II Encuentro
Migraciones e
internacionalización
del español

Fuentes para el estudio de la migración hispanohablante en Europa

Campus de Madrid-Princesa
Lunes, 17 de octubre de 2022

Fuentes para el estudio de la migración hispanohablante en Europa

Héctor Álvarez Mella

Charlotte Blattner

Ana Gómez-Pavón Durán

Índice

1 Introducción 4

2 La documentación en demografía de las lenguas migratorias 5

2.1 El concepto de hispanohablante con trasfondo migratorio 5

2.2 Demolingüística y metodología cuantitativa 6

3 Fuentes directas 7

3.1 La lengua en los censos 7

3.2 La lengua en los registros poblacionales 10

3.3 La lengua en las encuestas y las encuestas lingüísticas 11

4 Fuentes indirectas 12

4.1 Indicadores sociales para la competencia lingüística 13

4.2 Oficina Europea de Estadística (Eurostat) 14

4.3 Instituto Nacional de Estadística de España (INE) . 15

5 Conclusiones 16

6 Referencias bibliográficas 17

Resumen

Las diferentes instituciones europeas responsables de la recopilación de datos demográficos otorgan igualmente diferente importancia a los datos de carácter demolingüístico, por lo que una investigación como la presente no está exenta de complicaciones. Sin embargo, combinando estadísticas y datos poblacionales de carácter lingüístico, demográfico y migratorio, ha sido posible dar respuesta a las principales preguntas de la demolingüística de la migración hispanohablante en Europa.

Palabras clave: Demolingüística, migración, lengua materna, hispanohablante, Europa, sociodemografía, lengua regional.

1 Introducción

Las instituciones nacionales responsables de la recopilación de datos demográficos en Europa ofrecen una gran variedad de fuentes que resultan muy útiles para los estudios demolingüísticos. Sin embargo, dichas fuentes no están exentas de limitaciones que pueden suponer un reto para la investigación. En las páginas que siguen los autores se proponen ofrecer una guía para el aprovechamiento de dichas fuentes, con la intención de facilitar el trabajo de la demografía de las lenguas.

Uno de los principales retos de la demolingüística es la falta de disponibilidad de datos y, cuando están disponibles, la dificultad de su comparabilidad derivada de la falta de consenso y concreción conceptual en los asuntos culturales recogidos en las estadísticas poblacionales (Loureda *et al.* 2021). Un problema añadido supone el hecho de que, con frecuencia, los datos utilizados en los estudios demolingüísticos no provienen de fuentes directamente lingüísticas, sino que se crean a partir de "acciones y procedimientos administrativos, tanto en el nivel local, como en el nacional o el internacional" (Moreno Fernández, en prensa).

Por otra parte, el estudio de la población migrante en Europa, de creciente presencia y relevancia, resulta especialmente complejo desde el punto de vista demolingüístico. Las estadísticas recogen únicamente información sobre aquellas personas que aparecen registradas en el sistema legal del país, por tanto no reflejan totalmente la realidad del panorama lingüístico europeo, y en muchos casos no incluyen a las segundas y terceras generaciones de migrantes, que generalmente reciben la herencia lingüística de sus antecesores. Además, la concepción monolingüe predominante en el norte global, unida a la dificultad de recopilar en las estadísticas la diversidad lingüística derivada de la movilidad internacional, se refleja con frecuencia en los censos, registros y encuestas. Todo esto, unido a la rápida lectura de los datos demográficos, suele llevar a interpretaciones sesgadas de la complejidad de los grupos humanos.

Sin embargo, los estudios demolingüísticos aspiran a reconstruir una imagen lo más cercana posible a la realidad a partir de los datos disponibles. Desde el punto de vista de la lingüística, la demografía de las lenguas sirve para documentar y monitorizar la composición lingüística de

los países y regiones, los procesos de integración sociolingüística de las comunidades, así como el estado y la vitalidad de lenguas minoritarias/minorizadas y las lenguas de migración. Desde el punto de vista de la política y la gobernabilidad, las cuantificaciones y los análisis demolingüísticos sirven para asistir a las políticas y la planificación lingüística dirigidas a gestionar las condiciones de existencia de las lenguas (espacios de uso, lenguas y educación, derechos lingüísticos, etc.).

Tanto para los fines lingüísticos como para los políticos resulta crucial conocer la disponibilidad, el estado y la comparabilidad de las fuentes demográficas y qué información lingüística es posible obtener de ellas. El objetivo del trabajo que aquí se presenta es ofrecer una recopilación, clasificación y análisis crítico de las fuentes demolingüísticas para el estudio de la migración hispanohablante. La descripción y el comentario crítico de los tipos de datos, las fuentes, su disponibilidad y su comparabilidad es necesaria para evitar que el tratamiento estadístico y sintético de la información demolingüística transmita una imagen de los grupos lingüísticos y migratorios más homogénea de lo que es en realidad (Schwenke 2018).

2 La documentación en demografía de las lenguas migratorias

El análisis cuantitativo de la migración hispanohablante debe recurrir a datos migratorios y a datos lingüísticos. Se trata de un ejercicio de demolingüística de la migración, disciplina que se ocupa del estudio de los repertorios lingüísticos de las poblaciones migrantes. Su objetivo principal es describir y analizar la composición, la estructura y la dinámica de una población migrante que se caracteriza por el conocimiento o el uso de una lengua de una variedad lingüística (Moreno Fernández 2014; Loureda *et al*. 2020).

Tanto el proceso de documentación como los datos y las fuentes de la demolingüística de la migración están subordinados a su objeto de análisis, generalmente la lengua entre poblaciones migrantes, y a su metodología, de carácter cuantitativo y comparativo. Antes de ofrecer un desglose crítico de las fuentes para el estudio de la migración hispanohablante en Europa es necesario enmarcar el espacio epistemológico en el que se desarrolla.

2.1 El concepto de hispanohablante con trasfondo migratorio

La demolingüística de la migración hispanohablante en Europa se ocupa del grupo de personas que son capaces de hablar español en diferentes contextos migratorios. Varios criterios permiten identificar dichos contextos: “el nacimiento en otro país, la condición de extranjero –que excluye a los inmigrantes que han conseguido la nacionalidad del país de destino–, el haber residido anteriormente en otro país o, por último, el hecho de tener un trasfondo migratorio, lo que incluye personas con experiencia migratoria directa y también a los nacidos de migrantes en el país de destino” (Loureda *et al*. 2022: 10). En este último sentido, el concepto de trasfondo migratorio remite a las generaciones, es decir, considera a las personas que se han desplazado (primera generación de inmigrantes) y sus descendientes (segunda y tercera generación). Es evidente que los descendientes de inmigrantes no pueden ser considerados migrantes si se sigue la definición común de inmigrante internacional que remite al desplazamiento espacial. Sin embargo, cuando la demografía de la migración se ocupa de fenómenos culturales como la lengua, es necesario considerar los procesos de transmisión intergeneracional que ocurren en el seno de las comunidades.

Sobre la base de lo expuesto, la demolingüística de la migración hispanohablante estudia el grupo de personas con trasfondo migratorio que incluyen el español en sus repertorios lingüísticos. Para determinar quién es un ***hispanohablante con trasfondo migratorio*** es necesario definir los dos conceptos que componen el sintagma: a) trasfondo migratorio y b) hablante de español.

a) Una persona con ***trasfondo migratorio*** es aquella cuya biografía se ve o se ha visto influida decisivamente por la migración, incluyendo tanto a los migrantes que se han desplazado a otro país (migrantes internacionales[1]) como a sus descendientes independientemente de su lugar de nacimiento. El concepto de trasfondo migratorio pretende capturar el impacto de la migración en dinámicas sociales y culturales como el multilingüismo o la integración sociolingüística.

b) Un ***hablante de español*** es una persona en cuyo repertorio lingüístico se incluyen competencias comunicativas en lengua española. Su nivel de dominio del español puede variar en función de las competencias que habrá adquirido a lo largo de su biografía lingüística en distintos espacios de socialización y aprendizaje (familia, escuela, comunidad, etc). Atendiendo a este criterio, se puede distinguir entre hablante de dominio nativo, capaz de usar la lengua o variedad "de una manera apropiada en su forma y adecuada al contexto comunicativo" (Moreno-Fernández 2014: 23), y hablante de competencia limitada, con un dominio limitado sociolingüística o estilísticamente y restringido a ciertos temas o situaciones (Loureda *et al*. 2020: 25).

Considerando ambas definiciones, el concepto de hispanohablante con trasfondo migratorio es heterogéneo e incluye perfiles de hablantes muy distintos. Desde el punto de vista migratorio, se pueden identificar varios grupos de personas para las que el español forma parte de su biografía lingüística, bien con un dominio nativo, bien con competencias limitadas, e independientemente de modo de adquisición (socialización en la familia o en la comunidad hispanohablante, aprendizaje como lengua extranjera, inmersión lingüística en contextos migratorios, etc.):

- Personas procedentes de países hispanohablantes que residen en Europa con o sin la nacionalidad del país de acogida;
- Sus descendientes nacidos en el país hispanohablante de origen, en el país de acogida o en otro país;
- Inmigrantes de procedencia no hispanohablante residentes en Europa que han adquirido competencias de español independientemente del contexto;
- Europeos de procedencia no hispanohablante que han adquirido competencias de español en contexto migratorios.

Así entendida, la migración hispanohablante es un grupo poblacional que se define por un atributo lingüístico: ser conocedor y/o usuario del español. Se trata de una *realidad lingüística* o *comunicativa* que, como tal, no está identificada en estadísticas poblacionales, pues se deja abarcar por las categorías que sirven habitualmente de indicador lingüístico a la demografía: la lengua materna y el origen étnico.

1 Las Naciones Unidas (ONU) definen como *migrante internacional* a "toda persona que se traslada a un país distinto al de su nacionalidad o de residencia habitual, de manera que el país de destino se convierte efectivamente en su nuevo país de residencia habitual" (Naciones Unidas 2022). La ***categoría de migrante*** es, por tanto, principalmente espacial, al remitir a un desplazamiento, y político-administrativa, pues depende de los marcos legales y los regímenes de migración y nacionalidad de los estados que determinan qué personas se incluyen en las estadísticas de inmigrantes.

2.2 Demolingüística y metodología cuantitativa

La característica diferencial de la demolingüística respecto a otras disciplinas dedicadas al estudio de poblaciones lingüísticas como la etnolingüística es su enfoque cuantitativo. Aunque suela hablarse de recuentos poblacionales, los datos de la demografía están más cerca de la probabilidad que de la contabilidad. En efecto, los datos demográficos, de carácter probabilístico, permiten estimar el volumen de un grupo poblacional en un espacio, de su evolución en el tiempo, y la probabilidad de que sus miembros presenten ciertas características sociodemográficas y culturales, entre las que se encuentra la competencia lingüística.

Desde el punto de vista espacial, la demolingüística recurre a datos lingüísticos y poblacionales contenidos en las estadísticas poblacionales tanto en el nivel local, como en el nacional o el internacional. Este tipo de fuentes suelen tener varios problemas de a) disponibilidad y b) comparabilidad:

a) En primer lugar, no siempre se dispone de información lingüística directa en la mayoría de los casos. En la mayor parte de los países europeos no existe un censo lingüístico. Sin embargo, la lengua es uno de los atributos humanos que ha interesado la demografía. En efecto, los censos poblacionales y otras fuentes demográficas oficiales incluyen información lingüística de diversa índole como la identidad etnolingüística de los ciudadanos o sus usos y competencias lingüísticas.

b) En segundo lugar, los datos existentes no siempre son comparables (Poulain 2008). La información lingüística depende de la definición de las categorías, la elección de técnicas y herramientas de recopilación de datos, y de las prácticas censales y de documentación, que por lo general son muy variadas y no están coordinadas.

Ambos escollos obligan a la demolingüística a trabajar con base de datos construidas, por lo general, a partir de la combinación de varias fuentes estadísticas, de las cuales debe revisar su exactitud y su comparabilidad. De este modo, los procesos de documentación demolingüística recurren a fuentes directas, que contienen información sobre las lenguas de los ciudadanos, como los censos poblacionales y los registros civiles, y a fuentes indirectas que permiten cuantificar grupos de hablantes a partir de indicadores sociodemográficos que apuntan a determinados perfiles lingüísticos más o menos probables.

3 Fuentes directas

Las fuentes directas son aquellas que “aportan información de primera mano sobre el número de hablantes de una lengua en un territorio determinado” (Moreno Fernández 2022). Esta información es recopilada a partir de preguntas lingüísticas (ej. lengua materna, lengua del hogar, grado de dominio de la lengua local, etc.) que se incluyen en los censos, registros poblacionales o encuestas (Extra 2017).

La información lingüística de carácter directo o primario suele incluirse en los métodos de contabilidad poblacional con tres objetivos: definir los perfiles lingüísticos de la población; conocer el grado de integración sociolingüística de las personas migrantes y/o analizar el uso de las lenguas regionales y cooficiales en aquellos países donde conviven dos o más idiomas (o dialectos) locales.

El apartado presenta el estado actual de las fuentes demolingüísticas directas para el estudio de la migración: los censos lingüísticos, los registros y las encuestas lingüísticas.

3.1 La lengua en los censos

Los censos ofrecen recuentos periódicos (normalmente realizados cada 10 años) de la población y las viviendas de un país o región determinados, y recogen información sobre sus principales características (geográficas, demográficas, sociales y económicas, además de las características de los hogares y las familias). Su fin es ofrecer una base estadística para la gobernanza y la gestión de aspectos sociales de un país o región. A diferencia de los registros civiles (véase § 3.2), los censos recopilan sus datos a partir de una muestra que permite una extrapolación estadística más o menos ajustada de sus resultados al conjunto de la población.

Los censos poblacionales suelen incluir información relevante para la demolingüística de la migración en apartados dedicados a las lenguas y también a la identidad étnica, dado que determinados grupos étnicos suelen definirse por el uso compartido de una lengua o variedad lingüísticas. En el caso de la migración hispanohablante, sin embargo, la etnicidad no es relevante pues la comunidad lingüística del español no constituye una entidad étnica establecida como tal en las estadísticas demográficas de ninguno de los países europeos (Leeman, en prensa).

La información censal más valiosa para la demolingüística de la migración es la derivada de las preguntas lingüísticas en combinación con otras variables sociodemográficas de los ciudadanos. Un análisis de los cuestionarios censales de todos los países europeos durante los últimos veinte años permite clasificarlos en tres categorías según su contenido lingüístico (tabla 1):

a) *Censos sin pregunta lingüística*. De acuerdo con los datos disponibles, 16 de los 41 países europeos estudiados no incluyen cuestiones lingüísticas en sus censos poblacionales y no muestran interés por obtener estadísticas sobre los repertorios y los usos lingüísticos de la población.

b) *Censos con una sola pregunta*. La mayoría de los censos poblacionales, 10 de 41 países europeos, incluyen al menos una cuestión relacionada con la lengua, que suele ser la pregunta por la lengua materna. Esta pregunta permite estimar el número de personas que pertenecen al grupo de dominio nativo de una lengua. Los ciudadanos que se identifican como hablantes nativos de español en un censo pueden ser por lo general inmigrantes, inmigrantes nacionalizados o descendientes de inmigrantes, por lo que el número no suele coincidir con las estadísticas de los registros de extranjería. Otros países prefieren enfocarse hacia la lengua regional (España) o la lengua del hogar (Alemania, Austria).

c) *Censos con varias preguntas*. 15 de 41 países incluyen más de una pregunta lingüística en su censo poblacional. La mayoría de los países suele combinar la pregunta por la lengua materna con otras variables lingüísticas, como la lengua en el hogar o con la familia. En países multilingües, como Irlanda o el Reino Unido, se pregunta también por el uso de la lengua regional y el nivel de competencia en la lengua mayoritaria del país.

Tabla 1
Clasificación de los censos europeos según información lingüística

Tipo	Subtipo	Países
a) Censos sin pregunta lingüística		Bélgica, Dinamarca, Grecia, Islandia, Letonia, Liechtenstein, Malta, Mónaco, Noruega, Países Bajos, Polonia, Portugal, San Marino, Turquía
b) Censos con una sola pregunta	Lengua materna	Albania, Bosnia y Herzegovina, Croacia, Eslovaquia, República Checa, Serbia
	Lengua principal	-
	Lengua regional	España
	Lengua hablada con fluidez	Chipre
	Lengua del hogar	Alemania, Austria
c) Censos con varias preguntas	Lengua en el hogar/con la familia y amigos	Bielorrusia, Eslovenia, Hungría, Irlanda, Suiza
	Lengua principal	Reino Unido (Inglaterra), Suiza
	Lengua materna	Armenia, Azerbaiyán, Bielorrusia, Bulgaria, Eslovenia, Estonia, Hungría, Rumanía, Rusia, Ucrania
	Lengua en el trabajo/institución educativa	Suiza
	Otras lenguas habladas con fluidez	Armenia, Azerbaiyán, Bielorrusia, Estonia, Hungría, Rusia, Ucrania
	Lengua regional	Estonia, Irlanda, Reino Unido, (Gales, Escocia, Irlanda del Norte)
	Nivel de competencia en la lengua mayoritaria del país	Irlanda, Reino Unido (Inglaterra)

Fuente: United Nations Statistics Division.

La información de los censos presenta al menos dos limitaciones para la documentación y el análisis demolingüístico. En primer lugar, la heterogeneidad de preguntas y modalidades de documentar la información lingüística dificulta la comparación de los resultados de los censos nacionales entre sí. Este hecho tiene un primer efecto evidente para la demolingüística: no permiten una cuantificación agregada del número de hablantes. En segundo lugar, el carácter muestral de los censos y la forma de tabular y reportar la información no es sensible a grupos lingüísticos poco predominantes en una región. En efecto, un problema común de los censos es que las oficinas de estadística nacionales solo suelen reportar información sobre lenguas no nacionales más comunes en el país, pues las muestras de grupos lingüísticos de poco tamaño no permiten extrapolar los datos al conjunto de la población de manera ajustada.

Conscientes de las limitaciones reales de los censos existentes, varios organismos internacionales preocupados por aspectos de gobernanza internacional vienen desarrollando propuestas para la homogeneización de las preguntas lingüísticas. En sus recomendaciones recientes, Naciones

Unidas (2017) fija y define una serie de categorías lingüísticas más relevantes que deberían contener los censos poblacionales, y subraya la importancia de la inclusión de todas las personas residentes en el país a la hora de recoger datos lingüísticos en censos poblacionales:

- Lengua materna (lengua hablada en casa durante la infancia).
- Lengua principal (lengua en la que el hablante presenta un mayor dominio).
- Lengua de uso más común (lengua mayoritariamente hablada por el hablante).
- Capacidad de hablar una o más lengua(s) (incluyendo la lengua oficial de país de residencia).

Además, es importante señalar que la Unión Europea y otros organismos internacionales están promoviendo una renovación metodológica de las herramientas para la documentación demográfica en la que los registros cobran un mayor protagonismo, dado que cada vez se otorga más importancia a los aspectos regionales. La innovación, que se comenzó a aplicar en alguno de los censos de 2021, consiste en combinar datos muestrales procedentes de cuestionarios digitales con información sociodemográfica recogida en registros como el padrón o el registro civil que han ampliado la información requerida a los ciudadanos en sus procesos administrativos.

La utilidad de los censos para la demolingüística de la migración depende del tipo de preguntas lingüísticas que contenga y estos solo pueden ser usados para análisis referidos a los propios países. Desde el punto de vista de una cuantificación de los hablantes de español con trasfondo migratorio en Europa, la información lingüística de los censos sirve para ajustar de forma contrastiva datos demolingüísticos y poblacionales recogidos en otras fuentes. Su mayor potencial es que permiten analizar la información lingüística en función del resto de variables sociodemográficas que incluye el censo: sexo, edad, país de nacimiento del encuestado y de los padres, tiempo de permanencia en el país, estado civil y tipo de matrimonio (endogámico/exogámico), etc. (Clyne 2003; Pauwels 2016). Los institutos nacionales de estadística de los países suelen poner a disposición de los usuarios informes y estadísticas seleccionadas que cruzan la pregunta o las preguntas lingüísticas con alguna de las variables mencionadas. Amén de que estos datos pueden ser de gran utilidad para analizar las dinámicas demolingüísticas de la migración hispanohablante, un estudio más profundo de correlaciones requiere acceder a los microdatos de los censos que por lo general suelen ponerse a disposición de investigadores interesados a través de un formulario oficial de solicitud.

3.2 La lengua en los registros poblacionales

Los registros civiles suelen incluir información sobre realidades, eventos y procesos de los que los ciudadanos dan cuenta a las administraciones locales, regionales o nacionales como los nombres, los nacimientos y defunciones, el estado civil o las nacionalizaciones. El conocimiento lingüístico de los ciudadanos es una información de difícil registro como toda aquella relacionada con los repertorios culturales de los individuos. Sin embargo, existen instituciones administrativas en Europa que registran datos lingüísticos.

Un ejemplo paradigmático de país que registra información lingüística sobre sus ciudadanos es Finlandia. El sistema de información de la población finlandesa (Finnish Population Information Systems) es un registro nacional con información básica sobre los ciudadanos finlandeses y los ciudadanos extranjeros que residen en Finlandia de forma permanente o temporal. Entre la información personal registrada se encuentra la lengua materna (native language). El registro

nacional de Finlandia es interesante para el estudio de la demolingüística de la migración pues el formulario de registro de la población extranjera incluye una pregunta por la lengua materna. Como se observa en la figura, el formulario solo permite introducir una única lengua, excluyendo la posibilidad de representar los perfiles lingüísticos de hablantes con varias lenguas maternas.

Imagen 1.
Pregunta lingüística contenida en el registro poblacional de Finlandia

12. Country of birth	13. Place of birth
14. Native language (Only one can be registered)	
15. Contact language ☐ Finnish ☐ Swedish	If your native language is neither Finnish nor Swedish, please indicate the language in which you would like to receive official messages that are not available in your native language.
16. Citizenships	

Fuente: Finnish Population Information Systems.

Los registros tienen una evidente ventaja respecto a los censos: permiten cuantificar con gran exactitud un grupo lingüístico determinado. Sin embargo, el aprovechamiento de la ventaja derivada de una recolección y un tratamiento digital de datos y del carácter universal del registro depende del tipo y la calidad de la información lingüística que se recoge. Si, como en el caso del registro finlandés, los ciudadanos nacionales y extranjeros solo son preguntados por su lengua materna y su posibilidad de respuesta se reduce a una, el registro invisibiliza los repertorios multilingües de los ciudadanos que son especialmente comunes en el caso de las poblaciones inmigrantes.

3.3 La lengua en las encuestas y las encuestas lingüísticas

Las encuestas pueden ser realizadas o bien por las autoridades administrativas, o bien por otras entidades o iniciativas de un espacio determinado. El objetivo suele ser el conocimiento de una realidad, en este caso una realidad lingüística. Contienen preguntas más específicas que el censo, puesto que no están únicamente destinadas a cuantificar la presencia de diferentes lenguas, sino también a obtener información sobre uso, dominio, frecuencia, actitudes, etc. Las encuestas pueden ser complementadas con entrevistas o reuniones de grupos focales.

La demolingüística encuentra información relevante para el estudio de las lenguas migratorias en encuestas realizadas a nivel internacional europeo como el Eurobarómetro o la Labour Force Survey, y nivel nacional y regional en los distintos países en los que están disponibles.

El **Eurobarómetro** es la encuesta lingüística más importante de la Unión Europea. Aunque su principal objetivo es estudiar el conocimiento de lenguas extranjeras de los ciudadanos europeos, en sus tres versiones de 2000, 2005 y 2012 incluye una pregunta por la lengua materna. En sus dos primeras versiones, el cuestionario preguntaba únicamente por la lengua materna; en 2012 la formulación de la pregunta cambió a: “Thinking about the languages that you speak, which language is your mother tongue?”. Dado que la encuesta dispone de muestras para cada país, sus resultados permiten estimar cuántas personas tienen español como lengua materna en cada uno de los países no hispanohablantes de Europa.

La **Labour Force Survey** es la mayor encuesta europea por muestreo de hogares privados que se empezó a realizar a partir de 1983. Mediante los resultados trimestrales se mide la participación en el mercado laboral de las personas mayores de 15 años y sobre las personas que no forman parte de la población activa. Tras la realización de las encuestas (mediante cuestionarios y entrevistas), la selección de las muestras y la transmisión de los resultados a Eurostat por parte de los institutos nacionales de estadísticas de los distintos países (a día de hoy, participan 35 países), Eurostat armoniza los datos a nivel europeo. Dentro de la encuesta, existe un módulo dedicado a la situación laboral de los migrantes y su descendencia inmediata ("Labour market situation of migrants and their immediate descendants") que se incluyó en 2008, 2014 y 2021. El módulo pretende comprobar si el país de formación tiene alguna repercusión en la integración de los inmigrantes en el mercado laboral. Incluye las respuestas de ciudadanos extranjeros y sus descendientes mayores de 15 años a partir de tres preguntas:

1. Competencias en la lengua principal del país de acogida antes de migrar
2. Competencias actuales en la lengua principal del país de acogida
3. Participación en un curso de la lengua principal del país de acogida

La segunda permite identificar el conocimiento que tienen los inmigrantes de la lengua principal del país en el que residen. Para el estudio de la migración hispanohablante es útil pues permite conocer el nivel de competencia en lengua española de los migrantes que residen en España. La combinación con otras variables permite saber de qué generación se trata, la edad, el sexo y la procedencia.

Además de llevarse a cabo encuestas a nivel europeo, existen también iniciativas por parte de instituciones nacionales y regionales de realizar encuestas lingüísticas. Se realizan con el objetivo de conocer los usos y conocimientos lingüísticos de la población, tanto de las lenguas nacionales y regionales, como de las extranjeras.

Un ejemplo sería la encuesta sobre **Los ciudadanos y su tiempo libre ("I Cittadini e il Tempo Libero" (CTL))** que se realiza en Italia desde el 1995 y que forma parte de un sistema integrado de encuestas sociales. Aquí se incluyen preguntas sobre el uso del italiano y sus dialectos, el uso de las lenguas maternas de los residentes de origen extranjero en el hogar o en los círculos sociales más estrechos, y los conocimientos de lenguas extranjeras de la población (Instituto Nazionale di Stadistica 2017). Asimismo, el **Estudio del Panel Socioeconómico Alemán (SOEP)** del Instituto Alemán de Investigación Económica (DIW), que recoge anualmente información representativa sobre aspectos socioeconómicos de los hogares privados en Alemania, contiene preguntas relacionadas con los usos lingüísticos en el ámbito familiar, la lengua materna y el aprendizaje de lenguas extranjeras (Adler 2019).

En algunos países multilingües europeos se organizan encuestas para conocer las prácticas y usos relacionados con las distintas lenguas cooficiales. En Suiza se realiza desde 2014 un estudio quinquenal sobre lengua, cultura y religión de la población. Esta encuesta incluye preguntas sobre las lenguas más usadas en distintas etapas y contextos (en la infancia, con los descendientes, con la pareja, con las amistades, en el hogar, en el lugar de trabajo etc.); así como sobre las lenguas que la población desea aprender (Bundesamt für Statistik 2022). En España también existen encuestas sobre el conocimiento y los usos lingüísticos del gallego (Instituto Galego de Estadística), catalán (ej. Encuesta de usos lingüísticos de la población de la Generalitat de Catalunya) o euskera (Instituto Vasco de Estadística).

Dadas las modificaciones que genera la migración en la realidad lingüística de un país, la población migrante recibe en algunos países una especial atención por parte de las instituciones nacionales. En las encuestas del SOEP de Alemania los participantes de origen extranjero deben responder a preguntas relacionadas con el conocimiento de la lengua alemana, lengua materna y otras lenguas extranjeras y sobre los usos lingüísticos en diferentes ámbitos (familiar, laboral, de las amistades). Algunos estados consideran necesario incluso diseñar encuestas exclusivamente para este grupo poblacional, como la **Encuesta Nacional de Inmigrantes** (2007) de España llevada a cabo por el Instituto Nacional de Estadística, donde se recoge información sobre las lenguas maternas de los participantes.

4 Fuentes indirectas

Como se ha visto en el análisis de las fuentes directas, es común que las estadísticas poblacionales no incluyan información sobre las lenguas de los ciudadanos y que cuando están disponibles sean fragmentarias y/o presenten limitaciones conceptuales y metodológicas que dificultan los análisis demolingüísticos. Por ello, el estudio de las lenguas y sus poblaciones suele recurrir a fuentes indirectas de carácter demográfico que permitan deducir la información lingüística a partir de datos primarios, generalmente de tipo sociodemográfico (Moreno Fernández 2014; Moreno Fernández en prensa).

4.1 Indicadores sociales para la competencia lingüística

La principal variable demográfica que permite identificar todos estos grupos en las estadísticas de población es la procedencia a la que la demolingüística suele aproximarse a partir de dos indicadores sociales: la nacionalidad y el país de nacimiento. Eurostat ofrece una definición de las dos categorías que puede servir para cualquier fuente demográfica europea:

- La **nacionalidad** denota el vínculo legal entre un individuo y su estado, que se adquiere por nacimiento o tras un proceso de nacionalización basado en la declaración, la elección, el matrimonio u otras vías consideradas en las leyes de naturalización en cada país.
- El **país de nacimiento** es el país en el que tiene lugar el nacimiento, que suele coincidir con el país de residencia (considerando sus fronteras actuales) de la madre en el momento del nacimiento.

Los servicios nacionales de estadística de cada país recopilan información sobre la nacionalidad y el país de nacimiento que suele recogerse en distintos registros civiles como el padrón de residentes o el registro de extranjería. Ambos datos demográficos sirven como indicadores lingüísticos si se tiene en cuenta una serie de consideraciones, partiendo del hecho de que tener la nacionalidad de un país o haber nacido en él no implica necesariamente hablar una o varias de las lenguas del país en cuestión. Sin embargo, su utilidad para los estudios de demografía lingüística resulta evidente pues permite identificar un grupo de manera objetiva y comparar los datos entre países y regiones. La nacionalidad como indicador lingüístico presenta una ventaja sobre el país de nacimiento: mientras este último es un hecho coyuntural, sobre todo en familias transnacionales, e invariable que no refleja dinámicas sociales como la integración sociolingüística, la nacionalidad, combinada con estadísticas de nacionalizaciones, permite agrupar ciudadanos que pueden compartir la lengua asociada con la nacionalidad (Extra 2017; Leeman, en prensa).

En el caso de la población hispanohablante procedente de América Latina la combinación comparada de ambos indicadores es importante si se considera que algunos ciudadanos "pueden tener la nacionalidad de un país europeo, como es el caso de aquellos procedentes de Argentina, Uruguay y Venezuela con nacionalidad italiana" (Leeman, en prensa).

El análisis de la migración y sus características se basa en distintos datos disponibles en fuentes estadísticas de los dos espacios principales que se consideran: los datos de España proceden en su mayoría del Instituto Nacional de Estadística (INE) y los de los países europeos, de Eurostat. La lista completa de fuentes puede verse en la tabla 2.

Tabla 2
Fuentes para el estudio demográfico de la migración

Territorio	Fuente y marco temporal	Tipo de datos
España	INE Cifras de población (1998-2021)	Población residente de origen extranjero (nacionalizada o no) según procedencia, sexo y edad
	INE Encuesta continua de hogares (2013-2021)	Descendientes de personas extranjeras
Europa	Eurostat Cifras de población (1998-2020)	Población residente de origen extranjero (no nacionalizada) según procedencia, sexo y edad
	Eurostat Adquisición de nacionalidad (1998-2020)	Población extranjera que ha adquirido la nacionalidad de un país europeo según el año y el país de naturalización
	Eurostat Estadísticas de fertilidad (2011-2020)	Indicadores de natalidad según edad y país (permite estimar el grupo de segunda generación)

Las fuentes mencionadas permiten estudiar los distintos perfiles migratorios. Como se ha visto anteriormente, el grupo de personas hispanohablantes con trasfondo migratorio se compone de varios subgrupos de hablantes que pueden ser identificados con mayor o menor exactitud en las estadísticas de población. La tabla 3 desglosa los principales perfiles según las fuentes para su cuantificación:

Tabla 3
Fuentes para la cuantificación de los perfiles migratorios en Europa

Fuente	Perfil migratorio
Eurostat	Hispanohablantes con trasfondo migratorio residentes en Europa (sin España)
INE	Hispanohablantes con trasfondo migratorio residentes en España y migrantes europeos retornados a sus países de origen

4.2 Oficina Europea de Estadística (Eurostat)

El grupo de hispanohablantes con trasfondo migratorio que residen en Europa (sin España) puede reconstruirse a partir de los datos que proporciona la Oficina Europea de Estadística (Eurostat), que realiza el mayor ejercicio de compilación y unificación de datos poblacionales en Europa. En su base de datos (https://ec.europa.eu/eurostat/data/database) Eurostat recopila estadísticas

poblacionales aportadas por los estados miembro que reflejan los recuentos demográficos de cada país a 31 de diciembre del año de referencia y que en la base de datos central se publican referidas al 1 de enero del año siguiente.

La unidad de observación de las estadísticas poblacionales es el residente habitual (*usual resident*), es decir, la persona que "normalmente pasa en el lugar de residencia el período de descanso diario, independientemente de las ausencias temporales por motivos de ocio, vacaciones, visitas a amigos y familiares, negocios, tratamiento médico o peregrinaje religioso"[2] (Eurostat 2022b). Dos criterios permiten incluir a un ciudadano en estas estadísticas: que haya sido residente habitual en un país durante al menos doce años antes del recuento; o que haya llegado al país en los últimos doce meses con la intención de residir allí al menos durante un año. Este criterio excluye de los recuentos personas en condiciones de movilidad temporal breve como los turistas, los estudiantes de intercambio, los trabajadores temporeros o los empleados enviados por las empresas a sus delegaciones en el extranjero para estancias breves.

Entre sus estadísticas poblacionales, Eurostat publica dos series de datos que ofrecen información sobre los principales indicadores sociales de la migración: el país de origen (Serie: "Population on 1 January by age group, sex and country of birth") y la nacionalidad (Serie: "Population on 1 January by age group, sex and citizenship"). La base de datos poblacionales según nacionalidad permite identificar al grupo de población extranjera con nacionalidad de un país en el que el español es lengua oficial. Las variables sociodemográficas de la base de datos permiten describir el grupo y su evolución desde 1998:

- Variables sociodemográficas: el sexo y la edad. La edad es especialmente relevante para el estudio demolingüístico pues aporta información sobre el envejecimiento del grupo y sus tendencias generacionales.
- Distribución espacial. Dado que los datos se registran en cada país, se dispone también de información sobre el lugar de residencia. Eso permite analizar la distribución espacial. El portal de datos de Eurostat no ofrece las estadísticas según la distribución regional de la población dentro de cada país, aunque según la descripción metodológica los datos estarían disponibles. Para eso hay que recurrir a las fuentes nacionales. El lugar de residencia es relevante porque permite describir los espacios lingüísticos en los que se encuentra el grupo, las lenguas con las que entran en contacto los hablantes y también permite identificar patrones espaciales como algunas relaciones prioritarias entre origen y destino.

Como se ha visto, la principal ventaja del criterio *país de nacionalidad* es que se puede combinar con las estadísticas de naturalizaciones. Los inmigrantes que se han nacionalizado en el país de acogida dejan de aparecer en las estadísticas sucesivas de residentes con su antigua nacionalidad, incluso en los casos en los que el régimen de doble nacionalidad les permita mantenerla. La base de datos de Eurostat ("Acquisition and loss of citizenship" [migr_acqn]) recoge estadísticas nacionales de las adquisiciones de nacionalidad que tienen lugar cada año en los países europeos. El análisis de esta fuente permite identificar a un grupo de personas procedentes de países hispanohablantes que no aparecen en la base de datos de población con su nacionalidad anterior de un país hispanohablante.

2 Texto original: "normally spends the daily period of rest, regardless of temporary absences for purposes of recreation, holidays, visits to friends and relatives, business, medical treatment or religious pilgrimage" (Eurostat 2022b).

Eurostat no incluye estadísticas de descendientes de inmigrantes por lo que no es posible obtener una cuantificación directa a partir de sus bases de datos. Sin embargo, las estadísticas de natalidad combinadas con los datos de población extranjera según procedencia permiten estimar un grupo de personas con al menos un padre inmigrante teniendo en cuenta su procedencia.

4.3 Instituto Nacional de Estadística de España (INE)

Las estadísticas poblacionales de España ofrecen varias bases de datos útiles para el estudio demolingüístico de la migración hispanohablante. Concretamente, el Instituto Nacional de Estadística pone a disposición de los usuarios información demográfica que permite cuantificar y describir el grupo de personas con trasfondo migratorio que hablan español en España: por un lado, migrantes procedentes de países y regiones hispanohablantes; por otro, migrantes procedentes de países no hispanohablantes que habrán adquirido competencias de español a través de su experiencia migratoria.

La serie de datos del Instituto Nacional de Estadística (INE) que permite una agrupación homogénea de los orígenes migratorios se inicia en 2002 y está disponible en el apartado dedicado a cifras poblacionales. Las estadísticas de población residente permiten cuantificar los ciudadanos de origen extranjero (nacionalizada o no) según procedencia, sexo y edad. Al grupo de ciudadanos de procedencia de un país hispanohablante se les puede suponer con poco margen de error un dominio nativo del español. Para estimar las competencias lingüísticas del resto, las estadísticas se pueden combinar con los resultados del apartado lingüístico de la Encuesta de Población Activa-Módulo[3] (versión nacional de la Labour Force Survey, comentada ya en el apartado 3.3 de este trabajo), que ofrecen los porcentajes de migrantes según procedencia que consideran que disponen de un dominio nativo del español, competencias avanzadas, intermedias o de principiante.

La Estadística de Variaciones Residenciales[4] es una fuente de gran valor para estimar el grupo de personas con trasfondo migratorio que se han regresado a sus países de origen después de una estancia en España. Este grupo de retornados es relevante para la demolingüística de la migración hispanohablante pues se trata de personas de origen europeo que han adquirido competencias de español en su experiencia migratoria.

Por último, la Encuesta Continua de Hogares[5] permite acercarse a una descripción mínima del grupo de descendientes de inmigrantes hispanohablantes. La fuente ofrece información por países de procedencia que solo considera descendientes de madres extranjeras, no contabilizando, así, casos de descendientes de hombres extranjeros en matrimonios con una persona de nacionalidad española. Tampoco se registran los descendientes de tercera generación, un grupo de personas que tienden a diluirse sistemáticamente en las estadísticas poblacionales europeas.

3 https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176918&menu=resultados&idp=1254735976595#!tabs-1254736195124
4 https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736177013&menu=resultados&idp=1254734710990
5 https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176952&menu=resultados&idp=1254735572981

5 Conclusiones

A lo largo de las páginas que componen este trabajo se han revisado críticamente las fuentes para el estudio demolingüístico de la migración hispanohablante en Europa. La combinación de estadísticas y datos poblacionales de carácter lingüístico, demográfico y migratorio es el único camino para construir la base empírica para un análisis cuantitativo de los hablantes de una lengua. Las fuentes discutidas ofrecen material más que suficiente para dar respuesta a las principales preguntas de la demolingüística de la migración hispanohablante:

- ¿Cuántas personas con trasfondo migratorio son capaces de comunicarse en español en Europa?
- ¿Cuáles son sus perfiles lingüísticos y sociodemográficos?
- ¿Qué dinámicas determinan su distribución espacial y las condiciones de uso de la lengua?
- ¿Cuáles son sus escenarios demográficos futuros?

Aparte de mostrar las potencialidades de las fuentes lingüísticas y demográficas disponibles, su discusión crítica pone de manifiesto la poca sensibilidad de la contabilidad nacional de los países europeos hacia las lenguas migratorias. La falta de información sobre los repertorios lingüísticos de los ciudadanos con trasfondo migratorio es la consecuencia del predominio de regímenes lingüísticos monolingües (Gogolin 1997) en la mayoría de regiones de Europa. Incluso en países que se declaran naciones de migración como Alemania, los únicos datos lingüísticos que recogen las estadísticas van dirigidas a identificar grupos poblacionales migratorios que podrían tener dificultades para comunicarse en alemán, suponiendo un posible problema para su integración social (Adler 2019). Si bien este tipo de información lingüística es altamente útil para gestionar políticas sociales de integración, es insuficiente para el estudio demolingüistico de las sociedades europeas cada vez más diversas y multilingües.

6 Referencias bibliográficas

ADLER, A. (2019), “Sprachstatistik in Deutschland”, *Deutsche Sprache*, 3 (47). Berlín: Erich Schmidt Verlag, pp. 197-219.

BUNDESAMT FÜR STATISTIK (2022), *Erhebung zur Sprache, Religion und Kultur,* https://www.bfs.admin.ch/bfs/de/home/statistiken/bevoelkerung/erhebungen/esrk.html

CONSEJO ECONÓMICO Y SOCIAL DE ESPAÑA (2019). *La inmigración en España: efectos y oportunidades.* Informe 02/2019. https://www.ces.es/documents/10180/5209150/Inf0219.pdf

CLYNE, M. G. (2003), *Dynamics of Language Contact*. Cambridge: Cambridge University Press.

DEUTSCHES INSTITUT FÜR WIRTSCHAFTSFORSCHUNG (s.f.). *Socio-Economic Panel*. https://www.diw.de/de/soep

EUROBAROMETER (2000), *Special survey on languages (EB 54LAN).* https://www.gesis.org/en/eurobarometer-data-service/survey-series/standard-special-eb/study-overview/eurobarometer-54lan-za-3389-dec-2000

EUROBAROMETER (2005a), *European Union Enlargement, the European Constitution, Economic Challenges, Innovative Products and Services (EB 63.4).* https://www.gesis.org/en/eurobarometer-data-service/survey-series/standard-special-eb/study-overview/eurobarometer-634-za-4411-may-jun-2005

EUROBAROMETER (2005b), *Foreign Languages, Biotechnology, Organized Crime, and Health Items (EB 64.3).* https://www.gesis.org/en/eurobarometer-data-service/survey-series/standard-special-eb/study-overview/eurobarometer-643-za-4415-nov-dec-2005

EUROBAROMETER (2012), *Robotics, Civil Protection, Humanitarian Aid, Smoking Habits, and Multilingualism (EB 77.1).* https://www.gesis.org/en/eurobarometer-data-service/survey-series/standard-special-eb/study-overview/eurobarometer-771-za-5597-feb-mar-2012

EUROSTAT (2022a), https://ec.europa.eu/eurostat/web/main/home

EUROSTAT (2022b), *Migration and asylum*. https://ec.europa.eu/eurostat/web/migration-asylum/international-migration-citizenship/methodology

EXTRA, G. Y GORTER, D. (2008), "The constellation of languages in Europe: an inclusive approach". En G. Extra y D. Gorter (eds.), *Multilingual Europe: Facts and Policies*. Berlin, New York: De Gruyter Mouton, 2008, pp. 3-62. https://doi.org/10.1515/9783110208351.1.3

EXTRA, G. (2017), “The Constellation of Languages in Europe: Comparative Perspectives on Regional and Immigrant Minority Languages”. En O. Kagan, M. M. Carreira, y C. Hitchens Chik (eds.), *The Routledge Handbook of Heritage Language Education: From Innovation to Program Building (Routledge Handbooks in Linguistics),* Routledge, pp. 11-21.

FINNISH POPULATION INFORMATION SYSTEMS (s.f.). https://dvv.fi/en/population-information-system#:~:text=The%20Finnish%20Population%20Information%20System,projects%2C%20residences%20and%20real%20estate.

GOGOLIN, I. (1997), “The ‘monolingual habitus’ as the common feature in teaching in the language of the majority in different countries”, *Per Linguam,* 13(2), pp. 38-49.

ISTITUTO NAZIONALE DI STATISTICA (2015), *I cittadini e il tempo libero (CTL)*. https://www.istat.it/it/archivio/227593

LEEMAN, J. (en prensa), “(In)visible identities and inequities: The construction of Latinidad in European censuses”. En Marquez Reiter & A. Patiño-Santos (Eds.), *Language practices and processes among Latin Americans in Europe*. Routledge.

LOUREDA, Ó., MORENO FERNÁNDEZ, F., ÁLVAREZ MELLA, H. Y SCHEFFLER, D. (2020), “Alemania y sus hablantes de español”, en VV. AA., *El español en el mundo 2020. Anuario del Instituto Cervantes.* Madrid: Instituto Cervantes y Bala Perdida Editorial S. L., pp. 321-354.

LOUREDA, Ó., ÁLVAREZ MELLA, H. y BLATTNER, C. (2021), “La Europa hispanohablante: una mirada demolingüística”. En J. L. García Delgado (dir.), *El español lengua internacional: proyección y economía.* Pamplona: Universidad Nebrija, Civitas y Thompson Reuters, pp. 65-83.

LOUREDA, Ó., GUTIÉRREZ PALACIOS, R., ÁLVAREZ MELLA, H. y BLATTERN, C. (en prensa), La migración hispanohablante en Europa: perfiles demolingüísticos e integración social. Documento de Trabajo, Observatorio Nebrija del Español, Universidad Nebrija.

MORENO FERNÁNDEZ, F. (2014), “Fundamentos de demografía lingüística a propósito de la lengua española”. *Revista Internacional de Lingüística Iberoamericana,* Vol. 12, No. 2 (24), pp. 19-38.

MORENO FERNÁNDEZ, F. (en prensa), *Demography of languages.* Londres: Routledge.

NACIONES UNIDAS (2017), Principles and Recommendations for Population and Housing Censuses. En línea: https://unstats.un.org/unsd/demographic-social/Standards-and-Methods/files/Principles_and_Recommendations/Population-and-Housing-Censuses/Series_M67rev3-E.pdf [última consulta 30/09/2022]

NACIONES UNIDAS (2022), Definición migrante. En línea: https://refugeesmigrants.un.org/es/definitions#:~:text=definiciones%20de%20refugiados-,Migrante,utilizados%2C%20legales%20u%20otros%C2%BB [última consulta 30/09/2022].

PAUWELS, A. (2016), *Language Maintenance and Shift*. Cambridge: Cambridge University Press.

POULAIN, M. (2008). European migration statistics: definitions, data and challenges. En M. Barni y G. Extra (eds.), *Mapping Linguistic Diversity in Multicultural Contexts*. Berlin/New York: Mouton de Gruyter, pp. 43–66.

SCHWENKE, H. (2018), *Globale Migration zur Einführung.* Hamburg: Junius.

UNITED NATIONS STATISTIC DIVISION (s.f.), *World Population and Housing Census Programme*. https://unstats.un.org/unsd/demographic-social/census/document-resources/

OBSERVATORIO NEBRIJA DEL ESPAÑOL

www.nebrija.com
www.fundacionnebrija.org
www.nebrija500.es

FUNDACIÓN NEBRIJA

ANTONIO DE NEBRIJA 500 AÑOS

GOBIERNO DE ESPAÑA

MINISTERIO DE ASUNTOS EXTERIORES, UNION EUROPEA Y COOPERACION