

Repede, Doina

**Corpus oral del español de los migrantes residentes en la ciudad de Sevilla  
(Corpus.migra.sev) : construcción, codificación y exploración**

*Études romanes de Brno*. 2024, vol. 45, iss. 4, pp. 173-191

ISSN 2336-4416 (online)

Stable URL (DOI): <https://doi.org/10.5817/ERB2024-4-9>

Stable URL (handle): <https://hdl.handle.net/11222.digilib/digilib.81319>

License: [CC BY-SA 4.0 International](https://creativecommons.org/licenses/by-sa/4.0/)

Access Date: 20. 02. 2025

Version: 20250219

Terms of use: Digital Library of the Faculty of Arts, Masaryk University provides access to digitized documents strictly for personal use, unless otherwise specified.

---

## ***Corpus oral del español de los migrantes residentes en la ciudad de Sevilla (Corpus.migra.sev). Construcción, codificación y exploración***

### ***Oral Corpus of Spanish Migrants Residing in the City of Sevilla (Corpus.migra.sev). Construction, Encoding, and Exploration***

**DOINA REPEDE** [drepede@ugr.es]

Universidad de Granada, España

---

#### **RESUMEN**

*Corpus.migra.sev* es un corpus informatizado constituido por entrevistas semidirigidas, basadas en un guion conversacional, realizadas a hablantes peruanos, colombianos, nicaragüenses y ecuatorianos residentes en la ciudad de Sevilla, grabadas entre 2020 y 2023. Mediante el diseño y la elaboración de este corpus, el objetivo fundamental consiste en recopilar y codificar muestras orales representativas de las distintas comunidades hispanohablantes presentes en la ciudad hispalense, en consonancia con otros corpus ya recogidos, como el CORDIESIN (*Corpus dinámico del español de la inmigración*) en la comunidad de Madrid. Estos materiales permitirán obtener resultados sobre las características lingüísticas del español hablado por los migrantes, sobre su acomodación e integración sociolingüística, a la vez que posibilitará la realización de estudios contrastivos con el español de hablantes vernáculos y/o diferentes comunidades migrantes residentes en España. En este trabajo se exponen algunos aspectos relacionados con el diseño y construcción del corpus, su codificación, así como las diversas posibilidades de explotación de esta base de datos.

#### **PALABRAS CLAVE**

Corpus oral; variación; español; transcripción; migración; Sevilla

#### **ABSTRACT**

*Corpus.migra.sev* is a computerized corpus consisting of semi-directed interviews, based on a conversational script, conducted with Peruvian, Colombian, Nicaraguan, and Ecuadorian speakers residing in the city of Seville. These interviews were recorded between the years 2020 and 2023. The primary objective behind the design and creation of this corpus is to collect and encode representative oral samples from the various Spanish-speaking communities present in Seville. This effort is in line with other corpora currently collected, such as CORDIESIN (Dynamic Corpus of Spanish Immigration) in the Madrid community. The materials gathered through *Corpus.migra.sev* will enable researchers to obtain insights into the linguistic characteristics of Spanish spoken by migrants, as well as their ac-

commodation and sociolinguistic integration. Furthermore, it will facilitate the conduct of contrastive studies between the Spanish of native speakers and/or the migrant community residing in Spain. In this paper, various aspects related to the design and construction of the corpus are discussed, including its encoding, as well as the diverse possibilities for exploiting this database.

## KEYWORDS

Spoken corpus; variation; Spanish; transcription; migration; Seville

RECIBIDO 2024-03-11; ACEPTADO 2024-04-23

---

Proyecto de I+D+i, convocatoria 2019, *Agenda 2050. El español de Sevilla y Las Palmas de Gran Canaria: procesos de variación y cambio espaciales y sociales* (VARES-AGENDA50), Ref. PID2019-104982GB-C54, financiado por el Ministerio de Ciencia, Innovación y Universidades de España y por la Agencia Estatal de Investigación (MCIU/AEI/10.13039/501100011033).

## 1. Introducción

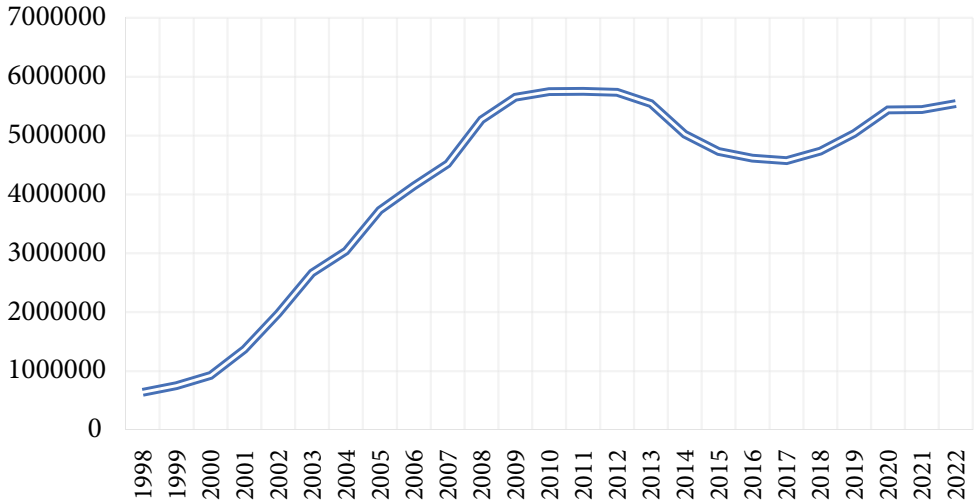
El fenómeno de *migración* se describe como el “desplazamiento geográfico de individuos o grupos, generalmente por causas económicas o sociales” (DLE 2023, versión online<sup>1</sup>), que ocurre en el interior de un país –migración interior– o se produce cuando las personas se desplazan a un país diferente del propio –migración exterior– (Bartolomé 2002). Si nos centramos especialmente en la migración internacional, este movimiento “ha alcanzado durante el siglo XX y el siglo XXI unas cotas desconocidas en épocas pasadas, tanto en términos cuantitativos como proporcionales, y todo indica que avanza a un ritmo elevado, mayor incluso que el previsto” (Paredes García 2020: 40). España, como parte de Europa, ha vivido un proceso semejante en cuanto a los flujos migratorios durante el siglo XX. Desde finales de los años 80, la migración en este país se ha convertido en un fenómeno de importancia demográfica, social y económica, si bien es verdad que es a partir de 1998 cuando la población extranjera alcanza “una gran notoriedad” en el país (Fernández de Molina 2019: 8; Colectivo Ioé 2002, 2010, García Cívico 2011) y “crece exponencialmente hasta 2008” (Paredes García 2020: 41), momento en el que llega a representar el 12,2 % de la población total. Este movimiento migratorio se mantiene estable hasta 2012, cuando empieza a disminuir la entrada y a acelerarse la salida de los extranjeros, hasta el 2017. No obstante, a partir de 2018, la entrada de los extranjeros vuelve a aumentar progresivamente casi en todas las comunidades autónomas<sup>2</sup> (Gráfico 1).

Atendiendo al perfil y el origen de las personas migrantes en España, estos difieren desde su comienzo hasta la actualidad. Los primeros migrantes hispanoamericanos, que dejaron su país de origen en los años 70 por cuestiones políticas, provenían de países bajo régimen dictatorial, como

---

1 <https://dle.rae.es/>

2 Población extranjera por comunidades y provincias, nacionalidad y sexo (2022). Disponible en línea: <https://www.ine.es/jaxi/Tabla.htm?path=/t20/e245/p04/provi/l0/&file=0ccaa002.px&L=0>.



**Gráfico 1.** Evolución población extranjera en España

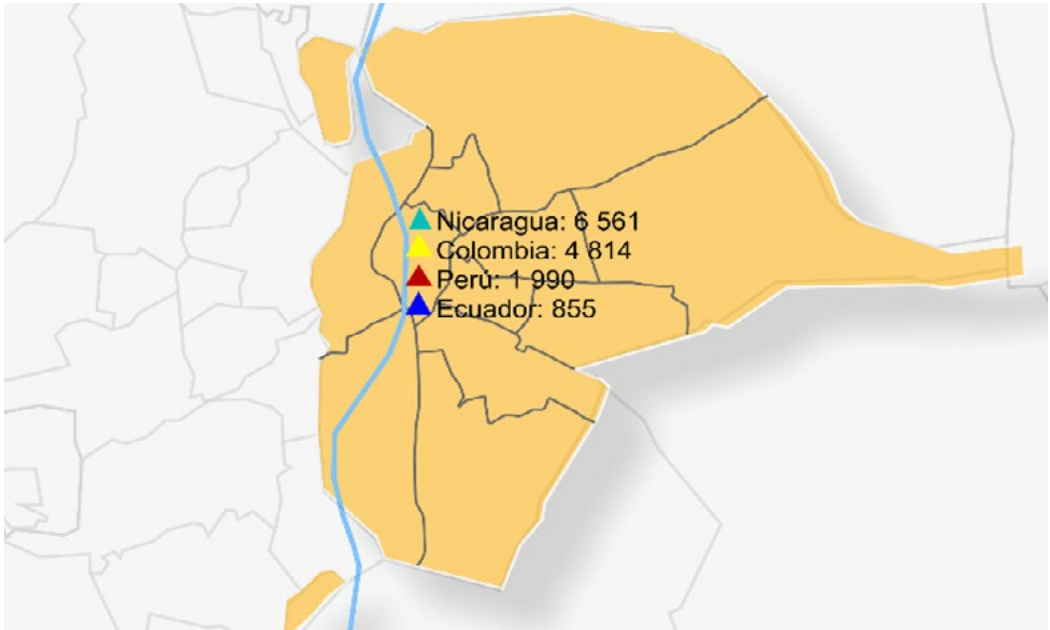
Cuba, Chile, Uruguay y Argentina, y eran individuos con un nivel alto de educación (Valero-Matas *et al.* 2014: 16). Más tarde, en la década de los años 90 se produce “un cambio de tendencia” en el perfil del migrante que llega a España, pues son los dominicanos y peruanos los que “desplazaron al resto de los latinoamericanos, convirtiéndose en el año 1995 en las principales colonias, fundamentalmente por la llegada de mujeres para trabajar en el servicio doméstico” (idem). A su vez, durante el siglo XX estos colectivos de migrantes latinoamericanos son superados por otros procedentes de América del Sur, como ecuatorianos y colombianos, “quienes pasaron a erigirse como los principales grupos” (Valero-Matas *et al.* 2014: 16; Martínez Buján & Golías Pérez 2005). Actualmente el mapa migratorio en España se completa con población procedente de los países de Europa del Este, del Magreb, de África Subsahariana y de Asia<sup>3</sup>.

De entre los distintos colectivos hispanohablantes residentes en España, nos interesan especialmente, como veremos más adelante, el grupo peruano, el colombiano, el ecuatoriano y el nicaragüense. Las tres primeras comunidades muestran una presencia mayor que la nicaragüense (Tabla 1) en todo el territorio español, mientras que este colectivo supera al ecuatoriano y el peruano en la comunidad andaluza.

País	España	Andalucía
Colombia	314 679	25 669
Perú	120 255	7 003
Ecuador	119 885	6 976
Nicaragua	60 681	10 471

**Tabla 1.** Reparto general de los colectivos nicaragüense, colombiano, peruano y ecuatoriano en España

3 Para más detalles sobre las características de la población extranjera en España, consúltese también Ferrer & Urdiales (2004).



**Figura 1.** Distribución de los grupos nicaragüense, colombiano, peruano y ecuatoriano en Sevilla

Asimismo, en la ciudad de Sevilla, receptora constante de hablantes de origen hispano, podemos observar en la Figura 1 que, según los datos extraídos del Instituto Nacional de Estadística (INE), la comunidad nicaragüense supera con creces a la población colombiana, peruana y ecuatoriana.

## 2. Bases de datos para el estudio de la integración inmigrante

La integración de inmigrantes en el país de acogida es un proceso continuo y multidimensional que requiere el compromiso y la colaboración de diferentes actores. Concretamente, y en palabras de Moreno Fernández (2009: 131), se trata de “un proceso de ajuste mutuo, por parte de una población inmigrante y de una población residente, que permite la construcción intersubjetiva de la realidad social de ambas poblaciones y que las lleva a compartir unos valores, sean los propios de la población residente, sean los de la población residente e inmigrante”. Este proceso de integración presenta un carácter “bidireccional de continua reproducción y renovación” (ídem), esto es, que requiere tanto el esfuerzo y la adaptación por parte de los inmigrantes como la voluntad y la capacidad de la sociedad de acogida para recibir y aceptar a los nuevos miembros. Por tanto, con el propósito de ver cómo se produce este tipo de integración<sup>4</sup>, los diferentes colectivos migrantes presentes en España han sido objeto de estudio de “numerosos investigadores de diversas

<sup>4</sup> Véase, por ejemplo, el trabajo de Paredes García & Sancho Pascual (2018) que se centra en la influencia de las expectativas de futuro de la población inmigrante de la Comunidad de Madrid en su proceso de integración social y sociolingüística.

disciplinas, que se han acercado al fenómeno con perspectivas y métodos también diversos” (Paredes García 2020: 43). En el ámbito de la lingüística, entre otros, se ha llegado a hablar de una *Lingüística de la migración* (Zimmerman & Morgenthaler García 2007) que se enfoca en el estudio de los fenómenos lingüísticos que ocurren como resultado de los movimientos migratorios de personas entre diferentes regiones, países o culturas<sup>5</sup> y que sirve para “determinar los efectos de la migración sobre la lengua y la comunicación entre los grupos migrantes y los de la sociedad receptora” (Paredes García 2020: 43); o también de una *sociolingüística de la migración* (Moreno Fernández 2009) en la que se aplican “los principios y métodos de la sociolingüística a las situaciones de contacto producidas por los movimientos migratorios” (Paredes García 2020: 43).

Para llevar a cabo esos estudios, se han recogido corpus lingüísticos relacionados con la migración, como el *Corpus Dinámico del Español de la Inmigración* que se diseñó en 2006 en la Universidad de Alcalá para estudiar la integración de la población migrante en España. Este corpus contiene “testimonios que permiten describir pautas de integración sociolingüística y comunicativa dentro de las diferentes comunidades de acogida” (Fernández López 2017). Se basa en “historias de vida y entrevistas” pertenecientes a “tres realidades lingüísticas bien diferenciadas, Madrid, Andalucía y Cataluña, y se estudian en cada punto de recogida de información grupos de inmigrantes hispanos y no hispanos” (idem). Además, este corpus se complementa con el proyecto IN.MIGRA2CM. *La población migrante de la Comunidad de Madrid: estudio multidisciplinar y herramientas para la integración lingüística* que recoge el corpus sociolingüístico de ecuatorianos, colombianos y dominicanos. También incluye materiales sobre la población migrante no hispanohablante residente en Alcalá de Henares (rumanos y polacos) y en la ciudad de Madrid (marroquíes, chinos y brasileños). En definitiva, tal y como subraya Molina Martos (2020: 84), estos materiales suponen “la base de estudios diversos que deberán permitir llegar a resultados específicos acerca de la integración sociolingüística”<sup>6</sup> como, por ejemplo, “determinar el significado social asociado a los usos lingüísticos y comunicativos” (Paredes García 2020: 43), o entender los comportamientos sociolingüísticos y los fundamentos de los mecanismos que los sustentan (Bravo García 2011).

Con este mismo objetivo se ha elaborado también el Corpus.migra.sev, constituido por muestras de entrevistas semidirigidas realizadas a hablantes de cuatro comunidades hispanas presentes en la ciudad de Sevilla, a saber, colombiana, ecuatoriana, nicaragüense y peruana<sup>7</sup>. Estos materiales permitirán examinar la diversidad lingüística en Sevilla y obtener información sobre las características del español utilizado por los migrantes, así como su proceso de adaptación e integración sociolingüística. Además, facilitarán comparaciones con el español hablado por

5 Dentro de este mismo campo, también se han llevado a cabo estudios sobre la presencia lingüística de los migrados procedentes de Hispanoamérica en textos escritos de lugares públicos en la ciudad de Sevilla (Pons Rodríguez 2011) o el paisaje lingüístico de la migración en el barrio de Ruzafa en Valencia (Gómez-Pavón Durán y Quilís Merín 2021), entre otros.

6 En este sentido contamos ya con estudios recientes “sobre la adecuación de emigrantes hablantes de español americano a comunidades españolas” (Manjón-Cabeza Cruz 2018: 69), esto es, ecuatorianos en Madrid (Sánchez Pascual 2014 y 2015) y Granada (Manjón-Cabeza Cruz 2018), dominicanos en Madrid (Molina Martos 2020), ecuatoguineanos en Granada (Manjón-Cabeza Cruz & Oyono Midje 2022), argentinos en Málaga (von Essen 2016, 2020, 2021, 2022, 2023) o peruanos en Sevilla (Repede 2022, 2023a, 2023b, 2024), pero también hablantes no hispanos, como rumanos, polacos y marroquíes en la ciudad de Granada (Manjón-Cabeza Cruz & Oyono Midje 2022, Fernández de Molina 2022, Sosinski 2022, Svetozarovová & Cruz Ortiz 2022 y Waluch de la Torre 2022).

7 Sobre este colectivo y sus particularidades sociolingüísticas, véase especialmente Repede (2022).

los sevillanos nativos, utilizando corpus similares, y contrastar estos datos con otros materiales previamente recolectados sobre la comunidad migrante en España.

Incluiremos a continuación una descripción general de las herramientas utilizadas para diseñar y construir el corpus: recogida de datos, transcripción y etiquetación de los materiales audio, clasificación social de los informantes, así como las diversas posibilidades para recuperar la información léxica, gramatical y contextual proporcionada por los materiales del corpus.

### 3. Corpus.migra.sev. Características generales

La estructura externa del Corpus.migra.sev (<https://corpusmigrasev.snlt.es/>) se compone de cinco partes generales, a saber: i) breve presentación del corpus, en la que se indican las comunidades de hablantes hispanos entrevistados (peruanos, ecuatorianos, colombianos y nicaragüenses) y los objetivos; ii) descripción lingüística, de las cuatro variedades presentes en la base de datos; iii) mapa con la distribución de los informantes según los diferentes barrios de Sevilla, iv) metodología con los criterios de elaboración del corpus, la relación de marcas aplicadas en la transcripción y el código asignado a cada informante en la base de datos; y, finalmente, v) el corpus propiamente dicho. Además, se añade un contacto a través de un formulario y/o correo electrónico.

El corpus está formado por 48 entrevistas semidirigidas, 12 por cada colectivo, de 40 minutos aproximadamente cada grabación. Las entrevistas corresponden a 24 mujeres y 24 hombres divididos en dos grupos de edad (1 = 20 a 44 años, 2 = 45 años en adelante) y tres niveles de estudios (primarios, medios y universitarios). En la versión accesible actualmente en internet (11 de abril de 2024) están disponibles 44 entrevistas y 12 transcripciones. Asimismo, hasta la fecha de redacción de estas líneas, el corpus recoge 258 412 palabras, distribuidas entre las cuatro muestras de habla hispana, con un número ligeramente menor en el subcorpus de entrevistas realizadas a informantes ecuatorianos (Tabla 3).

País	Palabras
Colombia	63781
Nicaragua	61243
Ecuador	47536
Perú	85852
Total	258412

**Tabla 2.** Recuento de palabras en el Corpus.migra.sev

Todas las entrevistas realizadas se han registrado en archivos en formato M4A, y ocasionalmente en MP3, mediante aplicaciones de grabación para dispositivo móvil. Asimismo, las grabaciones se han llevado a cabo en diferentes lugares de fácil acceso para los informantes (domicilio del encuestado, lugar de trabajo, tienda latina, cafetería, etc.), siempre procurando que no haya niveles elevados de ruido.

### 3.1 Recogida de datos

Tal y como se ha indicado previamente, el Corpus.migra.sev está constituido por entrevistas semidirigidas, basadas en un guion conversacional, realizadas a hablantes peruanos, colombianos, nicaragüenses y ecuatorianos residentes en la ciudad de Sevilla, grabadas entre los años 2020 y 2023. Las entrevistas se estructuran “en forma de relato de vida” (Paredes García 2020: 60) o historias de vida (Hernández Moreno 2009; Lucca Irizarry & Berríos Rivera 2009; Ruiz Olbuénaga 2012), esto es, “una perspectiva fenomenológica, la cual visualiza la conducta humana, lo que las personas dicen y hacen, como el producto de la definición de su mundo” (Chárriez Cordero 2012: 50). Mediante este tipo de entrevistas se busca reconstruir y transmitir los acontecimientos vividos por los individuos como actores y participantes en la vida social (Chárriez Cordero 2012). Específicamente en el Corpus.migra.sev, las entrevistas están orientadas a que los informantes expresen “sus sensaciones y opiniones acerca del proceso migratorio, desde su lugar de origen hasta su situación actual” (Paredes García 2020: 60), con la finalidad de tener una visión completa del escenario migratorio en el que están implicados los entrevistados.

Formalmente, se han tomado en consideración los siguientes tópicos generales de interacción:

Tópicos generales de conversación	Preguntas
Presentación	¿De dónde es? ¿De qué ciudad? ¿Cuántos años tiene?
Escenario tradicional y cultural del país/ciudad de origen	¿Qué recuerdos tiene de las fiestas que se celebran en su país y/o ciudad? ¿Cómo se celebra la Navidad, la Semana Santa, etc.? ¿Alguna fiesta típica de su ciudad? ¿Qué celebraciones familiares o culturales, tradiciones y rituales son importantes en su vida?
Diferencias y semejanzas entre el país de origen y España	¿En qué se diferencia la ciudad de origen de Sevilla? ¿Difiere el modo de celebrar las fiestas?
Fecha de llegada a Sevilla	¿Cuándo llegó a Sevilla? ¿Cómo? ¿Por qué eligió ir a España y no a otro país? ¿Por qué eligió venir a Sevilla y no a otra ciudad de España? ¿Vino para quedarse? ¿Suele ir de visita a su país? ¿Cada cuánto tiempo?
Trabajo	¿A qué se dedicaba en su país? ¿A qué se dedica ahora mismo? ¿Siempre ha trabajado en esto? ¿Le gusta su trabajo? ¿Tiene contrato de trabajo? / ¿cuándo ha estado trabajando / siempre ha tenido contrato o no? ¿Qué relación tiene con los compañeros (españoles, extranjeros)?
Estudios	¿Qué estudios tiene? ¿Qué estudió en su país? ¿Aquí en España ha podido completar / iniciar los estudios?
Familia	¿Está casada/o / está soltera/o? ¿Tiene hijas/hijos? ¿Cuántos años tienen más o menos? ¿A qué se dedican? ¿Estudian? ¿Trabajan? ¿Están aquí en Sevilla / España?



Tópicos generales de conversación	Preguntas
Barrio y vivienda	¿Dónde vive? ¿En qué barrio? ¿Descríbalo un poco? ¿Cómo se lleva con los vecinos? ¿Tiene piso o vive de alquiler? ¿Vive solo / con su familia o comparte piso?
Situaciones agradables y/o desagradables vividas en Sevilla / España	¿Qué opina sobre las leyes que hay en España / sobre el trabajo de los inmigrantes? ¿Están adaptadas las leyes o no? ¿Benefician al inmigrante o no? ¿Debería cambiar algo? ¿Cree que hay racismo en Sevilla / España? ¿Ha sufrido algún acto de racismo desde que vive en la ciudad?
Visión de futuro	¿Qué expectativas tiene de futuro? ¿Piensa quedarse en Sevilla / España o volver a su país? ¿Cómo se ve dentro de veinte años?

**Tabla 3.** Módulos temáticos de conversación en el Corpus.migra.sev

Se ha procurado en todo momento “crear una atmósfera segura y de confianza para cumplir con el propósito del estudio y evitar interrumpir al participante innecesariamente” (Chárriez Cordero 2012: 57), así como mantener una conversación que se vaya acercando progresivamente a los registros más espontáneos. Además, se ha intentado dar prioridad a los temas con los que el entrevistado se ha sentido más cómodo, si bien es cierto que, en el caso de que el investigador se ha percatado “de que el narrador demuestra poca profundidad o se ha desviado del propósito del estudio” (Chárriez Cordero 2012: 57) se ha recurrido “al uso de preguntas que le ayuden a redirigir el proceso para obtener la información deseada (Lucca Irizarry & Berríos Rivera, 2003)” (ibidem).

Siguiendo a Taylor & Bodgan (1984), Viruet (1997) y Atkinson (1998) en la realización de las entrevistas se han tenido en cuenta los siguientes principales aspectos (Chárriez Cordero 2012): a) explicación inicial del propósito y empleo de la entrevista por parte del entrevistador; b) el consentimiento informado de los participantes tanto previamente como posteriormente a la grabación; c) la retirada del participante en cualquier momento de la entrevista si así lo desea, además de garantizarle que la información recopilada no va a ser empleada con ninguna finalidad; d) el compromiso del equipo investigador a emplear los materiales tan solo con fines didácticos e investigadores y mantener el anonimato de los hablantes; e) la anonimización de la identidad de los informantes y de otras personas aludidas que puede aparecer en las entrevistas. En este caso, las referencias personales se han eliminado en el audio y se han sustituido por un silencio con la misma duración, empleando para ello el programa de tratamiento de sonido Wavepad<sup>8</sup>; mientras que en la transcripción se ha mantenido tan solo la inicial del nombre y el apellido del informante entrevistado o de terceros presentes o mencionados en la entrevista, como se muestra en el siguiente fragmento:

8 <https://www.nch.com.au/wavepad/es/index.html>



(1) E: bueno / ¿cómo te llamas / de dónde eres? / cuéntame

I: eh<alargamiento/> soy L L L / eh<alargamiento/> peruana / Trujillo de nacimiento / de la costa / eh<alargamiento/> del norte // me encanta mi país / llevo aquí diecisiete años / mmm pero me sigue gustando mi país / (Corpus – Corpus (snlt.es))

Este proceso “es laborioso pues la versión anonimizada no solo ha de mantener la coherencia interna que permita el seguimiento de la continuidad referencial del discurso sino que además ha de resultar verosímil en lo que se refiere a la información geográfica y demás indicaciones contextualizadoras” (Barcala *et al.* 2018: 219). Finalmente, mencionamos que no se han sometido al proceso de anonimización los nombres y apellidos de personas que pertenecen a la esfera pública, como políticos, escritores, etc., “siempre que las referencias no atañesen a su vida privada y no resultaran lesivas para su imagen ni atentasen contra su honor (Sampson 2000)” (Vázquez Rozas *et al.* 2020: 196).

## 3.2 La elaboración del corpus

### 3.2.1 Los metadatos

Otro componente fundamental en la configuración del Corpus.migra.sev es el conjunto de información que se ofrece con respecto a los participantes o la interacción. En este caso, cada grabación cuenta con dos tipos de fichas –una técnica y otra sociológica– que incluyen sistemáticamente información sobre las características de la grabación y del sujeto entrevistado. Concretamente, en la ficha técnica se registran aspectos como a) la situación del evento grabado (fecha, país, ciudad, lugar de la entrevista); b) el tipo de interacción: entrevista semidirigida; c) las propiedades del archivo (duración, formato, observaciones sobre la grabación); d) datos personales del entrevistador (nombre y apellidos, edad, sexo, estudios, profesión, país de origen), su relación con el informante (desconocidos previamente, relación de amistad, etc.) y el grado de formalidad (edad, estatus –superior, inferior, igualdad–); e) el proceso de transcripción: transcriptor/a y fecha de transcripción, revisor/a y fechas de revisión; y finalmente f) el número de palabras.

En la ficha sociológica se aportan datos de los sujetos participantes en las entrevistas, que permiten definir las principales características sociales de la muestra de informantes, así como hacer una aproximación al nivel de vida de cada uno de los encuestados. Se incluyen a) datos personales del informante (ciudad de origen, sexo, edad y nivel de instrucción); b) lenguas que dominan además de la lengua materna; c) profesión en su país de origen y en España (Sevilla); d) información sobre los ingresos mensuales; e) tiempo de residencia en España y especialmente en Sevilla; e) estado civil y posesión de la nacionalidad española; f) situación laboral y profesión; g) intención de volver a su país de origen, así como h) algunas observaciones relacionadas con el tipo de vivienda del entrevistado, viajes realizados, etc.

Ilustramos a continuación un ejemplo de cada una de las dos fichas explicadas previamente:

FICHA TÉCNICA DE LA GRABACIÓN	FICHA SOCIOLÓGICA DEL INFORMANTE
<b>Código del informante:</b> PSE_H212_004	<b>Origen:</b> Chimbote
<b>Fecha de grabación:</b> 10/08/2020	<b>Sexo:</b> hombre
<b>Lugar de la grabación:</b> tienda de la mujer del informante	<b>Edad:</b> 45 años
<b>Localidad:</b> Sevilla	<b>Hijos:</b> 1 (vive con sus padres en Sevilla)
<b>País:</b> España	<b>Estado civil:</b> casado (su mujer es peruana)
<b>Duración de la grabación:</b> 48'07"	<b>Grado de instrucción:</b>
<b>Tipo de archivo:</b> M4A	<ul style="list-style-type: none"> <li>▪ en Perú: estudios primarios</li> <li>▪ en España: ninguno</li> </ul>
<b>Observaciones sobre la grabación:</b> ruidos de fondo	<b>Idiomas:</b> ninguno
<b>Audiencia:</b> ninguna	<b>Tiempo de residencia en Sevilla:</b> 12 años
<b>Tipo de entrevista:</b> semidirigida	<b>Residencia en otras localidades de España:</b> no
<b>Datos personales del entrevistador:</b>	<b>Nacionalidad española:</b> sí
<ul style="list-style-type: none"> <li>▪ Nombre y Apellidos: Doina Repede</li> <li>▪ Sexo: Mujer</li> <li>▪ Edad: 36 años</li> <li>▪ País de origen: Rumania</li> <li>▪ Grado de instrucción: doctorado</li> <li>▪ Profesión: investigadora</li> </ul>	<b>Profesión:</b>
<b>Relación Entrevistador [E] / Informante [I]:</b>	<ul style="list-style-type: none"> <li>▪ en Perú: pesquero</li> <li>▪ en España: auxiliar de seguridad</li> </ul>
<ul style="list-style-type: none"> <li>▪ Formalidad (tenor y estatus): E &gt; I</li> <li>▪ Formalidad (tenor y edad): E &lt; I</li> <li>▪ Formalidad (tenor y grado de proximidad): desconocidos</li> </ul>	<b>Situación laboral:</b> parado
<b>Transcripción y revisión:</b> Doina Repede	<b>Ingresos mensuales:</b> menos de 1000 euros
<ul style="list-style-type: none"> <li>▪ Fecha de transcripción: 09/12/2020</li> <li>▪ Primera fecha de revisión: 27/02/2021</li> <li>▪ Segunda fecha de revisión: 17/04/2021</li> </ul>	<b>Tipo de vivienda:</b> propia
<b>Número de palabras:</b> 10 452	<b>Viajes por el extranjero:</b> no
	<b>Observaciones:</b> vive con familiares peruanos. No ha regresado a Perú y no tiene intención de volver a su país.

**Figura 2.** Muestra de ficha técnica y sociológica

Asimismo, en la ficha técnica de la grabación se indica el código que se le ha asignado al sujeto entrevistado dentro del corpus, que permite identificarlo según su comunidad de hablantes, sexo, edad, nivel de instrucción y tiempo de residencia en la ciudad de Sevilla. El código presenta las siguientes claves:

- a) Origen: C = Colombia; N = Nicaragua; E = Ecuador; P = Perú, seguido de la ciudad de residencia SE = Sevilla.
- b) Se añade un guion bajo ( \_ ) y el sexo/género: H = hombre; M = mujer.
- c) Grupo de edad: 1 = 20 a 44 años; 2 = 45 años en adelante.

- d) Nivel de educación: 1 = sin estudios / estudios primarios; 2 = estudios secundarios; 3 = estudios superiores.
- e) Tiempo de residencia: 1 = menos de 4 años de residencia; 2 = entre 5 y 14 años de residencia; 3 = más de 15 años de residencia.
- f) Se incluye otro guion bajo ( \_ ) y el número de la entrevista dentro del corpus (01, 02, 03, 04, 05...048).

Por ejemplo, PSE\_H212\_004 se interpreta como entrevista realizada a un hombre, de origen peruano, residente en Sevilla, del segundo tramo de edad, estudios primarios, con más de cinco años de residencia; el número 004 indica que se trata del hablante 4 de los 48 informantes entrevistados para el Corpus.migra.sev.

En líneas generales, tal y como afirma Vázquez Rozas (2014: 3), “los datos contextuales recopilados no solo sirven para identificar cada interacción sino que son imprescindibles para facilitar la posterior recuperación de la información, ya que permiten combinar propiedades del evento y elementos de la transcripción al formular los criterios de búsqueda”.

### 3.2.2 Transcripción y etiquetación del corpus

Como en cualquier otro corpus con las mismas características, “una parte crucial del tratamiento de los textos para su posterior utilización como fuente de datos lingüísticos consiste en su anotación (markup) y etiquetación (tagging)” (Barcala *et al.* 2018: 224). El sistema de transcripción del Corpus.migra.sev sigue la norma académica de escritura y acentuación del español con respecto a las unidades verbales, salvo para las mayúsculas, que se emplean únicamente para los nombres propios. Tal y como se ha mencionado previamente en cuanto a la anonimización de los materiales, los nombres de personas y de lugares que posibilitan la identificación del informante han sido sustituidos en la transcripción por alguna inicial mayúscula, sin punto detrás. Cada intervención del hablante –turno de habla o de apoyo– se ha considerado un turno diferente y se ha incluido en un párrafo aparte. Para indicar el papel de los interlocutores en la entrevista se ha empleado E para el entrevistador e I para el informante. Asimismo, se ha intentado reflejar adecuadamente los diversos elementos lingüísticos como “los marcadores de tipo fático e interjetivo, que cumplen una función interactiva fundamental para el desarrollo normal de las conversaciones” (Barcala *et al.* 2018: 222), pero que suelen plantear dudas de transcripción (*ibidem*). Es el caso de partículas como *ah*, *eh*, *uhm* o *uhum*, *pff*, etc. como se puede observar en los fragmentos (2) y (3). Cuando estas partículas manifiestan vacilación y funcionan como prolongadores, se ha optado por emplear la etiqueta <alargamiento/> (4).



- (2) eso lo<alargamiento/> tú vas por la esquina / y todo el mundo te conoce / el barrio / la gente / la juerga / las jergas / las jergas son<alargamiento/> fantásticas / tú para ir / dice <cita> no / tú a la esquina ¿por dónde vas </cita> / <cita> voy de frente </cita> / no allí no te dicen voy de frente / allí te dicen <cita> de fresa / de fresa </cita> || E: *ah* ¿sí? || I: sí / o sea ¿me entiendes? || E: sí sí || I: es <vacilación/> es campechano / es / me encanta //



(3) esos son <vacilación/> son costumbre que<alargamiento/> bueno tienes que aprender / pero bueno / a la larga bueno / mmm yo entré como ayudante / y me pusieron como encargado / porque era una persona / me ha dicho <vacilación/> porque iba todos los días a cumplir / además ese / argentinos <vacilación/> argentinos son un poco más<alargamiento/>s / ¿cómo te puedo explicar? / son un poco más vagos /



(4) allí encuentras más oportunidades de trabajo / de estudio / y<alargamiento/> de repente mejora también tu calidad de vida / y ese conglomerado hace que haya más tráfico / eh<alargamiento/> te mueres obviamente en <vacilación/> en cruzarte de tu piso al <vacilación/> al trabajo / que haya más informalidad en el tema del trabajo / sube el empleo / eh<alargamiento/> delincuencia

Cabe señalar también que en todo momento se han respetado las particularidades de cada una de las cuatro variedades presentes en los materiales del corpus, así como del nivel de instrucción de los informantes.

La transcripción de las grabaciones se ha realizado de forma manual y se ha incluido un conjunto limitado de marcas y etiquetas pertenecientes al Standard Generalized Markup Language (SGML) contenidas en el Text Encoding Initiative (TEI), que aportan información relevante sobre las características de algunos elementos lingüísticos, paraverbales e incluso contextuales (Tabla 4; ejemplo 5). Dado el carácter oral de los materiales, no se emplean signos de puntuación, sino una serie de marcas que sirven para señalar pausas de diferente duración. Aparte de lo indicado, se incluyen los signos de interrogación y exclamación, las etiquetas de ruidos (comunicativos y provenientes del ambiente más cercano), de risas, de forma fónica (alargamiento, palabra cortada, secuencia cortada, vacilación, solapamientos, transcripción dudosa, ininteligible), de léxico (término, siglas), de cita, una etiqueta de observaciones que recoge información contextual o explicaciones complementarias necesarias para comprender la transcripción y la grabación y, por último, el tiempo marcado generalmente cada dos minutos al comienzo del turno de habla del entrevistado, con la finalidad de agilizar y facilitar la consulta de la transcripción. No obstante, dado que en ocasiones el parlamento del encuestado dura más de dos minutos, se ha marcado el tiempo dentro de la intervención.

Marcas y etiquetas		Explicación
<i>Barra</i>	/	Pausa breve, similar a la coma en la escritura
<i>Barra doble</i>	//	Pausa algo más larga, similar al punto
<i>Signo de exclamación</i>	¡!	Indica los enunciados que, a juicio del transcriptor, se ajustan a los parámetros entonativos habituales de las exclamaciones
<i>Signo de interrogación</i>	¿?	Marca los enunciados interrogativos
<i>Silencio</i>	<silencio/>	Ausencia de emisión superior a la habitual en el parlamento del hablante

Marcas y etiquetas		Explicación
<i>Ininteligible</i>	<ininteligible/>	Reproducción de discurso que no se ha podido entender
<i>Transcripción dudosa</i>	<transcripción_dudosa> </transcripción_dudosa>	Transcripción dudosa para el transcriptor y revisores
<i>Solapamientos</i>	<simultáneo> </simultáneo>	Superposición de los fragmentos de discurso que el entrevistador y el informante han producido de forma simultánea
<i>Palabra cortada</i>	<palabra_cortada/>	El hablante deja una palabra inacabada y reformula su discurso
<i>Secuencia cortada</i>	<secuencia_cortada/>	Turno de habla interrumpido repentinamente por alguno de los interlocutores
<i>Alargamiento</i>	<alargamiento/>	En el proceso de construcción espontánea del discurso, se alarga la palabra o parte de ella
<i>Vacilación</i>	<vacilación/>	Breve titubeo o repetición de la misma palabra o estructura
<i>Sic</i>	<sic> </sic>	Errores de dicción que no deben interpretarse como descuido del transcriptor
<i>Ruido</i>	<ruido = ""/>	Ruidos, tanto emitidos por alguno de los interlocutores (tos, suspiro, carraspeo, etc.) como procedentes del entorno en el que se realiza la entrevista, que interfieren en la transcripción  <ruido = "chasquido_boca"/> ; <ruido = "moto"/>; <ruido = "palmadas"/>
<i>Risas</i>	<risas = ""/>	Incluye los casos de risas y entre risas y se señala al interlocutor que las protagoniza
<i>Cita</i>	<cita> </cita>	El hablante reproduce en estilo directo palabras propias o ajenas
<i>Interrupción de grabación</i>	<interrupción_grabación/>	Durante la entrevista se interrumpió puntualmente la recogida del audio
<i>Observación complementaria</i>	<observación complementaria/>	Se incluye información adicional relacionada con los gestos, movimientos, etc. realizados por el informante o por terceros presentes en la entrevista
<i>Tiempo</i>	<tiempo = "00:00"/>	Proporciona información del minuto y del segundo de la entrevista en el que se encuentra el parlamento que precede

**Tabla 4.** Lista de marcas y etiquetas usadas en el Corpus.migra.sev



(5) I: y la inmigración también de bastante venezolanos que ha influido también / lamentablemente en la delincuencia peruana / que ya la teníamos pero ahora se ha incrementado mucho más /

E: uhum

I: y<alargamiento/> el tema <ruido = “chasquido\_boca”/> en general del Perú / de Lima / el tema de la corrupción es que también nos <vacilación/> no<palabra\_cortada/> nos golpea bastante ¿no? / en general a todos //

E: uhum

I: <tiempo = “03:58”/> somos <vacilación/> somos / porque es obvio / que también me incluyo / este <vacilación/> es en estas sociedades somos muy buenos / muy hábiles para <ininteligible/> / pero también estamos de la mano con la corrupción / estamos con el tema de la delincuencia / entonces es una mezcla de <vacilación/> de muchas cosas //

Mencionamos también que este sistema de etiquetado y transcripción empleado en el Corpus.migra.sev se acerca al utilizado en el corpus PRESEEA-Sevilla<sup>9</sup> (Repede 2019a, 2019b, 2019c) que contiene entrevistas semidirigidas realizadas a informantes sevillanos (hombres y mujeres) con nivel de instrucción alta, media y baja, y diferentes tramos de edad (primera generación de los jóvenes, de 20 a 34 años; la segunda, de las personas adultas o de mediana edad, de 35 a 54 años; y la tercera, los mayores, de 55 años en adelante), con la finalidad de poder llevar a cabo estudios comparativos entre los hispanos residentes en la ciudad de Sevilla y los hablantes vernáculos.

### 3.3 Búsqueda de información en el Corpus.migra.sev

Para utilizar los materiales y datos que ofrece Corpus.migra.sev se ha diseñado una aplicación de consulta, a la que se accede mediante previo registro disponible en <https://corpusmigrasev.snlt.es/lista-corpus/>.

Esta aplicación permite a) realizar búsquedas según variables sociales (edad, nivel de estudios, sexo, ciudad y país de origen y tiempo de residencia en Sevilla); b) acceder directamente al audio alineado; c) consultar los fragmentos de audio correspondientes a los resultados de búsqueda; y d) descargar las transcripciones con las etiquetas incluidas en formato Word, Excel, pdf, html o incluso imprimirlas. Por otra parte, la aplicación permite combinar libremente las distintas variables sociales (Figura 3), abriendo “la posibilidad de trabajar con corpus virtuales a la medida del usuario” (Barcala *et al.* 2018: 230). Además, es posible elegir búsqueda de elementos léxicos o gramaticales que interesen en cada momento, cruzados con cualquiera de las variables sociales incluidas en el corpus (Figura 4). Todo ello permitirá la recuperación de formas lingüísticas, que aparecerán marcadas con color dentro del corpus, para más fácil manejo por parte del usuario, así como el cómputo total de cada una de ella dentro de cada transcripción.

9 <https://sociolingüísticaandaluza.snlt.es/corpus-preseea-sevilla>

Sexo = Hombre And (Edad = < 44 años) And (Estudios = Primarios) And (T.residencia = < 4 años) And (Ciudad = Lima) And (País = Perú)

Código	Sexo	Edad	Estudios	T.residencia	Ciudad	País
<input checked="" type="checkbox"/> PSE_H111_001	Hombre	< 44 años	Primarios	< 4 años	Lima	Perú

Transcripción:  
E: buenas tardes  
I: ¿qué tal?  
E: buenas tardes  
E: que calor hace ¿verdad?  
I: ¡sí! / está que quema / ... más

**Figura 3.** Ejemplo de combinación de parámetros sociales

Exportar Imprimir

ito

Código	Sexo	Edad	Estudios	residencia	Ciudad	País
<input type="checkbox"/> PSE_H111_001	Hombre	< 44 años	Primarios	< 4 años	Lima	Perú
<input type="checkbox"/> PSE_H131_002	Hombre	< 44 años	Superiores	< 4 años	Lima	Perú

Transcripción:  
... abierto / con la gente de afuera //  
E: ¿sí? no  
I: sí / eso / es muy bonito / muy bonito también  
E: uhum  
I: con todos sus contrastes... más

**Figura 4.** Ejemplo de búsqueda en el *Corpus.migra.sev*

## 4. Consideraciones finales

En este trabajo se ha presentado el *Corpus.migra.sev*, un corpus de lengua oral del español de los migrantes residentes en la ciudad de Sevilla, más concretamente, hablantes peruanos, colombianos, ecuatorianos y nicaragüenses. En el apartado §2 se ha reflexionado sobre la necesidad y la utilidad de un corpus de estas características, con un claro interés para los estudios de la investigación lingüística y sociológica de la migración. *Corpus.migra.sev* viene, en efecto, a unirse a otros corpus que registran las variedades del español de las diferentes comunidades hispanohablantes que residen en España.

El corpus presentado se ofrece en su totalidad, abierto a la consulta de cualquier persona interesada, que debe citar la procedencia de los datos. Sintetizamos a continuación las diversas posibilidades de empleo de este corpus en línea: a) está disponible para su descarga en formato textual y etiquetado; b) permite el acceso libre y directo a los audios correspondientes; c) posibilita el contraste entre el registro audio y la transcripción y etiquetación incluida, así como la consulta inmediata del fragmento de audio correspondiente a cada búsqueda; d) ofrece información sociolingüística de los hablantes e información técnica de cada una de las grabaciones realizadas;



y, por último, e) se puede manejar según las necesidades particulares de los usuarios combinando criterios sociales con elementos lingüísticos.

Actualmente seguimos trabajando en la introducción y la transcripción de entrevistas semidirigidas, pero también en la mejora de algunos aspectos de la codificación para mayor aprovechamiento de estos materiales por parte de los usuarios interesados.

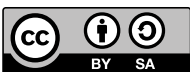
## Referencias bibliográficas

- Atkinson, R. (1998). *The life story interview*. London: Sage Publications.
- Barcala, M.; Domínguez, E.; Fernández, A.; Rivas, R.; Santalla, M.<sup>a</sup> P.; Vázquez, V.; & Villapol, R. (2018). El corpus ESLORA de español oral. Diseño, desarrollo y explotación. *Chimera. Romance Corpora and Linguistic Studies*, 5, 2, 217-237. <<https://revistas.uam.es/chimera/article/view/chimera2018.5.2.003>>
- Bartolomé, M. (2002). *El reto de las migraciones en sociedades multiculturales, desde la perspectiva de la educación*. Salamanca: Publicaciones Universidad Pontificia de Salamanca.
- Bravo-García, E. (2011). Con la mirada en el otro: implicaciones sociolingüísticas de la inmigración hispanoamericana en Andalucía. *Itinerarios: revista de estudios lingüísticos, literarios, históricos y antropológicos*, 13, 131-153.
- Chárriz Cordero, M. (2012). Historias de vida: Una metodología de investigación cualitativa. *Revista Griot*, 1, 50-67.
- Colectivo Ioé (2002). ¿Cómo abordar el estudio de las migraciones? Propuesta Teórico-Metodológica. In F. Checa (Ed.), *Las Migraciones a debate. IV Congreso de Inmigración Africana*, Almería 6-8 de abril de 2000 (pp.17-54). Barcelona: Icaria.
- . (2010). Discursos de la población migrante en torno a su instalación en España. Exploración cualitativa. *Opiniones y Actitudes*, 64. Madrid: Centro de Investigaciones Sociológicas.
- Fernández de Molina, E. (2019). Tabú e imagen social: estrategias comunicativas usadas por los ecuatorianos residentes en Granada para hablar del ámbito laboral. *Lengua y migración*, 11, 2, 7-34. <<https://ebuah.uah.es/dspace/handle/10017/41148>>
- . (2022). La pronunciación de la /s/ implosiva de los rumanos y polacos residentes en Granada. ¿Un caso de acomodación lingüística? *Círculo de Lingüística Aplicada a la Comunicación*, 92, 13-28. <<https://dx.doi.org/10.5209/clac.83918>>
- Fernández López, M.<sup>a</sup> C. (2017). Integración de la población inmigrante en los centros escolares de la Comunidad de Madrid: Análisis de necesidades formativas y curriculares. *Actas del III Congreso Internacional SICELE. Investigación e innovación en ELE. Evaluación y variedad lingüística del español*. <[http://cvc.cervantes.es/ensenanza/biblioteca\\_ele/sicele/019\\_fernandezlopez.htm](http://cvc.cervantes.es/ensenanza/biblioteca_ele/sicele/019_fernandezlopez.htm)>
- Ferrer, A.; & Urdiales, M.<sup>a</sup> E. (2004). Características de la población extranjera en España. *Geo Crítica / Scripta Nova. Revista electrónica de geografía y ciencias sociales*, 8, 157-180.
- García Cívico, J. (2011). Sobre el proceso de integración social del inmigrante en España. La cuestión de los indicadores. *Revista Electrónica del Instituto de Investigaciones "Ambrosio L. Gioja"*, 7, 50-77.

- Gómez-Pavón Durán, A.; & Quilis Merín, M. (2021). El paisaje lingüístico de la migración en el Barrio de Ruzafa en Valencia: una mirada a través del tiempo. *Cultura, Lenguaje y Representación*, 25, 135-154. <<http://dx.doi.org/10.6035/clr.2021.25.8>>
- Grupo UAH-LIVARES (2023). *Corpus Dinámico del Español de la Inmigración* (CORDIESIN). <<http://corpuscordiesin.uah.es>>
- Hernández, K. S. (2009). El método historia de vida: alcances y potencialidades. <<http://www.gestiopolis.com/economia/metodo-de-investigación-cualitativa.htm>>
- Lucca Irizarry, N.; & Berrios Rivera, R. (2009). *Investigación cualitativa. Fundamentos, diseños y estrategias*. Puerto Rico: Ediciones SM.
- Manjón-Cabeza, A. (2018). Una acomodación lingüística problemática: los diminutivos en los emigrantes ecuatorianos en Granada (España). *Itinerarios: revista de estudios lingüísticos, literarios, históricos y antropológicos*, 28, 67-80.
- Manjón-Cabeza Cruz, A.; & Oyono Midje, E. (2022). Uso de los apreciativos en cinco grupos de migrantes: adecuación local y prestigio global. *Círculo de Lingüística Aplicada a la Comunicación*, 92, 29-45. <https://dx.doi.org/10.5209/clac.83913>.
- Martínez Buján, R.; & Golías Pérez, M. (2005). La latinoamericanización de la inmigración en España. *Cuadernos geográficos*, 36, 51-64.
- Molina Martos, I. (2020). Inmigrantes colombianos en Madrid: actitudes lingüísticas y pautas de integración social. *Lengua y migración*, 12, 1. Monográfico, 83-102. <<http://ym.linguas.net/Download.axd?type=ArticleItem&id=249>>
- Moreno Fernández, F. (2009). Integración sociolingüística en contextos de inmigración: marco epistemológico para su estudio en España. *Lengua y Migración*, 1, 1, 121-156.
- Paredes García, F. (2020). Un modelo para el análisis de la integración sociolingüística de la población migrante: fundamentos, dimensiones e instrumentos. *Lengua y migración*, 12, 1. Monográfico, 39-81. <<https://erevistas.publicaciones.uah.es/ojs/index.php/lenguaymigracion/article/view/65>>
- Paredes García, F.; & Sancho Pascual, M.<sup>a</sup> (2018). Influencia de las expectativas de permanencia o retorno en la integración sociolingüística de la población migrante en la comunidad de Madrid. *RILL. Revista Internacional de Lingüística Iberoamericana*, 16, 1, 41-68.
- Pons Rodríguez, L. (2011). Hispanoamérica en el paisaje lingüístico de Sevilla. *Itinerarios: revista de estudios lingüísticos, literarios, históricos y antropológicos*, 13, 97-129.
- Repede, D. (Ed.) (2019a). *El español hablado en Sevilla. Corpus PRESEEA-Sevilla*. Vol. I. Hablantes de instrucción alta. Sevilla: Editorial Universidad de Sevilla.
- . (2019b). *El español hablado en Sevilla. Corpus PRESEEA-Sevilla*. Vol. II. Hablantes de instrucción baja. Sevilla: Editorial Universidad de Sevilla.
- . (2019c). *El español hablado en Sevilla. Corpus PRESEEA-Sevilla*. Vol. III. Hablantes de instrucción media. Sevilla: Editorial Universidad de Sevilla.
- Repede, D. (2022). *Corpus oral para el estudio sociolingüístico de hablantes peruanos en la ciudad de Sevilla (PERUSEV)*. Sevilla: Editorial Universidad de Sevilla.
- . (2023a). El marco introductor de discurso directo mediante los marcadores del discurso: estudio contrastivo entre PRESEEA-Sevilla y PERUSEV. *Oralia. Análisis del Discurso Oral*, 26, 1, 91-118. <<https://doi.org/10.25115/oralia.v26i1.8269>>

- . (2023b). Estudio piloto sobre los marcadores del discurso como introductores de cita directa en el corpus oral PERUSEV. *Lingüística*, 39, 2, 11-32. <<https://doi.org/10.5935/2079-312x.20230015>>
- . (2024). Estrategias introductoras y tipo de reproducción de las citas directas en el corpus oral PERUSEV. *Revista Española de Lingüística Aplicada*, 1-29. <<https://doi.org/10.1075/resla.22067.rep>>
- Ruiz Olabuénaga, J. I. (2012). Historias de vida. In *Metodología de la Investigación Cualitativa* (pp. 267-313). Bilbao: Universidad de Deusto.
- Sampson, G. (2000). CHRISTINE Corpus: Documentation. <<http://www.grsampson.net/ChrisDoc.html>>
- Sancho Pascual, M.<sup>a</sup> (2014). *Integración sociolingüística de los inmigrantes ecuatorianos en Madrid*. Alcalá: Universidad de Alcalá de Henares.
- . (2015). Integración sociolingüística de los inmigrantes ecuatorianos en Madrid: datos sobre el diminutivo según los corpus Ispie-Madrid y Preseea-Madrid. In A. M.<sup>a</sup>, I. Cestero, Molina, & F. Paredes (Eds.), *Patrones sociolingüísticos de Madrid* (pp. 155-176). Berna: Peter Lang.
- Sosinski, M. (2022). Realización de consonantes españolas por inmigrantes polacos en Granada. Estudio contrastivo. *Círculo de Lingüística Aplicada a la Comunicación*, 92, 63-86. <<https://doi.org/10.5209/clac.83916>>
- Svetozarovová, R.; & Cruz Ortiz, R. (2022). Acomodación sociolingüística de los inmigrantes rumanos a la norma local: la /d/ intervocálica. *Círculo de Lingüística Aplicada a la Comunicación*, 92, 87-107. <<https://dx.doi.org/10.5209/clac.83917>>
- Taylor, S. J.; & Bogdan, R. (1984). *Qualitative research method: The search for meanings*. New York: John Wiley.
- Valero-Matas, J. A.; Coca, J. R.; & Valero-Oteo, I. (2014). Análisis de la inmigración en España y la crisis económica. *Papeles de población*, 20, 80, 9-45. <[http://www.scielo.org.mx/scielo.php?script=sci\\_arttext&pid=S1405-74252014000200002](http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1405-74252014000200002)>
- Vázquez Rozas, V. (2014). ESLORA: Diseño, codificación y explotación de un corpus oral de español de Galicia. *II Workshop de Procesamiento Automatizado de Texto y Corpus* [en línea]. Valparaíso. <[http://gramatica.usc.es/~vvazq/pdf\\_public/corpus\\_eslora\\_pres.pdf](http://gramatica.usc.es/~vvazq/pdf_public/corpus_eslora_pres.pdf)>
- Vázquez Rozas, V.; Domínguez Noya, E.; Fernández Sanmartín, A.; Rojo, G.; & Santalla, M.<sup>a</sup> P. (2020). Codificación y anotación del habla en un contexto bilingüe: el corpus ESLORA de español de Galicia. In A. Gallego, & F. Roca, F. (Eds.), *Dialectología digital del español* (pp. 189-224). Santiago de Compostela: Universidad Santiago de Compostela.
- Viruet, E. (1997). El uso de la técnica de la historia de la vida en la investigación psicosocial. *Cruz Anasta*, 20, 33-43.
- von Essen, M.<sup>a</sup> C. (2016). Variedades del español en contacto: acomodación sociolingüística de una comunidad de inmigrantes argentinos en la ciudad de Málaga. Análisis acústico de las variantes alofónicas de /j/. *Lengua y migración*, 8, 2, 7-43.
- . (2020). On the different ways of being a bidialectal immigrant: The case of Argentineans in Spain. *Lengua y migración*, 12, 2, 7-43.
- . (2021). *Identidad y contacto de variedades. La acomodación lingüística de los inmigrantes rioplatenses en Málaga*. Berlín: Peter Lang.
- . (2022). La influencia del entorno fonológico en la caracterización fonética de la consonante ye en contextos de inmigración. *Lingüística en la red (LinRed)*, Estudios de variación lingüística: Homenaje a Juan Andrés Villena Ponsoda, 19, 1-20.
- . (2023). Identity, language socialization, and family language policies in dialect contact: the case of Argentinean immigrants in Malaga, Spain. *Journal of World Languages*, 9, 1, 133-160.

- Waluch de la Torre E. (2022). Realización de vocales españolas por hablantes nativos e inmigrantes polacos. *Círculo de Lingüística Aplicada a la Comunicación*, 92, 109-125. <<https://doi.org/10.5209/clac.83914>>
- Zimmerman, K.; & Morgenthaler García, L. (2007). Introducción: ¿Lingüística y migración o lingüística de la migración?: De la construcción de un objeto científico hacia una nueva disciplina. *RILLI. Revista Internacional de Lingüística Iberoamericana*, 2, 10, 7-19.



This work can be used in accordance with the Creative Commons BY-SA 4.0 International license terms and conditions (<https://creativecommons.org/licenses/by-sa/4.0/legalcode>). This does not apply to works or elements (such as images or photographs) that are used in the work under a contractual license or exception or limitation to relevant rights.