de indicadores estadísticos empresariales*. INE, ONE & FAEDPYME. https://www.ine.gob.gt/2024/03/21/informe-de-indicadores-estadisticos-empresariales-con-enfoque-de-genero/

National Institute of Statistics of Honduras (INE Honduras) (2017). *Encuesta Permanente de Hogares de Propósitos Múltiples (EPHPM)*. National Institute of Statistics. https://ine.gob.hn › category › población

National Institute of Statistics of Honduras (INE Honduras) (2024a). *Sistema Estadístico Nacional*. https://sen.ine.gob.hn/Home/Index

National Institute of Statistics of Honduras (INE Honduras) (2024b). BASEINE. *Base de datos en línea*. https://temp.ine.gob.hn/bases-de-datos/

National Institute of Statistics of Paraguay (INE Paraguay) (2022). *Primeros resultados finales del IV Censo Nacional de Población y Vivienda para Pueblos Indígenas*. https://www.ine.gov.py/censo2022/documentos/indigena/Resultados-Finales-Censo-Indigena.pdf

National Institute of Statistics of Paraguay (INE Paraguay) (2024a). *Población*. https://www.ine.gov.py/microdatos/index.php?cant=99&tema=TODOS

National Institute of Statistics of Paraguay (INE Paraguay) (2024b). *Proyecciones de la población nacional por sexo y edad al 2050*. INE Paraguay. https://www.ine.gov.py/publication-single.php?codec=266

National Institute of Statistics of Paraguay (INE Uruguay) (2023). *Censo de Población, Hogares y Vivienda*. https://www.gub.uy/instituto-nacional-estadistica/comunicacion/noticias/poblacion-preliminar-3444263-habitantes

National Institute of Statistics of Spain (INE España) (2021). *Encuesta de características esenciales de la población y las viviendas*. Año 2021. https://www.ine.es/dynt3/inebase/index.htm?padre=8981&capsel=9008

National Institute of Statistics of Spain (INE España) (2024a). *Estadística continua de población*. https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736177095&menu=ultiDatos&idp=1254735572981

National Institute of Statistics of Spain (INE España) (2024b). *Estadística del Padrón de españoles residentes en el extranjero*.-https://www.ine.es/dynt3/inebase/index.htm?padre=3669&capsel=4553-

National Institute of Statistics of Spain (INE España) (2024c). *Proyecciones de Población. Años 2024-2074*. https://www.ine.es/dyngs/Prensa/es/PROP20242074.htm

National Office of Statistics and Census of El Salvador (ONEC) (n.d.). Población y estadísticas demográficas. Accessed on 30 March 2025. https://onec.bcr.gob.sv/poblacion-y-estadisticas-demograficas/

National Population Council of Mexico (CONAPO) (2024). *Indicadores demográficos de la República Mexicana, en el año 2024*. https://portales.segob.gob.mx/es/Derechos_Humanos/Contexto_Demografico#:~:text=Se%20estima%20que%20en%202024,a%C3%B1os%2C%20representa%2029.4%20por%20ciento

National Population Council of Mexico (CONAPO) & Mexican Ministry of the Environment and Natural Resources (SEMARNAT) (2020). *Proyecciones de población*. https://apps1.semarnat.gob.mx:8443/dgeia/compendio_2020/dgeiawf.semarnat.gob.mx_8080/ibi_apps/WFServlet0eae.html

National Statistics Office of the Dominican Republic (ONE) (10 August 2023). *Boletín Censal*. https://www.one.gob.do/media/0hqd0f5b/boletín-censal-26.pdf

National Statistics Office of the Dominican Republic (ONE) (2025). *Estimaciones y proyecciones demográficas*. https://www.one.gob.do/datos-y-estadisticas/temas/estadisticas-demograficas/estimaciones-y-proyecciones-demograficas/

Nuadda (n.d.). *Menos del 1 % de los brasileños hablan español*. https://www.nuadda.com/menos-del-1-de-los-brasilenos-hablan-espanol/

Observatory of Linguistic and Cultural Diversity in the Internet (OBDILCI) (2024). *Indicadores de la presencia de lenguas en la Internet* (Version 5.2 - November 2024). https://www.obdilci.org/proyectos/principal/

Office for National Statistics United Kingdom (2022). *Language, England and Wales: Census 2021*. https://www.ons.gov.uk/peoplepopulationandcommunity/culturalidentity/language/bulletins/languageenglandandwales/census2021

Office of the Registrar General & Census Commissioner India (2011). *Concepts and Definitions*. https://censusindia.gov.in/nada/index.php/catalog/42561

Office of the Registrar General & Census Commissioner India (2021). *Population by bilingualism and trilingualism, India-2011*. https://censusindia.gov.in/nada/index.php/catalog/10262#metadata-description

Office of the Registrar General & Census Commissioner India (2022). *Population by mother tongue, India-2011*. https://censusindia.gov.in/nada/index.php/catalog/10191

Passel, J. S. & Krogstad, J. M. (22 July 2024). *What we know about unauthorized immigrants living in the U.S.* Pew Research Center. https://www.pewresearch.org/short-reads/2024/07/22/what-we-know-about-unauthorized-immigrants-living-in-the-us/#:~:text=Who are unauthorized immigrants%3F

Peruanos en India (n.d.). *Home* [Facebook page]. Facebook. Accessed on 30 March 2025. https://www.facebook.com/peruanosenindia/

Peruvian Ministry of Health (MINSA) (2024). *Población Estimada 2025*. REUNIS Repositorio Única Nacional de Información y Salud. https://www.minsa.gob.pe/reunis/?op=1&niv=5&tbl=1

Philippine Statistics Authority (2023). *Ethnicity in the Philippines (2020 Census of Population and Housing)*. https://psa.gov.ph/statistics/population-and-housing/node/1684059978

Philippine Statistics Authority (2024). *Population Projection Statistics*. https://psa.gov.ph/statistics/census/projected-population

Pottier, B. (1968). L'Espagnol. In A. Martinet (Ed.), *Le langage* (pp. 887-905). Gallimard.

Private Sector Council for Educational Assistance (COSPAE) (2025). *Programas y proyectos educativos*. https://cospae.org/

Pujol Riembau, Ò. (2020). El español en la India: Una lengua favorecida. In Various Authors, *El español en el mundo 2020. Anuario del Instituto Cervantes*. Instituto Cervantes & Bala Perdida, S. L. https://cvc.cervantes.es/lengua/anuario/anuario_20/pujol/p01.htm

R4V. Inter-Agency Coordination Platform for Refugees and Migrants from Venezuela (2024). *R4V Latin America and the Caribbean, Venezuelan Refugees and Migrants in the Region*. https://www.r4v.info/en/document/r4v-latin-america-and-caribbean-venezuelan-refugees-and-migrants-region-nov-2024-0

Republic of Equatorial Guinea (2016). *Censo de población 2015. Resultados preliminares*. Republic of Equatorial Guinea, UNDP, UNFPA, FAO, ILO and UNICEF. https://ireda.ceped.org/inventaire/ressources/gnq-2015-rec_resultados_preliminares.pdf

Rojas, J. C. (26 March 2017). Interview with Farida Loudaya. *El Tiempo*. https://www.eltiempo.com/mundo/africa/entrevista-con-farida-loudaya-embajadora-de-marruecos-en-colombia-71664

Rosenblat, Á. (1977). El futuro de nuestra lengua. In Á. Rosenblat, *Sentido mágico de la palabra*. Central University of Venezuela.

Santos Rovira, J. M. & Serrano Lucas, C. M. (2022). *Demolingüística del español en Portugal*. Instituto Cervantes, University of Heidelberg & University of Zurich. https://cvc.cervantes.es/lengua/espanol_europa/espanol_portugal/

Sato Center (2012). *Language Varieties. Papiamentu*. https://www.hawaii.edu/satocenter/langnet/definitions/papiamentu.html

Serrano Avilés, J. (2014). La enseñanza del español en África Subsahariana: documentación y propuestas. In J. Serrano Avilés (Ed.), *La enseñanza del español en el África Subsahariana* (Chapter 1, pp. 15-92). Embassy of Spain in Kenya, Spanish Agency for International Development Cooperation (AECID), Instituto Cervantes, Casa África & Los Libros de la Catarata. https://cvc.cervantes.es/lengua/eeas/default.htm

Spanish Ministry of Education and Vocational Training (MEFP) (2018). *El mundo estudia español 2018*. Technical Secretariat-General of the Spanish Ministry of Education. https://www.educacionyfp.gob.es/mc/redele/el-mundo-estudia-espa-ol/2018.html

Spanish Ministry of Education and Vocational Training (MEFP) (2020). *El mundo estudia español 2020*. Technical Secretariat-General of the Spanish Ministry of Education. https://www.educacionyfp.gob.es/mc/redele/el-mundo-estudia-espa-ol/2020.html

Spanish Ministry of Education and Vocational Training (MEFP) (2022). *El mundo estudia español 2022*. Technical Secretariat-General of the Spanish Ministry of Education. https://www.educacionyfp.gob.es/mc/redele/el-mundo-estudia-espa-ol/2022.html

State Data Center-Puerto Rico (2023). *Estimados Anuales Poblacionales*. SDC - PR. https://censo.estadisticas.pr/EncuestaComunidad

State Data Center-Puerto Rico (2025). *Estimados Anuales Poblacionales*. SDC - PR. https://censo.estadisticas.pr/

Statista (2024a). *Spoken languages of Bonaire, Saint-Eustatius and Saba 2021*. https://www.statista.com/statistics/1377799/caribbean-netherlands-spoken-languages-2021/#:~:text=Papiamento

Statista (2024b). *Total population of Venezuela*. https://www.statista.com/statistics/370820/total-population-of-venezuela/

Statistical Institute of Belize (SIB) (2022). *Belize Population and Housing Census 2022*. https://sib.org.bz/census/2022-census/

Statistics Canada (2024a). *Number of students in regular programs for youth, public elementary and secondary schools, by grade and sex*. https://doi.org/10.25318/3710000701-eng

Statistics Canada (2024b). *Postsecondary enrolments, by registration status, institution type, status of student in Canada and gender*. https://doi.org/10.25318/3710001801-eng

Statistics Canada (2025). *Mother tongue by single and multiple mother tongue responses: Canada, provinces and territories, census divisions and census subdivisions*. https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=9810018001

Statistics Norway (2021). *Population: Children, families and households*. https://www.ssb.no/en/befolkning/?de=Children%2C+families+and+households&innholdstype=publikasjon-artikkel

Statistik Austria (2024). *Fremdsprachenunterricht der Schüler: innen im Schuljahr 2022/23* [Dataset]. https://www.statistik.at/statistiken/bevoelkerung-und-soziales/bildung/schulbesuch/schuelerinnen

Statistisches Bundesamt (n.d.). Schüler mit Fremdsprachen-Unterricht. https://www-genesis.destatis.de/genesis//online?operation=table&code=21111-0006&bypass=true&levelindex=1&levelid=1709031181542#abreadcrumb

Stats New Zealand (2018). *Latin American ethnic group*. https://www.stats.govt.nz/tools/2018-census-ethnic-group-summaries/latin-american

Tano, M. (2023). Dinámicas de la lengua española en los espacios de la educación superior francesa. *Archiletras científica: revista de investigación de lengua y letras*, 10, 147-163.

The Bureau of Statistics and Plans. Government of Guam (2020). *The 2020 Census of Guam*. https://bsp.guam.gov/2020-census/

Ugarte Farrerons, V. (2012). El español en Japón. In Various Authors, *El español en el mundo. Anuario del Instituto Cervantes 2012*. Instituto Cervantes & Official State Gazette National Agency (AEBOE). https://cvc.cervantes.es/lengua/anuario/anuario_12/ugarte/p01.htm

UNICAMP (2024). *Inmigrantes internacionais registrados no Brasil. Observatório das Migrações em São Paulo*. https://www.nepo.unicamp.br/observatorio/bancointerativo/numeros-imigracao-internacional/sincre-sismigra/

United Nations (UN) (2018). *Sahrawi Refugees in Tindouf, Algeria: Total In-Camp Population*. United Nations High Commissioner for Refugees (UNHCR) https://wsrw.org/files/dated/2018-10-29/unhcr_2018.pdf

United Nations (UN) (2020). *International Migrant Stock*. Department of Economic and Social Affairs. Population Division. https://www.un.org/development/desa/pd/content/international-migrant-stock

United Nations (UN) (2022a) *Población mundial*. Department of Economic and Social Affairs. https://www.un.org/es/desa-es/la-poblaci%C3%B3n-mundial-llegar%C3%A1-8000-millones-en-2022

United Nations (UN) (2022b). *World Population Prospects 2022. Summary of Results*. Department of Economic and Social Affairs. Population Division. https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/wpp2022_summary_of_results.pdf

United Nations (UN) (2024a). *Peace, dignity and equality on a healthy planet*. https://www.un.org/en/global-issues/population#:~:text=The

United Nations (UN) (2024b). *Word Population Prospects 2024*. Department of Economic and Social Affairs. Population Division. https://www.un.org/development/desa/pd/world-population-prospects-2024

United Nations Development Programme (UNDP) (2024a). *Human Development Report 2023/2024*. https://hdr.undp.org/content/human-development-report-2023-24

United Nations Development Programme (UNDP) (2024b). *Human Development Report 2023/2024 technical notes*. https://hdr.undp.org/sites/default/files/2023-24_HDR/hdr2023-24_technical_notes.pdf

United Nations Economic Commission for Latin America and the Caribbean (ECLAC) (n.d.-a). Latin America and the Caribbean: Population estimates and projections. Accessed on 30 March 2025. https://www.cepal.org/en/topics/demographic-projections/latin-america-and-caribbean-population-estimates-and-projections

United Nations Economic Commission for Latin America and the Caribbean (ECLAC) (n.d.-b). REDATAM. Processing, analysis and dissemination of Statistics and Census around the world. Accessed on 30 March 2025. https://redatam.org/en

United Nations Economic Commission for Latin America and the Caribbean (ECLAC) (1983). *Venezuela: Estimaciones de proyecciones de población, 1950-2025*. Latin American Demographic Centre - Central Office of Statistics and Informatics. https://hdl.handle.net/11362/8005

United Nations Economic Commission for Latin America and the Caribbean (ECLAC) (2022). *Social Panorama of Latin America and the Caribbean*. https://repositorio.cepal.org/server/api/core/bitstreams/a1208761-efa2-4f3a-8be9-bc9368c370c0/content

United States Census Bureau (2020). *Guam Demographic Profile*. US Census Bureau Summary Tables. https://data.census.gov/table/DECENNIALDPGU2020.DP1?g=040XX00US66&d=DECIA%20Guam%20Demographic%20Profile&tid=DECENNIALDPGU2020.DP1

United States Census Bureau (2023). *National Population Projections Tables: Main Series*. US Census Bureau Summary Tables. https://www.census.gov/data/tables/2023/demo/popproj/2023-summary-tables.html

United States Census Bureau (2025). *U.S. and World Population Clock*. https://www.census.gov/popclock/

Venezolanos en India (n.d.). *Home* [Facebook page]. Facebook. Accessed on 30 March 2025. https://www.facebook.com/venezolanosenindia/

W3Techs (n.d.). *Usage statistics of content languages for websites*. Accessed on February 2025. https://w3techs.com/technologies/overview/content_language

World Bank (n.d.-a). *World Bank Open Data*. Accessed on 30 March 2025. https://data.worldbank.org/

World Bank (n.d.-b). *Population estimates and projections*. Accessed on 30 March 2025. https://databank.worldbank.org/Population/id/622a9444#

World Bank (2024). *Annual Report 2024*. https://www.worldbank.org/en/about/annual-report

World Bank (2025). *Urban Development*. World Bank Groupo. https://www.worldbank.org/en/topic/urbandevelopment/overview

Zieliński, A., Berta, T., Medveczká, M., Peskova, J., Štrbáková, R., Tatoj, C., Valero Fernández, P., Álvarez Mella, H. & Loureda Lamas, Ó. (forthcoming). *Demolingüística del español en Europa Central*. Instituto Cervantes, University of Heidelberg & University of Zurich.

II. COMPLEMENTARY ANALYSIS

THE ECONOMIC VALUE OF SPANISH IN A NEW INTERNATIONAL CONTEXT

Juan Carlos Jiménez Jiménez
UNIVERSITY OF ALCALÁ

1. Introduction

Spanish has made remarkable progress in recent decades in various domains, from its demographic influence worldwide to its role as a facilitator of international economic flows of people, goods, services and capital. These flows have coincided with widespread globalisation and technological transformation at the turn of the century as noted in *Emigración y lengua: El papel del español en las migraciones internacionales* [Emigration and language: The role of Spanish in international migration]. As argued here, an international language needs a global economy to fulfil its potential, and Spanish has proven this.

However, as the second quarter of this already turbulent century begins, all indications point to a shift in the world order, and not only in economic terms, characterised by nationalism and introversion. Rules are being swiftly replaced by discretion, multilateral institutions by the survival of the fittest, and within some democracies (and, surprisingly, the most powerful one), the rule of law is faltering, leading to chaos where there was once predictability, including in international relations. Language is as important for understanding in exchange as, if not more so, the confidence provided by respect for formal and informal rules, agreements and institutions. A shared language is, needless to say, a form of common currency. It produces its greatest results in the marketplace and in free competition.

The magnitude and speed of change are not yet clear, but fortunately, the direction of the pendulum swing is already evident. This shift in the international liberal order towards illiberal principles, despite being foreshadowed by numerous warning signs, is being backed by the very people, both benefactors and beneficiaries, who contributed most to creating and sustaining it for eight decades. This is not good news for the future of Spanish. Even less so because the United States, which is the epicentre of the upheaval, is precisely where many of the hopes for Spanish as the language of a rising global middle class were pinned. There, Spanish and its speakers, within a much more nuanced political and cultural battle, seem to have turned into the domestic enemy to be defeated in the face of the desired imperial dominance of English, both within and beyond its borders. The less important issue is whether or not it is an official language. More serious is the imposition of an *English-only* policy that harks back to times gone by. Today, this strategy appears to extend beyond language to reject the identifying characteristics of a substantial portion of the population and a large part of the territory where Spanish was spoken before English arrived.

This new situation must be considered in every analytical assessment. The authors of *Los futuros del español. Horizonte de una lengua internacional* [The Futures of Spanish: Horizon of an International Language] (Alonso Rodríguez, Jiménez Jiménez and García Delgado, 2023), published at the end of 2023, attempt to map out likely scenarios for our language in the medium (2030) and long (2050) term. However, as mentioned above, a huge wave is now cresting on the horizon, which should not be ignored and must be skilfully navigated to avoid losing balance. The aim here is to summarise what we know to date regarding the economic value of Spanish and, above all, to reassess its foreseeable scenarios based on realities that are not as promising as one might think (but not devoid of opportunities either). Therefore, forecasts for Spanish in terms of global GDP, trade flows and the digital economy will be the subject of the following sections. In the concluding section, the overview will be revisited to offer a summary proposal of what can

be done to maintain the vigour of Spanish as a lever for progress in the Spanish-speaking world.

2. Spanish speakers' purchasing power: from a "conservative" to a "pessimistic" scenario

Almost two decades ago, estimates for 2006 put the purchasing power of Spanish speakers worldwide at around 9% of global GDP (actually, in a range between 8.7% and 9.2%, based on two alternative hypotheses) (Jiménez Jiménez, 2007, 2009). The methodology relied on attributing the per capita income of their respective countries to these speakers, with the exception of North American Hispanics, for whom some adjustments had to be made. Recently replicated calculations based on 2019-2020 figures showed slightly higher orders of magnitude, between 9.1% and 9.6% of global GDP (Alonso Rodríguez, Jiménez Jiménez and García Delgado, 2023). Both figures illustrated how the Hispanic community in the United States is growing, both in terms of numbers and purchasing power within the world's largest economy.

The projections made for 2050 (and the intermediate thresholds of 2025 and 2030) were particularly conservative, given the many and highly unpredictable variables involved in this type of calculation. While up to 2025, the purchasing power of Spanish speakers was expected to continue to rise slowly, reaching between 9.4% and 9.9% of global GDP, the outlook for 2050 lowered these percentages by about half a point, to a range between 8.8% and 9.4%. This downward trend was only cushioned in the most optimistic scenario for North American Hispanics (Alonso Rodríguez, Jiménez Jiménez and García Delgado, 2023).

However, the percentages and decimals should not obscure what these statistics are actually indicating. Firstly, Spanish-speaking purchasing power compared to the global total does not fluctuate sharply, either in the previous two decades or in the next two and a half decades. It stands at around 9% at the beginning and end of the first half of this

century, and half a point higher at the peak midway point. Secondly, the slight decline in the proportion of Spanish speakers in the world that is expected to occur between now and 2050 (as the demographic trend is adamant) is not likely to be offset by an increase in their per capita income. Only the large-scale spread of Spanish beyond its current areas of native dominance in wealthier countries could change this, which is not foreseeable. Thirdly, any calculation depends heavily on assumptions made about Hispanic Americans, such as how their demography will develop, how their income and purchasing power will evolve, and how Spanish will survive among this group. The latter factor is particularly significant.

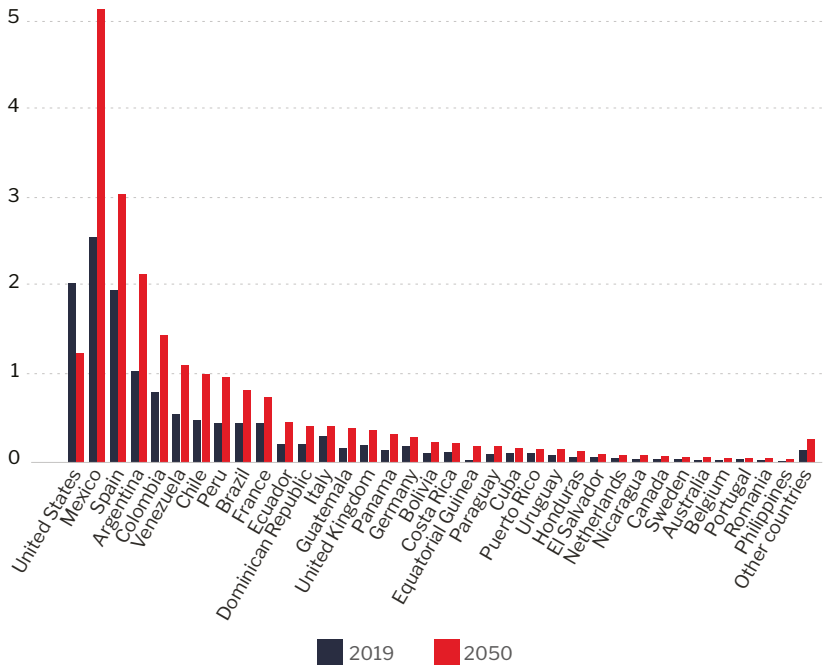
Today, changes in the status of Spanish in the United States seem more than likely, making it necessary to combine the projections mentioned above with others that are even more cautious, or rather, pessimistic, and therefore perhaps more realistic. The Hispanic American contribution to *Spanish-speaking GDP* was initially estimated for the 2023 calculations and projections up to 2050 based on two sets of assumptions. The first, concerning the per capita income of Hispanics, allowed for two options. The most optimistic assumed that their per capita income (in reality, the average income of Hispanic households) could be estimated at 81.9% of the national average in 2019 (according to estimates by the United States Census Bureau), although this proportion would then be revised upwards, converging with the average American standard of living by 2050 (option 1) . The other, more prudent scenario applied the same percentage that Hispanics represent in American purchasing power, according to the Selig Center's projections for 2050, to the US GDP for each year, which reduced the calculations by almost a third (option 2). The second hypothesis (common for both options 1 and 2) applied a correction factor that diluted the demographic projections for Hispanics from the US census based on the percentage of those who have Spanish language skills. This proportion was initially estimated at 71.6% of those who spoke Spanish at home in 2020, and was then reduced in line with the trend observed to date of a gradual, albeit much

more moderate than in other languages, loss of Spanish language skills among the Hispanic community.

Current circumstances nevertheless call for a scenario to be considered that is not radical, but is substantially alternative to the two just mentioned. "Option 3" would be[1]: What if the loss of language skills among Hispanic communities were more acute than previously estimated? A scenario could be envisaged in which only 25% of the Hispanic population in the United States would be fluent in Spanish by 2050. This is not a random percentage. It represents the current proportion of third-generation Hispanics who are bilingual in Spanish and English (the rest are monolingual in English) (Lacomba, 2023, with data from the Pew Research Center). If migration flows continue to decline even further over the next three decades and all Hispanic population growth is based on those born in the United States, i.e. when the majority of the group is third generation, the English-only policy, assuming it overcomes the current counterbalances, could entrench this declining proportion of speakers. In the most favourable of the previous options, Hispanic Americans reached a global purchasing power above that of any other country, ahead of Mexico, and softened the decline in Spanish-speaking purchasing power within global GDP. However, in this most unfavourable (but not implausible) option 3, without this slowdown, the significance of Spanish would fall to 7.7% of global GDP in 2050, i.e. one percentage point less than the most conservative of the previous estimates (option 2) for 2050 and almost two points less than the average estimates for 2019-2020. In fact, this is clearly below the most conservative estimates at the beginning of this century. Graph 1 shows how the ranking would look by country, with Mexico increasingly leading the way.

1. Compare our Graph 1 with the two that appear in Jiménez Jiménez and Narbona Moreno (2022, pp. 10-11): the most optimistic (alternative 1) and the (at that time) prudent (alternative 2).

GRAPH 1 PURCHASING POWER OF SPANISH SPEAKERS WORLDWIDE 2019-2050 (GDP IN PPP, BILLIONS OF USD AT INTERNATIONAL PRICES)



SOURCE: elaboración propia.

In summary, while initial calculations of Spanish speakers' purchasing power worldwide showed a gradual decline that could be offset by the economic progress of some Hispanic American communities that managed to preserve their language skills to a large extent over generations, a less optimistic hypothesis regarding the evolution of Spanish in the United States shows how that decline is becoming inevitable and much more acute.

However, the downward trend in Spanish language proficiency among Hispanics has so far been counterbalanced by some strong factors. These include frequent communication between the Hispanic diaspora in the United States and their communities of origin, particularly Mexico,

and the vitality of Spanish-language cultural expressions (music, audiovisual media, etc.), which appeal greatly to younger people. Most importantly, the perception over the last couple of decades that Spanish (breaking with an old stigma as a negative social marker) was valuable in the job market, and that for young Hispanics, balanced bilingualism in English and Spanish was the most profitable option, both in terms of access to higher education and, subsequently, more favourable job placement and higher salaries. Gándara (2015) wondered how it was possible that speaking only English was worth the same (or even more) than speaking English and Spanish. Indeed, it defies common sense, as well as educational economics, psychology, and neuroscience, that Hispanics would be better off (in terms of wages and access to employment) being monolingual in English than bilingual in English and Spanish.

Based on longitudinal databases, covering the entire school career of cohorts of bilingual and monolingual Hispanics in English until they entered the labour market, three studies have flipped this strange paradox of bilingualism in Spanish within the United States (Agirdag, 2014; Rumbaut, 2014; Santibañez and Zárate, 2014). They show that in the US today: a) people who are equally fluent in Spanish and English earn significantly more than monolinguals (or limited bilinguals) with whom they share "immigrant backgrounds" (between $2,000 and $3,200 more per year); b) balanced bilingualism yields a significant advantage, which increases with the level of bilingual proficiency, both in terms of lower school dropout rates and higher salaries and more sought-after occupations; and c) Spanish bilinguals are more likely to attend university.

Moreover, Hispanics who retain their Spanish language skills receive a "wage premium" in the United States, and non-Hispanics who learn Spanish (more than 7 million) also benefit from this premium. A recent study estimated that the difference between the average remuneration of bilinguals and those who are monolingual in English was close to $2,000–3,000 per year over the period 2014–2019, which amounts to an income premium equivalent to 3.3% (Martínez García and Martínez García, 2022). Although it is only circumstantial evidence that confirms

this perception, Martín Leralta, Jiménez Jiménez and Carrera Troyano (2024) found that Americans rate their Spanish certificates highly (the highest rating by nationality, double the overall average). This is particularly true of those who are already in the labour market, when examining the economic value of Spanish certification through a large sample of DELE and SIELE candidates from around the world.

Expectations seemed positive, and hopefully they will remain that way. However, a resurgence in the negative social stigma attached to speaking Spanish in the United States, in an effort to enforce English as the sole official language and amid a much broader wave of social regression, casts doubt on those advances and, with them, the very future of Spanish for several decades to come. Nevertheless, this is not the only threat looming over the Spanish language due to the winds of change blowing from across the Atlantic.

3. Trade and protectionism: closing and opening doors

Hardly any leading economist, or even one lacking prestige, would argue that "tariff", understood as customs duties, is the most beautiful word in the dictionary. However, once protectionism is triggered, it is insatiable. By eroding competition (and thus competitiveness) among protected local producers, it encourages other activities directly or indirectly affected by the higher costs of those producers to seek protection under its umbrella. It makes little difference whether it is just a transactional strategy (aimed at gaining advantages in other spheres) or part of a comprehensive defensive policy. The arbitrary discretion with which tariff threats are handed out causes uncertainty among stakeholders and hampers trade even more than the tariffs themselves. The favourable effect that various factors have on international trade, including the existence of a common language between the parties, is thus weakened, if not nullified.

This is a real challenge. Not since the inception of the General Agreement on Tariffs and Trade in 1947 has the world faced such a universal

and indiscriminate increase in this form of taxation on imports from other countries. The protectionist impulses of the United States, in a "beggar-thy-neighbour" policy that proved catastrophic in the 1930s and failed to bear fruit during Trump's first term in office, threaten international trade if they become widespread. For decades, trade has brought economic progress to much of the world (without negating certain downsides that are not relevant to discuss here) and, as far as language is concerned, it has enabled Spanish to play an important role in facilitating exchanges (and, not forgetting, capital movements) between the countries that share it (including, incidentally, the United States itself). This facilitating role is now seriously compromised.

Gravity models have so far served as the basic methodological tool for estimating the effect of language sharing on international trade. The data shows that, according to these models, a common language approximately doubles trade between countries that share it, while controlling for other factors that influence trade. When looking at Spanish for the period 1996-2019 (to circumvent the disruption caused by the health crisis), the effect is only slightly lower than that average. However, it is clearly lower than that obtained for the Spanish language until 2007, when it exceeded the average of other languages, as well as the effect of English, beyond its status as the international business lingua franca, among English-speaking countries (Jiménez Jiménez and Narbona Moreno, 2011, 2022).

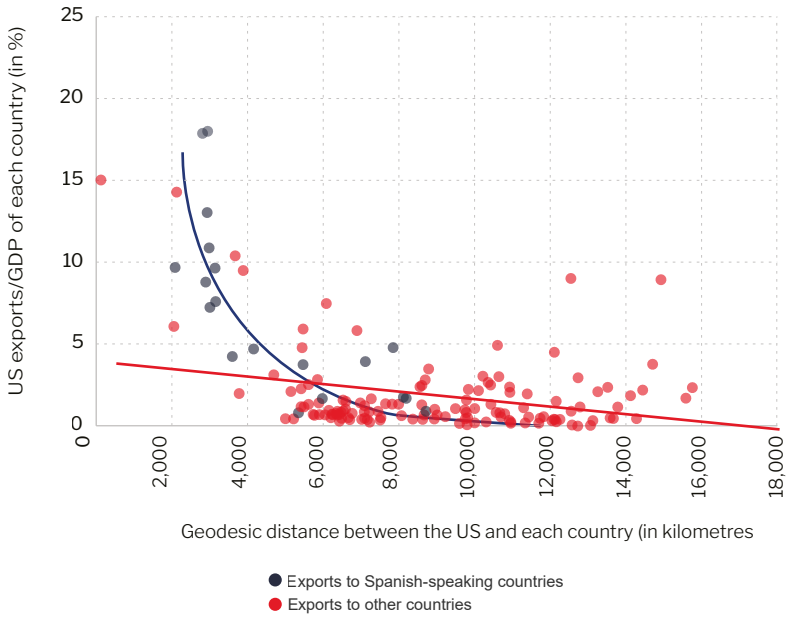
In other words, for Spanish, the knock-on effect of language on trade in goods continues to follow the general pattern seen for other languages. That is, it increases trade by around 100% between Spanish-speaking countries, although this contrasts with the results obtained in previous analyses, particularly before the Great Recession of 2008. Therefore, the facilitating effect of Spanish on trade, although much more significant in the past, continues to exist, but now within universal parameters. Nevertheless, one of the limitations with these calculations is that identifying each country (and therefore its total population) with its majority language, which is generally a reasonable assumption for

Spanish, clashes with the particular case of the United States, which is extremely important, especially for Mexico and the Central American nations.

The attempt by Martínez García and Martínez García (2022) and Martínez García et al. (2021) to quantify the extent to which US exports and imports are conditioned, if not favoured, by the fact that a large proportion of its population is Spanish-speaking is therefore very interesting. To do so, they correlate the geodesic distance of each country with a relative index that measures the intensity of exports from the United States to those countries and the imports that the United States receives from each of them. They conclude that the widespread knowledge of Spanish in the United States constitutes an important positive externality that facilitates its bilateral trade relations with Spanish-speaking countries. Specifically, the authors estimate that "US exports are $160 billion higher and imports are $109 billion higher than would be predicted if the US did not have a strong cultural and linguistic relationship with Spanish-speaking countries".

A closer examination of exports is needed. Graph 2, with some adjustments to the authors' base data and an exponential rather than logarithmic adjustment, clearly shows how US exports are heavily skewed towards Spanish-speaking countries (solid line) compared to the rest (dashed line). Our calculations suggest that Spanish-speaking countries export twice as much to the United States as other countries, and import almost three times as much from it. Not everything can be attributed solely to Spanish, but the fact that it is the most widely spoken language in the United States after English is undoubtedly a factor in this significant commercial inclination towards Spanish-speaking countries. Will it survive a tariff war that curtails trade in both directions (in the likely scenario that the aggrieved countries reciprocate)? The odds are that it will not, which adds another element of irrationality, and another disadvantage for the role of Spanish, to this policy of introversion.

GRAPH 2 US EXPORTS IN RELATION TO THE SIZE OF THE ECONOMY, DISTANCE AND REFERENCE LANGUAGE OF EACH TRADING PARTNER (2015-2019)



SOURCE: compiled by the author based on Martínez-García et al. (2021), using IMF data.

A radical halt to trade between the United States and the Hispanic world, led by Mexico, has a secondary negative impact on Spanish in that country. It weakens a fundamental source of its value in the labour market, namely companies' interest in hiring bilingual speakers to respond to the business opportunities opened up by economic relations with countries south of the border, as has been the case until now. This would depreciate a mutually beneficial asset, as the economic benefits of this increased activity extend even to those who do not speak Spanish.

Therefore, the current trend of economic nationalism paints a bleak picture for the future of Spanish if it continues to spread. This view sees not only trade of all kinds with other countries, but also languages, as part

of a zero-sum game where what one gains is always at the expense of another, thus disregarding the wealth that multilingualism brings. Unfortunately, even though the cognitive benefits of multilingualism have been proven (among countless accredited studies, Bialystok (2011) can be cited), they fail to resonate with the most obtuse minds that oppose it (and who, not coincidentally, tend not to trust scientific work either).

However, tariffs imposed by the main trading partner of many Spanish-speaking countries, such as the United States, may be a golden opportunity to reorient trade in the region for those products where this is feasible. Indeed, intraregional trade in Latin America, valid for its largely Spanish-speaking part, does not reach the proportions seen in other large areas of the planet. It has in fact been declining over recent decades, reaching no more than a mere 12% of the total at the beginning of this decade (United Nations Economic Commission for Latin America and the Caribbean [ECLAC], 2021). Excluding Spain within its regional market (the European Union), only Mexico has been integrated, for the time being, into a major agreement, but with its two northern neighbours. Mercosur, the Andean Community of Nations and the Central American Common Market have so far represented limited integration projects, which has enabled China to become the region's second largest trading partner in just a few years. If the former now takes a step back, an opportunity arises, which is also a challenge, for a renewed regionalism to allow Spanish-speaking countries (plus Brazil, benefiting from linguistic proximity) to make their common language the commercial stimulus that models suggest it has the potential to be. Trans-Latin American companies were perhaps the first to appreciate this. Cooperation and institutionality are the operative words.

In summary, Spanish for the Hispanic world, and all the shared substratum it carries, has long been a crutch to overcome other shortcomings and make it easier to communicate with each other. This opens up a way to alleviate some of the expected slowdown in global trade, especially from its major customer in the north. The regional market, which covers a large part of the American continent, is expanding, while capitalising

on the opportunities recently opened up by the Mercosur-European Union trade agreement, backed by the linguistic familiarity of Spanish and Portuguese on both sides of the Atlantic. It is all duly noted.

4. Challenges for Spanish in the digital economy

The previous section discussed trade in goods, which is historically the most significant on a global scale and has also been the focus of analyses carried out from the perspective of the economics of language to date. This is due to the availability of data, to examine the relationship between these transactions and the enjoyment of a common language. However, for decades, and particularly since the post-2008 recession, trade in services has become an ever-more important factor. In 2008, it accounted for barely a quarter of trade volume, but by 2023 (according to the latest World Bank data), it is close to a third. Consequently, exports of services are currently growing three times faster than exports of goods. From the perspective of the language economy and the economic value of Spanish, this is crucial. Given the number of Spanish speakers (and their purchasing power), Spanish finds itself in a very favourable position (at least prima facie) to capitalise on the network-related externalities afforded by its size.

Data series to replicate the analyses that have been carried out to reveal the quantitative relationship between language and trade (in goods) are still lacking. However, it is clear that in this growing area of trade, namely services, the facilitating role of having a common language is greater than in the case of goods (first advantage, as services particularly require human interaction and the "psychological proximity" that a common language provides. Similarly, an essential part of this trade in services is closely linked to the development of the digital economy itself. Not only does this economy facilitate the exchange of intangible goods, but it also serves as a channel, through e-commerce and its platforms, for a growing proportion of goods movements. The second advantage of a common language is that services (and e-commerce retail) are much more easily exempt from tariffs than goods.

The digital economy is the new playing field where language must assert its "club good" character, capable of spreading network externalities where this is precisely the form of interaction. Spanish has a privileged asset in this regard, as the number of speakers is crucial in exploiting this advantage. In fact, the digital economy's ability to transcend geography and borders more easily than the traditional economy, and the fact that digital connectivity is replacing face-to-face contact in many activities, does not diminish the importance of communication. On the contrary, it highlights the importance of having a common language that reduces the costs of reaching agreement in exchanges and creates the necessary bridges of trust.

The question now is whether Spanish speakers are prepared (or at least well positioned) to seize the advantage of having a language spoken by more than 600 million people in this reborn trade opened up by the digital economy, which transcends political borders, geographical features, costly distances and even capricious tariffs. The available information remains limited, yet it does offer a glimpse into the range of opportunities that could arise if actions are taken appropriately. It also highlights potential challenges if these actions are not implemented with the necessary urgency and ambition. This is what Rebeca Grynspan, Secretary-General of the United Nations Conference on Trade and Development (UNCTAD), recently said when referring to the role of Latin America, or rather the Spanish-speaking world of the continent, in the digitisation of trade: "The region can lead the way in digital platforms and connectivity if it invests heavily in technology, infrastructure and workforce training" (Grynspan, 2025). However, the current situation falls far short of this ideal.

Spanish-speaking countries still have considerable room for improvement in the global digital economy scene. Often, attention is drawn to the most favourable indicator, namely the number of Internet users. This figure is generally high in our countries due to the rapid transition to digital telephony and broadband from poor fixed-line equipment. Although estimates on this point are unstable and sometimes inconsistent (7-8% of the world total, according to various sources), the

proportion of Spanish-speaking Internet users, or the penetration or connectivity ratio, is around 75%. This is around ten percentage points above the global average (although major differences persist within the Spanish-speaking world, from the almost universal connectivity of North Americans, Spaniards and speakers in the Southern Cone to the much more limited connectivity of Central Americans and Equatorial Guineans). Applying the same criteria as before for the population as a whole, Spanish-speaking internet users account for 12.5% of the planet's "virtual" purchasing power[2]. This percentage is likely to decline in proportion to global purchasing power, depending on the scenarios discussed above regarding the persistence of Spanish in the United States.

However, neither the number of Internet users (or their relative size) guarantees that the advantages of the digital economy will be exploited. This is especially true if their role is more passive (users) than active (content producers) within the network, and if the development of the information society and ICTs in Spanish-speaking countries does not reach the levels required for this to happen. The key global indicators of the digital economy (ICT Development Index (IDI) and Networked Readiness Index (NRI)) reveal that Spanish-speaking countries, despite their heterogeneity, generally occupy a mid-to-low position in the respective rankings. According to the most recent NRI data from 2024, only five of the 16 Spanish-speaking countries included in the sample of 133 nations rank higher than average, which is barely a third[3]. Seven, on the other hand, fall into the bottom third of the scale, repeating a pattern that has been seen since 2017. This indicates little progress in this area, at least in comparison with the rest of the world.

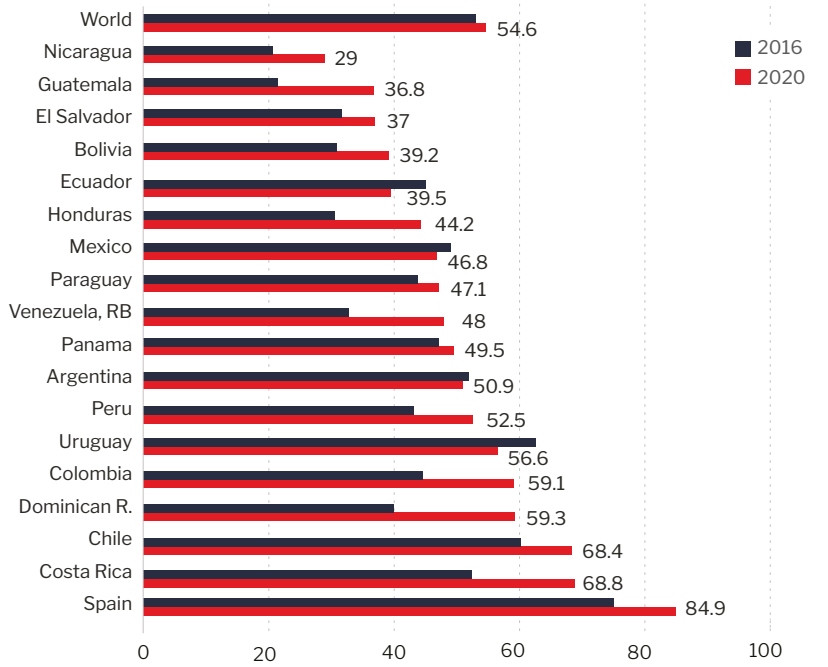
UNCTAD's B2C E-commerce Index is another index that assesses each economy's readiness to support online shopping (e-commerce) by its consumers, based on the simple average of four indicators (or "pillars"):

2. Given a global "multilingualism ratio" of 43%, the combined purchasing power of all languages would exceed 100%. See Pimienta (2021).

3. Calculated based on https://networkreadinessindex.org/.

(i) percentage of people with a bank account; (ii) proportion of people who use the internet; (iii) postal reliability index; and (iv) secure internet servers per million inhabitants. The data in Graph 3 illustrates once again how only one-third (six of the 18 Spanish-speaking countries considered) exceed the global average index, despite generally positive developments on this front. Despite the widespread presence of Spanish-speaking internet users, structural factors are undoubtedly hindering the potential expansion of e-commerce in these countries. Three factors summarise this situation: low levels of banking penetration (i.e., means of payment that can be used electronically), limited postal reliability (reliable delivery systems for purchased goods) and a lack of secure internet servers, which is an essential component for the development of the information society.

GRAPH 3 UNCTAD B2C E-COMMERCE INDEX FOR SPANISH-SPEAKING COUNTRIES 2016-2020 (0-100)



SOURCE: compiled by the author based on UNCTAD data. See: https://unctad.org/system/files/official-document/tn_unctad_ict4d17_en.pdf.

Ultimately, being connected is not as important as the purpose for which one is connected. There is little point in having many people connected to the internet if this only serves to passively browse and download content, watch television online, chat or share emails. If, due to various obstacles or shortcomings, this does not translate into access to services such as financial services, contacting public authorities, shopping online or uploading your own content, for example, then it becomes useless.

Once again, cooperation and institutionality are the operative words for overcoming many of the difficulties facing the future of Spanish in the new network and platform economy. Common governance of the digital ecosystem in Spanish-speaking countries is needed, providing impetus and coordination (two challenging goals in the Hispanic community) to some initiatives that have emerged in this sphere. There are two that must be mentioned here. Firstly, the Digital Agenda for Latin America and the Caribbean (eLAC), promoted by the United Nations Economic Commission for Latin America and the Caribbean (ECLAC) and the Development Bank of Latin America. Secondly, in the more specific field of language and the development of new digital tools for individuals and businesses, the Spanish Language and Artificial Intelligence (LEIA) project, led by the Royal Spanish Academy (Real Academia Española).

5. *Lingua est libertas*

Naturally, these aspects do not exhaust the risks that the new international relations scenario and the role of battering ram adopted by the United States pose to the consolidation of the economic value of Spanish. The migration issue is self-evident and is, in fact, at the root of this isolationist policy. Models carried out to quantify the effect of sharing a language (controlling for other variables) over the decision to emigrate highlight how linguistic proximity between two countries, the country of origin and the destination country, plays a highly positive role in stimulating migration flows between the two (Alonso Rodríguez, Jiménez Jiménez and García Delgado, 2023).

In particular, Spanish proficiency has a strong influence on both the decision to emigrate and the choice of destination. In the case of the large-scale emigration from Spanish America to Spain in the 2000s, this trebled the attraction figures and offered a "salary bonus" for language proficiency of up to 30% (Alonso Rodríguez and Gutiérrez Palacios, 2010). No less important is the fact that this linguistic community is an essential factor in explaining the success of migration, reducing costs and facilitating integration. Aside from the massive waves of Hispanic emigration to the United States in the past, driven by socio-economic factors but with Spanish paving the way, the reality is that nearly 60% of registered immigration in the vast majority of Spanish-speaking countries comes from another Spanish-speaking country. This mobility should be preserved, as it is a source of progress for the recipient countries and alleviates tensions and generates foreign exchange for the country of origin.

With headwinds blowing stronger than ever, Spanish cannot afford to remain idle. If the language seeks to increase its economic value, it must build on the economic progress of the countries where it is spoken. How can this be achieved? Through investment in factors that favour its integration into the digital economy (infrastructure, education, etc.), by improving institutional quality levels, which are as important as a common language in stimulating exchanges, and by safeguarding the free movement of factors (people and capital) and products. Meanwhile, nothing will better sustain the value of Spanish within the United States than the strength acquired, both externally and in every respect, by the rest of the Spanish-speaking world.

To paraphrase the father of economics, Adam Smith, language is what makes exchange possible and, with it, the division of labour, regardless of borders, which is the basis of a nation's prosperity. For that very reason, it needs freedom to unfold its potential. Applying Manuel Alvar's (1983) reflection to economics, "to possess a language is not to be walled up in a cell that imprisons us, but to have a key to discover the light". If only the discursive faculty of human beings could finally prevail over walls, tariffs and borders. Languages are better suited to freedom.

Referencias bibliográficas

Agirdag, O. (2014). The literal cost of language assimilation for the children of immigration: The effects of bilingualism on labor market outcomes. In R. M. Callahan & P. C. Gándara (Eds.), *The bilingual advantage. Language, literacy, and the U.S. labor market* (pp. 160-181). Channel View Publications.

Alvar López, M. (1983). *La lengua como libertad.* Cultura Hispánica.

Alonso Rodríguez, J. A. & Gutiérrez Palacios, R. (Dirs.) (2010). *Emigración y lengua: El papel del español en las migraciones internacionales* (Colección Fundación Telefónica). Ariel.

Alonso Rodríguez, J. A., Jiménez Jiménez, J. C. & García Delgado, J. L. (2023). *Los futuros del español.* Horizonte de una lengua internacional. Alianza Editorial.

Bialystok, E. (2011). Reshaping the mind: The benefits of bilingualism. *Canadian Journal of Experimental Psychology / Revue canadienne de psychologie expérimentale*, 65 (4), 229-235. https://criancabilingue.wordpress.com/wp-content/uploads/2013/09/cep-65-4-229.pdf

Gándara, P. C. (2015). Is There Really a Labor Market Advantage to Being Bilingual in the U.S.? *Policy Information Report and ETS Research Report Series RR-15-07.*

Grynspan, R. (7 January 2025). *For a more just global economy: The role of Latin America.* UNCTAD. https://unctad.org/news/more-just-global-economy-role-latin-america

Jiménez Jiménez, J. C. (2007). *El español: valor de un activo económico.* Caja de Burgos Foundation & Ortega y Gasset Research Institute.

Jiménez Jiménez, J. C. (2009). *El poder de compra del español en el mundo.* Revista de Occidente, 335, 21-36.

Jiménez Jiménez, J. C. & Narbona Moreno, A. (2011). *El español en los flujos económicos internacionales. Un instrumento de la internacionalización empresarial.* (Colección Fundación Telefónica). Ariel.

Jiménez Jiménez, J. C. & Narbona Moreno, A. (2022). *El español en los flujos de comercio internacional: horizonte 2030-2050.* Nebrija Observatory of Spanish. https://www.nebrija.com/catedras/observatorio-nebrija-espanol/pdf/espanol-flujos-comercio-intern.pdf

Lacomba, C. (2023). *Hispanic Map of the United States 2022.* Estudios del Observatorio / Observatorio Studies, 84. Instituto Cervantes at the Faculty of Arts and Sciences of Harvard University. https://cervantesobservatorio.fas.harvard.edu/sites/default/files/84_en_hispanic_map_2022_1.pdf

Martín Leralta, S., Jiménez Jiménez, J. C. & Carrera Troyano, M. (2024). *La certificación en español: un universo valioso. Cartografía y valor económico.* Instituto Cervantes & Nebrija Foundation.

Martínez García, E. & Martínez García, M. T. (2022). *The Economic Value of Spanish in the United States: Opportunities and Challenges for the Future.* Estudios del Observatorio / Observatorio Studies, 78. Instituto Cervantes at the Faculty of Arts and Sciences of Harvard University. https://cervantesobservatorio.fas.harvard.edu/sites/default/files/78_en_martinez_y_martinez_june_28_0.pdf

Martínez García, E., Martínez García, M. T., Coulter, J. & Grossman, V. (13 July 2021). Se habla español: U.S. yet to realize many benefits of a growing bilingual population, *Dallas Fed Economics Blog.* https://www.dallasfed.org/research/economics/2021/0713#:~:text=Growing%20Spanish%E2%80%93English%20bilingualism%20in,compared%20with%20monolingual%20English%20speakers.

Pimienta, D. (2021). *Enhanced and second version of an alternative approach to produce indicators of languages in the Internet.* Observatory of Linguistic and Cultural Diversity in the Internet. https://www.obdilci.org/wp-content/uploads/2024/01/ALI-V2-EN.pdf

Rumbaut, R. G. (2014). English plus: Exploring the socioeconomic benefits of bilingualism in Southern California. In R. M. Callahan & P. C. Gándara (Eds.), *The bilingual advantage. Language, literacy, and the U.S. labor market* (pp. 182-205). Channel View Publications.

Santibañez, L. & Zárate, M. E. (2014). Bilinguals in the U.S. and college enrollment. In R. M. Callahan & P. C. Gándara (Eds.), *The bilingual advantage. Language, literacy, and the U.S. labor market* (pp. 211-233). Channel View Publications.

United Nations Economic Commission for Latin America and the Caribbean (ECLAC) (2021). *International Trade Outlook for Latin America and the Caribbean 2021.* ECLAC. https://repositorio.cepal.org/handle/11362/47536.

ARTIFICIAL INTELLIGENCE AND THE SPANISH LANGUAGE: CURRENT STATUS AND FUTURE PROSPECTS

Elena González-Blanco García
MICROSOFT

1. Introduction

Rapid advances in artificial intelligence, especially following the arrival of generative AI, are transforming natural language processing around the world, and the Spanish language is no exception. With over 600 million speakers worldwide, Spanish represents a linguistic asset of enormous importance, whose future in the digital age offers great promise, but also significant challenges. This article analyses how AI is being used in Spanish and the Spanish-speaking world, including the public strategies that have stimulated this area, available linguistic resources, the emerging business ecosystem, the most important research groups, and the impact of these technologies on Spanish and its variants.

2. Artificial intelligence and language: history of a scientific discipline

Artificial intelligence (AI) and its application to natural language is one of the most fascinating transformations in the contemporary world. Over the last seven decades, this discipline has transformed from a philosophical utopia into an omnipresent technological reality. It now extends from language teaching and machine translation to conversational assistance and text generation.

Early developments in AI date back to the 1950s, when mathematician Alan Turing famously suggested a test to determine whether a machine

could "think" by conversing in natural language (Turing, 1950). Simultaneously, linguists such as Noam Chomsky laid the formal foundations of generative syntax, which enabled early computer systems to attempt to process human language using explicit rules (Chomsky, 1957). In 1954, a Georgetown University experiment successfully translated Russian sentences into English using a system of rules (Hutchins, 2014), leading to high expectations for machine translation. However, in 1966, a US government report concluded that these breakthroughs were small, and funding for natural language processing (NLP) was drastically cut, starting a period known as the "AI winter" (Automatic Language Processing Advisory Committee [ALPAC], 1966).

During the 1970s and 1980s, symbolic systems were the dominant paradigm, with software following human-programmed rules to understand or generate sentences. One famous example was ELIZA, a rudimentary chatbot designed to simulate a therapist (Weizenbaum, 1966). A more ambitious example was SHRDLU, which allowed commands to be given in English to a virtual world of blocks (Winograd, 1972). However, such systems only worked well in very limited contexts. The real game changer came in the late 1980s and 1990s, when increased computing power combined with the availability of large bilingual corpora (thanks, for example, to the European Parliament) made it possible to abandon rule-based systems and adopt statistical methods (Koehn, 2010). Machines began to "learn" linguistic patterns from large amounts of text for the first time, without direct human intervention. This gave rise to statistical translation systems and the first attempts to use neural networks for language (Hochreiter and Schmidhuber, 1997).

The 21st century ushered in major advances. In 2013, the concept of representing words as numerical vectors, capturing their meanings through mathematical proximity and vector representation, gained popularity (Mikolov et al., 2013). Shortly thereafter, Google introduced neural machine translation, which significantly outperformed previous systems (Wu et al., 2016). However, the turning point came in 2017 when the article "Attention Is All You Need" was published, suggesting a new

IT architecture called Transformer (Vaswani et al., 2017). It revolutionised language processing by enabling machines to understand entire text contexts much more efficiently, replacing previous sequence-based approaches.

Since then, the pace of progress has been dizzying. In 2018, Google introduced BERT, a vector-based model that anticipated deep text comprehension (Devlin et al., 2018). Meanwhile, OpenAI launched its series of GPT (Generative Pre-trained Transformer) models, which are trained with massive amounts of text in multiple languages (Radford et al., 2018). In November 2022, the launch of ChatGPT, based on GPT-3.5, marked a watershed moment, reaching over a million users in just five days and 100 million two months later (OpenAI, 2022). This model made it possible to hold fluid conversations, write texts, answer questions, and write code, all in natural language. In 2023, OpenAI introduced GPT-4, featuring even more sophisticated capabilities, such as image and graphic interpretation (OpenAI, 2023).

A global technology race ensued as these generative systems became increasingly popular. Google launched its Gemini series (formerly Bard), Meta published the open-source LLaMA models[1], and others emerged, such as Claude[2], developed by Anthropic, and the Phi series, specialising in small but efficient models. The race is relentless worldwide, as demonstrated by countries such as China, which has developed its own models (DeepSeek, Yi-34B, Moonshot, SparkDesk) cheaply and quickly.

Simultaneously, the Middle East has emerged as a major AI development hub, with initiatives led primarily by the United Arab Emirates. The Abu Dhabi Technology Innovation Institute (TII)[3], in collaboration with the G42 technology group, led the development of the Falcon family of models, including Falcon 40B and Falcon 180B, the latter launched in

1. See: https://www.llama.com/.

2. See: https://claude.ai/.

3. See: https://www.tii.ae/ai-and-digital-science.

2023 and recognised as one of the most powerful open-source models. Furthermore, in 2024, the company AI71[4] was launched, also promoted by G42 and TII, with a view to developing secure and customised AI solutions for governments and large institutions. This strategic commitment positions the Middle East at the cutting edge of generative AI development, promoting regional technological sovereignty and expanding the global innovation map beyond the traditional hubs in the US, Europe and China.

3. The Spanish AI outlook: corpora, resources and models

To develop and train artificial intelligence models, you need vast amounts of data and, therefore, linguistic resources, including textual and oral corpora, computational lexicons and terminology databases, to name a few. The Spanish language has a rich corpus compiling tradition, initiated by pioneering projects from the Royal Spanish Academy (RAE) and associated academies, and continued today by various open initiatives.

Classical reference corpora include those created under the auspices of the Royal Spanish Academy (RAE) for its lexicographical work. The Reference Corpus of the Current Spanish [Corpus de Referencia del Español Actual (CREA)][5] compiled electronic versions of Spanish texts produced between 1975 and 2004 to document the actual contemporary use of the language in both Spain and Latin America. CREA was launched in the 1990s and reached several hundred million words, covering a variety of genres (90% written texts, such as books, newspapers and blogs, and 10% oral transcriptions) with a 50/50 geographical balance between materials from Spain and Latin America. Meanwhile, the Diachronic Corpus of Spanish [Corpus Diacrónico del Español (CORDE)][6], designed to be used for historical purposes, brings together around 300 million

4. See: https://ai71.ai/.

5. See: https://corpus.rae.es/creanet.html.

6. See: https://www.rae.es/banco-de-datos/corde.

words from texts representing all periods of Spanish, from its medieval origins to the 20th century. Published in the early 2000s, CORDE includes some 5,500 complete works (literary, documentary, scientific, etc.) that have been carefully selected to trace over a thousand years of evolution in Spanish from the Iberian Peninsula and Latin America. These colossal databases, CREA and CORDE, have been essential tools for Spanish linguistic research and provided the foundation for works such as the Nuevo diccionario histórico del español (New Historical Dictionary of Spanish).

The RAE and the Association of Spanish Language Academies (ASALE) have undertaken a new large-scale synchronised corpus project in the 21st century: the Spanish Twenty-first Century Corpus [Corpus del Español del Siglo XXI (CORPES XXI)][7]. The CORPES project, which began in 2013, aims to collect 300 million spelling forms produced between 2001 and 2030, covering all Spanish-speaking countries and a variety of registers. Periodic updates to CORPES reflect its growth, with version 0.92 (published in June 2020) reaching 312 million orthographic forms from more than 300,000 documents (including, for the first time, a significant portion of oral transcriptions).

Likewise, Spanish corpus availability has been boosted by international web crawling and open data initiatives. For example, the esTenTen corpus (part of the TenTen series of internet corpora) reached 8 billion words in Spanish, retrieved from the web and including texts from 19 national varieties of the language (Jakubíček et al., 2013). Such massive corpora, compiled by European universities, are valuable for training machine learning systems that require vast amounts of text. "Parallel corpora" that include Spanish are also abundant in the multilingual field and are used primarily for machine translation, from the classic Europarl[8] (from the European Parliament) to collections such as OpenSubtitles, ParaCrawl, and WikiMatrix, with hundreds of

7. See: https://www.rae.es/corpes/.

8. See: https://www.statmt.org/europarl/.

millions of bilingual segments in Spanish-English, Spanish-French, etc. In fact, the NLP community has been building open repositories where Spanish is particularly well represented. This is illustrated by the Hugging Face Datasets platform[9], which as of 2025 features dozens of datasets in Spanish, ranging from benchmarking collections (e.g., Semantic Textual Similarity in Spanish, translations of the MMLU general knowledge questionnaire into Spanish, and the ArguAna set of translated argumentative queries) to specialised corpora such as MilaNLProc HONEST[10] (designed to evaluate bias in Spanish text generation). Most of this data has been compiled thanks to competitions and forums such as IberLEF (Iberian Languages Evaluation Forum)[11], where teams from Spain and Latin America share data for tasks such as sentiment analysis, offensive language detection, and automatic summarisation, etc.

One key resource, which is probably even less recognised in AI, is Spanish speech and phonetic corpora. They have become more important with the rise of AI speech recognition and speech synthesis. Global initiatives such as Mozilla's Common Voice[12] include large-scale audio contributions from volunteers in Spanish (as well as Catalan, Galician, Basque, etc.), providing tens of thousands of recordings that can be used to train speech-to-text and text-to-speech algorithms. Likewise, various universities have compiled corpora of spontaneous speech in different varieties (e.g., corpora of telephone conversations, interviews, or radio broadcasts in American Spanish. Although these resources tend to be less centralised than textual ones, their advantage lies in the fact that, due to their uneven distribution and lack of awareness, in most cases they have not been used to train LLMs or AI algorithms.

9. See: https://huggingface.co/.

10. See: https://github.com/MilaNLProc/honest.

11. See: https://sites.google.com/view/iberlef-2025.

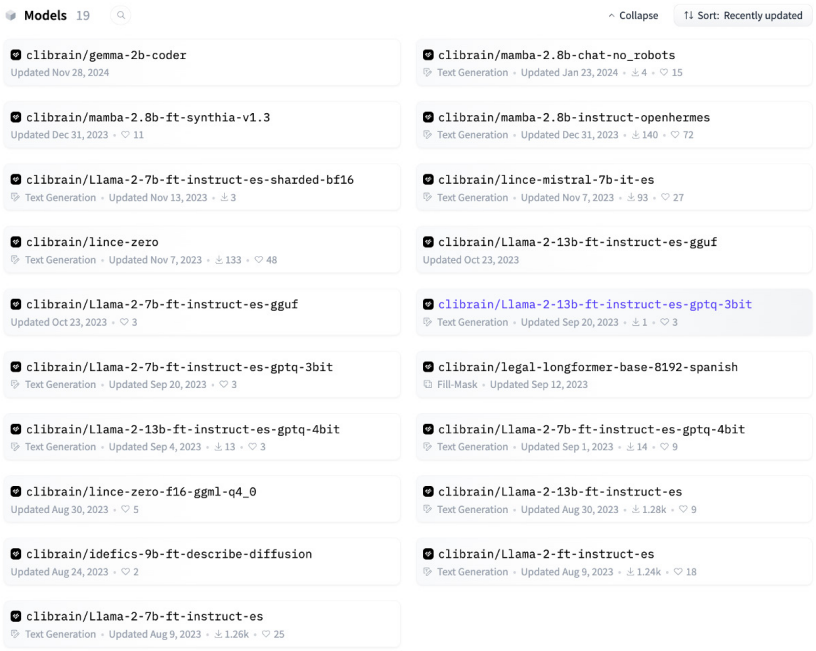
12. See: https://commonvoice.mozilla.org/.

In summary, the Spanish language resource ecosystem combines "large general corpora" (RAE, BNE, online) with modern "specific datasets" accessible in open repositories. This abundance of data, which was unimaginable just two decades ago, fuels current advances in language AI. Nevertheless, areas for improvement persist, such as the availability of high-quality labelled data for complex tasks (deep semantic analysis, coreference, etc.) or representative corpora of colloquial registers and social media. Collaborative efforts between institutions (academic, governmental, and business) are crucial to identifying gaps and creating new resources that further enrich the Spanish language outlook in the age of AI.

In the Spanish-speaking world, language models trained specifically in Spanish have appeared in recent years, representing a qualitative leap forward for applying artificial intelligence to the Spanish-speaking ecosystem. One of the first was the MarIA model, developed by the Barcelona Supercomputing Centre – National Supercomputing Center (BSC-CNS. It is considered the first massive language model trained entirely in Spanish, using the digital collections of the National Library of Spain as its basis (Gutiérrez Fandiño et al., 2021). It showed remarkable performance in tasks such as reading comprehension, text classification and content generation, being particularly useful in contexts that require cultural or regulatory sensitivity. After ChatGPT appeared, the first Spanish generative model, created by Spanish start-up Clibrain, was LINCE[13]. This family of models was specifically designed to cover Spanish and all its variants, achieving excellent results for Spanish in evaluation benchmarks. Commercially launched in 2023 and available on Huggingface in an open version with 7 billion parameters, LINCE achieved results comparable in quality to GPT-3 but optimised for Spanish, and stands out for its ability to effectively handle dialectal variants of Spanish.

13. See: https://techcrunch.com/2023/07/12/lince-llm/.

GRAPH 1 LINCE FAMILY MODELS, BY CLIBRAIN



SOURCE: https://huggingface.co/clibrain.

Likewise, evaluation forums such as IberLEF have tested models specifically optimised for Spanish in tasks such as sentiment detection, offensive language detection, and semantic analysis. Among these, BETO, a Spanish version of BERT, and RoBERTuito (Pérez et al., 2022), a variant of RoBERTa focused on Latin American Spanish, stand out, having achieved outstanding results in multiple benchmarks (Cañete et al., 2020).

However, this recent momentum would not have been possible in the absence of a solid foundation of computational tools and resources developed over the last two decades. Examples of some of these pioneering projects include FreeLing (Carreras et al., 2004) is a language analysis suite developed by the Universitat Politècnica de Catalunya.

It provides functionalities such as tokenisation, lemmatisation, morphosyntactic analysis, grammatical tagging, syntactic analysis and entity recognition, and has been widely used in multilingual industrial and academic environments. Apertium (Forcada et al., 2011) is an open-source platform specialising in related language pairs (such as Spanish-Catalan or Spanish-Galician) that is widely used in regional governments and in language processing education.

Deep syntactic analysis in Spanish has also benefited from annotated corpora such as AnCora (Taulé Delor, Martí Antonín and Recasens Potau, 2008), which was developed by the University of Barcelona and the Autonomous University of Barcelona and includes syntactic and semantic annotations in both Spanish and Catalan. The AnCoraPipe tool supports automatic annotation and has been widely used for training and evaluating parsers. Meanwhile, the IXA group at the University of the Basque Country has developed the IXA Pipes toolkit, which focuses on morphosyntactic analysis, disambiguation and semantic annotation in Spanish and Basque, within a multilingual vision of linguistic processing (Agerri Gascón, Bermúdez and Rigau Claramunt, 2014).

As for Latin America, the Linguistic Engineering Group at the National Autonomous University of Mexico (UNAM) has produced a series of specific tools for analysing Mexican Spanish, referred to as NLP Tools from the GIL group, applied both to NLP projects and computational linguistics studies (Galicia Haro, Gelbukh and Bolshakov, 2002). The TALP group at the Technical University of Catalonia has also advanced machine translation and information extraction tools under the name TALP Tools, focusing particularly on statistical and machine learning methods (Crego Clemente, 2008). Another key resource is UDPipe, which is a set of syntactic models trained for multiple languages (including Spanish) in accordance with Universal Dependencies standards that provide grammatical dependency analysis, lemmatisation, and POS tagging (Straka and Strakova, 2017). Lastly, special mention should be given to LinguaKit (Gamallo et al., 2018), a tool developed by the CITIUS group at the University of Santiago de Compostela, which combines syntactic

analysis, entity extraction and sentiment classification, and has been used in both academic research and business prototypes. Finally, the resources developed by Stanford University deserve recognition, which although mainly focused on English, also had specific libraries developed in Spanish based on NLTK systems[14].

Most of these resources were created in universities and public institutions and played a decisive role in the initial development of automated Spanish processing. While many have been superseded by generative AI models in terms of accuracy, they remain relevant in specialised tasks, creating controlled linguistic corpora, training students, and comparative evaluation of new systems.

4. AI and language research groups in the Spanish-speaking world

Progress in AI applied to Spanish has been achieved thanks to decades of work by academic and research groups specialising in natural language processing, artificial intelligence and machine learning. The Spanish-speaking world has an active and well-established scientific community, including centres of excellence that have contributed to the development of resources and algorithms, as well as the training of new researchers.

In Spain, many autonomous communities have at least one leading team in language technologies, and the most prolific have been those linked to the existence of co-official languages. One of the most prominent groups is the Language Technologies Unit at BSC-CNS in Barcelona. This group has led national projects such as Plan TL, MarIA and the ALIA models (Barcelona Supercomputing Center, 2025) and other European projects, which combine the power of supercomputing with computational linguistics.

14. See: https://www.nltk.org/index.html.

Universities have also played a key role in this ecosystem. The IXA group at the University of the Basque Country is known internationally for its work on multilingual machine translation (especially between Basque, Spanish, and English) and biomedical language processing. IXA has contributed to open-source tools (e.g., the Apertium translation engine for pairs such as Spanish-Basque) and has led clinical NLP projects in collaboration with regional health systems. In Catalonia, the Technical University of Catalonia has the TALP (Language and Speech Technologies) group, which has been researching statistical and neural machine translation and speech recognition for decades. Similarly, the Pompeu Fabra University is home to the TALN group, which works on automatic summarisation and semantic analysis. In Madrid, the Technical University of Madrid and the Open University of Spain (UNED) have active groups working on information retrieval and NLP. The former, for example, was a pioneer in Spain in the field of QA (question answering), and the UNED has been organising international competitions in Spanish for many years (CLEF tasks). Another benchmark is the PRHLT Research Center at the Technical University of Valencia, which specialises in text mining, hate speech and plagiarism detection, and has coordinated challenges in opinion analysis in Spanish. The Instituto de Ingeniería del Conocimiento (Institute of Knowledge Engineering, IIC), linked to the Autonomous University of Madrid, has been applying NLP to industry (banks, surveys, etc.) for years. The Spanish National Research Council (CSIC) has hosted projects at the Institute of Language, Literature and Anthropology focused on corpus linguistics and dialectology which, although not initially oriented towards NLP, are an invaluable source of information through their Atlas Lingüísticos (Linguistic Atlases). Furthermore, the recent opening of the ELLIS Alicante unit (European Laboratory for Learning and Intelligent Systems) brings an international AI focused hub that includes research into language models and deep learning, which could generate synergies with the field of linguistics.

Regarding scientific associations, the Spanish Society for Natural Language Processing (SEPLN), founded in the 1980s, should be highlighted.

It has maintained a cohesive academic community and organises an annual conference with significant scientific publications in the sector.

Latin America also has notable research groups in NLP, despite having more limited resources in some cases. Mexico is one of the most active countries, with the Natural Language Laboratory at the National Technical Institute (IPN) producing numerous works on semantics and computational linguistics. The National Autonomous University of Mexico (UNAM) is home to the Linguistic Engineering Group (GIL), which develops everything from electronic dictionaries to information extraction systems focused on Mexican Spanish. In South America, Argentina has the Argentinian Society of Human Language Technologies, which brings together local researchers; the University of Chile and the Pontifical Catholic University of Chile have teams collaborating on regional projects (e.g. development of spell-checkers for Chilean Spanish, corpus analysis of Spanish on Chilean social networks). Countries such as Colombia and Peru have increased their presence. For example, the University of the Andes (Colombia) has participated in machine translation competitions for Spanish, and Peru has begun to explore tools for translation between Spanish and Quechua using AI. Several of these projects are interconnected through Ibero-American events (the CICLing conference in Mexico or the IBERAMIA AI conference, which often includes NLP sessions, etc.).

International community initiatives promoted by young Spanish-speaking researchers deserve special mention, as they have emerged to democratise NLP in our language. One recent example is SomosNLP, which was launched in 2021 and brings together students and professionals from different countries to achieve fair representation of Spanish and its variants in the age of AI. Projects include #Somos600M, which aims to create corpora and models that reflect the rich dialectal diversity of Spanish spoken by over 600 million people worldwide. Initiatives such as these, on the fringes of formal institutions, demonstrate how passionate the younger generations are about promoting Spanish in global AI, with open and collaborative approaches

(e.g. international hackathons organised by SomosNLP with participants from dozens of countries, focused on developing free linguistic resources).

5. The Spanish linguistic AI business ecosystem

The start-up ecosystem has emerged as a major driver of innovation, given the necessary pace of development in this sector. However, despite the availability of significant European funding and the emphasis on the need to promote this type of venture, few of these start-ups have achieved a size or international relevance comparable to that of European leaders such as Mistral.ai, a French start-up considered a "unicorn" for having already reached a valuation of over $1 billion. This start-up was created just three years ago and has stood out not only for the quality of its LLMs and algorithms, but also for its speed in raising particularly large investment rounds and for receiving strong political support from the French government.

The Spanish ecosystem of the last two decades has produced much smaller milestones, but with relevant names that have brought dynamism to the sector, such as Sherpa.ai (based in Bilbao, focused on virtual voice assistants and predictive AI), Inbenta (a Barcelona-based chatbot and multilingual semantic search developer with a strong focus on Spanish that has expanded to Silicon Valley) and Narrativa (a Madrid-based start-up that applies natural language generation to automated journalism, producing financial news autonomously). Text editors also share a strong track record, with numerous examples such as Escribelo.ai, a Spanish writing assistant optimised for SEO.

Machine translation, especially for languages spoken on the Iberian Peninsula, is another area where interesting business ventures have emerged. For example, Pangeanic (based in Valencia) has excelled in machine translation and text anonymisation, participating in European projects and attending the presentation of the TL Plan in 2015 as a leading company in the sector. Other established firms include Indra

(and its subsidiary Minsait), which has developed language processing solutions for public administrations; Eleka/Elhuyar and Vicomtech, based in the Basque Country, focused on multilingual systems that integrate Basque; and Verbio, originally from Barcelona, a specialist in voice technologies (recognition and synthesis in Spanish and other languages) since the early 2000s. Telefónica has also launched its Aura virtual assistant with Spanish language capabilities, which has led to collaborations with the RAE to ensure correct use of the language. Likewise, large technology consulting firms (IBM, NTT Data, Atos, etc.) have established NLP units in Spain to develop chatbots and semantic analysis for local clients, adapting global technologies to Spanish.

The new wave of generative AI, combined with increased funding and a focus on the creation of technology-based companies, has triggered the emergence of AI companies specialising in specific areas (such as SoyLuzia for education and MAIte for the legal sector. However, it is still too early to determine which of these companies will grow large enough to survive in this technological race. A typical example of the difficulties facing the ecosystem is Clibrain, which is a promising company founded in Madrid in 2022 that launched the first large language model entirely trained in Spanish, LINCE, in 2023. Despite their ambition and impressive technical models, which offered results comparable to GPT-3 in Spanish and excelled in their dialectal coverage, the company ceased operations less than a year after its launch. On the positive side, Magnific.ai, a Murcia-based start-up pioneering AI image enhancement, was sold to Freepik in record time, while Multiverse, based in the Basque Country, offers an interesting LLM compression system.

Another encouraging feature of this environment is the significant increase in venture capital investment in Spanish start-ups at various stages, from early-stage funding to large rounds where they compete with European and international companies.

6. Public strategies in Spain and the European framework

The Spanish government has developed a series of strategic plans to promote language technologies and AI applied to Spanish over the last decade, in line with broader European policies. One early milestone was the launch of the Plan to Promote Language Technologies (LT Plan) in October 2015. This five-year programme (2016-2020) involved a public investment of €90 million to boost the natural language processing (NLP) industry in Spanish.

Following the completion of the LT Plan in 2020, the national strategy was integrated into broader AI initiatives. In December 2020, the Spanish government approved the National Artificial Intelligence Strategy (ENIA), which recognised the Spanish language as a strategic asset in the digital ecosystem. In 2021, the ambitious Strategic Project for Economic Recovery and Transformation (PERTE) called "Nueva economía de la lengua" (New Economy of Language) was launched as part of the Next Generation EU plan. Officially unveiled in 2022, this PERTE mobilised €1.1 billion in public investment (partly from European funds) with a view to promoting Spanish and co-official languages as a factor for economic growth.

Meanwhile, Spain has been actively involved in the European AI and language regulatory and funding framework, supporting the proposed European AI Regulation (AI Act). One critical issue in the Hispanic context is how the linguistic AI revolution impacts Spain's co-official languages (Catalan, Galician, and Basque) and the many dialectal variants of Spanish spoken on the peninsula and in the Americas. These languages and varieties embody the multilingual and cultural richness of Spain and Latin America. Therefore, ensuring their presence in the AI era is essential to prevent digital gaps and the loss of linguistic heritage.

In the case of Spain's co-official languages, their technological development has always varied. Catalan, with more than 10 million speakers,

has had machine translation, spell checkers and voice assistants (e.g. Catalan versions of Siri or Alexa) for many years, driven in part by its similarity to other Romance languages and by the institutional support provided by the Generalitat de Catalunya (Catalan regional government). Recent initiatives, such as Projecte AINA (launched in 2022), aim to compile large Catalan oral and written corpora to train specific AI models and ensure the digital survival of Catalan in the future. Similarly, Basque (a non-Indo-European language spoken by around 750,000 people) has required specific developments. For example, Basque institutions and companies such as Elhuyar and Vicomtech have created neural translators and even assistants (e.g. Zurekin for Basque voice) based on bilingual corpora. Galician, which shares similarities with Portuguese and is spoken by around 2 million people, also has tools such as spell checkers (the Xunta de Galicia (Galician Regional Government) released one integrated into LibreOffice) and is working on voice synthesisers for Galician.

One of the most ambitious initiatives in the field of linguistic AI in Spanish is the development of the ALIA-40B model, led by the BSC-CNS. This model contains 40 billion parameters and has been trained on 2.5 TB of textual data in Spanish and co-official languages. Its Mixture-of-Experts architecture makes it possible to activate only the parts of the model necessary for each task, thus optimising computational resources. The project is active as of 2025 and has released some of its versions and weights through Hugging Face[15]. Meanwhile, the reduced version, Salamandra-7B[16], designed for low-latency environments, has shown particularly competitive results in minority languages. Both models are in constant evolution and being applied in sensitive domains such as legal and medical, using enhanced learning techniques.

Concurrently, the Basque Language Technology Centre (HiTZ), made up of the IXA and Aholab groups at the University of the Basque Country

15. See: https://huggingface.co/BSC-LT/ALIA-40b.

16. See: https://huggingface.co/BSC-LT/salamandra-7b-instruct.

(UPV/EHU), has gained recognition for its advances in applied linguistic technologies. Its recent developments include voice synthesis systems adapted to the dialectal variants of Basque (with an average intelligibility rate of 94%), semantic disambiguation tools for legal texts in Spanish, and linguistic models for the early detection of neurodegenerative diseases through the analysis of speech patterns. Their collaboration with Osakidetza (the Basque Health Service) has led to the implementation of automatic triage systems with 92% accuracy in classifying medical emergencies, representing a pioneering example of technology transfer in public health.

Furthermore, the TAN-IBE (Neural Machine Translation for Iberian Languages) project, coordinated by the Open University of Catalonia (UOC), is focusing on an inclusive multilingual strategy for machine translation. Using transfer learning techniques and Transformer-XL neural networks, the consortium has worked with parallel corpora of over 85 million sentence pairs for seven Romance languages spoken on the Iberian Peninsula: Spanish, Catalan, Galician, Portuguese, Asturian, Aragonese and Aranese. The combination of semantic alignment using knowledge graphs and cross-training has improved the BLEU score for translations between minority languages by 37%, thus significantly reducing reliance on large bilingual corpora (Oliver González, 2024). This approach constitutes a significant breakthrough in the accessibility and digital preservation of Spain's regional languages.

7. Conclusion

The Spanish language has an undeniable competitive advantage in linguistic AI, such as the enormous availability of data and corpora in Spanish and its many geographical and social variants. However, this abundance is limited by one key structural difficulty, namely that access to these resources is often restricted, be it due to legal, technical or interoperability obstacles. While various sources of funding and multiple parallel initiatives exist, the fragmented ecosystem limits the impact of current efforts. It would be desirable to have greater inter-institutional

and public-private collaboration to enable the launch of large-scale projects, which can compete in terms of funding, speed and quality with major international developments, especially those promoted in the United States and China. In addition, the richness and variety of the Spanish language presents a unique opportunity, whose diversity can only be fully addressed by Spanish-speaking communities. The application of AI to different industrial fields and, in particular, linguistic and cultural specialisation through diatopic, diastratic and diaphasic variants has yet to be addressed in the current development of artificial intelligence. The creation of an ecosystem where political momentum fosters collaboration between businesses, academia and cultural institutions is the only way to achieve these objectives.

Bibliographic references

Agerri Gascón, R., Bermúdez, J. & Rigau Claramunt, G. (2014). IXA pipes: Efficient and ready to use multilingual NLP tools. In N. Calzolari, K. Choukri, T. Declerck, H. Loftsson, B. Maegaard, J. Mariani, A. Moreno, J. Odijk, & S. Piperidis (Eds.), *Proceedings of the Ninth International Conference on Language Resources and Evaluation* (LREC'14) (pp. 3.823-3.828) UPV/EHU. European Language Resources Association (ELRA). http://www.lrec-conf.org/proceedings/lrec2014/pdf/775_Paper.pdf

Automatic Language Processing Advisory Committee (ALPAC) (1966). *Languages and Machines. Computers in Translation and Linguistics*. National Academy of Sciences.

Barcelona Supercomputing Center (21 January 2025). ALIA, *Europe's first public, open and multilingual AI infrastructure*. https://www.bsc.es/news/bsc-news/alia-europes-first-public-open-and-multilingual-ai-infrastructure

Cañete, J., Chaperon, G., Fuentes, R., Ho, J.-H., Kang, H. & Pérez, J. (2020). Spanish Pre-Trained BERT Model and Evaluation Data. https://arxiv.org/abs/2308.02976

Carreras X., Chao, I., Padró, L. & Padró, M. (2004). FreeLing: An open-source suite of language analyzers. In M. T. Lino, M. F. Xavier, F. Ferreira, R. Costa & R. Silva (Eds.), *Proceedings of the Ninth International Conference on Language Resources and Evaluation* (LREC'04). European Language Resources Association (ELRA). http://www.lrec-conf.org/proceedings/lrec2004/pdf/271.pdf

Crego Clemente, J. M. (2008). *Architecture and Modeling for N-gram-based Statistical Machine Translation* [PhD diss., Technical University of Catalonia]. https://nlp.lsi.upc.edu/papers/phd_jmcrego.pdf

Chomsky, N. (1957). *Syntactic Structures*. De Gruyter Mouton.

Devlin, J., Chang, M-W., Lee, K. & Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers. arXiv:1810.04805

Forcada, M. L., Ginestí-Rosell, M., Nordfalk, J., O'Regan, J., Ortiz Rojas, S., Pérez Ortiz, J. A., Sánchez Martínez, F., Ramírez Sánchez, G. & Tyers, F. M. (2011). *Apertium: A free/open-source platform for rule-based machine* translation. *Machine Translation*, 25 (2), 127-144.

Galicia Haro, S, N., Gelbukh, A. Bolshakov, I. A. (2002). Compilation of a Mexican Spanish text corpora. *WSEAS Press*. https://www.gelbukh.com/CV/Publications/2002/WSEAS-2002-Corpora.pdf

Gamallo Otero, P., García González, M., Piñeiro Pomar, C. A., Martínez Castaño, R & Pichel CAmpos, J. C. (2018). LinguaKit: A Big Data-Based Multilingual tool for linguistic annotation and information extraction. *IEEE* Xplore. https://ieeexplore.ieee.org/document/8554689

Gutiérrez Fandiño, A., Armengol Estapé, J., Pàmies Massip, M., Llop Palao, J., Silveira Ocampo, J., Pío Carrino, C., Armentano Oller, C., Rodríguez Penagos, C., González Agirre, A. & Villegas Montserrat, M. (2022). MarIA: Spanish Language Models. *Procesamiento del Lenguaje Natural*, 68, 39-60. http://journal.sepln.org/sepln/ojs/ojs/index.php/pln/article/view/6405/3820

Hitz Center UPV/EHU (2025). *Medical and Legal Domains*. https://www.hitz.eus/en/domains

Hochreiter, S. & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, 9 (8), 1.735-1.780.

Hutchins, J. (2014). The history of machine translation in a nutshell. https://aclanthology.org/www.mt-archive.info/10/Hutchins-2014.pdf?

Koehn, P. (2010). *Statistical Machine Translation*. Cambridge University Press.

Mikolov, T., Chen, K., Corrado, G. & Dean, J. (2013). Efficient Estimation of Word Representations in Vector Space. arXiv:1301.3781

Oliver González, A. (2024). TAN-IBE Participation in the Shared Task: Translation into Low-Resource Languages of Spain. En B. Haddow, T. Kocmi, P. Koehn y C. Monz (Eds.), *Proceedings of the Ninth Conference on Machine Translation* (pp. 871-877). https://aclanthology.org/2024.wmt-1.83/

Open AI (14 March 2023). *GPT-4*. https://openai.com/research/gpt-4

Open AI (30 November 2022). *Presentamos ChatGPT*. https://openai.com/es-ES/index/chatgpt/

Pérez, J. M., Furman, D. A., Alonso Alemany, L. & Luque, F. M. (2022). RoBERTuito: a pre-trained language model for social media text in Spanish. En N. Calzolari, F. Béchet, P. Blache, K. Choukri, C. Cieri, T. Declerck, S. Goggi, H. Isahara, B. Maegaard, J. Mariani, H. Mazo, J. Odijk & S. Piperidis (Eds.), *Proceedings of the Thirteenth Language Resources and Evaluation Conference* (LREC) (pp. 7.235-7.243). http://www.lrec-conf.org/proceedings/lrec2022/pdf/2022.lrec-1.785.pdf

Radford, A., Narashiman, K., Salimans, T. & Sutskever, I. (2018). Improving Language Understanding by Generative Pre-Training. *OpenAI*. https://cdn.openai.com/research-covers/language-unsupervised/language_understanding_paper.pdf

Straka, M. & Straková, J. (2017). Tokenizing, POS Tagging, Lemmatizing and Parsing UD 2.0 with UDPipe. En J. Hajic (Ed.), *Proceedings of the CoNLL 2017 Shared Task. Multilingual Parsing from Raw Text to Universal Dependencies* (pp. 88-89). The Association for Computacional Linguistics. https://universaldependencies.org/conll17/proceedings/K17-3.pdf

Taulé Delor, M., Martí Antonín, M. A. & Recasens Potau, M. (2008). AnCora: Multilevel annotated corpora for Catalan and Spanish. In N. Calzolari, K. Choukri, B. Maegaard, J. Mariani, J. Odijk, S. Piperidis & D. Tapias (Eds.), *Proceedings of the Sixth International Conference on Language Resources and Evaluation* (LREC'08). European Language Resources Association (ELRA). http://www.lrec-conf.org/proceedings/lrec2008/pdf/35_paper.pdf

Turing, A. M. (1950). Computing Machinery and Intelligence. *Mind*, 59 (236), 433-460. https://phil415.pbworks.com/f/TuringComputing.pdf

V. Jakubíček, M. Kilgarriff, A., Kovář, V, Rychlý, P. & Suchomel, V. (23-26 June 2013). *The TenTen Corpus Family*. 7th International Corpus Linguistics Conference (CL-2013), Lancaster, Reino Unido. https://www.sketchengine.eu/wp-content/uploads/The_TenTen_Corpus_2013.pdf

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L. & Polosukhin, I. (2017). Attention Is All You Need. NIPS 2017. https://arxiv.org/pdf/1706.03762

Weizenbaum, J. (1966). ELIZA-A Computer Program for the Study of Natural Language Communication Between Man and Machine. *Communications of the* ACM, 9 (1), 36-45.

Winograd, T. (1972). *Understanding Natural Language*. Academic Press.

Wu, Y., Schuster, M., Chen, Z., Le, Q. V., Norouzi, M., Macherey, W., Krikun, M., Cao, Y., Gao, Q., Macherey, K., Klingner, J., Shah, A., Johnson, M., Liu, X., Kaiser, L., Gouws, S., Kato, Y., Kudo, T., Kazawa, H., ... Dean, J. (2016). Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation. arXiv:1609.08144

TERMINOLOGY IN SPANISH AND DIGITAL MULTILINGUALISM

Elea Giménez Toledo
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS
(SPANISH NATIONAL RESEARCH COUNCIL (CSIC)
(INSTITUTE OF LANGUAGE, LITERATURE AND ANTHROPOLOGY)

1. Introduction

Languages currently encounter the most fertile and stimulating environment for growth, recognition, and development in the digital realm, thanks to language technologies. However, not all languages can take advantage of this enormous opportunity presented by the digital space, as they do not all enjoy the same conditions for entering this territory. The concept of digital linguistic equality has been defined as "the state of affairs in which all languages have the technological support and situational context necessary for them to continue to exist and to prosper as living languages in the digital age" (Gaspari et al., 2021).

Digital multilingualism has been championed in the field of research and in certain public policies for some time, following the observation of linguistic biases derived from the prevailing monolingualism of English (Spence, 2021).

A multilingual digital ecosystem requires, for one, a vision that is not strictly technological, but rather one that integrates the social and cultural perspectives represented by multilingualism (Spence, 2021). There is a need to consider the significance of the presence of multiple languages, at all levels, in the digital environment. Multilingualism brings with it a diversity of perspectives, visions and analyses. Therefore, as Horvath et al. (2024) argue, through collaborative effort and a commitment to language diversity and inclusivity, we can strive toward "infrastructures that challenge conventional thinking through their ability to

support and enrich the multilingual and multifaceted nature of contemporary scholarship". They further note that the existence of multilingual infrastructures should not be left to large companies, which is a sentiment shared in the Disrupting Digital Monolingualism report (Spence, 2021), but rather should be addressed with solutions provided by digital humanists.

The conviction that a multilingual digital ecosystem is needed must be followed by vision, strategy, resources and a substantial amount of work from each of the languages. Digital multilingualism can only occur by supporting the development of each language in the digital environment, encompassing all aspects of the concept (Giménez Toledo, 2024): public policies and scientific diplomacy actions; content circulation on the internet in different languages, duly labelled (e.g. Céspedes et al., 2025); availability of corpora and data in formats suitable for machine readability; language models; lexical resources which, in the case of open data, should follow the FAIR principles so that they can be found, are accessible, interoperable and reusable; existence of infrastructures, networks and data spaces; and availability of digital platforms, tools and algorithms that work in and with different languages.

This chapter is specifically aimed at analysing the influence of Spanish in the linguistic resources collected in European digital infrastructures and explores the relevance and usefulness of Spanish terminology as a key linguistic resource for the development of different language technologies.

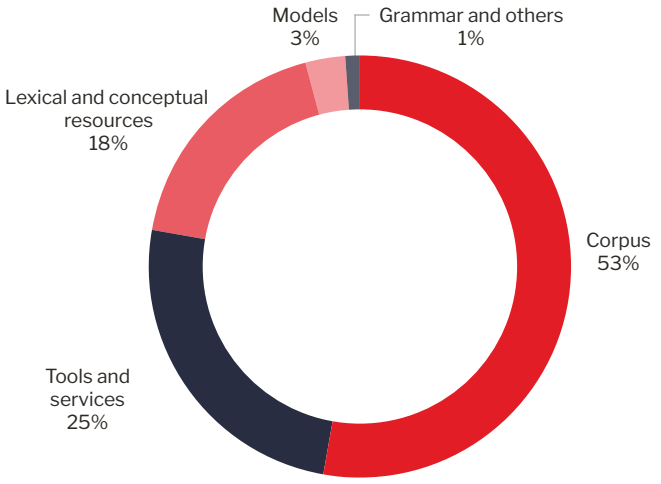
2 . Spanish language resources in digital data infrastructures

Language resources and technologies are critical for language development in the digital environment and must be recognised, accounted for and analysed as a valuable asset.

From corpora to linguistic data, transcription and translation tools, they are all crucial for developing artificial intelligence applications such as virtual assistants, voice searches and conversational systems. This is why the European Commission launched the European Language Data Space, with a view to sharing language resources in the public and private sectors, thereby enabling the development of tools and services. This data space is used to share linguistic data and connect data providers and users, all of which takes place in an environment that complies with European legislation. The reason for creating this inherently multilingual data space is the need to feed it with data from different languages.

Within the European Language Data Space, the European Language Grid (ELG) has been created, whose catalogue contains language resources and technologies that are intended to be made visible, shared and reused. These resources include textual and multimodal corpora, language models, transcription and translation systems, lexical resources, annotated data collections used to train models, platforms for speech synthesis or recognition, named entity recognition systems, morphological analysers and lemmatisers, to name but a few. The ELG catalogue is a valuable source for analysing the state of languages in terms of language resources and technologies. In addition to the data collected, thanks in part to funded calls for proposals, it takes data from various sources such as CLARIN, ELRA, and the Zenodo and Hugging Face repositories.

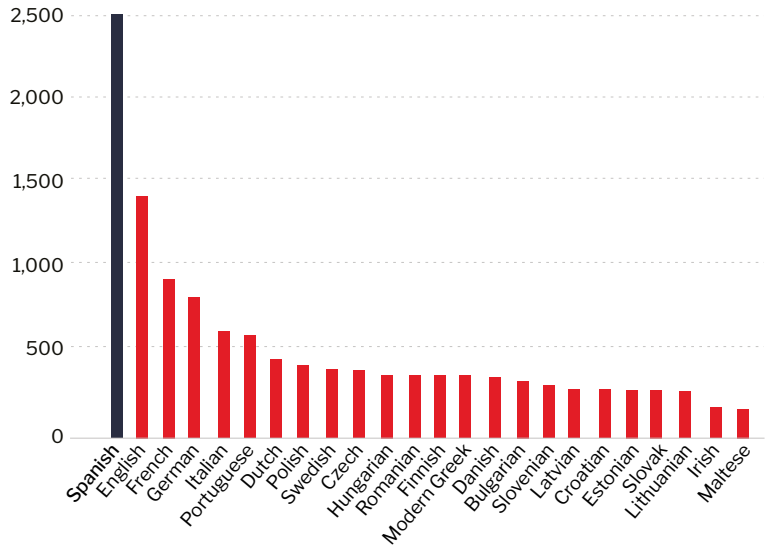
GRAPH 1 TYPES OF LANGUAGE RESOURCES AND TECHNOLOGIES (MULTILINGUAL) IN ELG (MARCH 2025)



SOURCE: compiled by the author based on data from ELG (2025).

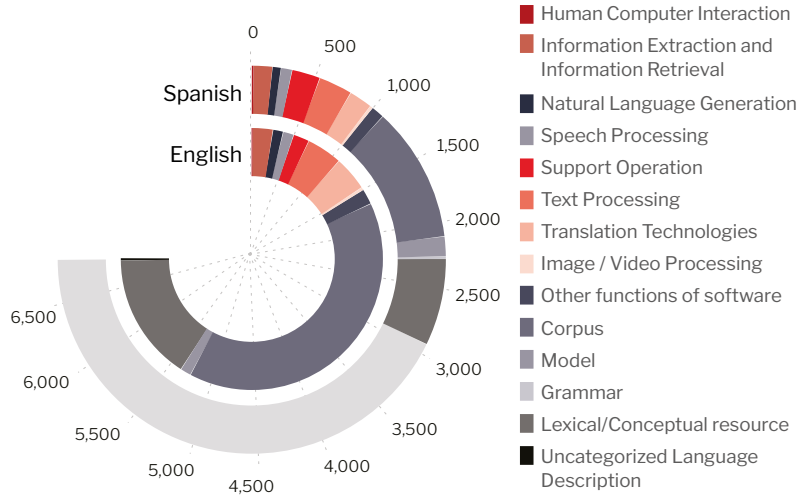
The Spanish language has the most resources in this catalogue (2,567), ahead of English (1,466), French (963), German (852) and Italian (644), meaning not only that Spanish resources and technologies are well developed, but also that the language is well positioned in the catalogue of benchmark resources in Europe (Graph 2). However, the gross figure must be clarified. On the one hand, the ranking of Spanish compared to other European languages varies depending on the type of linguistic resource. Accordingly, based on the number of lexical or conceptual resources that each language has in the ELG, i.e. the weight of each language in the total lexical resources (Table 1, last column), Spanish ranks second, accounting for 23% of the resources, but far behind the 50% contributed by English. French and German follow, providing approximately 18% of lexical resources. On the other hand, only 25.6% of the 2,567 language resources in Spanish are lexical resources, which is far behind the relative weight in other languages. This percentage indicates an area that needs to be developed or strengthened.

GRAPH 2 LANGUAGE RESOURCE AND TECHNOLOGY DISTRIBUTION BY LANGUAGE IN ELG (MARCH 2025)



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

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



SOURCE: compiled by the author based on data from ELG (2025).

TABLE 1 LEXICAL RESOURCES IN THE EUROPEAN LANGUAGE GRID

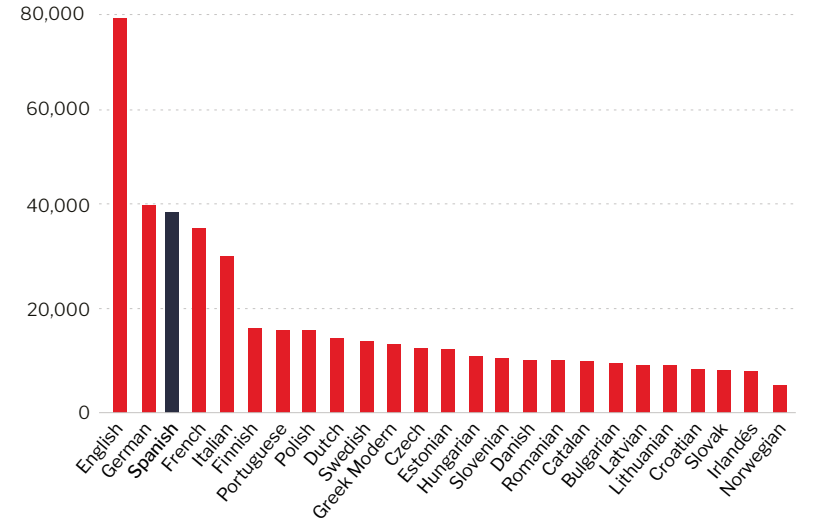
European Languages	No. of resources in ELG catalogue	No. of lexical resources in ELG catalogue	% lexical resources (relative)	% lexical resources (absolute)
English	1,466	1,428	97.41	50.05
Estonian	292	243	83.22	8.52
Slovenian	317	236	74.45	8.27
Lithuanian	284	208	73.24	7.29
German	852	518	60.80	18.16
Latvian	294	166	56.46	5.82
French	963	541	56.18	18.96
Finnish	377	202	53.58	7.08
Swedish	415	193	46.51	6.76
Modern Greek	377	164	43.50	5.75
Croatian	294	105	35.71	3.68
Czech	409	141	34.47	4.94
Italian	644	222	34.47	7.78
Irish	184	55	29.89	1.93
Bulgarian	347	101	29.11	3.54
Dutch	476	134	28.15	4.70
Polish	440	123	27.95	4.31
Maltese	176	48	27.27	1.68
Danish	365	98	26.85	3.43
Spanish	**2,567**	**657**	**25.59**	**23.03**

European Languages	No. of resources in ELG catalogue	No. of lexical resources in ELG catalogue	% lexical resources (relative)	% lexical resources (absolute)
Portuguese	620	150	24.19	5.26
Romanian	378	85	22.49	2.98
Hungarian	383	81	21.15	2.84
Slovak	292	55	18.84	1.93

SOURCE: compiled by the author based on data from ELG (2025).

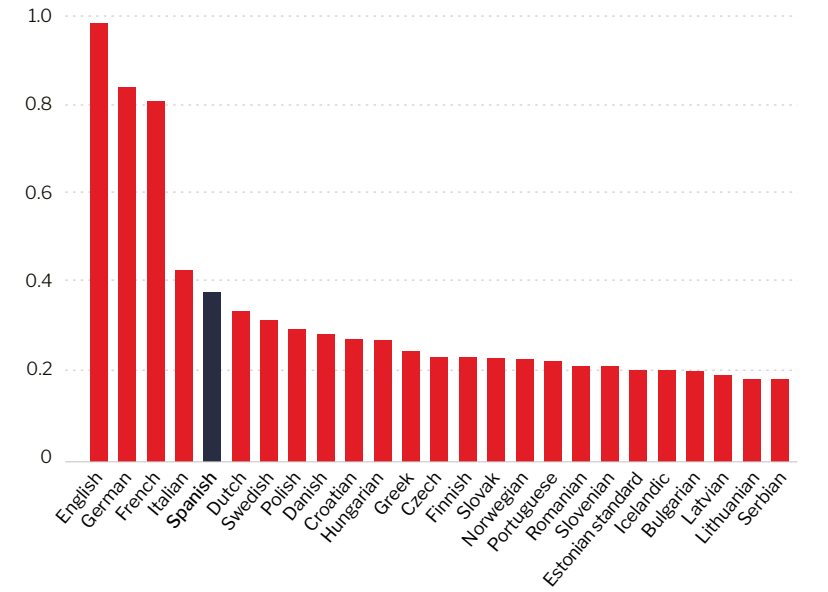
The relative importance of Spanish is also seen when considering DLE (Digital Language Equality) indicators. These indicators measure the weight of each language according to technological and contextual factors. Regarding technological factors, the indicator used considers the number of linguistic resources available for each language in ELG, and assigns them a specific weight based on variables such as whether they are annotated resources, specialised in a particular subject, open source or proprietary, etc. The formula used to calculate this expresses how readable and reusable the linguistic resources are in each language. Graph 4 shows how Spanish ranks third according to this indicator, behind English (more than 40 points behind) and German.

GRAPH 4 DIGITAL LANGUAGE EQUALITY SCORES BASED ON TECHNOLOGICAL FACTORS



SOURCE: compiled by the author based on data from ELG (2025).

GRAPH 5 DIGITAL LANGUAGE EQUALITY SCORES BASED ON CONTEXTUAL FACTORS



SOURCE: compiled by the author based on data from ELG (2025).

CLARIN (Common Language Resources and Technology Infrastructure) is one of Europe's major research infrastructures. It is an open science initiative dedicated to gathering and providing data, tools and services to support research based on linguistic resources. It works with national consortia from 24 European countries, including Spain, which constitute a network of institutions that produce linguistic resources, including both data and tools. The resources section brings together corpora, tools and lexical resources; among the latter, those covering Spanish are scarce. There is only one monolingual lexical inventory in Spanish. In twenty-one other multilingual lexical resources (language models, lexicons, dictionaries or glossaries), Spanish is one of the languages covered. A significant difference can be seen between the monolingual resources provided for other languages such as Estonian, Slovenian or Swedish and those for Spanish or French, which, despite being prominent languages in ELG, have an almost negligible presence in CLARIN.

This situation can be linked to at least two factors. Firstly, it could be due to a lack of awareness of the infrastructure among those who produce corpora, terminology resources, or other types of linguistic resources and tools. This would suggest that the infrastructure should be brought closer to the specialists and vice versa, and that there should be greater dissemination of what they represent, namely spaces built by everyone for a common purpose. Secondly, the lack of resources in Spanish could be related more to a shortage of readability than to their non-existence.

The Virtual Language Observatory launched by CLARIN makes it possible to track linguistic resources in each language from different catalogues. A search for lexical resources in Spanish yields no more than twenty results. Again, the low output may be related to the absence of Spanish language resources in the various catalogues or repositories, but also to technical issues such as missing names, keywords, broken links or incomplete metadata (corpus size, linguistic annotation, etc.) (Fišer, Lenardič and Erjavec, 2018). These shortcomings must be considered by language resource producers.

3. Terminology in Spanish

Terminology has been the subject of special attention in the quantitative analyses presented in the previous sections, among the wide variety of language resources and technologies that exist. This is because terms, specialised words from different scientific and professional fields, properly described by specialists in each domain and carefully crafted from a linguistic perspective, provide the basis for a multitude of language-based digital tools. Thesauri, vocabularies, ontologies, lexicons, lists of terms or named entities (proper names) are all part of what is generically referred to as terminology. Key uses of terminology in the digital environment include:

- efficient retrieval and discoverability of information in systems and search engines;
- automatic indexing or sorting of high volumes of documents, as is the case in the largest database of medical scientific output (National Library of Medicine, 2021; Chen, Bullard and Giustini, 2023; Moore, Yaqub and Sampat, 2024);
- linking to other terminologies in other languages thanks to linked linguistic data technology, making it possible, inter alia, to search for a term in one language and retrieve documents on that subject in other languages (e.g. GoTriple);
- automatic translation systems, as the very existence of carefully chosen and/or standardised terms facilitates the training of these systems and improves the quality of translations;
- the training of algorithms so that machines can understand scientific language in Spanish, generate texts containing technical terms and "speak" correct Spanish;
- support for plain language systems to enable difficult-to-understand texts due to their technical complexity (medical, legislative,

regulatory, etc.) to use alternative expressions that make their content more understandable to their target audience.

Beyond the digital environment, terminology is a fundamental tool to teach scientific and professional Spanish correctly, to communicate between doctors and patients through graphic medicine or infographics, and in the field of translation and interpreting, where specific resources have been created to meet highly demanding professional needs that require precision and immediacy.

4. Spanish in multilingual terminology resources

Many international institutions, from UNESCO to the UN and the European Union, are committed to multilingualism. This is reflected through their policies promoting and protecting different languages, as well as establishing working languages within the institutions themselves and languages for publishing their various documents. The protection of multilingualism entails respect for the identity of each nation, but also legal safeguards for citizens. For example, laws and regulations that affect people living in European Union countries are published in each of the 24 official languages of the EU to ensure they are accessible and understandable to all. Consequently, there is a significant volume of translations requiring technical precision, as the legislation that concerns EU citizens cannot contain inaccuracies or discrepancies between what is stated in one language and another. This need for precision applies above all to terminology. In English, the adjective "green" can be translated as "verde" (colour) or "ecológico" (concept). They are almost treated as synonyms, but they have different meanings, so that increasingly, in institutional translations, "verde" is translated as "green" and "ecológico" as "organic" (Arias Rodríguez, 2024).

These subtleties around vocabulary, searching for greater accuracy and consistency, require considerable work with words, especially technical terms. This is why multilingual terminology databases have

been developed, which are essential for the work carried out by institutional translators and interpreters. The volume of terms they contain is constantly growing, and expert validation and standardisation work is ongoing.

UNTERM is a UN database that contains the terminology used by its six official languages (Arabic, Chinese, Russian, Spanish, English and French). This resource compiles the terminology needed (hundreds of thousands of words) to carry out the work within the organisation and other international bodies that depend on the UN, such as the WHO and UNESCO. Scientific terms are part of this collection because the topics covered often involve direct connections with the field of science.

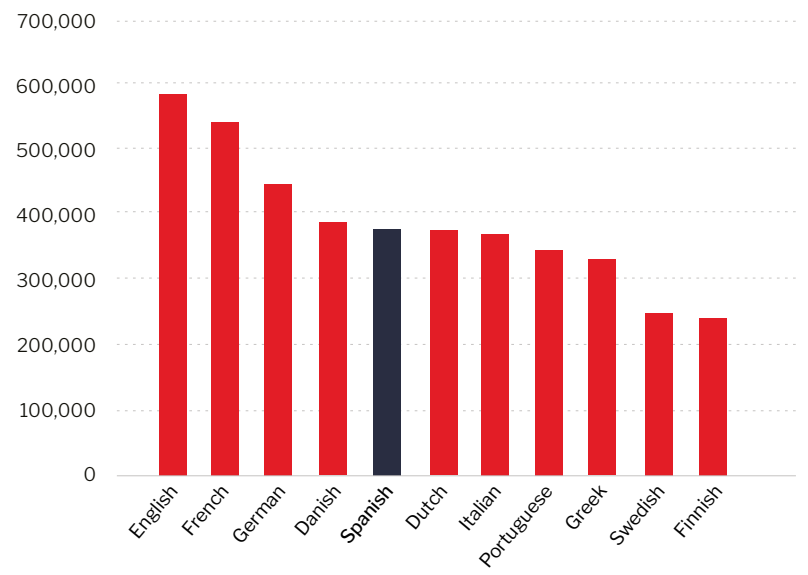
Meanwhile, IATE European Union Terminology (IATE, n.d.) is the world's largest multilingual terminology database, which was created in 2004 to "collect, disseminate and manage European Union-specific terminology". It contains almost seven million terms and covers all 24 official languages of the European Union, including Spanish. As an open resource, it is available to all users and indispensable for language professionals, especially translators and interpreters.

Despite being a multilingual database, the level of coverage for each language varies. On the one hand, this can be explained by the history of the European Union itself, as the number of official languages has grown as new countries have joined, and terms in those languages have therefore been added to IATE. On the other hand, translators' requirements also play a role, as they contribute to the database based on what they need to translate.

IATE statistics reveal some relevant indicators that illustrate the importance of languages. For example, this includes linguistic coverage, which refers to the number of entries or records for each language (Graph 6). Spanish is the fifth most widely covered language. English, French and German are working languages of the European Union,

which explains why they occupy the top positions. Danish and Dutch also rank highly due to their long history as official languages of the European Union.

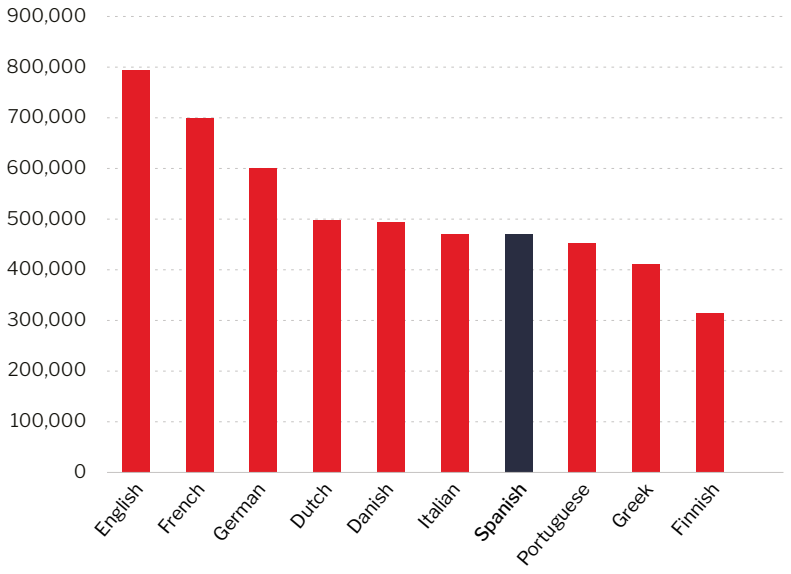
GRAPH 6 LINGUISTIC COVERAGE IN IATE (MARCH 2025))



SOURCE: compiled by the author based on data from IATE (n.d.).

The number of terms in this terminology database is another indicator to measure the presence of Spanish. There are around 470,000 terms, which is a sizeable number that places Spanish in seventh position. English and French top the list with almost 800,000 and 700,000 terms, respectively.

GRAPH 7 NUMBER OF TERMS PER LANGUAGE IN IATE (MARCH 2025)



SOURCE: compiled by the author based on data from IATE (n.d.).

These figures need contextualising. The volume of terminology data is associated with the number of years that different countries have been in the EU, together with the needs of translators, which sometimes involve more work in some languages than in others. It is useful to look at the data from another perspective, namely that terminology work has progressively evolved and the quality of the information provided for each term increases according to how recent the entries are. Languages that were incorporated into IATE later, such as Slovak, Croatian, Maltese, Estonian and Romanian, have more complete entries for each term, with definitions, notes and contextual information. Accordingly, Table 2 shows how the most recently added languages add more value to the database (more complete entries), while languages with greater coverage, including Spanish, have lower percentages of "rich" entries.

TABLE 2 WEALTH OF TERMINOLOGICAL INFORMATION IN IATE (MARCH 2025)

Language	Added value by language (%)
Slovak	80.15%
Croatian	78.20%
Maltese	78.13%
Estonian	75.68%
Romanian	73.41%
Hungarian	72.85%
Czech	72.42%
Bulgarian	71.97%
Polish	69.69%
Lithuanian	68.14%
Slovenian	67.04%
Latvian	65.63%
English	44.26%
German	38.16%
French	36.68%
Finnish	33.57%
Swedish	32.58%
Italian	32.47%
Dutch	32.03%
Greek	25.47%
Portuguese	25.40%
Danish	24.57%
Spanish	**24.22%**
Gaelic	21.12%

SOURCE: compiled by the author based on data from IATE (n.d.).

Recent developments in IATE, which now has a REST API, make it easier to access and reuse terminology data, which represents a key benefit for language professionals such as translators and interpreters.

5. Terminology in Spanish: TeresIA

The digital transformation of linguistic resources is key to improving how existing terminologies are interconnected, which are scattered and, together, would offer really valuable information. This is one of the objectives of TeresIA, a research project that aims to build a meta-search engine to retrieve terms found in different Spanish terminology resources. One of the first steps taken in the project was to compile a census of terminological resources (TeresIA, 2024), which identified around 400 resources, either in Spanish (monolingual) or including Spanish (multilingual). They cover a range of specialisations, come from various types of institutions and are in different formats. The next phase involves reaching agreements with producers regarding their inclusion in the TeresIA meta-search engine, analysing the structure of each resource, and converting the data to make it readable by the meta-search engine. Many of them lack interoperability. The enormous task of conversion comes from the need to render readable and retrievable the vast amount of terminological data found in different sources, which are often not easily accessible. The conversion will also entail creating terminological data sets that can be stored in linguistic data spaces. These will serve as a resource for the Spanish language, contributing to developing language technology in our language, and as an asset for open science because it implies sharing data and making it accessible and reusable.

In turn, new terminology neologisms continue to emerge. Scientific innovations and discoveries generate new words that must be added to Spanish terminology collections and validated by experts, given linguistic approval, and converted into machine-readable formats. To detect emerging terminological neologisms, a terminology extraction algorithm (KeyCare) has been developed within the TeresIA framework, which

can detect them from scientific text corpora. This new terminological data will be incorporated into linguistic data spaces, thereby generating the dynamic of producing open Spanish-language information that is available for reuse by different types of language professionals, companies, and language technology developers. Language technologies require infrastructure and data. The EU has put its efforts into creating the former. Researchers, technicians, professionals and institutions must focus on generating large data sets in Spanish.

6. Conclusion

SSince language technologies are omnipresent nowadays and require linguistic resources in all languages, and over 80% of all artificial intelligence is based on language, it becomes imperative to work with and on Spanish linguistic resources if language strength in the digital environment is to be enhanced, thereby contributing to multilingualism.

The data presented in this chapter leads to the general conclusion that Spanish linguistic data is well covered in both data spaces and terminology databases. However, this coverage varies in terms of both the types of resources and the infrastructure. While Spanish carries significant weight in ELG, the presence of resources in our language is scarce in CLARIN. Furthermore, in ELG, some types of linguistic resources in Spanish predominate over others. The existence of terminology resources could be improved, which will be attempted through TeresIA's data conversion project.

On the one hand, existing infrastructures and data spaces, including national ones such as INESdata, need to be better disseminated among terminology producers and academic institutions. On the other hand, linguistic data is required from all fields, but terminological data (i.e. data relating to scientific disciplines or professions) must be carefully produced and maintained if Spanish is to be promoted in science and as a working language in international organisations. Simultaneously, the digitisation of existing content in our language must be promoted

and a corpus of open academic texts in Spanish with Creative Commons licences must be created.

High-quality Spanish terminology that has been converted to web data formats, follows FAIR principles and can be linked (Linked Open Data) is becoming a critical line of work so that language technologies can also be developed in Spanish.

Acknowledgements

This chapter falls within the scope of TeresIA, a research project funded by the European Union's Next GenerationEU/PRTR programme through the Spanish Ministry for Digital Transformation and the Civil Service.

The author would like to thank the Spanish language community at the European Commission's Directorate-General for Translation for their collaboration in analysing Spanish in IATE.

Bibliographic references

Arias Rodríguez, I. (2024). *Green*: ¿verde o ecológico? *Puntoycoma*, 184, 5-6. https://ec.europa.eu/translation/spanish/magazine/es_magazine_es.htm

Céspedes, L., Kozlowski, D., Pradier, C., Sainte-Marie, M. H., Shokida, N. S., Benz, P., Poitras, C., Boudreau Ninkov, A., Ebrahimy, S., Ayeni, P., Filali, S., Li, B. & Larivière, V. (2025). Evaluating the linguistic coverage of OpenAlex: An assessment of metadata accuracy and completeness. *Journal of the Association for Information Science and Technology*, 1-12. https://doi.org/10.1002/asi.24979

Chen, E., Bullard, J. & Giustini, D. (2023). Automated indexing using NLM's Medical Text Indexer (MTI) compared to human indexing in Medline: a pilot study. *Journal of the Medical Library Association*, 111 (3), 684-694. https://doi.org/10.5195/jmla.2023.1588

European Language Grid (ELG) (2025). *European Language Grid. Catalogue*. ELG Consortium. https://live.european-language-grid.eu/catalogue/dashboard

Fišer, D., Lenardič, J. & Erjavec, T. (2018). CLARIN's key resource families. In N. Calzolari, K. Choukri, C. Cieri, T. Declerck, S. Goggi, K. Hasida, H. Isahara, B. Maegaard, J. Mariani, H. Mazom, A. Moreno, J. Odijk, S. Piperidis & T. Tokunaga (Eds.), *Proceedings of the Eleventh International Conference on Language Resources and Evaluation* (LREC 2018). European Language Resources Association (ELRA). https://aclanthology.org/L18-1210.pdf

Gaspari, F., Way, A., Dunne, J., Rehm, G., Piperidis, S. & Giagkou, M. (2021). *European Language Equality. Digital Language Equality (preliminary definition)* (Working Paper). ELE Consortium. https://european-language- equality.eu/wp-content/uploads/2021/05/ELE_Deliverable_D1_1.pdf

Giménez Toledo, E. (8 March 2024). El valor del multilingüismo en la comunicación de la ciencia. *The Conversation*. https://theconversation.com/el-valor-del-multilinguismo-en-la-comunicacion- de-la-ciencia-225363

Horváth, A., Wagner, C., Wrisley, D., Bernardi, J., Chadha, A., Garcia, D., Grallert, T., Ghosh, S., Ishida, Y., Kaye, A., Meden, A., Nagasaki, K., Palmer, D., Scheithauer, H., Świetlik, M., Tharsen, J. & Wang, Y. (2024). Multilinguality in Action: Towards Linguistic Diversity and Inclusion in Digital Humanities. *Magazén*, 5 (2), 179-196. http://doi.org/10.30687/mag/2724-3923/2024/02/001

IATE European Union terminology (n.d.). *IATE European Union terminology*. Accessed on March 2025. https://iate.europa.eu/home

National Library of Medicine (1 December 2021). MEDLINE 2022 Initiative: Transition to Automated Indexing. *NLM Technical Bulletin*, 443, e5. https://www.nlm.nih.gov/pubs/techbull/nd21/nd21_medline_2022.html

Moore, D. A. Q., Yaqub, O. & Sampat, B. N. (2024). Manual versus machine: How accurately does the Medical Text Indexer (MTI) classify different document types into disease areas? *PLOS One*, 19 (3). https://doi.org/10.1371/journal.pone.0297526

Spence, P. (2021). *Disrupting Digital Monolingualism: A report on multilingualism in digital theory and practice*. Language Acts and Worldmaking. https://zenodo.org/records/5743283

TeresIA (2024). *Informe del censo de recursos terminológicos de TeresIA*. https://proyectoteresia.org/wp-content/uploads/2024/09/TeresIA-ICRTT-1.0.0-20240405.pdf

THE HISPANIC VOTE IN THE 2024 US PRESIDENTIAL ELECTION

Francisco Javier Pueyo Mena
INSTITUTO CERVANTES OBSERVATORY AT HARVARD UNIVERSITY

1. Introduction

Regular elections are undeniably one of the fundamental pillars of democracy, ensuring that elected political candidates represent the will of the citizens called upon to vote. The transmission of political information (or misinformation) can also affect citizens' perceptions about the reality of the communities targeted by the media during elections. The latest US presidential elections, on 5 November 2024, are no exception. During the run-up to the election and after the results were announced, the Hispanic vote was at the centre of media attention. Media coverage in the United States, Spanish-speaking Latin America and Spain provided plenty of examples of this:

> Perhaps most surprising was Trump's 14-point increase among voters who identified as Hispanic, according to an exit poll conducted by Edison Research. Some 46% of self-identified Hispanic voters picked Trump, up from 32% in the 2020 election (Lange, Erickson and Heath, 2024, Reuters)[1].

> Trump won an astonishing 46% of the Latino vote in this election. This is the highest figure ever obtained by a Republican, even higher than that of George W. Bush in 2004 (Montanaro, 2024, NPR).

1. Quotes taken from texts originally published in Spanish have been translated into English by the editorial team

> Trump made historic inroads with voters of colour, immigrants and Latinos (Brammer, 2024, Los Angeles Times).
>
> An exit poll analysis conducted by The Washington Post reveals that Latino men propelled Trump to the White House, with 54% voting for him and 44% for Harris (El Universal, 2024).
>
> Latino voters increased their support for candidate and now President-elect Donald Trump by 13 points in this election, to 45% of that electorate, breaking the record set by Republican presidential candidate George W. Bush in 2004 (El Tiempo, 2024).
>
> Trump gained the support of 45% of Latino voters nationwide, compared to 53% for Harris, according to NBC News exit polls. Trump's support is much greater than in 2020, when he won 32% compared to Joe Biden's 65% (Lugones, 2024, Clarín).
>
> Trump has swept to victory by a wider margin than in the 2020 election, reducing the Democratic lead on the Latino vote from 33 to 8 points. This has happened despite the tycoon's tough immigration agenda, who on Thursday reiterated his plan to launch a large-scale deportation programme (Herrerías, 2024, El País).
>
> The Hispanic vote: the electoral earthquake that threatens the future of the Democratic Party (Ansorena, 2024, ABC).

The message conveyed by the media worldwide, as can be seen from the examples cited above, is that the Hispanic vote has shifted "surprisingly" towards Republican candidate Donald Trump in this election and away from Democratic candidate Kamala Harris. This reversal breaks with the traditional voting behaviour of the Hispanic community for as long as records have existed. One key factor has shaped public opinion transmitted to citizens by the media, namely the widespread use by all newspapers, television stations, radio stations and digital

media outlets worldwide of a single exit poll, the National Election Pool (NEP)[2], as a source of information on Hispanic voting behaviour, both nationally and in the decisive swing states. No media outlet considered other important polls, such as the 2024 American Electorate Voter Poll (2024-AEVP) (African American Research Collaborative and BSP Research, 2024a, 2024b)[3], whose minority population data differ substantially from those provided by the aforementioned Edison Research study for the mainstream media. The potential methodological problems with the NEP survey in measuring voting trends in minority communities such as Hispanics, Indigenous peoples, Asians, and African Americans were not factored in, and the few who later drew attention to this factor were unlikely to influence the state of opinion that had already spread globally:

> Unfortunately, exit polls have a poor track record and tend to create misleading narratives on election night. In particular, early exit polls have a poor track record when it comes to measuring Black, Latino and younger voters (Rogers, 2024, The New York Times).
>
> Latino organisations such as UnidosUS, Voto Latino, La Brega y la Fuerza Fund and the Hispanic Federation have joined forces with other minority representatives, such as Asian American Advancing Justice and AARC, to expose how exit polls fail to

2. The NEP is a US media consortium created in 2003 to provide exit poll information for US elections. The sampling system yielded biased results in the 2004 and 2016 US presidential elections. As of 2018, Edison Research was hired by a consortium of companies including ABC News, CBS News, CNN, and NBC News to conduct exit polls. Since 2020, Reuters has used the NEP to report the results of US presidential elections. Fox News and the Associated Press were previously part of the NEP, but left in 2017 and have since used AP VoteCast data.

3. The 2024 American Electorate Voter Poll is a large-scale survey conducted among African American, Latino, Asian American, Native American, and white voters across the country (in all 50 states), including the key swing states that decided the 2024 presidential election. Scientists and polling experts from BSP Research, African American Research Collaborative, and Harvard University conducted the survey. The poll was conducted between 18 October 2024 and 4 November 2024 and included a sample of over 9,000 voters, of whom 3,750 were Hispanic.

> accurately represent minorities and misrepresent the reality of these communities (Caro, 2024, El País).
>
> Did the Latino vote give Trump the victory? A new study suggests not (Telemundo, 2024).

Therefore, a more balanced view should be taken, using both pre-election estimates and tracking of votes once cast by voters, as their methodologies, in contrast to those of the NEP, are more careful when assigning demographic and linguistic weights to minority communities:

> Our survey continues to offer a considerable advantage over national exit polls, where each racial group is weighted to match demographics for their specific group within each state. The national exit polls only provide state level, or national weights, and do not apply weights to racial groups. This has led to a socioeconomic bias in Black and Latino samples (African American Research Collaborative and BSP Research, 2024b).

A determining factor in the methodology used by the two main exit polls is the language in which the interviews were conducted. While 2024-AEVP offered respondents a wide choice of languages "in English, Spanish, Chinese, Korean, Tagalog or Vietnamese, at the discretion of the respondent" (African American Research Collaborative and BSP Research, 2024b), the NEP survey was conducted exclusively in English. As discussed later, the language spoken at home (English or Spanish) constitutes a significant variable in the voting preferences of Latino communities in the United States.

The data used in this chapter has been compiled from a comprehensive and detailed report developed by the Observatory of the Spanish Language and Hispanic Cultures in the United States of the Instituto Cervantes at Harvard University (Vega Hübner and Pueyo Mena, 2025). This study's findings show that, even though a progressive increase in Latino support for the Republican Party has been observed,

no surprising or decisive shift in the traditional voting behaviour of Latino communities has taken place. This trend follows the pattern seen in previous elections and mirrors that of other voter communities. When the general public shifts towards a particular political option (Republican or Democrat), a similar trend is also observed in Latino communities, particularly among those sectors of the Hispanic population that are more integrated into the country's media and linguistic culture. This concept refers to those who speak exclusively English at home and therefore receive political information through the same media and social networks as their fellow citizens in the majority community:

> Increased use of English-language news sources has been driven by a rise in the proportion of Hispanics who report getting their news exclusively in English. According to the 2021 survey, one-third (32%) of Hispanic adults claim this to be the case, compared to 22% in 2006 (Lopez and Gonzalez-Barreda, 2013).
>
> The Hispanic millennial generation uses English-language news sources more than older generations, with 91% in 2016 stating that they get at least some of their news in English, compared to 68% who said they consume at least some of their news in Spanish (Flores and Lopez, 2018).

The Latino population that speaks primarily Spanish at home receives information from a wider variety of sources, including both US media in Spanish and English and media from their countries of origin, which naturally report in Spanish:

> Latinos born abroad, by contrast, prefer Spanish-language news sources. In 2016, 89% reported getting at least some of their news in Spanish, and 70% said they got at least some of it in English (Flores and Lopez, 2018).

2. Size and geographical distribution of the Hispanic electorate

The Hispanic population is the largest minority in the United States. The US Census Bureau estimates that it exceeds 65 million and represents about 20% of the country's total population (US Census Bureau, 2023a, 2024). Therefore, the Hispanic vote is potentially decisive in determining the outcome of presidential elections.

The number of Hispanics eligible to vote in the 2024 elections was over 36 million (Krogstad et al., 2024). According to the Pew Research Center, some 36.2 million Hispanics were eligible to vote in 2024, representing 14.7% of the total electorate, which is an all-time high. In the 2020 elections, the figure was around 32.3 million[4] (Krogstad et al., 2024). An increase of nearly 4 million people is symptomatic of the growth of the Hispanic population, which in 2023 increased by 1.8% over the previous year (United States Census Bureau, 2024). This is mainly due to natural population growth (more Hispanic people are being born than are dying), although immigration continues to be a significant factor.

The Pew Research Center's estimates of the number of Hispanics eligible to vote in the 2020 and 2024 elections are consistent with the increase in the number of Hispanic citizens of voting age reported by the Census Bureau: 32.7 million in 2021 compared to 34 million in 2023 (US Census Bureau, 2021, 2023b).

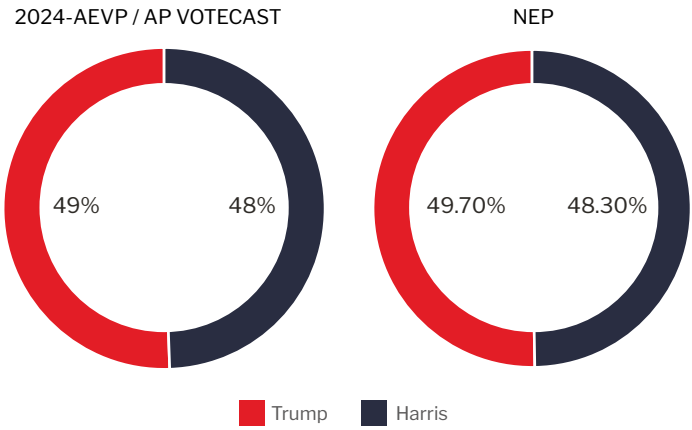
The states in which the population of Hispanic voters exceeds one million are California (8.5 million), Texas (6.5 million), Florida (3.5 million), New York (2.2 million), Arizona (1.3 million) and New Jersey (1.06 million). More than two-thirds of the total Hispanic electorate lives in these six states (Krogstad et al., 2024; United States Census Bureau, 2023b).

4. The Hispanic electorate increases at a rate of approximately 1.4 million people each year, and Hispanics accounted for 50% of the growth in the U.S. electorate from 2020 to 2024 (Krogstad *et al.*, 2024).

3. Overall results in the Hispanic community

The overall results of the 2024 US presidential election, without distinguishing between population groups, gave 49.8% support to Republican candidate Donald Trump and 48.3% to Democratic candidate Kamala Harris (Federal Election Commission, 2025). The exit poll results mentioned above accurately predicted these official results, within their statistical margins of error, as shown in Graph 1.

GRAPH 1 OVERALL RESULTS ACCORDING TO EXIT POLLS

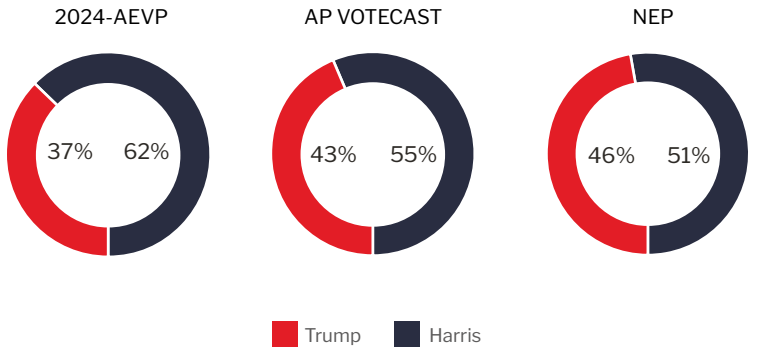


SOURCE: compiled by the author based on data from 2024-AEVP (African American Research Collaborative and BSP Research, 2024a), AP VoteCast[5] (Lohdi *et al.*, 2024) and NEP (CNN, 2024).

However, when focusing on the survey data when its results are broken down by voter communities, the discrepancies in measuring the vote in the Hispanic community could not be more striking.

5. AP VoteCast is a survey conducted by NORC at the University of Chicago for Fox News, *PBS NewsHour, The Wall Street Journal* and Associated Press. The poll was conducted between 28 October and 5 November 2024 with a sample of nearly 120,000 voters, 10% of whom were Hispanic (Associated Press, 2024).

GRAPH 2 OVERALL RESULTS (HISPANIC ELECTORATE)



SOURCE: compiled by the author based on data from 2024-AEVP (African American Research Collaborative and BSP Research, 2024a), AP VoteCast (Lohdi *et al.*, 2024) and NEP (CNN, 2024).

As shown in Graph 2, support for Kamala Harris ranges from 62% reported by 2024-AEVP to 51% reported by NEP (a difference of 11 percentage points). Support for Republican candidate Donald Trump varies between 37% reported by 2024-AEVP and 46% by NEP (a difference of 9 percentage points).

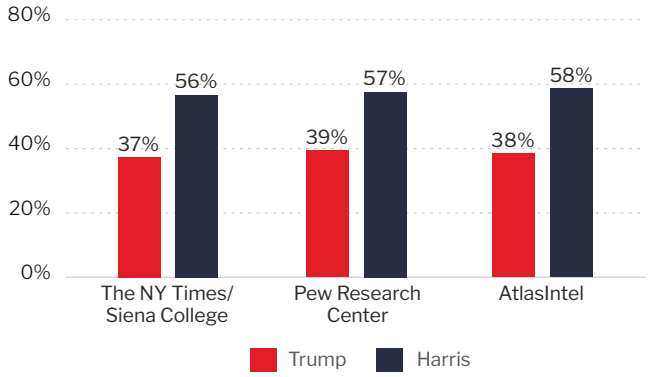
It is also worth noting the three main voting intention polls conducted in the weeks leading up to the election by the Pew Research Center[6], The New York Times/Siena College[7] and AtlasIntel[8], which show voting intentions halfway between the final results reported by 2024-AEVP and the data provided by NEP, as shown in Graph 3. The average of the three polls would be 57% support for the Democratic candidate and 39% for the Republican candidate.

6. This poll was conducted among US citizens across the country between 26 August and 2 September 2024. The sample included 691 Hispanics registered to vote (Lopez and Noe-Bustamante, 2024).

7. *The New York Times* and Siena College poll (*The New York Times*, 2024) took place between 29 September and 6 October 2024 with a sample of 2,516 voters from across the country, 310 of whom were Hispanic.

8. This national poll was conducted between 1 and 2 November 2024 with a sample of 2,463 voters, 12.1% of whom were Hispanic (AtlasIntel, 2024).

GRAPH 3 SUPPORT FROM THE HISPANIC ELECTORATE FOR CANDIDATES BASED ON PRELIMINARY VOTING INTENTION POLLS

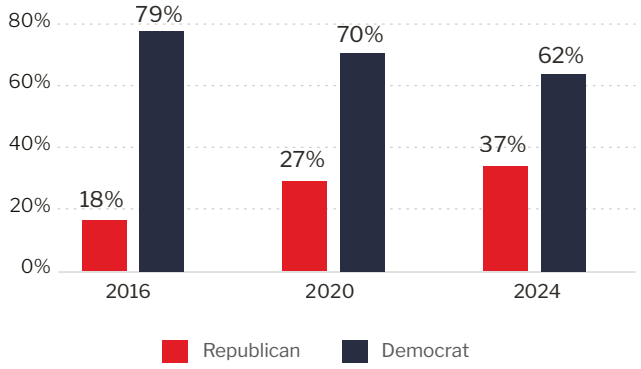


SOURCE: compiled by the author based on data from *The New York Times* / Siena College (*The New York Times*, 2024), Pew Research Center (Lopez and Noe-Bustamante, 2024) and AtlasIntel (2024).

The data from 2024-AEVP shows a steady increase in support for Donald Trump among the Latino community compared to the results of the same poll in the 2016 and 2020 elections. According to this source, Donald Trump's 37% support among voters represents a 10-point increase over the last four years, albeit still far from the historic 44% obtained by George W. Bush in 2004[9].

9. Subsequent studies by the Pew Research Center suggest that Hispanic voter support for the Republican Party in the 2004 elections was probably not as high as the 44% estimate based on exit polls indicated, as an over-representation of pro-Republican Cuban voters and other demographic discrepancies may have occurred (Suro, Fry and Passel, 2005).

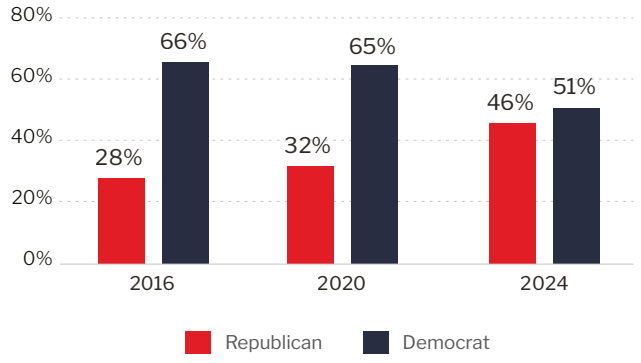
GRAPH 4 HISPANIC VOTING TRENDS BETWEEN 2016 AND 2024, ACCORDING TO THE AEVP



SOURCE: compiled by the author based on AEVP data.

Looking at the data provided by the NEP for the last three presidential elections, support for Trump would have risen from 28% in 2016 to 32% in 2020 and finally to 46% in 2024 (CNN, 2016, 2020, 2024). This would represent an increase of 14 percentage points between 2020 and 2024.

GRAPH 5 HISPANIC VOTING TRENDS BETWEEN 2016 AND 2024, ACCORDING TO THE NEP



SOURCE: compiled by author based on NEP data (CNN, 2016, 2020, 2024).

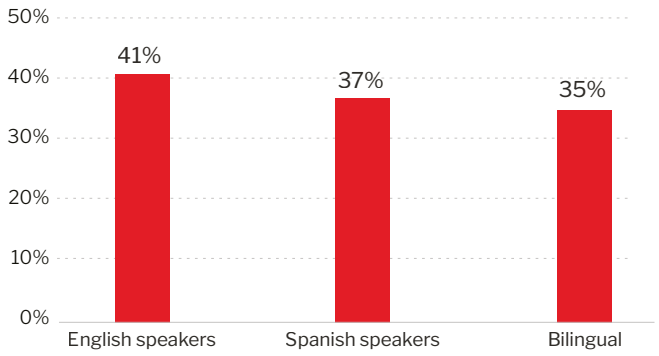
The 2024 result would therefore be a record, as it would exceed George W. Bush's historic 44% of the Hispanic vote in the 2004 presidential election by 2 points.

3.1 The Hispanic vote based on the language spoken at home

As mentioned in the introduction, the two exit polls differ in how they use (2024-AEVP) or do not use (NEP) Spanish during interviews, depending on voters' preferences. Similarly, differences arise between the two surveys in terms of whether to include the question of the primary language spoken at home. The 2024-AEVP survey alone considers this factor and classifies the sample into three groups, i.e., those who only speak Spanish at home, those who are bilingual in Spanish and English, and those who only use English at home.

Graph 6 illustrates that support for Donald Trump is higher among English-speaking Hispanics (41%) than among bilinguals (35%) and Spanish speakers (37%). A difference of 4 percentage points can be observed between English speakers and Spanish speakers, and 6 points between English-speaking households and bilingual households (African American Research Collaborative and BSP Research, 2024a).

GRAPH 6 HISPANIC VOTER SUPPORT FOR DONALD TRUMP BY LANGUAGE SPOKEN AT HOME



SOURCE: compiled by the author based on data from 2024-AEVP (African American Research Collaborative and BSP Research, 2024a).

All three groups have similar rates of first-time voters (13-14%), suggesting that language does not significantly influence electoral debut. However, voting history reveals that 67% of English speakers voted for the first time before 2018, compared to 60% of bilinguals and 64% of Spanish speakers, which suggests greater consolidation of voting habits among the former.

Spanish speakers and bilinguals are more likely to trust Democrats to address their priorities (56%), compared to 51% of English speakers. The latter favour Republicans (36%) to a greater extent than those who speak Spanish at home and bilinguals (both 32%), indicating a more conservative lean among English-speaking households.

The cost of living and inflation, followed by the economy and employment, are the most important issues for all three groups, while health concerns Spanish speakers (32%) more than bilinguals (26%) and English speakers (20%).

The most notable difference is in the priority given to reproductive rights: 29% among English speakers, 19% among bilinguals and 15% among Spanish speakers.

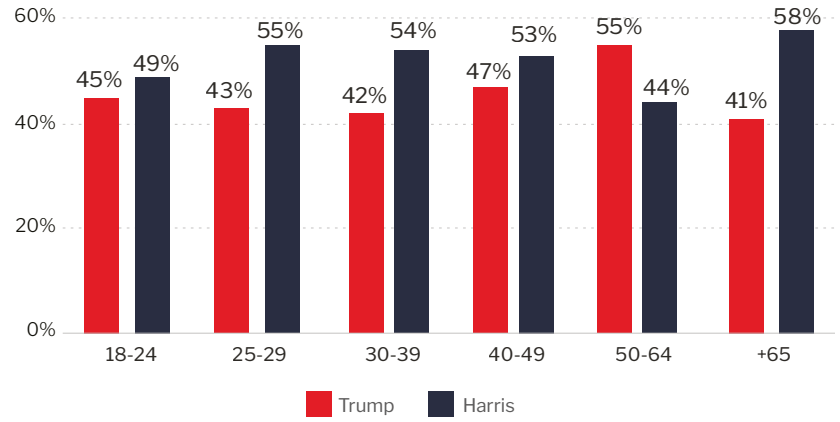
Regarding economic and social policies, support for housing construction and subsidies is high in all groups (from 85% to 91%), although English speakers show greater opposition (15%). On immigration reform, backing for undocumented immigrant regularisation is higher among Spanish speakers (92%) and bilinguals (90%) than among English speakers (85%).

Based on this data, it becomes apparent that the language spoken at home reflects social differences that influence electoral stance. English speakers are more critical of certain social measures and trust Democrats less, while Spanish speakers and bilinguals tend to prioritise inclusive economic and social policies. These observations underscore how important it is to consider language differences when formulating political strategies aimed at Hispanic voters, and when analysing the voting behaviour of the diverse Hispanic community.

3.2 The Hispanic vote by age

Data from the NEP poll (CNN, 2024), which divides participants into six age groups, shows that support for Trump was particularly high among Hispanics aged 50 to 64. Graph 7 shows that more Hispanics in this age bracket supported the Republican candidate than voted for his opponent (55% for Trump; 44% for Harris).

GRAPH 7 HISPANIC VOTER SUPPORT FOR THE TWO MAIN PRESIDENTIAL CANDIDATES BY AGE, ACCORDING TO THE NEP POLL

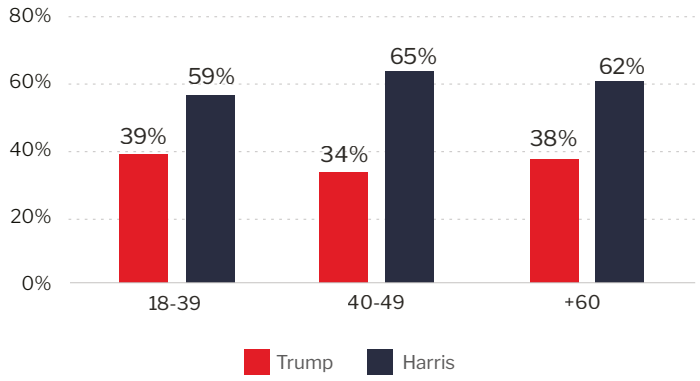


SOURCE: compiled by the author based on data from NEP (CNN, 2024).

Furthermore, this poll reveals greater support for the Republican candidate among 18-24 year olds (45% in favour of Trump) than among 25-29 year olds (43%), although in both cases Harris enjoys majority support. This may suggest a greater tendency among younger Hispanic voters, especially males, to vote Republican, as discussed in the following sections.

However, looking at the 2024-AEVP poll (African American Research Collaborative and BSP Research, 2024a), which also provides data on the Hispanic vote by age, substantial differences in the reported data are once again apparent. Graph 8 shows that, according to this survey, there is no Republican majority in any of the three age groups studied:

GRAPH 8 HISPANIC VOTER SUPPORT FOR THE TWO MAIN PRESIDENTIAL CANDIDATES BY AGE, ACCORDING TO THE 2024-AEVP POLL



SOURCE: compiled by the author based on data from 2024-AEVP (African American Research Collaborative and BSP Research, 2024a).

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

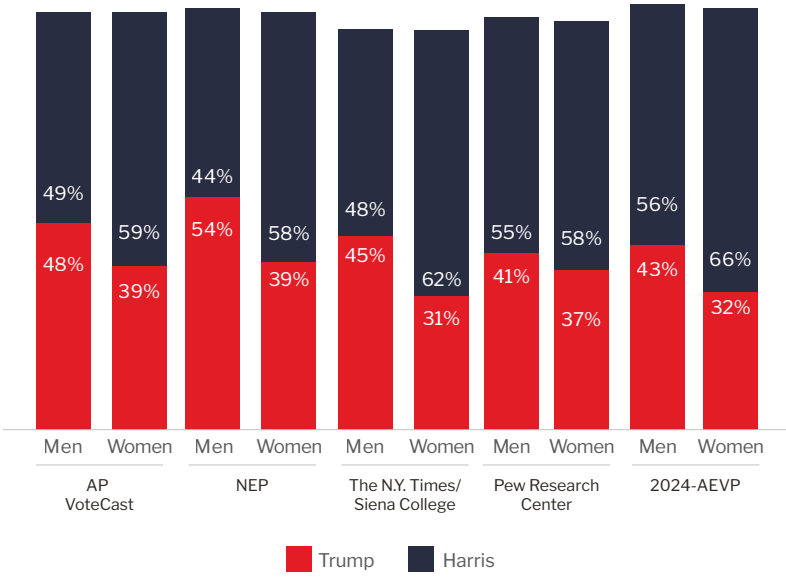
3.3 The Hispanic vote by gender

According to data from the Pew Research Center, Harris is the preferred candidate among both Hispanic women (58%) and Hispanic men (55%). Donald Trump, meanwhile, is the favourite candidate of 41% of men and 37% of women. These figures show greater support for Harris among women than among men (Lopez and Noe-Bustamante, 2024).

This difference between the electoral trends of Hispanic men and women is even more pronounced when examining data from other polls reviewed for this report. All of them indicate that support for the Republican candidate is significantly higher among men than among women. Graph 9 shows that the percentage of Hispanic men who support Trump always exceeds 40%, while the percentage of women in favour

of the Republican candidate never reaches that figure, ranging between 31% and 39%. Likewise, the percentage of men in favour of Harris only exceeds 55% in one of the estimates, and by just one percentage point, while the Hispanic female vote always hovers around 60% and even exceeds it according to two of the polls.

GRAPH 9 HISPANIC VOTER SUPPORT FOR THE TWO MAIN PRESIDENTIAL CANDIDATES BY GENDER ACCORDING TO DIFFERENT POLLS



SOURCE: compiled by the author based on data from AP VoteCast (Lohdi *et al.*, 2024), NEP (CNN, 2024), *The New York Times*/Siena College (*The New York Times*, 2024), Pew Research Center (Lopez and Noe-Bustamante, 2024) and 2024-AEVP (African American Research Collaborative and BSP Research, 2024a).

According to the NEP poll, Trump would even have won the majority of the Hispanic male vote with 54%. The 2024-AEVP data again reveals a much different reality, with an 11-percentage-point discrepancy: Trump would not have won the majority of the Hispanic male vote, receiving 43% compared to the 56% of Hispanic men who supported Harris. Although women undoubtedly show greater support for Harris in both polls, the percentages of that support also differ by 8 points (from 58% reported by the NEP to 66% reported by 2024-AEVP).

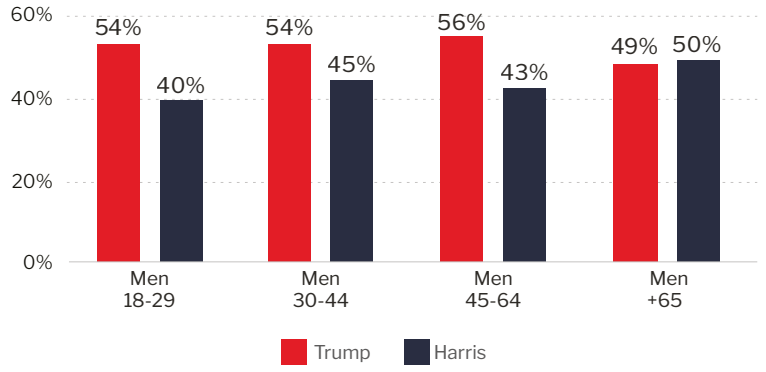
It is interesting to note that women's participation rate in elections is traditionally higher than that of men. This extends to all racial and ethnic subgroups, as well as to the Hispanic minority. In the 2020 elections, women accounted for 53.7% of Hispanic voters, and men for 46.3% (Fabina and Scherer, 2022).

3.4 The Hispanic vote by age and gender

Both the NEP and the 2024-AEVP provide data on voting by age group and gender. They show that female voters favour the Democratic Party more than male voters in all age groups (CNN, 2024; African American Research Collaborative and BSP Research, 2024). However, once again, substantial discrepancies are found between the two polls in their percentages.

The NEP data would indicate majority support for Trump among men in three of the four age groups: 18-29, 30-44 and 45-65 (54%, 54% and 56% respectively), as shown in Graph 10. Only men over 65 would have voted 50% for Harris and 49% for Trump.

GRAPH 10 HISPANIC MALE SUPPORT FOR THE TWO MAIN PRESIDENTIAL CANDIDATES BY AGE AND GENDER, ACCORDING TO THE NEP POLL

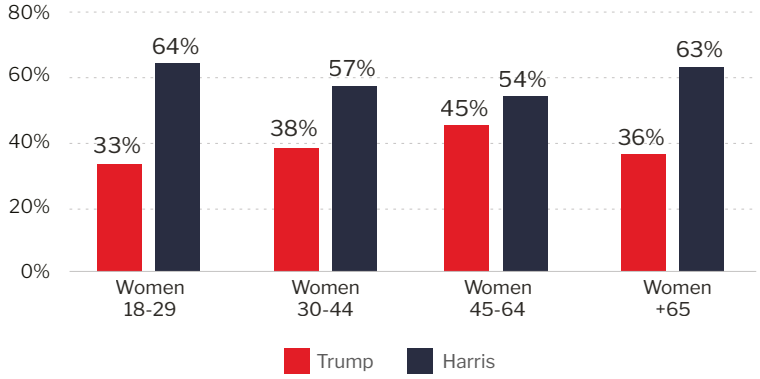


SOURCE: compiled by the author based on data from NEP (CNN, 2024).

As for women, Graph 11 indicates that Harris would have won the majority of the female vote in all age categories; support for the Republican

candidate would only have exceeded 40% among Hispanic women aged 45 to 64 (45% in favour of Trump and 54% in favour of Harris).

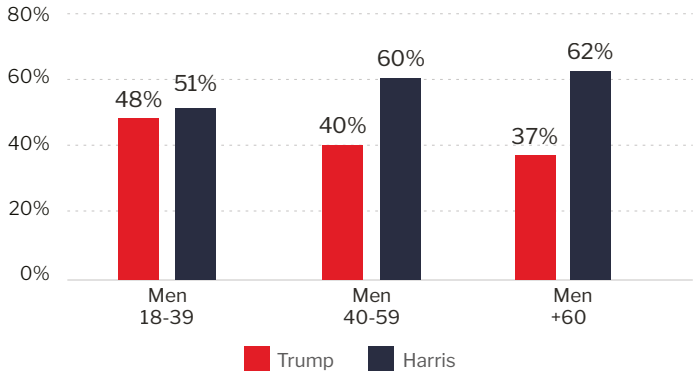
GRAPH 11 HISPANIC FEMALE SUPPORT FOR THE TWO MAIN PRESIDENTIAL CANDIDATES BY AGE AND GENDER, ACCORDING TO THE NEP POLL



SOURCE: compiled by the author based on data from NEP (CNN, 2024).

Furthermore, the 2024-AEVP survey, as shown in Graph 12, reports that no age group among men would have supported Trump by a majority. The greatest support for the Republican candidate would be among men aged 18 to 39 (48% for Trump and 51% for Harris).

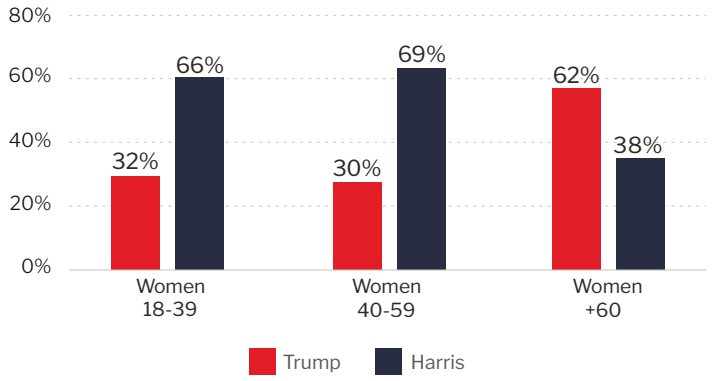
GRAPH 12 HISPANIC MALE SUPPORT FOR THE TWO MAIN PRESIDENTIAL CANDIDATES BY AGE AND GENDER, ACCORDING TO THE 2024-AEVP POLL



SOURCE: compiled by the author based on data from 2024-AEVP (African American Research Collaborative and BSP Research, 2024a).

Female support for Kamala Harris across all age brackets consistently exceeds that of men, particularly among younger Hispanic women, where there are differences of between 9 and 15 percentage points in their support for Harris compared to men, as shown in Graph 13.

GRAPH 13 HISPANIC FEMALE SUPPORT FOR THE TWO MAIN PRESIDENTIAL CANDIDATES BY AGE AND GENDER, ACCORDING TO THE 2024-AEVP POLL



SOURCE: compiled by the author based on data from 2024-AEVP (African American Research Collaborative and BSP Research, 2024a).

While the figures vary significantly from poll to poll, they suggest what could be a trend, with Republican voting among young Hispanics, who are traditionally and predominantly Democratic, appearing to be on the rise, especially among young men.

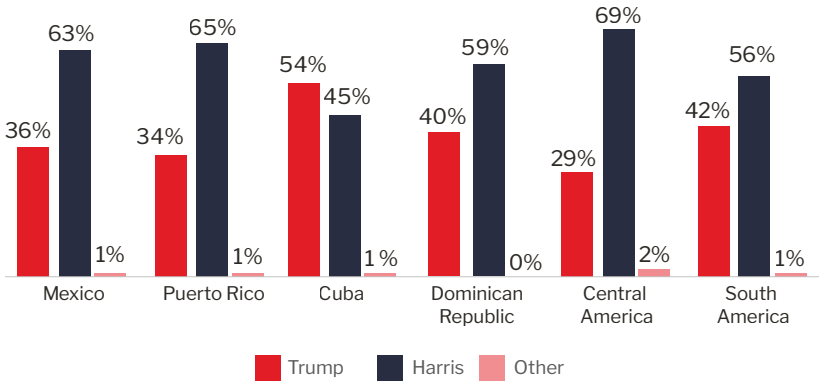
3.5 Voting bias based on voters' ancestry or origin

EIn the introduction, the Hispanic community was described as diverse, reflecting several sociodemographic variables. A key factor is undoubtedly the origin or ancestry of the different Hispanic peoples who constitute this group. Reporting on the Hispanic vote in the 2020 presidential election for The Hispanic Council, Güenechea Rodríguez and Ureña Uceda (2020) warn against the risk of viewing the Hispanic population as a "monolithic community". They point out that the different

cultural backgrounds shared by people from the same country of origin or ancestry are reflected in their electoral attitudes.

The only national survey that provides data on the electoral behaviour of the sample by territory of origin or ancestry within the Hispanic domain is the 2024-AEVP, which divides respondents into six subgroups: Mexicans, Puerto Ricans, Cubans, Dominicans, Central Americans and South Americans.

GRAPH 14 HISPANIC VOTER SUPPORT FOR THE MAIN PRESIDENTIAL CANDIDATES BY ORIGIN OR ANCESTRY



SOURCE: compiled by the author based on data from 2024-AEVP (African American Research Collaborative and BSP Research, 2024a).

The most remarkable data to be drawn from the graph refers to the vote among Hispanics of Cuban origin. Traditionally, most Cubans vote for the Republican Party. This community carries more sway in the state of Florida, where Cubans are the largest group within the Hispanic community (United States Census Bureau, 2023a). The following section on swing states shows that in 2024, the strength of this community in Florida tipped the overall Hispanic vote for Trump.

The 2024-AEVP survey also separates those born in the United States from those born in Puerto Rico or other countries. The data here shows greater support for the Republican Party among voters born in

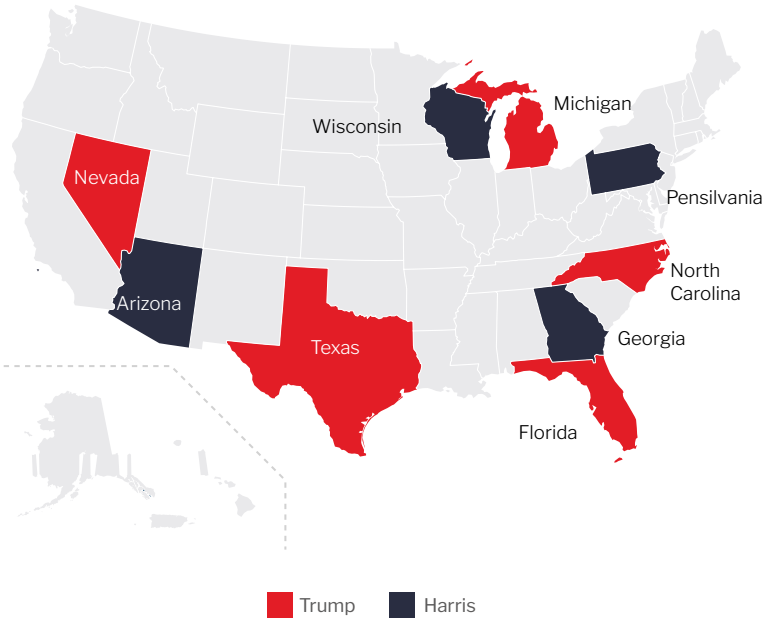
the United States than among those born in the U.S. insular area or other countries: 38% of those born in the United States support Trump, while 60% support Harris. In contrast, only 34% of Hispanics not born in the United States vote for Trump, while 65% support Harris.

4. Results in the seven swing states and those with the largest Hispanic populations

Only two polls offer data on the Latino vote in the seven swing states of 2024: the NEP poll and the 2024-AEVP poll. Similarly, both examine Hispanic voting behaviour in some states where the percentage of the Hispanic population is decisive for the outcome of the election. According to the population census, more than 15 million Hispanics reside in California (25% of the entire Latino population in the country), 12 million Hispanics live in the state of Texas (representing 19% of the total Latino population), and Florida has more than 6 million Latino residents, or nearly 10% of all Hispanics in the United States. (United States Census Bureau, 2023c). These three states together account for more than 50% of the entire Latino population in the United States. Therefore, to understand the Hispanic vote in 2024, research into their behaviour in these elections is essential.

The results of the two polls could not be more divergent. The NEP poll gives Republican Donald Trump the majority of the states studied: four of the seven swing states (Michigan, North Carolina, Nevada and Wisconsin) and two of those with the largest Hispanic populations (Florida and Texas), which is why the idea has spread that the Hispanic vote was a decisive factor in his victory. The map in Graph 15, based on the data in Table 1, provides a visual representation of the image conveyed by the media globally.

GRAPH 15 CANDIDATE WITH THE MOST HISPANIC VOTES IN THE SEVEN SWING STATES AND IN TWO HIGH HISPANIC POPULATION STATES, ACCORDING TO NEP



SUPPORT PERCENTAGE

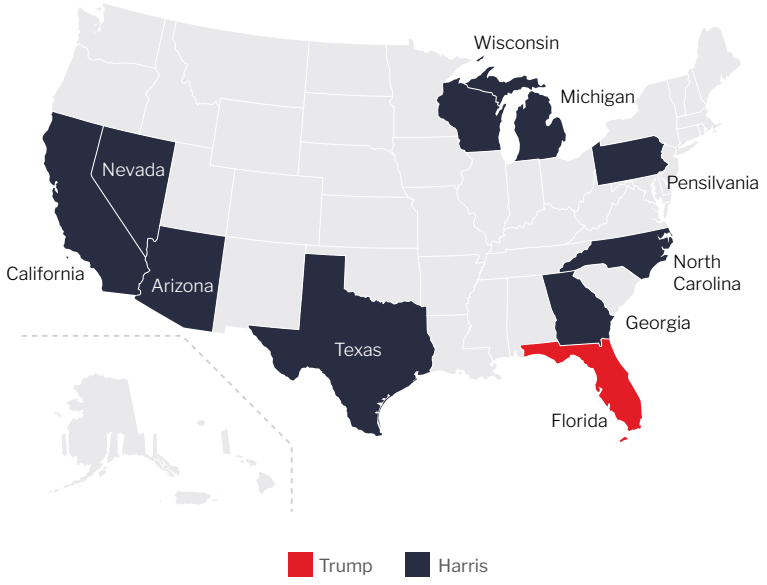
Swing states	Kamala Harris	Donald Trump
Arizona	54%	44%
Georgia	59%	41%
Michigan	37%	58%
North Carolina	49%	50%
Nevada	47%	49%
Pensilvania	59%	41%
Wisconsin	60%	38%

States with the largest Hispanic populations	Kamala Harris	Donald Trump
Florida	42%	58%
Texas	45%	55%

SOURCE: compiled by the author based on data from NEP (CNN, 2024)[10].

However, the 2024-AEVP poll indicates that the Latino population did not vote overwhelmingly for Donald Trump in any of the seven swing states or in two of the states with the largest Hispanic populations (California and Texas). Only Florida, with a high percentage of Cuban-born residents, opted for Trump, as has traditionally been the case. The map in Graph 16 provides a much different visualisation than the previous one, which has never been made public.

GRAPH 16 CANDIDATE WITH THE MOST HISPANIC VOTES IN THE SEVEN SWING STATES AND IN THE THREE STATES WITH THE LARGEST HISPANIC POPULATIONS, ACCORDING TO 2024-AEVP



10. The NEP survey does not collect data on the Hispanic vote in the state of California.

SUPPORT PERCENTAGE

Swing states	Kamala Harris	Donald Trump
Arizona	66%	33%
Georgia	64%	35%
Michigan	64%	33%
North Carolina	63%	35%
Nevada	64%	35%
Pensilvania	70%	28%
Wisconsin	64%	34%

States with the largest Hispanic populations	Kamala Harris	Donald Trump
California	66%	33%
Florida	43%	56%
Texas	60%	38%

SOURCE: compiled by the author based on data from 2024-AEVP (African American Research Collaborative and BSP Research, 2024a).

4. Conclusion

The apparent discrepancy between the NEP's exit poll (widely used by the global media), earlier polls, and other less publicised exit polls when analysing the Hispanic vote is troubling. It highlights the media's lack of interest in accurately reflecting the heterogeneity and diversity of minorities in the United States. The media ought to be extremely scrupulous, checking various sources, to show the social and cultural reality of these minority communities, as well as when portraying their political image in a society as polarised as the US. The accurate representation of the Hispanic population in the political arena should have received particular attention during this election cycle. Indeed, this community

has received unusually intense media attention, both before and after the elections, in relation to highly controversial issues such as illegal immigration, mass deportations, language use, and even an alleged increase in crime.

The use of methodological and linguistic biases and misinformation has created a state of public opinion that does not reflect the complex reality of the Hispanic vote's direction and evolution in the US presidential elections. This analysis underlines why institutions like the Instituto Cervantes' Observatory of the Spanish Language and Hispanic Cultures at Harvard University need to continue dedicating resources to rigorous analysis, using verified data, and basing their research on the demographic, cultural, linguistic, and political diversity of Hispanic communities.

The Hispanic vote continues to follow the same trend observed in previous election cycles, with the Hispanic community leaning mostly toward the Democratic Party, although there continues to be a growing trend of support for the Republican Party.

It is crucial to analyse how the distribution of the Latino vote varies according to language use at home, while also considering the age of voters. First-generation Hispanic immigrants mostly use Spanish, while second-generation immigrants are bilingual. Yet, for the third generation and beyond, the predominant trend is towards monolingualism in English, leading these generations to obtain political information from the same sources as most of their fellow US citizens. As mentioned above, bilingual and Spanish-dominant voters tend to lean more strongly toward the Democratic vote, while among monolingual English-speaking Hispanics, the Republican vote is more pronounced.

The evidence gathered in this study suggests that, although the Hispanic vote in the United States remains largely Democratic, significant variations exist based on age and gender. Generationally, the strongest support for the Republican Party is among those over 50, whereas the most notable redistribution of votes is observed among young people.

Furthermore, there is a gender gap, with Hispanic men showing a greater preference for the Republican candidate than women, who maintain strong support for Harris. When both factors are considered together, the shift towards the Republican vote is more pronounced among young men than women of the same generation. Although these trends require further analysis, they could have significant implications for future presidential elections.

Bibliographic references

African American Research Collaborative & BSP Research (2024a). *2024 American electorate voter poll: Latino voters*. https://2024electionpoll.us/wp-content/uploads/2024/11/7.-Latino-crosstab.pdf

African American Research Collaborative & BSP Research (2024b). *2024 American electorate voter poll. Methodology*. https://2024electionpoll.us/methodology/

Ansorena, J. (10 November 2024). El voto hispano: el seísmo electoral que amenaza el futuro del Partido Demócrata. *ABC*. https://www.abc.es/internacional/voto-hispano-seismo-electoral-amenaza-futuro-partido-20241109202349-nt.html

Associated Press (4 January 2024). *How we survey American voters with AP VoteCast*. https://apnews.com/ap-votecast-faq

AtlasIntel. (2 November 2024). *US Presidential Election 2024*. https://www.atlasintel.org/polls/general-release-polls

Brammer, J. P. (8 November 2024). Latinx files: Reckoning with the 2024 election results. *Los Angeles Times*. https://www.latimes.com/delos/newsletter/2024-11-08/latinos-trump-reckoning-deportation-right-shift-latinx-files

Caro, P. (12 November 2024). Las organizaciones latinas contradicen los datos sobre el apoyo hispano a Trump. *El País*. https://elpais.com/us/2024-11-13/las-organizaciones-latinas-contradicen-los-datos-sobre-el-apoyo-hispano-a-trump.html

CNN (2016). *Exit polls*. https://www.cnn.com/election/2016/results/exit-polls

CNN (2020). *Exit polls*. https://www.cnn.com/election/2020/exit-polls/president/national-results/7

CNN (2024). *Election 2024: Exit polls*. https://www.cnn.com/election/2024/exit-polls/national-results/general/president/7

El Tiempo (7 November 2024). Nuevo hito: Donald Trump rompe el récord de apoyo latino alcanzado por George W. Bush en las elecciones de Estados Unidos. https://www.eltiempo.com/mundo/eeuu-y-canada/trump-supero-el-record-de-w-bush-en-apoyo-entre-los-votantes-latinos-en-las-elecciones-3397482

El Universal (7 November 2024). Hispanos y mujeres, clave en el triunfo... de Trump. https://www.eluniversal.com.mx/mundo/hispanos-y-mujeres-clave-en-el-triunfo-de-trump/

Fabina, J. & Z. Scherer (February 2022). Voting and Registration in the Election of November 2020, *Current Population Survey Reports*, P20-585. U.S. Census Bureau. https://www.census.gov/library/publications/2022/demo/p20-585.html

Federal Election Commission (16 January 2025). *Official 2024 Presidential General Election Results*. https://www.fec.gov/resources/cms-content/documents/2024presgeresults.pdf

Flores, A. & Lopez, M. H. (11 January 2018). *Among U.S. Latinos, the internet now rivals television as a source for news*. Pew Research Center. https://www.pewresearch.org/short-reads/2018/01/11/among-u-s-latinos-the-internet-now-rivals-television-as-a-source-for-news/

Güenechea Rodríguez, J. I. & Ureña Uceda, D. (2020). *El papel del voto hispano en las elecciones presidenciales de EEUU 2020*. The Hispanic Council. https://www.hispanicouncil.org/wp-content/uploads/THC-Voto-hispano-Final.pdf

Herrerías, A. Á. (8 November 2024). El gran giro de los latinos hacia Trump: del auge en los estados clave al crecimiento entre hombres jóvenes. *El País*. https://elpais.com/us/2024-11-08/el-gran-giro-de-los-latinos-hacia-trump-del-auge-en-los-estados-clave-al-crecimiento-entre-hombres-jovenes.html

Krogstad, J. M., Passel, J. S., Budiman, A. & Natarajan, A. (10 January 2024). *Key facts about Hispanic eligible voters in 2024*. Pew Research Center. https://www.pewresearch.org/short-reads/2024/01/10/key-facts-about-hispanic-eligible-voters-in-2024/

Lange, J., Erickson, B. & Heath, B. (7 November 2024). Trump's return to power fueled by Hispanic working-class voter support. *Reuters*. https://www.reuters.com/world/us/trumps-return-power-fueled-by-hispanic-working-class-voter-support-2024-11-06/

Lodhi, H., Cheng, S., Kaufmann, P., Barria Urenda, P. & Fox, E. J. (2024). *AP VoteCast: How America voted in 2024*. The Associated Press. https://apnews.com/projects/election-results-2024/votecast/

Lopez, M. H. & Gonzalez-Barrera, A. (23 July 2013). *A growing share of Latinos get their news in English*. Pew Research Center. https://www.pewresearch.org/race-and-ethnicity/2013/07/23/a-growing-share-of-latinos-get-their-news-in-english/

Lopez, M. H. & Noe-Bustamante, L. (24 September 2024). *In Tight U.S. Presidential Race, Latino Voters' Preferences Mirror 2020*. Pew Research Center. https://www.pewresearch.org/race-and-ethnicity/2024/09/24/in-tight-u-s-presidential-race-latino-voters-preferences-mirror-2020/

Lugones, P. (6 November 2024). ¿Por qué ganó Donald Trump? Algunas claves del contundente giro a la derecha en Estados Unidos. *Clarín*. https://www.clarin.com/mundo/gano-donald-trump-claves-contundente-giro-derecha-estados-unidos_0_jG4ar8sTsG.html

Montanaro, D. (8 November 2024). Why Trump won – 9 takeaways from the 2024 election. *NPR*. https://www.npr.org/2024/11/08/g-s1-33274/2024-election-how-trump-won-takeaways

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

Suro, R., Fry, R. & Passel, J. (2005). *Hispanics and the 2004 Election: Population, Electorate and Voters*. Pew Research Center. https://www.pewresearch.org/race-and-ethnicity/2005/06/27/hispanics-and-the-2004-election/

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

The New York Times (13 October 2024). Cross-Tabs: October 2024 Times/Siena Poll of the Hispanic Likely Electorate. https://www.nytimes.com/interactive/2024/10/13/us/elections/times-siena-poll-hispanic-likely-electorate.html

United States Census Bureau (2021). *American Community Survey. S2901/Citizen, Voting-Age Population by Selected Characteristics*. https://data.census.gov/table/ACSST1Y2021.S2901?q=S2901

United States Census Bureau (2023a). *American Community Survey. DP05/ACS Demographic and Housing Estimates*. https://data.census.gov/table/ACSDP1Y2023.DP05

United States Census Bureau (2023b). *American Community Survey. S2901/Citizen, Voting-Age Population by Selected Characteristics*. https://data.census.gov/table/ACSST1Y2023.S2901?q=S2901

United States Census Bureau (2023c). *American Community Survey. S0201/Selected Population Profile in the United States*. https://data.census.gov/table/ACSSPP1Y2023.S0201?t=400&g=010XX00US_040XX00US06,12,48

United States Census Bureau (27 June 2024). *New Estimates Highlight Differences in Growth Between the U.S. Hispanic and Non-Hispanic Populations, CB24-109*. https://www.census.gov/newsroom/press-releases/2024/population-estimates-characteristics.html

Vega Hübner, B. & Pueyo Mena, F. J. (Dirs.). (2025). *The Hispanic Vote in the 2024 U.S. Presidential Elections*. Estudios del Observatorio/Observatory Studies, 95. Instituto Cervantes at the Faculty of Arts and Sciences of Harvard University. https://cervantesobservatorio.fas.harvard.edu/es/informes/el-voto-hispano-en-las-elecciones-presidenciales-de-los-estados-unidos-de-2024

III. CLOSING REMARKS AND RECOMMENDATIONS

The reality of the entire international dimension of the Spanish language in 2025 includes data that reflects its significance in demography, human development, culture, and technology. The excess of 500 million native speakers and 630 million potential speakers points to a vigorous and thriving linguistic community within a diverse landscape.

As in previous reports, the data compiled here provides a clear picture of the reality within the Spanish-speaking community and highlights areas for growth and improvement. As such, the conclusions of this report highlight certain data and outline courses of action that should be considered by authorities, organisations and individuals involved in activities that affect languages and their speakers:

a. The creation and maintenance of demolinguistic data and the collection of information on the demographic, geographical and social dynamics of Spanish and its communities of speakers, are tasks that should be supported and reinforced. This becomes even more evident when considering the numerous Spanish speakers who, to some extent, represent hidden populations, ranging from those learning Spanish through informal education and digital spaces to the dynamics of second- and third-generation Spanish-speaking migrants (heritage speakers).

b. The Spanish-speaking community is steadily growing within Hispanic countries themselves, keeping pace with demographics and attracting speakers of other languages. In view of this situation, countries need to offer high-quality Spanish language teaching that is compatible with bilingual or multilingual education programmes and geared towards sustainable language policies regarding linguistic and cultural diversity.

c. The development of the limited competence Spanish speaker group (the most significant of the groups considered) requires the creation and deployment of language maintenance programmes for the migrant population and ongoing support for lifelong Spanish language learning in communities where it is a minority language, particularly in the United States, Europe, and Brazil.

d. In the coming decades, the relative weight of speakers with limited Spanish competence will increase within the Spanish-speaking community. This calls for strong support for language education to improve Spanish language learning and usage, especially outside the borders of territories where it is the official language. The goal is to maintain knowledge and usage of Spanish among heritage speakers and to develop speakers with higher proficiency.

e. Spanish as a foreign language will hinge on its presence in established education systems (like in the US, Brazil, Western Europe, and West and Central Africa) and its institutionalisation or consolidation in new educational spaces. This requires Spanish language diplomacy guided by two strategies. The first is aimed at sustaining, improving and expanding existing structures and agreements, while the second is geared towards creating new structures and bilateral agreements specific to each region.

f. The Spanish language's international outlook reveals the need to strengthen the supply of Spanish courses to make it possible to study the language wherever demand exists. To achieve this, international education systems must be readjusted through training and refresher programmes for Spanish teachers.

g. The technological development within the Spanish language requires tools designed specifically for Spanish and implemented from Spanish to adapt better to its cultural, stylistic and thought idiosyncrasies.

h. The conclusions drawn from our complementary analyses are highly relevant to the sectors in question.

   - Regarding economics, investment must be made to promote the integration of Spanish-speaking communities into the digital economy (infrastructure, education, etc.), improve institutional quality, and safeguard the free movement of people, capital, and goods.

   - As for artificial intelligence, Spanish has a competitive advantage in terms of data and corpus availability. However, AI's current development still has a long way to go before it can be applied to different industrial fields, particularly in relation to its social, stylistic and geographical varieties.

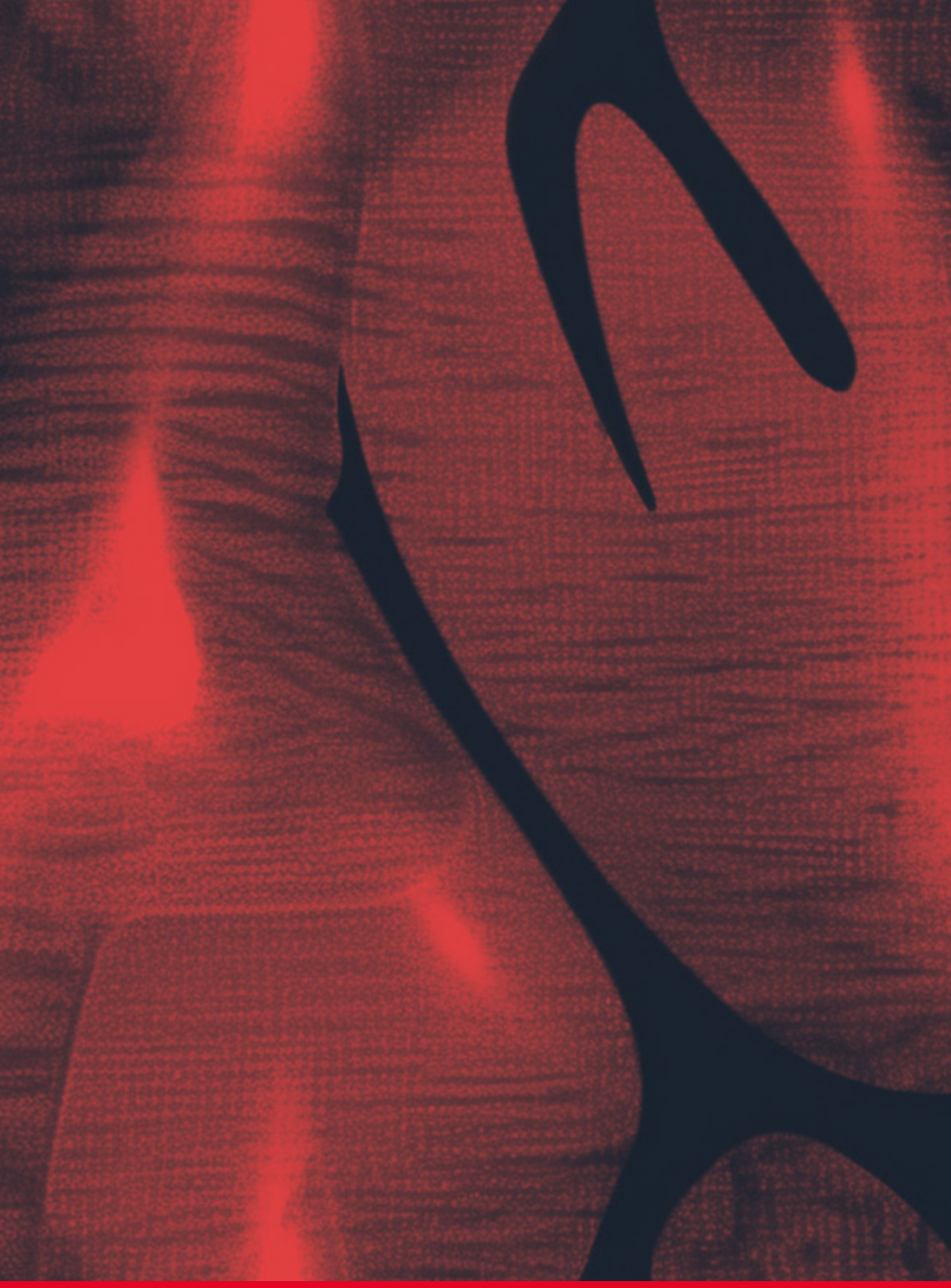
   - As far as language technologies are concerned, the Spanish language has the most resources in the European Language Grid catalogue, although its coverage is uneven in terms of both resource types and infrastructure.

   - As for the Hispanic vote in the last US elections, its distribution according to language use at home is found to be crucial, as well as the age of voters. First-generation Hispanic immigrants mostly use Spanish, while second-generation immigrants are bilingual. Yet, for the third generation and beyond, the predominant trend is towards monolingualism in English, leading these generations to obtain political information from the same sources as most of their fellow US citizens.

i. Diplomatic efforts remain indispensable in promoting the institutional and public use of Spanish, especially in international forums, with a view to its official recognition. Many negotiations must be held to ensure that Spanish is available to those who want or need it in public education systems, international organisations, initiatives and procedures, as well as diplomatic missions of all kinds.

j. The creation of cultural and scientific content in Spanish, disseminated through specialised channels and the media, plays a decisive role in the international appreciation of the Spanish language and the nations that use it to communicate.

All these tasks must be appropriately contextualised and adapted to the linguistic repertoire and communication needs of each country or territory. Furthermore, Spanish must take a responsible approach to diversity in its use and teaching, promoting the cultural values embodied in this language in all its manifestations. As the Nicaraguan poet Rubén Darío said, "To you my language should not be foreign". Nor will it be, provided it is used decisively and competently to serve international communities.



Instituto
Cervantes