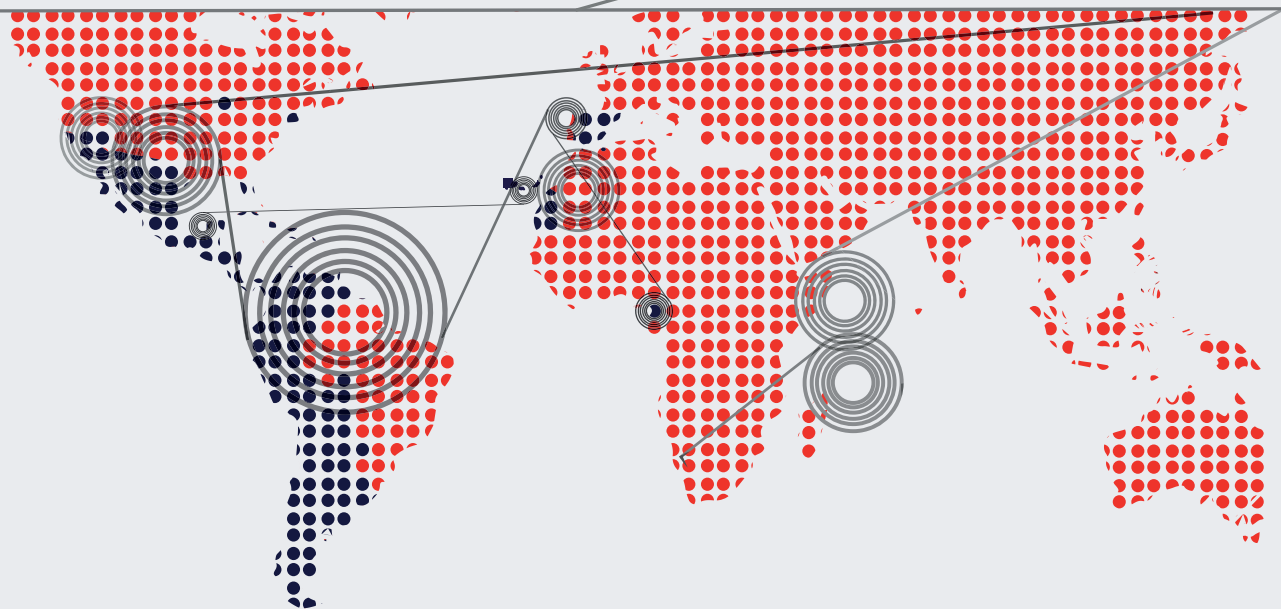


# La sociolingüística del español en España y América: 25 años de estudios coordinados



MONOGRAFÍAS  
HUMANIDADES 98

Isabel Molina Martos  
Pedro Martín Butragueño  
Gema Cuesta Chorro  
Nicolás Moreno Martín de Nicolás  
(eds.)

UAH



**La sociolingüística del español  
en España y América:  
25 años de estudios coordinados**

**UAH MONOGRAFÍAS  
HUMANIDADES 98**



**Comunidad  
de Madrid**

Dirección General  
de Investigación  
e Innovación Tecnológica

CONSEJERÍA DE EDUCACIÓN,  
CIENCIA Y UNIVERSIDADES

Acción financiada por la Comunidad de Madrid en el marco del Convenio Plurianual con la Universidad de Alcalá en la línea de actuación *Estímulo a la Excelencia para Profesores/as Universitarios/as Permanentes* (EPU-DPTO/2020/010)



MINISTERIO  
DE CIENCIA, INNOVACIÓN  
Y UNIVERSIDADES

**el** AGENCIA  
ESTATAL DE  
INVESTIGACIÓN

Esta publicación es parte del proyecto de I+D+i *Agenda 2050. El español del centro-norte de España: procesos de variación y cambio espaciales y sociales* (PID2019-104982GB-C51), financiado por MICIU/AEI/10.13039/501100011033

# La sociolingüística del español en España y América: 25 años de estudios coordinados

Edición de  
Isabel Molina Martos  
Pedro Martín Butragueño  
Gema Cuesta Chorro  
Nicolás Moreno Martín de Nicolás



Universidad  
de Alcalá

EDITORIAL  
UNIVERSIDAD DE ALCALÁ

La colección de Monografías de Humanidades de la Editorial de la Universidad de Alcalá ha sido distinguida con el Sello de Calidad en Edición Académica - Academic Publishing Quality (CEA-APQ).



El contenido de este libro no podrá ser reproducido, ni total ni parcialmente, sin el previo permiso escrito del editor. Todos los derechos reservados.

© De los textos: sus autores.  
© De las imágenes: sus autores.  
© De la ilustración de portada: Ronda Vázquez Martí  
© Editorial Universidad de Alcalá, 2025  
Plaza de San Diego, s/n  
28801 Alcalá de Henares  
[www.uah.es](http://www.uah.es)

I.S.B.N.: 978-84-10432-52-9

Composición: Solana e Hijos, A. G., S.A.U.  
Impresión y encuadernación: Solana e Hijos, A.G., S.A.U.  
Impreso en España

*A Francisco Moreno Fernández,  
que durante tres décadas, con gran dedicación y el tesón que lo caracterizan,  
nos ha movido a conocer la naturaleza variable y social de la lengua,  
como se refleja en los trabajos recogidos en este libro,  
otro nuevo fruto de su PRESEEA*

*y a los maestros ausentes:*

*Humberto López Morales  
José Antonio Samper Padilla  
Clara Eugenia Hernández Cabrera  
Paola Bentivoglio*



## ÍNDICE

<b>Introducción: 25 años de investigación sociolingüística coordinada en España y América</b> <i>Pedro Martín Butragueño e Isabel Molina Martos</i> .....	11
<b>Evolución lingüística en el curso de la vida. A propósito de PRESEEA</b> <i>Francisco Moreno Fernández</i> .....	25
<b>What does it take to be a native speaker (and who gets to say so)</b> <i>Dennis R. Preston</i> .....	51
<b>Variación e invariación morfosintácticas del español. Apuntes desde PRESEEA</b> <i>Carmen Silva-Corvalán</i> .....	75
<b>PRESEEA EN AMÉRICA</b>	
<b>Construcción del PRESEEA-Buenos Aires y análisis de datos</b> <i>Claudia Borzi</i> .....	97
<b>PRESEEA-La Habana: breve historia y resultados</b> <i>Ana María González Mafud</i> .....	111
<b>Elisión de /d/ intervocálica: aspectos de una tendencia hispánica</b> <i>María Claudia González-Rátiva</i> .....	123
<b>Hacia un modelo de la ciudad lingüística: estructuración sociolingüística básica de Ciudad de México a inicios del siglo XXI</b> <i>Pedro Martín Butragueño</i> .....	135
<b>El equipo de PRESEEA de Santiago de Chile: situación actual y desafíos a futuro</b> <i>Abelardo San Martín Núñez</i> .....	157

## PRESEEA EN ESPAÑA

<b>PRESEEA-Pamplona. Presentación y primeros estudios</b>	
<i>María José Gallucci y Carmela Pérez-Salazar</i> .....	169
<b>El PRESEEA en Valencia. Historia y resultados</b>	
<i>José Ramón Gómez Molina</i> .....	185
<b>Procesos de variación y cambio en Madrid: el papel de las identidades en la formación de regiolectos</b>	
<i>Isabel Molina Martos</i> .....	203
<b>El grupo PRESEEA-Granada: estudios, resultados y sugerencias hacia el futuro</b>	
<i>Juan Antonio Moya Corral</i> .....	217
<b>Una mirada al pasado, presente y futuro del proyecto PRESEEA en las Palmas de Gran Canaria</b>	
<i>Marta Samper Hernández</i> .....	225
<b>PRESEEA-Sevilla: 10 años de sociolingüística en la capital hispalense</b>	
<i>Juana Santana Marrero</i> .....	239
<b>Contribución del grupo PRESEEA-Málaga a la construcción del mapa sociolingüístico de variación y cambio en España desde un enfoque comparativo y multidimensional</b>	
<i>Juan Andrés Villena y Antonio Manuel Ávila</i> .....	255

**INTRODUCCIÓN:**  
**25 AÑOS DE INVESTIGACIÓN SOCIOLINGÜÍSTICA COORDINADA**  
**EN ESPAÑA Y AMÉRICA**

Pedro Martín Butragueño  
*El Colegio de México*

Isabel Molina Martos  
*Universidad de Alcalá*

En 2024 ha transcurrido más de un cuarto de siglo desde que se pusiera en marcha la iniciativa de abordar la investigación sociolingüística del español de España y América de una manera coordinada. Ya en 1992 Carmen Silva Corvalán había puesto sobre la mesa la necesidad de contar con un corpus de datos que sirviera como base para llevar a cabo un estudio sociolingüístico del español en toda su extensión y diversidad geográfica (Silva-Corvalán, 1994), pero fue Moreno Fernández quien formalizó públicamente esta iniciativa en abril de 1993, en el X Congreso Internacional de la Asociación de Lingüística y Filología de la América Latina (ALFAL) (Moreno Fernández, 1993). Tres años más tarde, en el Congreso de la ALFAL celebrado en las Palmas de Gran Canaria, se presentaba el primer borrador de la metodología que sentaría las bases para el desarrollo del macro corpus que hoy conocemos bajo las siglas de PRESEEA: «Proyecto para el Estudio Sociolingüístico del Español de España y América» (Moreno Fernández, 1996). En los años transcurridos desde que PRESEEA diera sus primeros pasos, se ha consolidado como el mayor corpus de datos del español de las principales capitales del mundo hispanohablante recogido sobre la misma base metodológica de preestratificación sociolingüística. Es de justicia destacar los nombres de Francisco Moreno Fernández, Carmen Silva-Corvalán y Humberto López Morales como responsables de esta idea de colaboración y de investigación conjunta que no ha dejado de dar frutos a lo largo del tiempo. PRESEEA se concibió como un proyecto coordinado que persigue la comparabilidad

de los estudios y el intercambio de materiales e informaciones. Esta idea de cooperación para estudiar la lengua española hoy día parece evidente, cuando el avance del conocimiento ya no puede concebirse desde la individualidad y cada vez más se reivindica el acceso abierto a la información científica, pero debe reconocerse que estos planteamientos entonces no eran tan evidentes. Cuando la Sociolingüística del español comienza a dar sus primeros pasos entre 1970 y 1980, lo hace en forma de monografías referidas a una comunidad de habla, de una manera semejante a como habían surgido, un siglo antes, los primeros trabajos dialectológicos. La Dialectología avanzó desde los estudios monográficos, del tipo: «el habla de X», hasta trabajos de amplia extensión geográfica, como fueron los atlas lingüísticos de gran dominio. Unas décadas después, la Sociolingüística ha tenido que hacer un recorrido similar, evolucionando desde el estudio de comunidades de habla aisladas a las investigaciones coordinadas como las que se ofrecen en el marco del PRESEEA, que conciben la investigación como una red en la que cada equipo aporta su investigación y a cambio puede recibir la del resto de los participantes en el estudio coordinado.

Esta iniciativa surgida con el propósito de desarrollar un macrocorpus sociolingüístico del español ha experimentado una evolución en sus objetivos y en el modo de abordar sus trabajos a lo largo del tiempo transcurrido. Así, si bien la primera etapa de PRESEEA fue empírica y descriptiva y consistió en construir un macro corpus sociolingüístico<sup>1</sup> de las variedades del español siguiendo la metodología establecida por PRESEEA, en la segunda etapa se planteó el objetivo de establecer los patrones sociolingüísticos de las comunidades de habla urbanas. Para ello, los equipos que ya habían cubierto la primera fase y contaban con un corpus de su comunidad de habla pasaron a analizar las variables sociolingüísticas fónicas, morfosintácticas, léxicas y discursivas más significativas con el propósito de dar cuenta de esos patrones<sup>2</sup>. Se analizaron así diferentes procesos variables característicos del mundo hispánico, como son las consonantes en proceso de erosión, tanto en el ataque como en la coda silábica: /d/ intervocálica, /s/ distensiva; la escisión de las dentales fricativas en andaluz o el yeísmo. En el nivel de la morfosintaxis, en el marco del PRESEEA se han estudiado procesos tales como la variación de los pronombres personales átonos en función de complemento; el queísmo y el dequeísmo; la personalización de *haber* o la expresión del sujeto nominal y pronominal. Se han abordado, asimismo, estudios comparativos de la variación léxica disponible en determinados grupos sociales; la convergencia y atrición del léxico dialectal; el léxico del tabú y el eufemismo o la fraseología. Por último, también se ha dedicado una atención especial a algunas variables pragmáticas como los mecanismos de atenuación; los apéndices interrogativos de control de contacto; la expresión del tabú o los usos y funciones del diminutivo.

---

<sup>1</sup> El tamaño del corpus depende de cada ciudad, pero como mínimo reúne 54 entrevistas semidirigidas, estratificadas por género, edad y nivel de instrucción y recogidas con una metodología sociolingüística básica. La web incluye una muestra de 18 entrevistas de cada uno de los corpus PRESEEA, disponibles en línea (<https://preseea.uah.es/>).

<sup>2</sup> Pueden consultarse las referencias de los trabajos publicados en <https://preseea.uah.es/equipos>.

Una vez establecidas las variables prioritarias para hacer estudios coordinados, el propósito es explicar los patrones sociolingüísticos hallados para crear un modelo predictivo del cambio lingüístico. El objetivo de este último es la creación de un mapa dinámico sociogeolectal del español y la predicción del cambio de la lengua española en el avance hacia el siglo XXII. En suma, las reflexiones teóricas que se derivan de los análisis pretenden la revisión y el esclarecimiento de aspectos fundamentales para el desarrollo de la teoría sociolingüística y dialectal. Llegados a este punto, se plantea como una meta ineludible la conexión de las investigaciones en curso a ambos lados del Atlántico con el propósito de favorecer la coordinación entre todas las variedades del español.

Algunos de los resultados relativos a cada uno de estos objetivos se reflejan en los diversos capítulos que componen el volumen, que aparece estructurado en tres partes bien diferenciadas. En primer lugar, los capítulos de Francisco Moreno Fernández, Carmen Silva Corvalán y Paul Preston abordan temas novedosos de una forma abarcadora, abriendo nuevas vetas para el trabajo en sociolingüística variacionista. Una segunda sección muestra algunas de las tareas desarrolladas a partir de diversas experiencias en América Latina en el seno de PRESEEA. Por fin, un tercer bloque de trabajos revisa diversas aportaciones llevadas a cabo por los equipos españoles asociados al proyecto.

El hilo conductor de los capítulos de Moreno Fernández, Preston y Silva Corvalán podría resumirse en que tocan aspectos vinculados a la madurez del proyecto PRESEEA, en la medida en que se refieren al aprovechamiento transversal de los materiales recogidos y de los estudios realizados, o al trabajo comparativo con otras familias de datos que, en sentido estricto, van más allá del planteamiento original del proyecto en sus líneas fundacionales. Otra lección fundamental es la de seguir mejorando e integrando las metodologías tradicionales con otras más modernas, en permanente búsqueda de datos sólidos y penetrantes. Conviene detenerse en cada uno de estos tres ensayos.

La colaboración de Francisco Moreno Fernández se dedica a la «Evolución lingüística en el curso de la vida. A propósito de PRESEEA». Como es bien sabido, una de las herramientas básicas en el trabajo variacionista es la hipótesis del tiempo aparente: del mismo modo que un corte geológico en el terreno permite ver los estratos correspondientes a diferentes épocas, la comparación de personas de diversas edades permite proyectar el presente hacia el pasado, como ha desarrollado Labov (1994) de modo magistral, entre otros autores. Las últimas décadas han visto, sin embargo, el análisis en tiempo real —método básico en lingüística histórica— como un fértil derrotero también para la sociolingüística. Así, los datos de PRESEEA se han comparado con los de otros muestreos, como los llevados a cabo en el marco del proyecto de la norma culta (Lope Blanch, 1986). Dentro del propio «Proyecto para el estudio sociolingüístico del español de España y América» se han iniciado ya algunos nuevos acopios de datos, en especial volviendo a grabar, cuando es posible, a algunos de los antiguos colaboradores. La comparación de momentos a lo largo de la vida de

una persona se vuelve una tarea fascinante, en la medida en que permite atisbar el comportamiento individual y sus efectos sobre la variación y el cambio lingüísticos. Si los estudios sobre adquisición lingüística y sobre deterioro han venido aportando ya algunas ideas, el análisis de las etapas estables puede enseñarnos mucho sobre la variación en lengua española. El trabajo de Moreno Fernández considera el comportamiento de la riqueza léxica y el de la construcción de *lo que es* en datos de Madrid y de Ciudad de México, para ofrecer un interesante acercamiento, al tiempo que una nueva perspectiva metodológica para este tipo de análisis.

El capítulo de Dennis R. Preston, «What Does It Take To Be A Native Speaker (And Who Gets To Say So)?», constituye una sugerente reflexión sobre la necesidad de incluir todo tipo de participantes en los estudios de variación lingüística geográfica y social, de modo que las muestras sean representativas de la compleja realidad demográfica, evitando visiones idealizadas que no corresponden al presente de los territorios. Si esto es válido para los datos de producción, tanto más lo es para los de percepción. La percepción de la realidad lingüística, a la sazón, se ha venido convirtiendo en un pujante objetivo en muy diferentes tipos de estudios –es ya, por ejemplo, una necesidad imperiosa en los trabajos de sociofonética– y, ciertamente, la arquitectura de la variación necesita integrarla cada vez en mayor medida. Preston enfatiza, en sus planteamientos, el papel de las poblaciones de herencia hispana en la conformación del inglés y de su percepción. Ante la evanescencia de postular un inglés americano general –quizá lo mismo podría plantearse para lenguas de otras latitudes–, un enfoque más realista obliga a trabajar con el inglés regional de la zona que se esté estudiando y a incluir de manera sistemática el efecto de los grupos exógenos con presencia más estable en esa misma zona. Solo así tendrá sentido discutir, valga el caso, las soluciones vocálicas de compromiso desarrolladas por ciertos hablantes, como revelan los planteamientos sociofonéticos sobre hablantes de diferentes herencias lingüísticas. Otro tanto cabe plantear acerca de los correlatos subjetivos, bajo diferentes enfoques, resumidos por Preston en lingüística popular, sociología lingüística y antropolingüística, hilo en que el autor reflexiona sobre la importancia del proyecto PRECAVES (Cestero Mancera y Paredes García, 2015, 2018), como pareja indispensable para el desarrollo y la perspectiva de PRESEEA. El estudio de la producción y de la percepción deben hermanarse si queremos crecer en el entendimiento de los aspectos lingüísticos y de la cara sociolingüística de la variación.

Por su parte, Carmen Silva Corvalán desarrolla el estudio «Variación e invariación morfosintácticas del español. Apuntes desde PRESEEA». El trabajo tiene numerosas virtudes. Una de ellas es la lectura transversal de la variación, en la medida en que esta cobra pleno sentido al comparar los resultados del análisis de una misma variable en varias ciudades. Aquí el papel de la coordinación y de las guías de trabajo que se han ido elaborando en el seno del proyecto son de suma importancia, pues ciertamente las diferencias metodológicas, en especial las que tienen que ver con el establecimiento del contexto de variación –cómo se ordena la variable dependiente y qué datos se acogen dentro de ella– y con la conformación de la estructura de variación puesta a examen

—las variables independientes y las variantes que se definen dentro de ellas— pueden ser lo bastante diferentes como para que las comparaciones sean difíciles o imposibles de realizar. Una segunda enseñanza, también de suma importancia, son las dosis de invariación que atraviesan numerosos fenómenos. A decir verdad, el sesgo interpretativo que nos lleva a veces como analistas a subrayar los aspectos variables de un proceso nos hace olvidar, o al menos a poner en segundo plano, aquellas dimensiones que no son variables. Esto es claro, por otra parte, al comparar el comportamiento de distintas variables, pues en muchos casos se puede observar que los patrones se repiten entre ciudades, incluida la estructura de la variación lingüística y social. De tal manera, es posible tener variación sociolingüística en una ciudad, pero invariación dialectológica al comparar diversas comunidades. Silva Corvalán examina en detalle, a propósito de estas cuestiones, la expresión del sujeto pronominal, la concordancia del verbo *haber* con su argumento plural y la alternancia entre *deber* y *deber de*. Por fin, en las conclusiones, anota un conjunto de sugerencias a propósito de cambios en curso, como la selección de *comiera* o *comiese* o la variación entre *revisar las publicaciones* y *revisar a las publicaciones*. Su caracterización se verá enriquecida por medio de trabajos que utilicen metodologías coordinadas.

El objetivo de la sección dedicada a los trabajos de PRESEEA sobre el español en América —y, en consecuencia, el sentido de la fracción de este libro que acopia los capítulos procedentes de ella— era ofrecer una muestra de las aportaciones realizadas en ese continente a partir de los datos recogidos en el proyecto sociolingüístico. Por supuesto, muchos otros equipos podrían haber formado parte de esa sucinta revisión, pero confiamos en que lo expuesto antes y ahora tenga suficiente representatividad como para dar una idea de los logros, limitaciones y posibilidades del proyecto en su desarrollo en los años siguientes. Existen diferencias sociocientíficas importantes entre la parte europea y la parte americana del proyecto. Estas diferencias no son solo una coyuntura del proyecto PRESEEA —que es lo que ahora nos interesa—, sino que forman parte de situaciones estructurales del desarrollo de la ciencia y de las instituciones científicas, que van mucho más allá de la propia sociolingüística y que, ciertamente, merecerían una reflexión bastante más detallada, imposible en este momento. Conviene observar, con todo, que el primer aspecto que debería considerarse es que al establecer una construcción España-América se está cruzando un país con un continente, cuando en realidad las diferencias lingüísticas, sociales y de desarrollo científico son muy diferentes entre los varios países americanos de habla española o castellana. Valga un ejemplo. Uno de los rasgos distintivos de la investigación en España dentro del proyecto PRESEEA es la densidad de equipos y las relaciones intensas entre ellos: subproyectos coordinados, encuentros frecuentes, intercambio de ideas, financiamiento relativamente sostenido, etc., son algunas características que permiten generar una institucionalización firme del proyecto. En contraste, en una parte de los equipos americanos no se da ninguna, o al menos no todas estas circunstancias: los subproyectos tienden a ser solo los del proyecto general; los en-

cuentros y el intercambio de ideas no se dan demasiado en reuniones nacionales; el financiamiento *ad hoc* es escaso o inexistente, etc. La situación conduce a que, cuanto más se dan estas últimas circunstancias, mayor es también la dificultad para institucionalizar PRESEEA, de modo que se depende casi exclusivamente de la valía y entusiasmo de los investigadores —enormes, cabe decir, sin duda, como en muchas otras investigaciones relacionadas con las humanidades—. Otros países latinoamericanos, sin embargo, vienen desarrollando, a su propio ritmo, redes de investigación semejantes a las europeas, en la medida en que con el paso de los años ha surgido una nómina de equipos que se ocupan de los enclaves urbanos más importantes del país. Además, esas redes nacionales han ido encontrando espacios propios en reuniones científicas nacionales e internacionales, así como en publicaciones específicas. Otro rasgo de las redes nacionales es la conformación de agendas propias, complementarias a las generales del proyecto PRESEEA, de forma que emergen más claramente problemas sociolingüísticos relevantes para las comunidades estudiadas y, además, se subrayan los asuntos y los análisis que más interesan a los equipos participantes. Estas redes nacionales emergentes, sean incipientes o consolidadas, se caracterizan por recibir ciertas dosis de financiamiento específico, así sea este variable de equipo a equipo o relativamente intermitente a lo largo del tiempo.

De ninguna manera se quiere decir con esto que América Latina tenga un papel secundario en el proyecto PRESEEA. Todo lo contrario. Basta ver la presencia constante de representantes de universidades y centros de investigación americanos en las reuniones generales del proyecto, la nómina de autores en las guías de trabajo o la inmensa cantidad de publicaciones e investigaciones realizadas, para comprobar que ha habido y hay una intensísima participación latinoamericana desde la fundación de PRESEEA y a lo largo del cuarto de siglo posterior. Hemos querido subrayar, sin embargo, las diferencias entre continentes e interamericanas por considerar que estas deben tenerse en cuenta para historiar adecuadamente el capítulo de desarrollo factual y analítico que supone PRESEEA dentro de la lingüística hispánica.

Claudia Borzi escribe la colaboración titulada «Construcción del PRESEEA-Buenos Aires y análisis de datos». Como en otros casos, la aportación es un interesante ejemplo de crecimiento de un equipo en lo que toca al estudio sociolingüístico de una ciudad clave en el contexto latinoamericano. El proyecto bonaerense arrancó en 2009 y desde entonces ha mostrado una ingente actividad, en diálogo con investigaciones previas, como los materiales derivados del estudio de la norma culta y sustentado, en especial, en el enfoque teórico derivado del cognoscitivismo de prototipos. Tras dedicar una sección a repasar la historia y las características socio-demográficas de la capital argentina, el texto de Borzi revisa la muestra obtenida, conformada por 108 entrevistas y presenta, a continuación, algunos de los intereses del equipo investigador. Algunos de los objetivos puestos en práctica son el análisis de las funciones de las preposiciones *de* y *por*; la prosodia de las condicionales con *si*; la variación acentual (*por ahí* / *por ahi*); el uso de ciertos verbos; marcadores discursivos como *después*, *che*, *luego*, *entonces*, *así que* y otros, el caso de *ahora*;

la posición del sujeto, etc. Como puede verse, un amplio conjunto de objetivos y de resultados, especialmente en sintaxis, discurso y prosodia.

La colaboración de Ana María González Mafud, «PRESEEA-La Habana: breve historia y resultados», es otro excelente ejemplo del tesón con el que vienen trabajando muchos de los equipos, en particular en los casos en que la adhesión al proyecto sociolingüístico se realizó cuando la red de investigación ya llevaba algunos años funcionando. El equipo habanero se sumó a PRESEEA en 2009 y ha desplegado una intensa actividad, reseñada en este capítulo. Para dar cuenta de ello, la exposición empieza con una breve recapitulación de los estudios afines llevados a cabo en los años anteriores y señala a continuación las circunstancias en que se desarrolla la integración en el proyecto internacional de sociolingüística del español, realizando en muy poco tiempo las 108 entrevistas de la capital cubana. El catálogo de investigaciones desarrolladas en esta última década y media, redondeando las cifras, es realmente muy amplio y se ha dirigido a todos los niveles lingüísticos, desde la fonética segmental al análisis discursivo, pasando por la gramática, la pragmática, la lexicografía y la prosodia. Como el propio texto señala, se ha concedido un muy importante papel a los estudios que involucran la dimensión fónica, pero lo mismo puede decirse, por ejemplo, del trabajo con problemas morfológicos y sintácticos o de los análisis sobre marcadores discursivos. Queda clara la sólida participación de los datos de PRESEEA en la descripción minuciosa del español hablado en Cuba.

El capítulo de María Claudia González Rátiva, «Elisión de /d/ intervocálica: aspectos de una tendencia hispánica», aborda un problema clásico en sociofonética del español, la caracterización del comportamiento de la /d/ situada entre vocales. Los datos de PRESEEA se prestan muy claramente al estudio de esta variable, por las condiciones de muestreo y por su abundante presencia en las entrevistas. Se trata, además, de un segmento que muestra una activa variación a lo largo y lo ancho de las variedades del español, por lo que la comparación entre comunidades –es decir, la dimensión geolingüística– es especialmente fructífera (véase asimismo Martín Butragueño, 2024). La aportación de González Rátiva permite ver también la madurez fonética del grupo de trabajo de Medellín, que ha permitido mostrar la predominancia de las soluciones aproximantes vocalizadas en dicha ciudad, a partir de una amplia muestra de 108 personas, cuyos datos se han medido instrumentalmente. Los datos proporcionados por la instrumentación, en reemplazo de la tasación cualitativa de la percepción del analista, es ya el presente, y sin duda el futuro, en el estudio de la variación segmental –otro tanto cabe decir de la socioprosodia–. Dentro del proyecto internacional de sociolingüística del español, el estudio de la /d/ intervocálica se beneficia de la cuidadosa guía elaborada por Samper Padilla, Malaver Arguinzones y Samper Hernández (2021), que permite avanzar con pasos firmes en el estudio comparado de los procesos que afectan a este fonema dental sonoro, cuya variación suele estar ligada, además de a un complejo conjunto de efectos fónicos, a una rica estela de valoraciones sociolingüísticas. Dadas las similitudes de comportamiento entre diferentes variedades, como señala González Rátiva, podría pensarse que la

variación de /d/ intervocálica sería también candidata a esa invariación en los predictores que señala Silva Corvalán en su propia colaboración a este libro (véanse los comentarios expuestos *supra*).

El objetivo de Pedro Martín Butragueño, en «Hacia un modelo de la ciudad lingüística: estructuración sociolingüística básica de Ciudad de México a inicios del s. XXI», enfatiza la necesidad de considerar de forma comparada diversas variables en una misma ciudad, pues ese tipo de paralelismos debería hacer posible comprender de mejor manera la estructura sociolingüística de una ciudad. Para ello, el capítulo analiza con metodología semejante los patrones sociales básicos de cinco variables lingüísticas –morfológicas, sintácticas, léxicas y prosódicas– en la Ciudad de México: el caso de los apreciativos en *-ito, a*; el problema de la concordancia del *haber* existencial y su argumento plural; la expresión o no de los sujetos pronominales; el empleo de indoamericanismos; y la resolución entonativa de los enunciados aseverativos. La idea es, precisamente, incluir variables de distinta naturaleza para poder tener testigos confiables y diversos de las soluciones preferidas por cada grupo de hablantes. Se toma como herramienta comparativa básica el cálculo de ANOVA de un factor y se busca una interpretación general al analizar paralelamente los predictores sociales significativos en cada caso, la jerarquización entre los predictores según la magnitud de su significación y una ponderación de la dispersión de los predictores sociales significativos para cada variable lingüística –en este último caso, se espera que la mayor dispersión sugiera un papel más relevante de los individuos–. Por fin, al revisar los límites de la interpretación general, se propone una hipótesis normalizadora en lo que toca al papel sociolingüístico de la capital mexicana. Se ensaya asimismo una definición del concepto de *ciudad lingüística*. La relevancia de profundizar en este concepto, objeto de la sociolingüística urbana, debería descansar en el análisis, con las mismas herramientas, de variables lingüísticas de distinta naturaleza.

Por fin, la exposición de Abelardo San Martín Núñez se dedica a «El equipo del PRESEEA de Santiago de Chile: situación actual y desafíos a futuro». El capítulo presenta la investigación realizada por el equipo chileno dentro del proyecto, para luego centrarse, específicamente, en la geolingüística y la sociolingüística de los marcadores del discurso. El equipo de Santiago de Chile se sumó al proyecto en 2007, para llevar a cabo las correspondientes entrevistas en un plazo breve, entre ese año y 2009. Subdividido en dos equipos en 2016, los temas más trabajados en el seno de la investigación de la capital chilena han sido los que tocan a la gramática y al discurso: las narraciones personales, la estratificación de los marcadores, el discurso referido, la atenuación, los clíticos, el voseo y el sujeto pronominal, según refiere San Martín Núñez. Debe también destacarse la participación en varias de las guías redactadas para estabilizar los análisis que se realizan dentro de PRESEEA y fomentar las comparaciones precisas (Guerrero González y Arriagada Anabalón, 2021; Guerrero González y Manjón-Cabeza Cruz, 2022; San Martín Núñez *et al.*, 2022). En cuanto a los marcadores del discurso –terreno de estudio que ha resultado

ser uno de los más fértiles dentro de PRESEEA—, algunos de los más estudiados han sido *¿cachái?, o sea, igual, al final, por ser, onda, por último, che, güey, ¿sabes?*, de forma que la contribución expone los diferentes valores y empleos que adoptan, al tiempo que se ofrecen diversos ejemplos. Una de las reflexiones más útiles, con respecto a los procesos de variación relacionados con los marcadores, tiene que ver con los posibles cruces entre zonas geográficas de uso, funciones desempeñadas, frecuencias, factores sociales y formalidad de la situación.

El distinto modo en que se ha articulado el proyecto en España y en América lleva a reflexionar sobre la dificultad de establecer convergencias plenas sobre su desarrollo a uno y otro lado del Atlántico. De partida, son más evidentes las divergencias, pues difícilmente pueden compararse las dinámicas que ha desencadenado el proyecto en territorios tan distintos en extensión y diversidad interna. Las diferencias son muchas y en buena parte se derivan de la vastedad del español americano, lo que ha determinado la distinta configuración de los equipos y redes de investigación. En España, los grupos integrantes de PRESEEA se articulan en el contexto de un mismo estado nacional y ello ha favorecido la financiación pública y el desarrollo coordinado de la investigación. Asimismo, la homogeneidad del territorio y de la realidad social española han permitido el más fácil encaje de una metodología de investigación común. Por último, las divergencias entre la investigación sociolingüística en uno y otro continente son sin duda mucho mayores en lo que se refiere a los análisis que se han derivado de la comparación de los corpus.

Por las razones aducidas, los equipos españoles integrados en PRESEEA han podido avanzar de manera conjunta durante muchos años y ello ha tenido como consecuencia una investigación más unitaria y resultados también más homogéneos. Pese a todo, la muestra de equipos españoles representados en esta publicación da cuenta de que incluso en España el avance del proyecto ha sido desigual en muchos sentidos, pero especialmente por lo que se refiere al momento en que los equipos se han incorporado a esta investigación coordinada. En este sentido, cabe distinguir, por una parte, los que desde mediados de la década de 1990 se incorporaron al proyecto y consiguieron coordinarse gracias a la financiación obtenida a través de instituciones públicas. Entre estos se cuentan Valencia, Madrid y Alcalá de Henares, Málaga, Granada, Las Palmas de Gran Canaria, Lérida y Sevilla, todos ellos representados en este volumen, con la excepción de Lérida.

En una segunda etapa, otros equipos españoles han avanzado con más dificultades por falta de financiación o se han incorporado más tarde: Oviedo, Gijón, Zaragoza, Palma de Mallorca, Santiago de Compostela, Pamplona, Cádiz y Santander.

En esta publicación, se ofrece una selección de siete trabajos centrados en algunas de esas comunidades de habla: Granada, Sevilla y Málaga en Andalucía; Las Palmas en el archipiélago canario; Madrid en el centro de la meseta castellana; y Valencia y Pamplona en regiones bilingües del oriente y del norte peninsular, respectivamente. Todas ellas tienen una larga trayectoria en el proyecto, salvo Pamplona, que se incor-

poró en 2020 y se incluye como muestra de una investigación en sus etapas iniciales. En el capítulo titulado «PRESEEA-Pamplona. Presentación y primeros estudios», M.<sup>a</sup> José Galluci y Carmela Pérez-Salazar dan cuenta de los primeros pasos de los trabajos de Pamplona, que suponen el inicio de la sociolingüística cuantitativa en esta comunidad navarra, una antigua encrucijada geográfica que fuera «territorio nuclear de los antiguos vascones». En este territorio de poblaciones mezcladas, el euskera y los romances han convivido determinando la singularidad de sus actuales rasgos lingüísticos, pues el español de Navarra es heredero del antiguo romance navarro y del castellano, y producto de la convivencia multisecular de ambos con el euskera. Las autoras explican en detalle cómo han adaptado las directrices generales de la metodología del PRESEEA a las peculiaridades de Pamplona y dan cuenta de hasta dónde han avanzado en la elaboración del corpus sociolingüístico de la ciudad, un trabajo que permitirá conocer los fenómenos de variación y los procesos de cambio lingüístico del español de Navarra, comunidad rica y compleja desde el punto de vista histórico-lingüístico.

En el otro extremo de la cronología del desarrollo de PRESEEA en España se encuentran las investigaciones sociolingüísticas de Valencia, Málaga, Granada, Las Palmas, Alcalá de Henares y Madrid, cuyos equipos participaron en el proyecto desde sus comienzos. En Valencia, el responsable durante más de dos décadas fue José Ramón Gómez Molina, autor del capítulo «El PRESEEA en Valencia. Historia y resultados», donde da cuenta de las etapas cubiertas a lo largo de este tiempo: desde la elaboración y la publicación del corpus hasta el análisis de múltiples variables sociolingüísticas seleccionadas por su relevancia para definir los patrones sociolingüísticos de la ciudad, cuyas características principales se desglosan y resumen en el texto. En Valencia, como en Pamplona, el castellano ha mantenido un contacto secular con otra lengua, el catalán valenciano, provocando una transferencia de influencias recíprocas que explican la divergencia de algunos procesos respecto a la evolución del castellano en las regiones monolingües.

La aportación de Juan Antonio Moya, «El grupo PRESEEA-Granada: estudios, resultados y sugerencias hacia el futuro», se organiza de manera semejante: un estado de la cuestión del avance del proyecto en la capital de la Andalucía oriental. En este caso, Granada contaba con una tradición de investigaciones sociolingüísticas anteriores al nacimiento de PRESEEA, cuyos resultados han contribuido al enriquecimiento de la investigación variacionista de la ciudad y han permitido, ya desde los años 90, estudios de contraste en tiempo real y un conocimiento profundo de los procesos de variación y cambio de un territorio que ha experimentado una intensa transformación desde la segunda mitad del siglo XX. La contribución de Moya pone el énfasis en la diversidad de proyectos afines y próximos a PRESEEA que están dando a conocer los procesos de cambio lingüístico y social en Granada y hace una propuesta de las direcciones que deberían seguir las investigaciones futuras. Desde el conocimiento que hoy tenemos del habla urbana de Granada, Moya considera necesario aproximarse al estudio de las hablas rurales de la región desde esta misma

perspectiva sociolingüística. Cuestiones como la desruralización, el abandono y el aislamiento del campo están haciendo surgir formas divergentes de las urbanas que es necesario conocer para lograr una comprensión integral de dos mundos que avanzan en direcciones aparentemente opuestas.

Marta Samper, en su capítulo «Una mirada al pasado, presente y futuro del proyecto PRESEEA en Las Palmas de Gran Canaria», ofrece una panorámica de la evolución del proyecto en el archipiélago, una pieza clave para conocer la difusión de los procesos evolutivos del español a ambos lados del Atlántico. El equipo de Las Palmas, como los de Valencia, Granada o Málaga, también se cuenta entre los primeros en unirse al proyecto y en participar de la planificación de las múltiples direcciones de su desarrollo. Los investigadores que lideraron la investigación, José Antonio Samper y Clara Eugenia Hernández, también fueron protagonistas de la incorporación de Las Palmas en las investigaciones del *Proyecto de la Norma Culta Hispánica* y en otras investigaciones panhispánicas que han proporcionado una información valiosísima para conocer la ciudad en muchas de sus dimensiones variables. Gracias a la intensa labor de estos investigadores, las Palmas cuenta con un rico acervo de datos que la sitúa en una posición privilegiada para el estudio sociolingüístico. Como Valencia y Granada, el equipo de Las Palmas ha avanzado notablemente en el conocimiento de los patrones sociolingüísticos urbanos, algunas de cuyas variables se detallan en este trabajo. Sin embargo, como señala la autora, queda mucho trabajo por hacer en una línea de investigación que es significativa en sí misma, pero además por el papel de puente de las Canarias entre España y América. En la investigación sociolingüística coordinada entre ambos continentes, los datos que arrojen las ciudades del archipiélago son de suma importancia para entender la evolución del español en el siglo XXI.

De nuevo en Andalucía, el capítulo de Juan A. Villena Ponsoda y Antonio M. Ávila, «Contribución del grupo PRESEEA-Málaga a la construcción del mapa sociolingüístico de variación y cambio en España desde un enfoque comparativo y multidimensional» adopta un enfoque diferente respecto de los anteriores, pues se centra en las aportaciones de la investigación malagueña a la elaboración de un mapa sociolingüístico de variación y cambio en España. Este estudio nos sitúa en la última de las etapas en las que ha desembocado la investigación coordinada: un enfoque comparado y multidimensional del estudio de la variación y el cambio sociolingüístico; en suma, una contribución al desarrollo de una sociolingüística comparada del español y del análisis de procesos de cambio lingüístico en tiempo aparente y real. El análisis se apoya en algunos aspectos de la fonología, la morfosintaxis, la variación léxica y la pragmática mediante la interpretación de algunas variables de estos niveles que se han analizado coordinadamente en el contexto del PRESEEA. Los datos recogidos en los últimos treinta años están permitiendo desarrollar una auténtica sociolingüística comparada del español cuya estructura los autores esbozan en esta contribución.

La incorporación del equipo sevillano al PRESEEA fue un poco más tardía, como señala Juana Santana en «PRESEEA-SEVILLA: 10 años de sociolingüística en la ca-

pital hispalense», pues no fue hasta 2013 cuando la capital de Andalucía occidental se incorporó al proyecto. Desde entonces, y a paso increíblemente acelerado, el equipo sevillano, liderado por la autora, ha cubierto los hitos principales hasta situarse en la vanguardia de la investigación coordinada. En tan solo diez años, han completado las labores de campo, han transcrito, etiquetado y publicado la totalidad del corpus y están realizando análisis variacionistas de los fenómenos lingüísticos priorizados en el proyecto para su comparabilidad con otras comunidades de habla y para la descripción de los patrones sociolingüísticos de Sevilla. De algunas de esas variables se da cuenta en estas páginas, con una descripción que se estructura en niveles de análisis: fonético-fonológico, morfosintáctico, léxico-semántico y pragmático-discursivo. La autora consigue ofrecer una imagen sucinta de hasta qué punto la incorporación de su equipo está consiguiendo aportar la información relativa a esta pieza tan valiosa de las hablas urbanas meridionales que contribuye a completar nuestro conocimiento del mosaico andaluz en particular y de la investigación sociolingüística panhispánica.

Por último, el estudio de Molina Martos, titulado «Procesos de variación y cambio en Madrid: el papel de las identidades en la formación de regiolectos», parte de los análisis realizados en Madrid para ejemplificar las dinámicas de variación y cambio características de la sociolingüística urbana de las hablas castellanas. La autora reflexiona sobre la evolución que ha experimentado la sociolingüística, tanto para dar cuenta de los cambios sociales que han tenido lugar en las ciudades, como de las unidades de análisis empleadas, en un proceso paralelo al de las tres olas de la sociolingüística (Eckert, 2012). El análisis de la consonante /-s/ en coda silábica sirve en estas páginas para ejemplificar cómo funciona la variación individual en la configuración de identidades sociolingüísticas, una orientación característica de los estudios de la tercera ola que los corpus PRESEEA también permiten abordar. El trabajo concluye que la variable sociolingüística no es portadora de un significado social fijo y estable, sino que evoca valores indexicales ya existentes, al tiempo que tiene la capacidad de forjar otros nuevos. Esta idea de indexicalidad resulta especialmente útil para abordar la cuestión de la construcción de identidades en interacción y superar la visión tradicional de la comunidad de habla estable y homogénea.

A la luz de lo expuesto en los trabajos de este libro, cabe reflexionar en los logros fácticos de PRESEEA en referencia al análisis de lo lingüístico, lo dialectológico y lo sociolingüístico. En el largo camino que se atisba por delante caben numerosas trayectorias: nuevos muestreos que permitan combinar el tiempo aparente y el tiempo real, más y mejores comparaciones entre ciudades, comparaciones entre variables dentro de una misma comunidad, mayor integración entre los datos de producción y de percepción, desarrollo de gramáticas variables completas, análisis más perspicaces de las estructuras sociolingüísticas, entre un largo etcétera. Quizá conviene detenerse un instante en este último punto, pues la inmensa mayoría de los trabajos realizados solamente examinan el papel de la edad, el sexo y el nivel de estudios. No debe olvidarse que esta triada es la base del muestreo por cuotas, y que, aunque en sí misma es muy

reveladora, nada impide analizarlas de otra forma o considerar otras variables sociales. Pueden considerarse de otra forma convirtiéndolas en continuas, como puede hacerse con la edad; enriqueciéndolas, por ejemplo, al considerar el género y no el sexo; o reordenándolas, como es posible con el nivel de estudios, entre muchas otras posibilidades. Y hay mucho que decir sobre el modo de vida, las profesiones, el nivel de ingresos, la zona de la ciudad en que se vive, las redes sociales, el papel de los individuos... Si bien el proyecto PRESEEA tiene enormes utilidades dialectológicas, no debe olvidarse que lleva el apellido de «sociolingüístico» ya en su propio nombre, así que es posible pensar, dada la enormidad de datos de que disponemos, en la generación de modelos que nos permitan entender el uso social del lenguaje, las razones para construir identidades o las motivaciones escondidas detrás de la variación y el cambio lingüístico. Todo ello es posible en el seno de una inmensa red científica difundida a lo largo y ancho del mundo hispano-hablante.

## REFERENCIAS

- Cestero Mancera, Ana M., y Florentino Paredes García (2015). «Creencias y actitudes hacia las variedades normativas del español actual: primeros resultados del Proyecto PRECAVES-XXI», *Spanish in Context*, 12(2): 255-279. <https://doi.org/10.1075/sic.12.2.04ces>
- Cestero, Ana M., y Florentino Paredes (2018). «Creencias y actitudes hacia las variedades cultas del español actual: el proyecto PRECAVES XXI», *Boletín de Filología*, 53(2): 11-43. <https://boletinfilologia.uchile.cl/index.php/BDF/article/view/51940>
- Eckert, Penelope (2012). «Three Waves of Variation Study: The Emergence of Meaning in the Study of Sociolinguistic Variation», *Annual Review of Anthropology*, 41(1): 87-100. <https://doi.org/10.1146/annurev-anthro-092611-145828>
- Guerrero González, Silvana, y Silvana Arriagada Anabalón (2021). *Guía PRESEEA de estudio del presente narrativo*. Alcalá de Henares: Universidad de Alcalá. <https://preseea.uah.es/documentos-preseea-de-investigacion>
- Guerrero González, Silvana, y Antonio Manjón-Cabeza Cruz (2022). *Guía PRESEEA de estudio de la expresión del sujeto pronominal*. Alcalá de Henares: Universidad de Alcalá. <https://preseea.uah.es/documentos-preseea-de-investigacion>
- Labov, William (1994). *Principles of Linguistic Change*. Vol. 1: *Internal Factors*. Oxford: Blackwell.
- Lope Blanch, Juan M. (1986). *El estudio del español hablado culto. Historia de un proyecto*. Ciudad de México: Universidad Nacional Autónoma de México.
- Martín Butragueño, Pedro (2024). «La variación en las consonantes oclusivas orales», En J. Gil y J. Llisterri (dirs), *Fonética y fonología descriptivas de la lengua española*. Vol. 1: Nivel segmental. Washington: Georgetown University Press, 377-411.

- Moreno Fernández, Francisco (1993). «Proyecto para el estudio sociolingüístico del español de España y América (PRESEEA)», *Lingüística*, 5: 268-271.
- Moreno Fernández, Francisco (1996). «Metodología del ‘Proyecto para el estudio sociolingüístico del español de España y América’ (PRESEEA)», *Lingüística*, 8: 257-287.
- Samper Padilla, José Antonio, Irania Malaver Arguinzones y Marta Samper Hernández (2021). *Guía PRESEEA de estudio de la /d/ intervocálica*. Alcalá: Universidad de Alcalá. <https://preseea.uah.es/documentos-preseea-de-investigacion>
- San Martín Núñez, Abelardo, Cristian Rojas Inostroza, Natalia Ruiz González y Patricia Córdova Abundis (2022). *Guía PRESEEA de estudio de los marcadores discursivos de reformulación*. Alcalá de Henares: Universidad de Alcalá. <https://preseea.uah.es/documentos-preseea-de-investigacion>
- Silva Corvalán, Carmen (1994). «Direcciones en los estudios sociolingüísticos de la lengua española», en *Actas del Congreso de la Lengua Española*, Sevilla, 7-10 octubre de 1992. Madrid: Instituto Cervantes, pp. 399-415.

## EVOLUCIÓN LINGÜÍSTICA EN EL CURSO DE LA VIDA. A PROPÓSITO DE PRESEEA\*

Francisco Moreno Fernández  
*Universität Heidelberg*

### 1. INTRODUCCIÓN

Dentro del gran marco de investigación que es PRESEEA, sus objetivos y propuestas han ido experimentando cambios en una trayectoria de enriquecimiento. Los objetivos iniciales se resumían en una idea central: construir un corpus de lengua española hablada a partir de una metodología sociolingüística básica y común (Moreno Fernández, 2021). Este objetivo, evidentemente, se mantiene, con la novedad de que lo que pretendía ser un corpus se está convirtiendo en un megacorpus, si consideramos los parámetros habituales en el estudio de la lengua hablada.

Este primer y fundacional objetivo conoció la adición de otros fines conforme avanzaban las investigaciones de campo. Así, a partir de 2010 (PRESEEA, 2010), se marcaron como objetivos axiales: a) estudiar la evolución de la lengua española, en la línea del estudio de la variación y el cambio; b) analizar la variación geosociolingüística en el macrodominio hispanohablante; c) realizar aportes a la teoría general de la variación y el cambio. Podría decirse que estos nuevos objetivos estaban implícitos en el primero, pero la adición fue hecha cuando se tuvo constancia de que podían abordarse de una manera sistemática y profunda, por la cantidad y la calidad de los materiales reunidos y previstos.

Unos años después, PRESEEA está en condiciones de dar un paso más en sus aspiraciones y sumar a sus objetivos uno derivado tanto del desarrollo del propio proyecto como de la evolución de los intereses en materia sociolingüística. El nuevo objetivo es analizar la variación y el cambio lingüístico en el curso de la vida (ing. *linguistic change in the life-span*). Esta propuesta tiene como base y razón una rea-

lidad de los actuales trabajos sociolingüísticos de la red PRESEEA: la realización de nuevas campañas de entrevistas y recolección de materiales de lengua hablada; campañas que ya han iniciado varios equipos en España y en América y que prometen abrir el macrocorpus a una nueva dimensión investigadora.

## 2. EL ESTUDIO DE LA VARIACIÓN Y EL CAMBIO EN ESPAÑOL

Es evidente que la sociolingüística del español no comienza ni termina con el proyecto PRESEEA. La lengua española cuenta con un conjunto de materiales de gran valor para el estudio de la variación y del cambio lingüístico. Entre esos materiales, destacan, como pioneros y como referencia para muchos trabajos posteriores, los corpus de hablantes de educación superior procedentes de diversas ciudades del territorio hispanohablante. Esa colección de materiales se conoce como proyecto PILEI o «Proyecto para el estudio de la norma culta de las principales ciudades hispanohablantes», y consiguió reunir materiales procedentes de Barcelona, Bogotá, Buenos Aires, Caracas, La Habana, La Paz, Las Palmas, Lima, Madrid, México, Miami, Montevideo, Panamá, San José de Costa Rica, San Juan de Puerto Rico, Santiago de Chile y Sevilla (Lope Blanch, 1986). Esos materiales, reunidos mediante entrevistas entre las décadas de los sesenta y de los ochenta del pasado siglo, son los más antiguos de que se dispone con un potencial para el estudio sociolingüístico. De hecho, muchos de ellos pueden compararse (y se han comparado) con materiales más recientes, por lo que resultan de gran utilidad para realizar estudios sobre cambios en tiempo real.

Asimismo, la sociolingüística hispánica se viene cultivando, sobre todo en la línea variacionista, desde los años setenta y ochenta, cuando se abordaron los estudios del español de Panamá o de San Juan de Puerto Rico. Efectivamente, la tesis doctoral de Henrietta Cedergren (1973) sobre el español de Panamá fue uno de los primeros trabajos alineados con la sociolingüística de William Labov y en ella se abordaba el análisis cuantitativo de variables fonéticas como la aspiración de /s/ final de sílaba, la elisión de diversas consonantes finales o la lenición del fonema prepalatal sordo. El corpus analizado estuvo formado por entrevistas realizadas a 79 informantes panameños. En una línea semejante se inscribe el estudio de Humberto López Morales (1983) sobre San Juan de Puerto Rico, que sirvió de base metodológica a otros trabajos, como el estudio de José Antonio Samper (1990) sobre aspectos fonéticos y niveles socioculturales del español de Las Palmas de Gran Canaria.

En cuanto a la construcción de corpus, han sido numerosas las experiencias sociolingüísticas de gran valor para el conocimiento de la variación y el cambio en español, más allá del proyecto PILEI. Entre ellos merecen mencionarse los trabajos realizados en Venezuela, concretamente en Caracas (Rosenblat y Bentgiovoglio, 1979; Bentgiovoglio y Malaver, 2012; Guirado, 2014; Galucci, 2020) y en México (Rodríguez Flores y Rodríguez Alfano, 1996; Rodríguez Alfano, 2005; Martín Bu-

tragueño, 2014). En este país, los materiales reunidos han servido para la elaboración de estudios sobre cambios lingüísticos, especialmente los derivados del contacto entre variedades y por la incidencia de los líderes lingüísticos (Martín Butragueño, 2006, 2012).

Naturalmente, entre los corpus de lengua hablada destacan los materiales aportados por el proyecto PRESEEA, que han servido como motor para el estudio sociolingüístico de numerosas comunidades hispanohablantes de Europa y de América desde el año 1996. Como es sabido, las muestras de PRESEEA para cada comunidad investigada están organizadas por cuotas con afijación uniforme, atendiendo a las variables sexo (H: Hombre; M: Mujer), edad (1: <35; 2: 36-55; 3: >56) y nivel educativo (1: Primario; 2: Medio, 3; Superior). Los tamaños de referencia para cada muestra oscilan entre 54 y 108 informantes, dependiendo de la complejidad social del núcleo urbano estudiado. Si el conjunto de comunidades investigadas con la misma metodología PRESEEA alcanza la cifra de 50, como se espera, será muy superior a 2500 el volumen total de hablantes cuyo discurso formará parte de este macrocorpus (Moreno Fernández y Cestero, 2020).

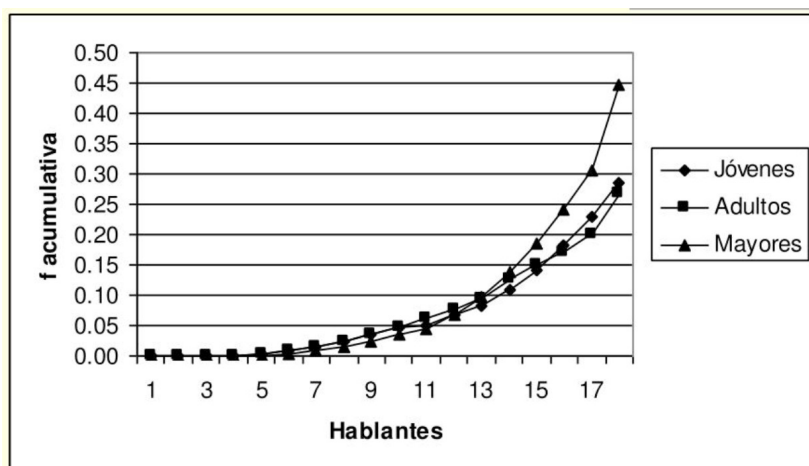
Con estos antecedentes, dado que la recogida de materiales por parte de varios equipos de PRESEEA se completó antes del año 2000, dos décadas después se han iniciado los trabajos para la recogida de nuevas muestras. Así se está haciendo para las comunidades de México, Málaga o Madrid, como se hizo previamente en Caracas y en Monterrey. De este modo, la continuidad en la construcción de corpus de lengua hablada está haciendo posible que hoy dispongamos de materiales sociolingüísticos procedentes de campañas realizadas en distintos momentos y que estemos en condiciones de abordar estudios de cambios en tiempo real, tanto estudios de panel como estudios de tendencia. Como es sabido, los estudios de tendencia analizan lingüísticamente una misma comunidad en dos momentos sucesivos, pero sin recurrir a los mismos informantes en cada ocasión; por su lado, los estudios de panel analizan lingüísticamente una misma comunidad en dos momentos sucesivos recurriendo a los mismos informantes, en la medida de lo posible (Sheppard, 2020). Aún existe un tercer tipo de estudio longitudinal, que es el estudio de cohortes. En ellos, se identifica una categoría de personas de interés (por ejemplo, de una misma generación, promoción o etapa) y posteriormente se encuesta periódicamente a personas que pertenecen a esa categoría.

Las cohortes más comunes que pueden interesar a los investigadores son: personas de determinadas generaciones o nacidas en el mismo periodo de tiempo; promociones; personas que empezaron a trabajar en un sector determinado al mismo tiempo; o quizás personas que tienen alguna experiencia vital específica en común. Esto significa que, si el comportamiento u otro fenómeno que interesa al investigador cambia, ya sea por algún acontecimiento mundial o porque las personas envejecen, el investigador podrá captar esos cambios.

En relación con el estudio del cambio lingüístico en español, no es necesario insistir en que los proyectos sociolingüísticos realizados en numerosas comunidades, dentro y

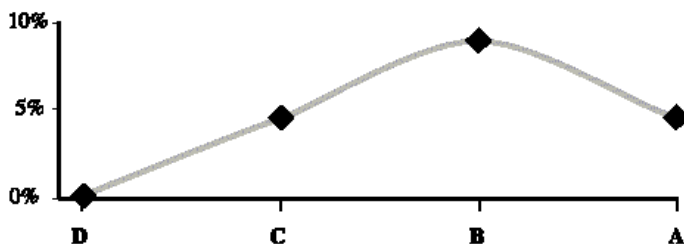
fuera de PRESEEA, han permitido abordar análisis de procesos de cambio en tiempo aparente (Sankoff, 2006). El cambio en tiempo aparente, como es sabido, supone analizar los usos lingüísticos de generaciones sucesivas y proyectarlos en una línea temporal, como se ha mostrado en infinidad de estudios, como los de Martín Butragueño (2006) en México o los realizados sobre el español de Caracas (Asratíán, 2007).

**GRÁFICO 1. FRECUENCIA RELATIVA ACUMULATIVA DE ASIBILACIÓN POR HABLANTES Y POR GRUPOS DE EDAD EN LA CIUDAD DE MÉXICO**



Fuente: Martín Butragueño, 2006.

**GRÁFICO 2. EVOLUCIÓN DE LA FORMA -SE EN EL TIEMPO APARENTE EN CARACAS (1987)**



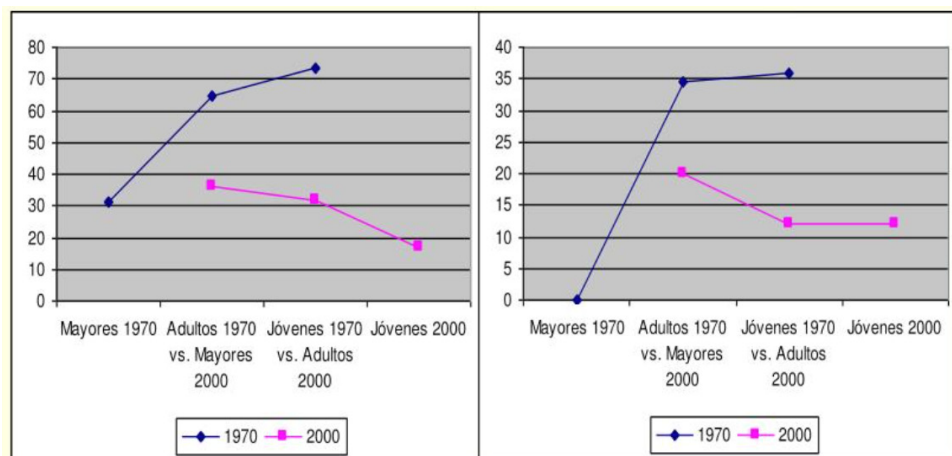
Fuente: Asratíán, 2007.

La hipótesis del tiempo aparente, sin embargo, da por hecho que el sistema lingüístico goza de una gran estabilidad a lo largo de la vida: esto es, dar por supuesto que los usos de cada generación van a tener continuidad a lo largo del tiempo, pero

este extremo no ha sido comprobado empíricamente y los estudios de panel están en condiciones de hacerlo.

Sin embargo, también ha habido algunas aproximaciones al cambio en tiempo real en español desde los estudios de tendencia. Podemos destacar los trabajos realizados sobre el español de México en los que se comparan los materiales procedentes del *Atlas Lingüístico de México* (Lope Blanch, 1990-2000) con los corpus publicados posteriormente. Esto permitió, por ejemplo, el análisis de la asibilación de (-r) y de (r) en la ciudad de México en 1970 y en 2000.

**GRÁFICO 3. DISTRIBUCIÓN EN TIEMPO APARENTE (EDAD) Y EN TIEMPO REAL (1970 vs. 2000) DE LA ASIBILACIÓN DE (ɾ) -IZQUIERDA- Y DE (r)- DERECHA, EN LA CIUDAD DE MÉXICO**



Fuente: Martín Butragueño, 2014.

Como el estudio de Martín Butragueño, existen otros que comparan corpus o materiales procedentes de momentos distintos, como el pionero de Henrietta Cedergren sobre Panamá (1973), el de Irania Malaver (2017) en Caracas sobre el uso de *ahorita*, el de Luz Marcela Hurtado (2016) sobre la expresión de impersonalidad en el español de Bogotá o el de María Eugenia Flores (2012) sobre *bastante* en Monterrey. Sirvan estos trabajos simplemente como referencia porque no aspiro en este punto a presentar una nómina exhaustiva.

Con planteamientos metodológicos diferentes, existen otros estudios de tendencia, como el trabajo de Antonio Briz (2023) sobre la localidad española de Requena, una comunidad del interior de la península dedicada a la actividad vitivinícola principalmente. Los cambios observados por Briz apuntan hacia tres hipótesis: la nivelación dialectal, la deslocalización de algunos fenómenos y lo dialectal como

marca identitaria. Otro ejemplo de estudio de tendencia es el realizado por Hernández Campoy (2003) sobre el español de la región de Murcia (España) a partir de materiales radiofónicos emitidos entre los años setenta y el 2000. En ese estudio se comprobó que el uso de rasgos estándares del castellano se va extendiendo paulatina y consistentemente por la región murciana y entre sus distintos sustratos sociales en detrimento de los rasgos lingüísticos locales meridionales. Desde el punto de vista geolingüístico, esta lenta pero constante erosión de los rasgos locales sigue una estructura jerárquica de difusión, desde los centros urbanos más grandes a los más pequeños. Finalmente, otro ejemplo de estudio de tendencia es el análisis del léxico disponible realizado sobre Santo Domingo por Orlando Alba, comparando la disponibilidad léxica analizada en 1990 y en 2008. En este caso no se trata de entrevistas, sino de cuestionarios, pero la intención de estudio de tendencia es la misma.

Esto en cuanto a los estudios de tendencia. El estudio de cambios a lo largo de la vida a partir de una metodología de panel; es decir, analizando a los mismos informantes, más allá de incipientes experiencias dentro de PRESEEA, cuenta con la experiencia del estudio de García Marcos (2021) sobre el español en la costa de Granada. García Marcos realizó su tesis doctoral sobre materiales analizados entre 1985 y 1987, y procedentes de 140 informantes. En 2017, el mismo investigador reunió materiales de 17 municipios de la misma zona y pudo reencontrar a 73 de los antiguos informantes, configurando una muestra mixta, de panel y de tendencia, aunque no se profundizara en las diferencias entre una y otra. En 1987 se trabajó con un total de 10,65 horas, con textos de 4,5 minutos de grabación de promedio, y en 2017, con 12 horas, con textos de 4,8 minutos de promedio (García Marcos 2021: 30).

### 3. EVOLUCIÓN Y CAMBIO LINGÜÍSTICO: PRINCIPIOS FUNDAMENTALES

Los análisis del cambio lingüístico en tiempo real y específicamente los estudios de cambios a lo largo de la vida –o en el curso de la vida (Hunt, 2005; Lowry, 2022)– a partir de materiales procedentes de paneles constituyen una línea fundamental de trabajo a partir de materiales reunidos en dos momentos o periodos más o menos distantes en el tiempo. A este respecto, hay un marco epistemológico que ha de tenerse muy en cuenta y que nos es otro que el de la sociología y la sociolingüística específicamente interesadas por la evolución a lo largo de la vida.

Desde una perspectiva sociológica, partimos de la base de que el «curso de la vida» se refiere a acontecimientos, transiciones y trayectorias de la vida de una persona –en este caso, referidos a sus prácticas lingüísticas– que se desarrollan en secuencias y dinámicas particulares a lo largo de las etapas de la vida (Lowry, 2022: 35). Asimismo, aceptamos los principios cardinales de la perspectiva del curso de la vida, a saber:

- Principio 1. El curso de la vida de una persona está implicado en y moldeado por los momentos históricos que vive.

- Principio 2. El efecto de un incidente biográfico mayor se verá moldeado por la edad y la etapa de la vida en que el acontecimiento se produzca.
- Principio 3. Ninguna historia de vida puede comprenderse independientemente de las historias de vida de los demás.
- Principio 4. Las condiciones, acontecimientos y experiencias en cualquier edad o etapa de la vida no pueden comprenderse independientemente de las condiciones, acontecimientos y experiencias de edades o etapas vitales anteriores.
- Principio 5. Los individuos hacen elecciones y realizan acciones dentro de los límites de su universo social.

Por otro lado, es importante la distinción entre la vida como fenómeno biológico y la vida como fenómeno sociocultural, aunque obviamente exista una fuerte interrelación entre ambas dimensiones. Las condiciones biológicas incluyen las funciones cognitivas y sus correlatos neurofisiológicos (Kemper, 2015). Desde 2015, la Organización Mundial de la Salud considera jóvenes a las personas con una edad comprendida entre los 25 y los 44 años; a los que tienen entre 44 y 60 años se los considera de edad mediana; la tercera edad se sitúa entre los 60 y los 75 años; la senilidad, entre los 75 y los 90 años; y a las personas mayores de 90 años se las considera longevas (Dyussenbayev, 2017). La dimensión sociocultural de la edad está vinculada a factores externos y sociales, como la educación o el trabajo. Además, debe tenerse en cuenta que el proceso de envejecimiento como realidad social está fuertemente correlacionado con la forma en que se percibe el propio envejecimiento (Rowe y Kahn, 1997) por lo que, en gran medida, se trata de un proceso construido individual y socialmente.

Desde una perspectiva sociolingüística, la mayor parte de los estudios sobre cambios lingüísticos a lo largo de la vida realizados hasta el momento se han centrado en los niveles fonético (Sankoff y Blondeau, 2007, 2013; Reubold y Harrington, 2015) y gramatical (Wagner y Sankoff, 2011), con menor atención a cuestiones pragmáticas o de discurso (Buchstaller, 2015). Se trata, pues, de una línea de estudio que aún ofrece muchas posibilidades para profundizar en otros niveles, como el léxico, que cuentan con menos tradición en los estudios de panel (Gerstenberg, 2015). Asimismo, las lenguas que han recibido un trato preferente dentro de esta especialidad han sido el inglés (Sankoff, 2018b) y el francés (Blondeau, 2001, 2006; Sankoff, 2018a), con atención parcial a otras lenguas como el danés (Jensen, 2017), el portugués (da Conceição De Paiva, Duarte y Guy, 2021) o el alemán suabo (Beaman y Tomaschek, 2021), por ejemplo. Sin embargo, la experiencia acumulada en la investigación sociolingüística del espacio hispanohablante hace del español una lengua idónea para el estudio del cambio lingüístico en tiempo real, aunque hasta el momento esa experiencia es mínima y prácticamente inexistente.

La sociolingüística del cambio a lo largo de la vida está siendo liderada por investigadores procedentes de los Estados Unidos y de Alemania. Entre los nombres propios más destacados deben mencionarse William Labov (1978, 1994, 2001), Nikolas Coupland (2001; Coupland *et al.*, 1991), Gillian Sankoff (2018b, 2019), Suzanne Wagner (2012),

Isabelle Buchstaller y Suzanne E. Wagner (2017). La perspectiva desde la que se trabaja es la del estudio de la variación y el cambio en tiempo real, centrándose en la conducta lingüística de los hablantes durante un periodo de tiempo y usando como metodología básica el estudio de panel, que consiste, como ya he mencionado, en estudiar a un grupo de hablantes individuales en un momento determinado, para volver a estudiarlo en un momento posterior de tal manera que el objeto de los análisis no son las comunidades, sino los hablantes concretos.

Las experiencias acumuladas por los estudios de panel sobre el cambio lingüístico a lo largo de la vida son todavía escasas y limitadas, pero han demostrado su clara vinculación tanto con otras perspectivas lingüísticas, como con otros aportes de la propia sociolingüística. En relación con otras perspectivas lingüísticas, es evidente la complementariedad de los estudios sobre el cambio a lo largo de la vida con la lingüística del desarrollo, la adquisición de lenguas, los estudios de la tercera edad (Abrams y Farrell, 2011) e incluso los análisis del deterioro y las patologías lingüísticas. En conjunto, estos estudios permiten comprobar bien la labilidad o bien la estabilidad del sistema lingüístico en el curso de la vida.

En cuanto a la relación de este tipo de estudios con otras propuestas sociolingüísticas para el análisis del cambio lingüístico, es muy clara su conexión con los estudios del cambio en tiempo aparente, no solamente por ofrecerles una alternativa, sino por la posibilidad de cruzarlos o combinarlos, por servir para poner a prueba los resultados de los análisis del cambio en tiempo aparente o incluso para verificar la hipótesis del cambio en tiempo aparente en su conjunto.

Por otro lado, los estudios de panel a lo largo de la vida permiten comprobar cómo se imbrican los cambios comunitarios y los cambios individuales. Cuando se detecta un cambio, existe la posibilidad de interpretarlo como propio de los individuos del panel —o, tal vez, de su respectiva generación— o como rasgo de la comunidad en su conjunto. De las complicadas intersecciones de los cambios individuales y los comunitarios ya dio cuenta William Labov en su *Principios del cambio lingüístico* (1994). Efectivamente, se trata de contrastar los resultados sobre los estudios del cambio individual con el cambio comunitario, pero también de complementarlos con los estudios de tendencia y de entender cómo los individuos responden a los cambios comunitarios y cómo las comunidades responden a los cambios individuales. Este es uno de los grandes retos de la sociolingüística del cambio.

La forma en que los estudios del cambio lingüístico a lo largo de la vida se han abordado se ha basado principalmente en la recolección y análisis de lengua hablada dentro de una comunidad, con la ventaja de que estos contextos suelen estar muy bien caracterizados y permiten ahondar en las identidades de los hablantes. En Montreal, Thibault y Vincent (1990) entrevistaron en 1984 a 60 hablantes de francés de los 120 que habían sido entrevistados en 1971 (Sankoff y Lessard, 1975). Dando continuidad a la cadena, Vincent *et al.* (1995) volvieron a entrevistar una tercera vez a 12 de los informantes de la primera muestra, junto a dos informantes de los entrevistados en 1984.

Aparte de esta metodología de corte sociolingüístico, combinada en ocasiones con aproximaciones de naturaleza más pragmática, también se han hecho estudios a partir de

datos procedentes del dominio público, como el realizado por Gillian Sankoff para estudiar cambios fonológicos a partir de una serie de documentales británicos o como el ya citado de Juan Manuel Hernández Campoy con materiales radiofónicos. Y a estas posibilidades podría añadirse la combinación de materiales procedentes de estudios de panel con otros basados en tendencias (Sundgren, 2009), procedimiento que aporta interesantes argumentos para comprender la intersección entre lo individual y lo comunitario.

Como puede fácilmente suponerse, las dificultades metodológicas que supone abordar un estudio de panel son distintas de las que se experimentan en los estudios sociolingüísticos de periodo fijo. Una de ellas consiste en que conforme una generación envejece, se ve menguada progresivamente en su número de integrantes, cuya localización, además, también resulta complicada. Por otro lado, las comunidades también experimentan transformaciones en el curso de la vida, debidas a la movilidad social o al desplazamiento de poblaciones, lo que puede afectar a la representatividad de las muestras. En este caso, sin embargo, la caracterización de los hablantes individuales puede paliar las dificultades derivadas de los cambios sociales o demográficos de la comunidad.

Las publicaciones aparecidas a lo largo de los últimos diez años sobre el cambio lingüístico en el curso de la vida han ido sumando aportaciones, conocimientos y conclusiones que sirven de explicación de lo analizado en cada caso, aunque también funcionan como hipótesis de posible verificación en análisis referidos a otras comunidades o a otras lenguas. Una de las aportaciones teóricas fundamentales ha consistido en proponer que las tendencias de cambio a lo largo de la vida pueden responder a distintas pautas o producirse de distintas maneras.

La pauta evolutiva más simple es la «continuidad» o la «estabilidad», en la que los análisis no arrojan datos sustancialmente diferentes entre los distintos momentos de las entrevistas. La segunda pauta es la «regresión» y consiste en recuperar en edades avanzadas usos que habían sido característicos de edades anteriores; pongamos por ejemplo la recuperación de rasgos dialectales abandonados durante cierto periodo de la vida por conveniencia laboral o por desplazamiento a otras comunidades. La tercera pauta es la «gradación de procesos por edades» (*age-grading*) y en ella las distintas etapas de la vida revelan una modificación o cambio de los usos lingüísticos de los hablantes hacia soluciones diferentes. Dentro de esta pauta de *age-grading*, Gillian Sankoff ha introducido recientemente una interesante matización al distinguir el *age-grading* propiamente dicho (cambios de los usos de los hablantes en cada generación), del «cambio de vida», referido a aquellos cambios que los hablantes adoptan en consonancia con cambios en progreso en el conjunto de la comunidad (Labov, 1994: 83; Sankoff, 2019). En las innovaciones léxicas, por mucho que la generación de los *boomers* en España se conciba como usuaria de expresiones viejunas como *guay del Paraguay* o *cantidubi* ‘mucho’ no es óbice para que las abuelas les pregunten a sus nietos por su *crush* o que les digan que algo *mola mazo*, aunque sea poco frecuente.

Otra de las conclusiones alcanzadas por estos estudios revela que todos los hablantes manifiestan algún cambio de conducta lingüística a lo largo del tiempo, si bien las contribuciones de los individuos a los cambios comunitarios son más dé-

biles que las innovaciones que proceden de la irrupción de una nueva generación. Esto se debe en parte a que los individuos ofrecen una diversidad de soluciones y comportamientos particulares, que se acomodan de modo diferente a los cambios comunitarios. Asimismo, estos estudios del cambio han comprobado que los niveles o subsistemas de la lengua se comportan de forma diferente en su evolución (Sankoff *et al.*, 2001, 2019). Así, la fonología se muestra como el nivel más estable y el léxico como el menos estable en el curso de la vida. Ello contribuye a reforzar la idea de la labilidad de la lengua de los individuos a lo largo del tiempo. Además, la conducta lingüística, especialmente el uso de formas más o menos cultas o estándares, puede verse afectada por factores sociales externos, a menudo relacionados con la educación y el género de los hablantes. Existen variables lingüísticas que disfrutan de una mayor relevancia (*salience*) social, construida desde la escuela, los roles sexuales o los entornos laborales, que frecuentemente se asocian a cambios desde arriba. En este sentido, se convierte en especialmente interesante la comparación de los efectos individuales de los cambios desde arriba y de los cambios desde abajo.

Para concluir esta revisión de los estudios del cambio a lo largo de la vida, quisiéramos introducir una reflexión desde una perspectiva histórica. Y es que los rasgos de las lenguas pueden tener un comportamiento diferente en su desarrollo o evolución, según su naturaleza. En líneas generales, al proceso evolutivo se le da el nombre de «cambio lingüístico», pero, dentro de tal proceso, puede distinguirse entre modificaciones de naturaleza cuantitativa y de naturaleza cualitativa. Los cambios cuantitativos suponen un aumento o una reducción de la frecuencia de una condición determinada o en el uso o aparición de un rasgo determinado, en cuyo caso podría hablarse de una «evolución». Los cambios cualitativos suponen la modificación, alteración o sustitución de un rasgo o característica por otro, como se documenta en los procesos de elisión o asimilación fonéticas, en los casos de gramaticalización o en la adopción de préstamos o neologismos.

#### 4. EL ESTUDIO DEL CAMBIO LINGÜÍSTICO DESDE PRESEEA

Un proyecto de la envergadura de PRESEEA no puede permanecer ajeno a la irrupción en sociolingüística de los estudios del cambio a lo largo de la vida. Y esto es así por muchas razones, entre las que destaca la responsabilidad de incorporar decididamente la lengua española a esta gran corriente de estudios. Como se ha comentado previamente, no son muchos los equipos de PRESEEA en condiciones de abordar estudios de panel: Caracas, Monterrey, Ciudad de México, Málaga y Madrid. Pero sí son suficientes para que no haya excusa en su desarrollo, a la espera de que se reúnan más materiales disponibles. Expondré dos muestras de este tipo de análisis, desde la consciencia de las muchas limitaciones metodológicas y técnicas que se han hallado para su desarrollo.

#### 4.1. La evolución de la riqueza léxica a lo largo de la vida: estudio sobre Madrid

Un primer análisis procede de un estudio de panel de la evolución de la riqueza léxica a lo largo de la vida (Ávila Muñoz, 2014; García Marcos 2021), en el que se han utilizado materiales procedentes de la lengua española. Se trata de un análisis parcial y experimental, dado que trabaja solamente con 6 informantes procedentes de la ciudad española de Madrid, con los que se ha mantenido el mismo tipo de entrevista semidirigida en 2001 y en 2022 (Moreno Fernández, 2023).

A partir de las 12 entrevistas realizadas en 2001 y 2022, procedentes de 6 informantes, un primer paso obligado para el análisis de la riqueza léxica es la descripción cuantitativa de los textos, incluida su extensión o longitud, su número de palabras. Para ello se ha utilizado con herramienta el programa *T-Lab*, orientado al análisis y la minería de textos (Lancia, 2020). En los textos analizados, se observa que su longitud es bastante pareja. Esto se debe al hecho de que las entrevistas de 2001 y de 2022 repetidas con los mismos informantes se realizaron siguiendo las mismas directrices metodológicas. A cada una de las unidades o vocablos que forman un texto se le da el nombre de «*token*», de tal manera que el número total de *tokens* de un texto es su número total de palabras; esto es, su extensión total incluidas todas las posibles apariciones múltiples de cada palabra.

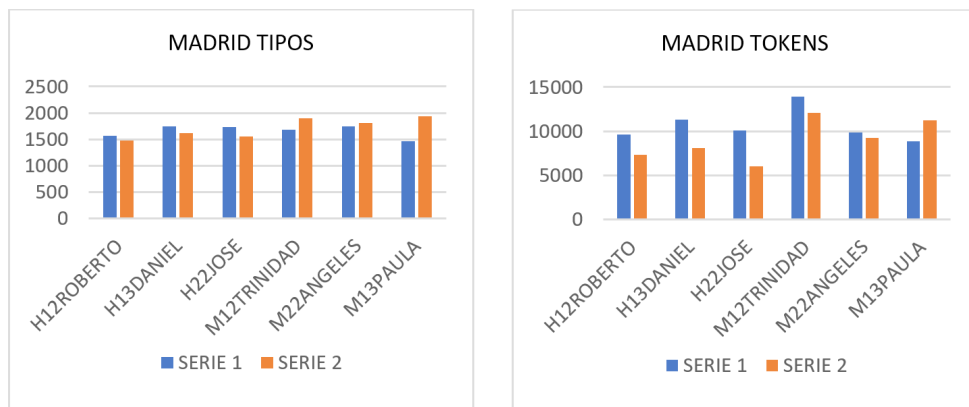
Otra medida de cuantificación es la de los tipos de palabras (*types*). Se considera que un tipo es cada palabra diferente, desconsiderando sus posibles repeticiones. Así, el número de tipos dentro un texto será el número de palabras diferentes que contenga, mientras que el número de total de vocablos vendrá expresado en forma de número de tokens. El Cuadro 1 muestra el número de tipos y de vocablos de cada uno de los entrevistados.

**CUADRO 1. FRECUENCIAS ABSOLUTAS DE TIPOS Y DE TOKENS EN LOS 6 INFORMANTES ENTREVISTADOS EN MADRID EN 2001 Y EN 2022, CON INDICACIÓN DEL CÓDIGO DE IDENTIFICACIÓN SOCIOLINGÜÍSTICA DE CADA HABLANTE**

TYPES	H12 ROBERTO	H13 DANIEL	H22 JOSE	M12 TRINIDAD	M22 ANGELES	M13 PAULA
2001	1572	1742	1733	1686	1747	1469
2022	1484	1616	1549	1904	1805	1934
TOKENS						
2001	9635	11306	10114	13958	9895	8867
2022	7366	8093	6046	12061	9237	11215

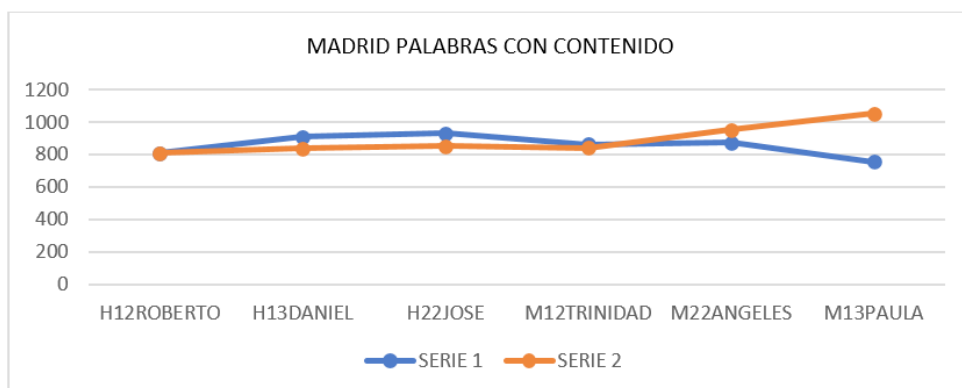
Los gráficos 4 y 5 muestran la cercanía de las entrevistas en cuanto al número de tipos y de vocablos utilizados por los informantes. Este factor cuantitativo es relevante por cuanto el análisis de la riqueza léxica es sensible a las diferencias significativas que pueda haber en la longitud de los discursos o textos analizados.

**GRÁFICOS 4 Y 5. HISTOGRAMAS DE NÚMERO DE TIPOS Y DE NÚMERO DE TOKENS POR ENTREVISTA EN 2001 (SERIE 1) Y EN 2022 (SERIE 2)**



Otra medida interesante para la caracterización cuantitativa de las entrevistas analizadas es el número de palabras con contenido; es decir, aquellos tipos de palabras que tienen significado léxico y no gramatical. Quedan excluidos, por tanto, de este recuento artículos, pronombres, preposiciones, conjunciones y otros elementos sin significado léxico. En este caso, no observamos diferencias relevantes entre 2001 (Serie 1) y 2022 (Serie 2), excepto en el caso de la mujer con estudios superiores: PAULA M13 > M23.

**GRÁFICO 6. NÚMERO DE TIPOS CON CONTENIDO EN LAS ENTREVISTAS REALIZADAS A 6 HABLANTES DE MADRID EN 2001 (SERIE 1) Y 2022 (SERIE 2)**



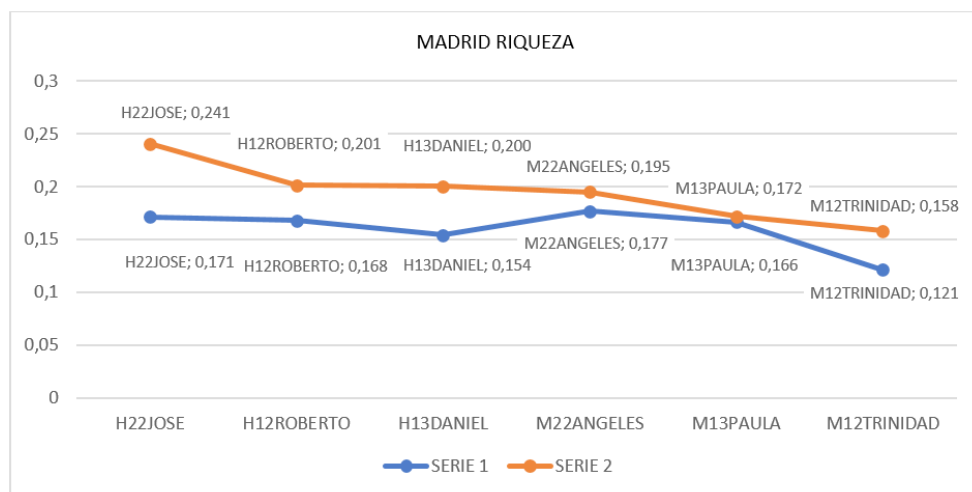
	<b>H12 ROBERTO</b>	<b>H13 DANIEL</b>	<b>H22 JOSE</b>	<b>M12 TRINIDAD</b>	<b>M22 ANGELES</b>	<b>M13 PAULA</b>
<b>2000</b>	808	909	930	864	873	754
<b>2022</b>	807	841	850	842	952	1056

El cálculo de la riqueza léxica de los hablantes madrileños analizados se realizó aplicando la ratio Tipos/Tokens (TTR). Los resultados del cálculo del TTR se ofrecen en el Cuadro 2 y su representación visual en el Gráfico 7.

**CUADRO 2. RESULTADOS DE CÁLCULO DE TTR POR HABLANTE Y PERIODO DE TIEMPO**

	H12 JOSE	H22 ROBERTO	H13 DANIEL	M22 ANGELES	M13 PAULA	M12 TRINIDAD
2001	.171	.168	.154	.177	.166	.121
2022	.242	.201	.200	.195	.172	.158

**GRÁFICO 7. RESULTADOS DE CÁLCULO DE TTR POR HABLANTE Y PERIODO DE TIEMPO**



En la observación de estos resultados se aprecia que, en todos los casos, las entrevistas realizadas en 2022 arrojan un índice de riqueza léxica mayor que las entrevistas realizadas en 2001. Solamente merece comentario el hecho de que la informante M13, Paula, que había mostrado una mayor diferencia en las frecuencias de palabras con contenido, ofrece una menor distancia entre la riqueza léxica de 2001 y 2022.

Este análisis partía de dos fundamentales preguntas de investigación relativas a la riqueza léxica en la lengua hablada dialogada. La primera se refería a la posible existencia de un incremento de la riqueza léxica a lo largo de la vida, comprobada a partir de la lengua hablada en situación de entrevista en dos periodos de tiempo. La segunda se refería a la incidencia de factores sociolingüísticos en el posible incremento de la riqueza léxica.

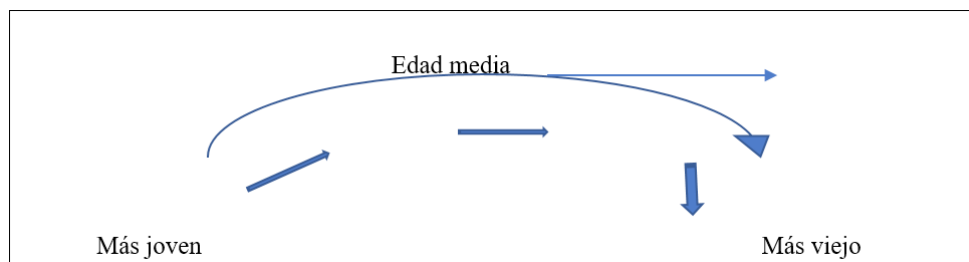
En relación con la primera cuestión, los datos procedentes de 6 informantes de la ciudad española de Madrid, en entrevistas realizadas en 2001 y en 2022, indican que en todos los casos se observa un incremento de la riqueza léxica, medida a tra-

vés del índice TTR. Este resultado se alinea con el de otros muchos estudios sobre la evolución del léxico en hablantes sanos en el transcurso de su vida (Gerstenberg, 2015). Asimismo, las observaciones individuales nos permiten afirmar que tal proceso puede interpretarse de forma individual, puesto que los incrementos observados son variables; de hecho, estudios anteriores han señalado que el comportamiento individual puede ser inestable en el transcurso del tiempo (Sankoff, 2005).

Desde una perspectiva sociolingüística, la significación del factor «edad» es redundante con el incremento general de la riqueza léxica, por lo que serían los factores sexo/género y nivel educativo los que merecerían una discusión específica. En términos generales, cabe aceptar la hipótesis de que los factores biográficos y ambientales, expresados a través del nivel educativo y la clase socioprofesional, podrían estar correlacionados con los indicadores de riqueza léxica (García Marcos 2021). Así lo comprobaron también Sankoff y Lessard (1975) para el francés de Montreal. Tales factores biográficos, sin embargo, no se han revelado como significativos en el nivel comunitario entre los informantes de Madrid.

De acuerdo con estas observaciones, la riqueza léxica evoluciona de forma inversa al uso de la lengua estándar explicado por Allan Bell (2014) (cambio en forma de U), según el cual los más jóvenes y los mayores tienden a utilizar formas no estándar, mientras que los de mediana edad, tienden a un mayor uso del estándar. En el caso del léxico, habría un proceso de enriquecimiento continuo, en forma de asíntota, que solo decaería motivado por un deterioro cognitivo, si este llega a producirse. De este modo, en el tramo más avanzado de la vida se puede encontrar una doble posible trayectoria: 1) el nivel de riqueza léxica puede mantenerse o incluso incrementarse, según lo observado; 2) el nivel de riqueza léxica puede decaer. Esta propuesta acepta, pues, que la competencia léxico-semántica es dinámica a lo largo de la vida.

**GRÁFICO 8. ESQUEMA DE EVOLUCIÓN DE LA RIQUEZA LÉXICA, CON INCREMENTO CONTINUADO Y POSIBLE CAÍDA POR DETERIORO COGNITIVO**



Si ponemos en relación nuestro análisis con las propuestas de modelos teóricos referidos a los procesos de cambio a lo largo de la vida, concluiríamos que el incremento de la riqueza léxica observado podría interpretarse como una gradación por edad (*age-grading*). Esta evolución condicionada por la edad supone nece-

sariamente cambios individuales a lo largo de la vida, pero no supone un cambio de la lengua en la comunidad. Esto es: el incremento de la riqueza léxica en el transcurso de la vida de cada hablante no significa que el conjunto de la comunidad incremente su riqueza léxica. En todo caso, podría interpretarse tal incremento como una evolución generacional hacia el estándar dominante (Bell, 2014), lo que no significaría necesariamente que toda la comunidad experimente tal evolución.

#### 4.2. Evolución de *lo que es* en la Ciudad de México

Con esta misma intención experimental, se ha abordado otro análisis de posible cambio en tiempo real a través de un estudio de panel. El objeto de estudio en este caso ha sido la expresión «lo que es», caracterizada por Lastra y Martín Butragueño (2016) como perífrasis informativa de relativo. Sirvan los siguientes ejemplos como muestras de lo que se consideran usos canónicos de «lo que es», dado que existen otras posibles formas y diferentes valores pragmáticos para esa expresión.

nos tocó rascar **lo que es** el frente de / de la casa / todo lo que es el frente  
ME-009-33H-97 (CSCM 2011)

entonces / todo eso te va formando también ya como que cierto escepticismo hacia  
**lo que es** // el amor / ¿no?»  
ME-106-21H-00 (CSCM 2011)

es un pueblo que está muy cercano como a una media hora/ de **lo que es** el centro  
de O  
ME-110-23M-22 (CSCM 2022)

normalmente nosotros/ hacemos **lo que es** la obra negra/ la cimentación  
ME-263-22H-22 (CSCM 2022)

Martín Butragueño y Lastra analizaron en su estudio de esta perífrasis sus propiedades sintácticas, semánticas y pragmáticas, así como su dispersión social. El estudio concluía que podía defenderse la existencia de una construcción específica, la perífrasis informativa de relativo (PIR), que resalta constituyentes que desempeñan ciertas funciones informativas, diferente a otras construcciones de relativo.

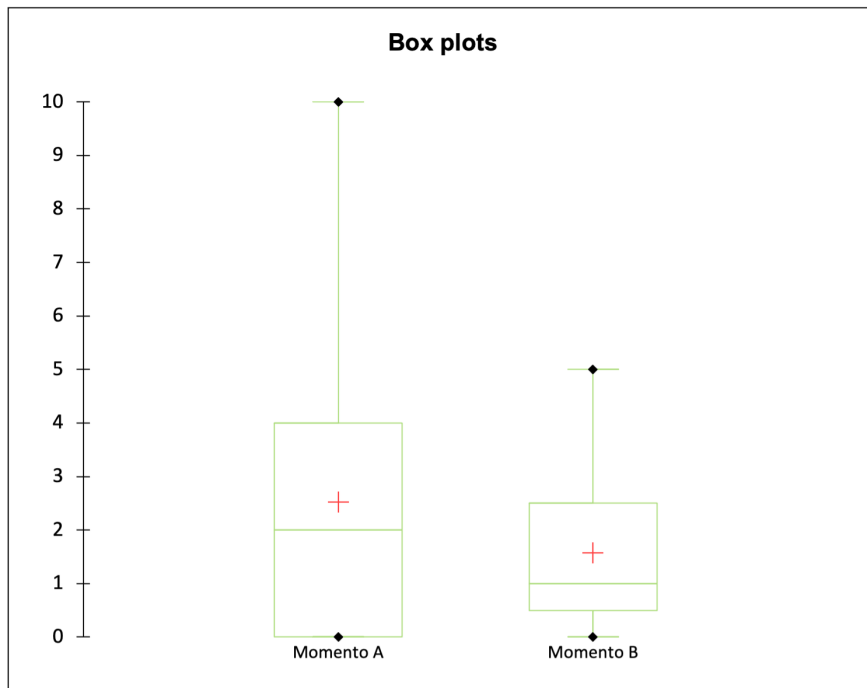
Entre los rasgos lingüísticos de las PIR canónicas se encuentran su eliminabilidad sin alterar las condiciones de verdad, el empleo obligatorio del relativo neutro y la presencia de un verbo débil (*ser*), que aparece preferentemente en presente y singular de tercera persona. En lo que se refiere a la distribución social de esta perífrasis, resultó que en México aparece mucho más en hombres que en mujeres y más en personas jóvenes. Estos últimos datos sugieren explorar la posibilidad de que se trate de un cambio en curso, que avanza por el camino de la difusión de las formas me-

nos marcadas y la estigmatización de las más marcadas (como *lo que viene siendo*). Para ello, sería necesario comparar los datos obtenidos con los de otras variedades geográficas, con datos procedentes de otros cortes temporales y, en especial, con materiales de diferentes registros.

Los datos sugieren en conjunto un patrón patrocinado por los hombres y los jóvenes. Hará falta más investigación en tiempo real y aparente para poder establecer si se trata de un cambio en curso, teniendo en cuenta las redes sociales, las ocupaciones y los eventos comunicativos específicos, pues llama la atención la elevada presencia de las PIR en algunas circunstancias. Así, es especialmente frecuente en ciertos casos, como en las crónicas radiofónicas de reporteros desplazados al lugar de la noticia (Estamos en lo que es la plaza de Tlatelolco), o en la descripción de la carta por parte de un mesero (Las enchiladas suizas traen lo que es queso y pollo), o en una agencia de viajes de la ciudad de México (le entrego los que viene siendo el boleto de los es su viaje). (Lastra y Martín Butragueño, 2016)

¿Qué está ocurriendo con la perífrasis informativa *lo que es* en la Ciudad de México una década después? En la búsqueda de una respuesta, hemos trabajado con materiales procedentes de PRESEEA – Ciudad de México (Corpus sociolingüístico de la ciudad de México) (Martín Butragueño y Lastra, 2011-2012), concretamente con los 19 informantes que fueron entrevistados, primero en la campaña de 1997-2005, con materiales publicados en 2011, y después en una nueva campaña realizada entre 2021-2022. La metodología utilizada en ambos momentos es básicamente la misma, si bien las entrevistas más recientes son, en conjunto, un tercio más cortas que las de 1997, por lo que la cifra absoluta de casos ha podido verse afectada.

El número de casos de *lo que es* canónico encontrados en estas entrevistas son un total de 78 (48 en 1997; en 20 en los 2022). Los perfiles sociológicos de los informantes incluyen hombres y mujeres de las tres generaciones previstas en PRESEEA y de los tres niveles educativos. Ocurre, sin embargo, que el bajo número total de entrevistas impide contar con una distribución proporcional de los perfiles. Si atendemos al conjunto de los datos obtenidos para cada uno de los momentos analizados, observamos que el número de casos *de lo que es* en 1997 es superior a lo encontrado dos décadas después. En el momento A, la frecuencia de uso es mayor, así como la dispersión de los datos. Ello nos llevaría a pensar que el cambio en marcha atisbado por un mayor uso de *lo que es* entre los jóvenes de principios de siglo no parece haber tenido continuidad.

**GRÁFICO 9. FRECUENCIA Y DISPERSIÓN DE LO QUE ES EN LA CIUDAD DE MÉXICO (1997-2022)**

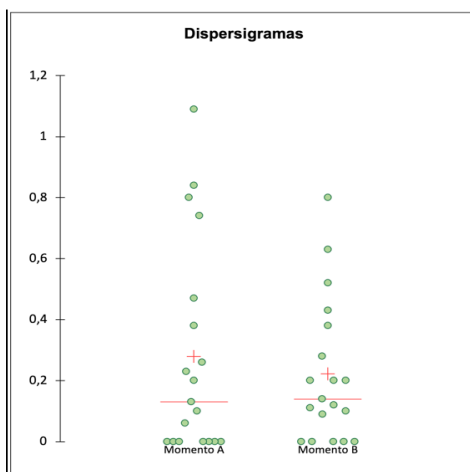
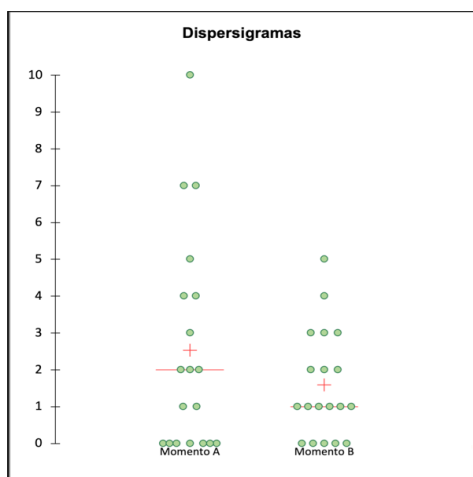
El problema que encontramos es que la muestra de datos manejada es pequeña y ofrece una distribución irregular. La prueba de Shapiro-Wilk nos dice que los datos no ofrecen una distribución normal. Asimismo, la prueba *t* de Student para dos muestras independientes nos lleva a confirmar que existen diferencias entre las muestras.

La precariedad de los datos obligó a explorar otro camino. Y es que, si este tipo de usos lingüísticos son sensibles al volumen total de materiales analizados (esto es, a la longitud de las entrevistas), podría observarse cuál es la ratio de uso de *lo que es* por mil palabras, de tal modo que las muestras de ambos momentos pudieran ser comparables sobre esta base. Para ello se procedió a calcular el número de casos por mil palabras en cada entrevista y en cada momento, para así analizar la diferencia entre ambos conjuntos de materiales (momento A y momento B). El resultado ofreció efectivamente una diferencia a la baja entre la ratio del momento A (0,24) y la del momento B (0,18). Esta diferencia no resultaba tan llamativa como las frecuencias absolutas nos hacían ver.

**CUADRO 3. NÚMERO DE PALABRAS Y RATIO DE PALABRAS POR INFORMANTE EN PANEL DE MÉXICO**

Informante	Nº pal. 97 (A)	Nº pal. 22 (B)	n A	n B	r A	r B	rA - rB
ME 002 31M 97	11334	7772	0	0	0	0	=
ME-005-22M-97	9459	8259	7	0	0,74	0	-0,74
ME-006-32H-97	9964	10730	2	1	0,20	0,09	-0,11
ME-008-32M-22	6548	9764	0	2	0	0,20	+0,2
ME-009-33H-97	13010	7372	5	3	0,38	0,38	=
ME-026-22M-98	11235	7789	0	1	0	0,12	+0,12
ME-106-22H-22	14406	9677	2	1	0,13	0,10	-0,13
ME-107-31M-00	9587	10708	1	0	0,10	0	-0,1
ME-110-23M-22	11240	7585	0	4	0	0,52	+0,52
ME-156-33H-01	7567	6823	0	3	0	0,43	+0,43
ME-197-31H-01	22268	11372	2	1	0,08	0,08	=
ME-213-33M-22	8387	7065	4	2	0,47	0,28	-0,19
ME-242-33M-04	6393	9751	7	2	1,09	0,20	-0,89
ME-245-33H-05	16066	7316	1	0	0,06	0	-0,06
ME-254-33H-21	15275	14452	4	3	0,26	0,20	-0,06
ME-255-32M-05	12757	8743	3	1	0,23	0,11	-0,11
ME-262-13H-05	8407	7116	0	1	0	0,14	+0,14
ME-263-22H-22	11896	7923	10	5	0,84	0,63	-0,21
ME-264-33M-05	12110	7906	0	0	0	0	=
TOTAL	217909	168123	48	30	0,24	0,18	-0,06
HOMBRES						0,21	0,22
MUJERES						0,26	0,14

**GRÁFICOS 10 Y 11. DISPERSIGRAMAS DE FRECUENCIAS ABSOLUTAS Y DE RATIO CASOS / 1000 PALABRAS**

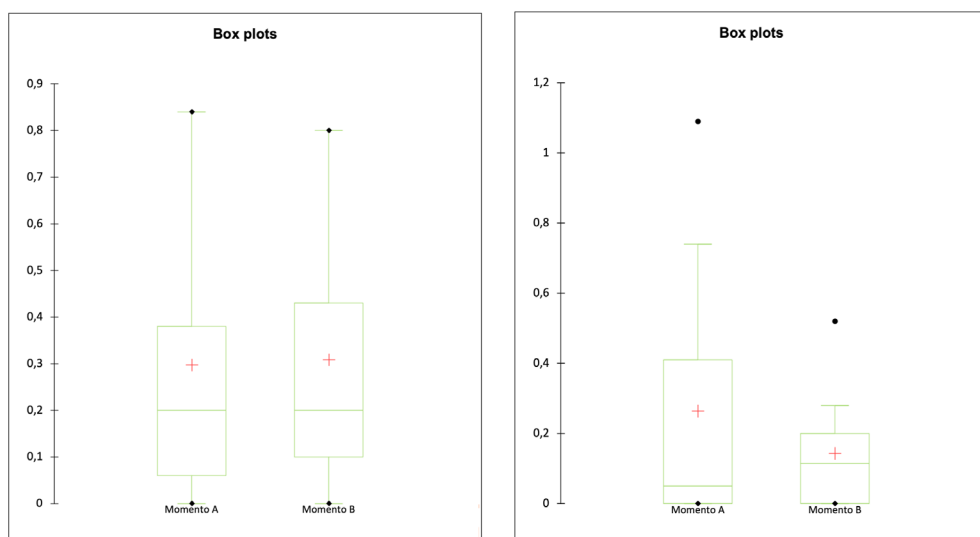


Por otro lado, el estudio de Martín Butragueño y Lastra apuntaba a la significación del factor sexo y, de un modo secundario, a la del factor edad en el uso de *lo que es*. Este último factor no ha permitido obtener conclusiones válidas por la irregular distribución de los informantes por su edad, pero el análisis de la ratio ofrece algunas pistas que pueden ser interesantes. Si los hombres eran los que en 1997 patrocinaban el uso de *lo que es*, los informantes analizados no lo demuestran, pero sí revelan que la ratio se ha mantenido en los hombres a lo largo de las últimas dos décadas, mientras que, en las mujeres, ha podido descender. Esto nos apuntaría a una posible pauta de estabilidad. Quedaría pendiente el análisis particularizado de los hablantes más destacados por sus frecuencias de uso de *lo que es*, así como de la naturaleza pragmática y semántica de los usos más y menos afectados por una posible evolución.

Si analizamos los datos de hombres y mujeres por separado, dado que el sexo fue la variable más significativa en el análisis de 2016, observamos que el uso de *lo que es* en los hombres no muestra significativas diferencias entre 1997 y 2022, mientras que sí hay reducción de su uso en las mujeres.

Las conclusiones que obtenemos de esta primera aproximación al posible cambio de *lo que es* entre 1997 y 2022 podrían apuntar a una cierta reducción de su uso en la Ciudad de México, aunque dentro de una pauta general de estabilidad. Atendiendo a los individuos, los gráficos nos muestran el uso destacado de *lo que es* en un hombre H12 > H22, que destaca frente a los demás y que muestra una reducción de su uso en el segundo momento, y una mujer M23 > M33, que muestra una mayor reducción proporcional.

### GRÁFICOS 12 Y 13. BLOX PLOTS DE *LO QUE ES* EN HOMBRES Y MUJERES POR LA RATIO CASOS/PALABRAS



## 5. CONCLUSIONES

La limitación de los datos aquí manejados, en cuanto a su volumen y en cuanto a distribución sociolingüística, aconseja enfatizar unas conclusiones de naturaleza eminentemente metodológica. La primera conclusión sería que es necesario contar con unos datos de panel cuantitativa y cualitativamente apropiados para el estudio del cambio en tiempo real. Los trabajos realizados con el francés de Canadá manejan muchos más datos de los que aquí se han podido utilizar. En cualquier caso, nos quedamos con la positiva experiencia de haber trabajado por primera vez con paneles de hispanohablantes y de poder contar con PRESEEA para contribuir al seguimiento de esta vía de estudio.

Por otro lado, pensando específicamente en PRESEEA, se hace necesario mantener la aplicación de las pautas metodológicas seguidas en etapas anteriores en todos sus aspectos, incluida la duración de las entrevistas. Acabamos de presentar algunos datos y resultados de Madrid y de la ciudad de México, pero debe decirse que hemos intentado realizar comparaciones de los mismos objetos de estudio en otras comunidades de PRESEEA que también ofrecen paneles y las diferencias en cuanto a la duración de las entrevistas han sido significativas. Probablemente, esto no habría sido inconveniente si los objetos elegidos para el estudio hubieran sido de naturaleza fonética o morfológica, pero PRESEEA aspira a ofrecer materiales aptos para muy variados tipos de análisis y es importante que la homogeneidad metodológica se mantenga en lo posible. Martín Butragueño y Lastra hablaban en 2016 de la necesidad de comparar datos de diversas variedades geográficas y PRESEEA asumió este como uno de sus compromisos, pero para ello hemos de garantizar la base metodológica que tan buenos resultados ha dado hasta el momento. Confiamos en que la calidad y la cantidad de los materiales permita su aplicación al ámbito del cambio en tiempo real, del que aquí se ha presentado un simple escarceo.

La sociolingüística –concretamente, la sociolingüística variacionista– se encuentra en unas condiciones inmejorables para abordar la evolución del habla en tiempo real y la caracterización del habla adulta no condicionada por patologías. Tengamos en cuenta que para 2050, una cuarta parte de la humanidad tendrá una edad igual o superior a los 65 años (Kowal, He, Goodkind, 2016). En este sentido, la sociolingüística tiene mucho que contribuir para el conocimiento de este perfil de hablantes, en complementariedad con los estudios de gerontología y de trastornos del lenguaje y, dentro de la lingüística, con los de lingüística histórica o los estudios del deterioro y la pérdida de lenguas. La virtud de la sociolingüística de los cambios a lo largo de la vida está en que contribuye a conocer por dentro la gran heterogeneidad lingüística que encierra el grupo de los adultos que no muestran un declive cognitivo, así como a detectar, prevenir o contrarrestar situaciones de edadismo.

El proyecto PRESEEA, dentro del ámbito hispanohablante, reúne las condiciones necesarias para poder convertirse en fuente y referencia para todo tipo de estudios del cambio en tiempo real, tanto en análisis de paneles como de tendencias. Estos

pueden abordarse de manera autónoma, pero también pueden ponerse en contraste con las conclusiones obtenidas del estudio del cambio en tiempo aparente. Al mismo tiempo, es importante analizar los cambios lingüísticos que se producen en el nivel comunitario, pero también en el nivel individual, así como –y esto es lo más complicado– analizar las imbricaciones entre los cambios individuales y los cambios comunitarios.

Para todo ello es necesario emprender nuevas campañas en las comunidades que cuentan con materiales reunidos hace dos décadas. Pero, es igualmente necesario que las nuevas campañas den continuidad metodológica a las líneas maestras sobre las que se ha construido PRESEEA. Ya resulta de por sí complicado mantener una mínima homogeneidad entre los materiales reunidos en casi cinco decenas de comunidades; pues más complicado resulta mantenerla cuando las campañas se reiteran. Necesitamos mantener las orientaciones metodológicas que han vertebrado nuestro proyecto, incluida la duración y las condiciones de las entrevistas. Es evidente que las dificultades que hay que superar son muchas, comenzando por que los paneles de campañas sucesivas van reduciendo sus tamaños, como muestra el caso de Montreal. Esta circunstancia nos obliga a un esfuerzo de investigación que de momento no es capaz de resolver ninguna inteligencia artificial de manera satisfactoria. De hecho, uno de los grandes retos para las tecnologías del lenguaje y para la inteligencia artificial es y será el manejo de la lengua hablada.

## REFERENCIAS

- Abrams, Lise y Meagan T. Farrell (2011). «Language processing in normal aging». En J. Guendouzi, F. Locke, y M.J. Williams (eds.), *The handbook of psycholinguistic and cognitive processes: perspectives in communication disorders*. Nueva York: Psychology Press, pp. 49-73.
- Asratían, A. (2007). «Variación -ra/-se en el español hablado en caracas». *Boletín de Lingüística*, 19(27): 5-41.
- Ávila Muñoz, Antonio M. (2014). «Patrones sociolingüísticos de la riqueza léxica. Estudio basado en una propuesta original para el cálculo del índice de la densidad léxica virtual de los hablantes». *Lingüística Española Actual*, 36: 171-194.
- Beaman, Karen y Fabian Tomaschek (2021). «Loss of historical phonetic contrast across the lifespan: articulatory, lexical, and social effects on sound change in Swabian». En K. Beaman, e I. Buchstaller (eds.), *Language variation and language change across the lifespan: theoretical and empirical perspectives from panel studies*. Londres: Routledge, pp. 209-234.
- Bell, Allan (2014). *The guidebook to sociolinguistics*. Oxford: Wiley.
- Bentivoglio, Paola e Irania Malaver (2012). «Corpus sociolingüístico de Caracas: PRESEEA Caracas 2004–2010. Hablantes de instrucción superior». *Boletín de Lingüística*, 24(37-38): 144-180.

- Blondeau, Helene (2001). «Corpora comparability and changes in real time within the paradigm of the personal pronouns in Montreal French». *Journal of Sociolinguistics*, 5(4): 453-474.
- Blondeau, Helene (2006). «Panel studies and language variation». En K. Brown (ed.), *Encyclopedia of language and linguistics*. Amsterdam: Elsevier, pp. 150-152.
- Briz, Antonio (2023). «Tres hipótesis explicativas del cambio dialectal: el caso del habla de la comarca de Requena-Utiel». *Homenaje a José Antonio Samper que le brinda la Academia Canaria de la Lengua* (en prensa). Academia Canaria de la Lengua.
- Buchstaller, Isabelle (2015). «Exploring linguistic malleability across the life span: age-specific patterns in quotative use». *Language in Society*, 44: 457-496.
- Buchstaller, Isabelle, y Suzanne E. Wagner (2017). «Introduction: Using panel data in sociolinguistic study of variation and change». En S. E. Wagner e I. Buchstaller (eds.), *Panel studies of language variation and change*. Londres: Routledge.
- Cedergren, Henrietta (1973). *The interplay of social and linguistic factors in Panama*. [Tesis doctoral]. Cornell University.
- Coupland, Nikolas (2001). «Age in social and sociolinguistic theory». En N. Coupland, S. Sarangi y C. Candlin (eds), *Sociolinguistics and Social Theory*. Harlow, UK: Pearson Education, pp. 185-211.
- Coupland, Nikolas, Justine Coupland y Howard Giles (1991). *Language, Society and the Elderly: Discourse, Identity and Ageing*. Oxford: Basil Blackwell.
- Da Conceição De Paiva, Maria, María Eugenia L. Duarte, y Gregory R. Guy (2021). «Alignment of individuals with community trends: subjects from the Portuguese». En K. Beaman e I. Buchstaller (eds.), *Language variation and language change across the lifespan: theoretical and empirical perspectives from panel studies*. Londres: Routledge, pp. 56-76.
- Dyussenbayev, Akhmet (2017). «Age Periods of Human Life». *Advances in Social Sciences Research Journal*, 4(6): 258-263.
- Flores, María E. (2012). «Polisemia y Funciones de ‘bastante en El Habla de Monterrey». En L. Rodríguez y M. Flores (eds.), *Estudios lingüístico-discursivos en emisiones lingüísticas reales. Aplicaciones al corpus de El habla de Monterrey*. Monterrey: UANL, pp. 239-258.
- Gallucci, María José (2020). «Paola Bentivoglio (1932–2019) y la lingüística en Venezuela». *Lingüística*, 36(1): 173-184.
- García Marcos, Francisco (2021). *Variación y cambio sociolingüístico en tiempo real. El español en la Costa Granadina (1987-2017)*. Granada: UJA Editorial.
- Gerstenberg, Annette (2015). «A sociolinguistic perspective on vocabulary richness in a seven-year comparison of older adults». En A. Gerstenberg y A. Voeste (eds.), *Language development: the lifespan perspective*. Amsterdam: John Benjamins, pp. 109-128.
- Guirado, Kristel (2014). «Corpus diacrónico del habla de Caracas 1987/2013». *Boletín de Lingüística*, 24(41–42): 19-44.

- Hernández Campoy, Juan Manuel (2003). «Complementary approaches to the diffusion of standard features in a local community». En D. Britain y J. Cheshire (eds.), *Social dialectology: in honour of Peter Trudgill*. Amsterdam: John Benjamins, pp. 23-37.
- Hunt, Stephen (2005). *The life course. A sociological introduction*. London/New York: Palgrave.
- Hurtado, Luz Marcela (2016). «Dinámica social en la expresión de impersonalidad en Bogotá: un cambio lingüístico evidente en tres periodos». *Lingüística y literatura*, 69: 177-191. DOI: <https://doi.org/10.17533/udea.lyl.n69a07>
- Jensen, Torbel Juel (2017). «Generic ‘du’ in time and context: a study of intra-individual variation and change». En D. Duncker y B. Perregaard (eds.), *Creativity and continuity: perspectives on the dynamics of language conventionalization*. Copenhagen: Copenhagen University Press, pp. 209-234.
- Kemper, Susan (2015). «Language production in late life». En A. Gerstenberg y A. Voeste (eds.), *Language development: the lifespan perspective*. Amsterdam: John Benjamins, pp. 59-75.
- Kowal, Paul, Daniel Goodkind y Wan He (2016). *An Aging World: 2015*, International Population Reports. U.S. Government Printing Office, Washington DC. <http://www.census.gov/library/publications/2016/demo/P95-16-1.html>
- Labov, William (1978). «On the use of the present to explain the past». En P. Baldi y R. Werth (eds.), *Readings in historical phonology*. Pennsylvania: Pennsylvania State University Press, pp. 275-312.
- Labov, William (1994). *Principles of linguistic change. Vol. I: Internal factors*. Oxford: Blackwell.
- Labov, William (2001). *Principles of linguistic change. Vol. II: Social factors*. Oxford: Blackwell.
- Lancia, Franco (2020). «T-LAB Plus 2020 – User’s Manual». *T-LAB*. <https://www.tlab.it/bibliography/>. Accedido el 6 de diciembre de 2022.
- Lastra, Yolanda y Pedro Martín Butragueño (2016). «Allá llega a lo que es el pueblo de San Agustín. El caso de la perífrasis informativa con lo que es en el «corpus sociolingüístico de la ciudad de México»». *Lingüística y literatura*, 69: 269-293.
- Lope Blanch, Juan M. (1986). *El estudio del español hablado culto. Historia de un proyecto*. México: Universidad Nacional Autónoma de México.
- Lope Blanch, Juan M. (1990-2000). *Atlas lingüístico de México*. 6 vols. México: El Colegio de México – Fondo de Cultura Económica - UNAM.
- López Morales, Humberto (1983). *Estratificación social del español en San Juan de Puerto Rico*. Ciudad de México: Universidad Nacional Autónoma de México.
- Lowry, Deborah (2022). *Aging and life course. Social and cultural contexts*. Rowman & Littlefield.
- Malaver, Irania (2017). «Ahorita: lexicalización y cambio lingüístico en la comunidad de habla caraqueña». *Nueva Revista de Filología Hispánica*, 65(1). <https://doi.org/10.24201/nrfh.v65i1.2828>

- Martín Butragueño, Pedro (2006). *Líderes lingüísticos: estudios de variación y cambio*. México: El Colegio de México.
- Martín Butragueño, Pedro (2012). «Variación y cambio lingüístico en el español mexicano». *Español Actual*, 98: 11-38.
- Martín Butragueño, Pedro (2014). *Fonología variable del español de México. I: Procesos segmentales*. México: El Colegio de México.
- Martín Butragueño, Pedro y Yolanda Lastra (2011-2012). *Corpus sociolingüístico de la Ciudad de México*. México: El Colegio de México.
- Moreno Fernández, Francisco (2023). «Language Evolution over the Lifespan: Experimental approach to Lexical Richness in Spanish». *Roczniki Humanistyczne*, 71(6): 165-184. DOI: <https://doi.org/10.18290/rh237106.8s>
- Moreno Fernández, Francisco (ed.) (2021). *Metodología del «Proyecto para el estudio sociolingüístico del español de España y de América. (PRESEEA)»*. Universidad de Alcalá (Documentos PRESEEA de trabajo).
- Moreno Fernández, Francisco y Ana M. Cestero Mancera (2020). «El proyecto PRESEEA: desarrollos analíticos». En A.J. Gallego y F. Roca Urgell (eds.), *Dialectología digital del español* (Verba Anexo 80). Servicio de Publicaciones e Intercambio Científico Campus Vida, pp. 119-138.
- PRESEEA (2010). «Documento de conclusiones». *V Encuentro Internacional PRESEEA*. Comillas. <https://preseea.uah.es/sites/default/files/2022-02/conclusiones%20Comillas%202010.pdf>
- Reubold, Ulrich y Jonathan Harrington (2015). «Disassociating the effects of age from phonetic change: a longitudinal study of formant frequencies». En A. Gerstenberg y A. Voeste (eds.), *Language development: the lifespan perspective*. Amsterdam: John Benjamins, pp. 9-38.
- Rodríguez Alfano, Lidia (coord.) (2005). *Investigación sociolingüística. El habla de Monterrey. Su trayectoria en una página electrónica*. Monterrey: Universidad Autónoma de Nuevo León.
- Rodríguez Flores, Dora E. y Lidia Rodríguez Alfano (1996). «El habla de Monterrey: historia de un proyecto». En D.E. Rodríguez Flores y L. Rodríguez Alfano (eds.), *Lenguaje y sociedad. Metodologías y análisis aplicados al habla de Monterrey*. Monterrey: Trillas, pp. 11-34.
- Rosenblat, Ángel (dir.) y Paola Bentivoglio (ed.) (1979). *El habla culta de Caracas. Materiales para su estudio*. Caracas: Facultad de Humanidades y Educación, Universidad Central de Venezuela.
- Rowe, John W. y Robert L. Kahn (1997). «Successful aging». *The Gerontologist*, 37(4): 433-440.
- Samper, José Antonio (1990). *Estudio sociolingüístico del español de Las Palmas de Gran Canaria*. La Caja de Canarias.
- Sankoff, Gillian (2005). «Cross-sectional and longitudinal studies in sociolinguistics». En P. Trudgill (ed.), *Sociolinguistics: an international handbook of the science of language and society*. Berlin: De Gruyter, pp. 1003-1013.

- Sankoff, Gillian (2006). «Age: Apparent time and real time». En K. Brown (ed.), *Elsevier Encyclopedia of Language and Linguistics* (2 ed.). Amsterdam: Elsevier s.v.
- Sankoff, Gillian (2018a). «Before there were corpora: the evolution of the Montreal French Project as a longitudinal study». En S.E. Wagner e I. Buchstaller (eds.), *Panel Studies of Variation and Change*. Londres: Routledge, pp. 21-52.
- Sankoff, Gillian (2018b). «Language Change across the Lifespan». *Annual Review of Linguistics*, 4: 297-316.
- Sankoff, Gillian (2019). «Language change across the life-span: three trajectory types». *Language*, 95(2): 197-229.
- Sankoff, David y Réjean Lessard (1975). «Vocabulary richness: a sociolinguistic analysis». *Science, New Series*, 190(4215): 689-690.
- Sankoff, Gillian y Helene Blondeau (2007). «Language change across the lifespan: /r/ in Montreal French». *Language*, 83(3): 560-588.
- Sankoff, Gillian y Helene Blondeau (2013). «Instability of the [r] ~ [R] alternation in Montreal French: an exploration of stylistic conditioning of a sound change in progress». En L. Spreafico y A. Vietti (eds.), *Rhotics: new data and perspectives*. Bolzano: Free University of Bozen-Bolzano Press, pp. 249-265.  
DOI: 10.13124/9788860461025
- Sankoff, Gillian, Helene Blondeau y Anne Charity (2001). «Individual roles in a real-time change: Montreal (r>R) 1947–1995». En H. Van de Velde y R. van Hout (eds.), *R-atics: Sociolinguistic, Phonetic and Phonological Characteristics of /r/*. Bruselas: ILVP, pp. 141-157.
- Sheppard, Valerie (2020). *Research Methods for the Social Sciences: an Introduction*. British Columbia Campus. <https://pressbooks.bccampus.ca/jibcresearchmethods/>
- Sundgren, Eva (2009). «The varying influence of social and linguistic factors on language stability and change: The case of Eskiltuna». *Language Variation and Change*, 21: 97-133.
- Thibault Pierrette y Diane Vincent (1990). *Un corpus de français parlé. Montréal 84: historique, méthodes et perspectives de recherche*. Québec: Département de langues et linguistique, Université Laval.
- Vincent, Diane, Marty Laforest y Guylaine Martel (1995). «Le corpus de Montréal 1995: Adaptation de la méthode d'enquête sociolinguistique pour l'analyse conversationnelle». *Dialangue*, 6: 29-46.
- Wagner, Suzanne E. (2012). «Real-time evidence for age grad(ing) in late adolescence». *Language Variation and Change*, 24: 179-202.
- Wagner, Suzanne E. y Gillian Sankoff. (2011). «Age grading in the Montréal French inflected future». *Language Variation and Change*, 23: 1-39.

\* He de expresar mi agradecimiento a las personas que, de un modo u otro, me han proporcionado materiales, datos, referencias y opiniones para la redacción de este texto. En México, he contado con la siempre valiosa ayuda de Pedro Martín Butragueño y, muy especialmente, con la de Dinorah Montiel, que me ha hecho

llegar unos materiales muy valiosos. En Málaga, Antonio Ávila y Matilde Vida me han proporcionado tanto materiales, como referencias y opiniones. En Heidelberg, he contado con el apoyo de Jana Weckesser; para los materiales de Madrid, he contado con la ayuda de Nicolás Moreno Martín de Nicolás. Las limitaciones de este texto son responsabilidad exclusiva del autor.

# WHAT DOES IT TAKE TO BE A NATIVE SPEAKER (AND WHO GETS TO SAY SO)?

Dennis R. Preston  
*University of Kentucky*

## 1. INTRODUCTION

This paper deals with problems of identity in the selection of respondents in dialectology and sociolinguistics, with special reference to Hispanic heritage populations in the United States, and considers both linguistic and social concerns in making such selections.

The Proyecto para el Estudio Sociolingüístico del Español de España y de América (PRESEEA) has taken a broad and inclusive approach in determining who should be included in regional sociolinguistic investigations of Spanish under its umbrella as have many other country and region specific studies of Spanish varieties in the Americas. This liberal approach, however, has not completely answered the question of who is to be included in such studies, nor if the answer should come from a primarily linguistic or social point of view.

I will not review influences on Spanish in the United States, such as those carefully studied in Silva-Corvalán (1994) and many others that show the influence of American English on Spanish from the phonetic to the semanto-pragmatic and every level in between. An equally impressive review of indigenous language influences on Spanish could be given, ranging from the reconfiguration of the vowels under the influence of the Quechua allophonic system (Pasquale, 2001) to the semanto-pragmatic influence of Quechua grammaticalized evidentials on Bolivian morphosyntax and pragmatics (Babel, 2009).

But I want to look at the influence of Spanish on English and turn first to the linguistic side and provide justification for its primacy. Forty years ago, Bell (1984)

suggested that stylistic variation (informal to formal) is always weaker than status variation (low to high). He concluded that style is derived from but is a weaker reflection of status. In a review of a large number of quantitative studies, I have also shown that the linguistic influences on variation are more powerful than those of status (Preston, 1991).

TABLE 1. DOMINICAN /s/-DELETION

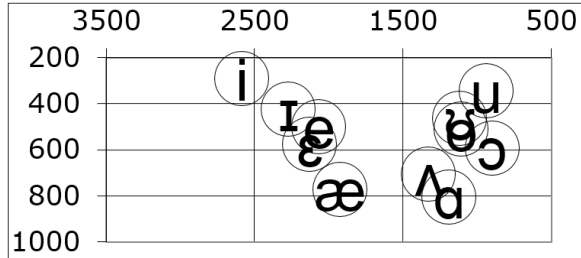
Factors	Log odds	Tokens (N)	Proportion
<b>Sex</b>			
Female	0.85	1871	0.361
Male	-0.85	1542	0.149
<b>Literacy</b>			
Literate	0.971	1982	0.347
Semiliterate	-0.971	1431	0.152
<b>Position + following context</b>			
Internal_N	3.804	38	0.500
Internal_T	1.714	810	0.406
Sandhi_V	0.569	507	0.323
Pre-pausal	0.538	829	0.238
Internal_F	-0.340	10	0.100
Sandhi_T	-0.695	442	0.208
Sandi_D	-0.760	419	0.112
Sandi_N	-0.841	216	0.176
Sandhi_L	-1.133	87	0.126
Internal_D	-1.171	21	0.048
Sandhi_F	-1.684	34	0.147

Source: Bullock *et al.*, 2014.

Table 1 shows the results of a study of [s]-lenition in Dominican Spanish. Stylistic variation was not significant, but the status characteristics of gender and literacy were, both facts in keeping with Bell's prediction. The linguistic factors of position in the word and following segment, however, were most powerful, as both the log-odds and proportions show. I will look first at some phonological evidence from Spanish heritage speakers in the US that supports this primacy and hope to show how such findings also address the question of inclusiveness in sociolinguistic and dialectological research.

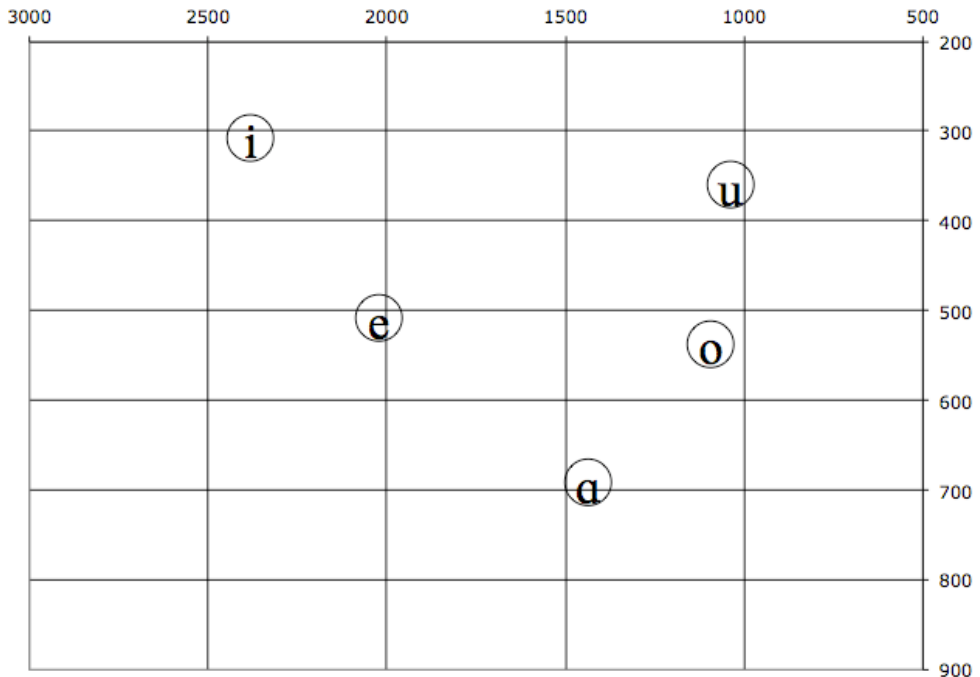
2. DIALECTOLOGY, SOCIOLINGUISTICS, AND RESEARCH: THE SELECTION OF RESPONDENTS

FIGURE 1. NORMALIZED F1/F2 MEANS SCORES OF AMERICAN ENGLISH VOWEL DATA



Source: Peterson and Barney, 1952; Stevens, 1998, for /e/ and /o/.

FIGURE 2. NORMALIZED F1/F2 MEANS SCORES OF TWO MEXICAN AMERICAN (SOUTH TEXAS) SPANISH SPEAKERS' VOWEL DATA PROVIDED IN NORM



Source: Thomas and Kendall, 2007.

**FIGURE 3. THE SPANISH FIVE-VOWEL REALIZATION OF THE “GENERAL AMERICAN” SYSTEM**

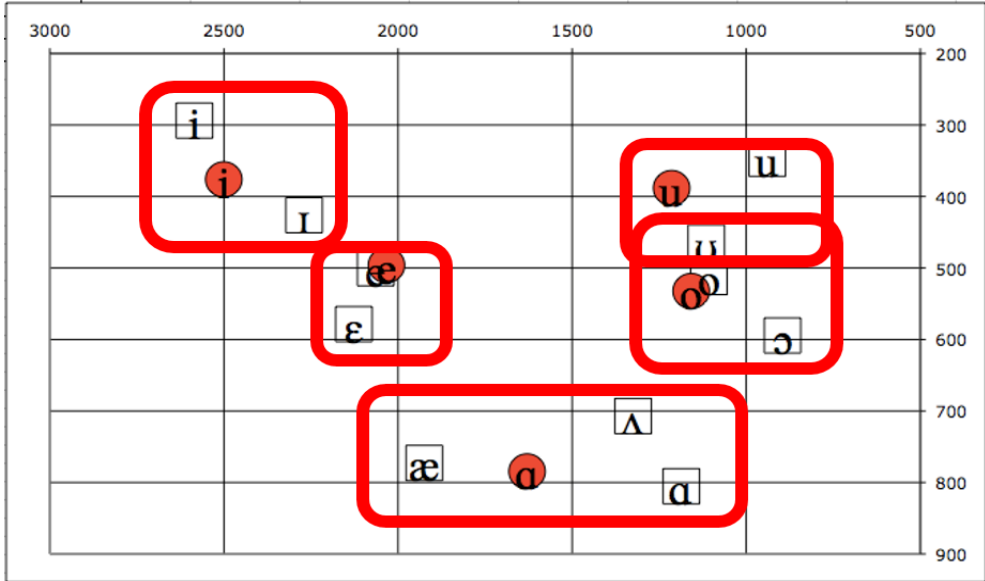
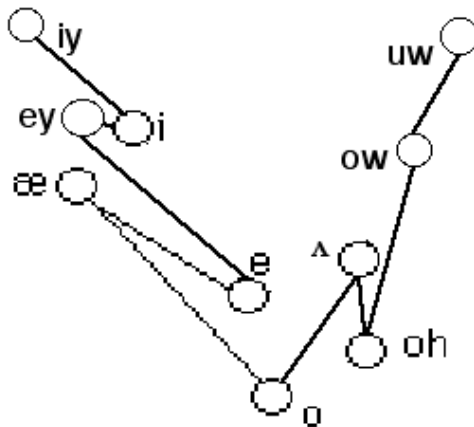


Figure 1 displays the vowel system of US English (ignoring its three diphthongs), and figure 2 shows the Spanish five vowel system. Figure 3 shows the two systems combined and displays the compromise (or “interference”) solutions that Spanish might impose. I believe this is nonsense. This so-called “General American” (GA) system does not exist (Van Riper, 1986) and therefore offers Spanish speakers no opportunity to acquire it.

**FIGURE 4 JANICE R., FEMALE, 14, DETROIT (MI)**

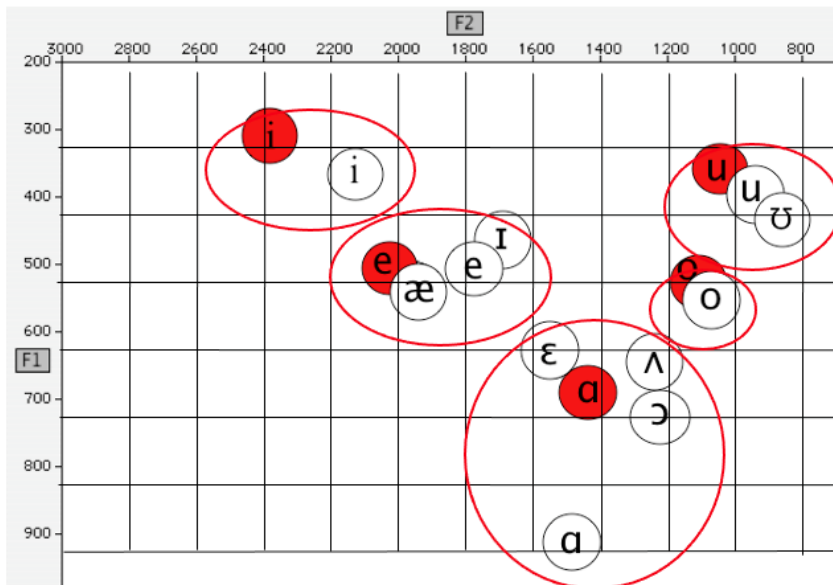


Source: adapted from Labov, 1996, Figure 2.

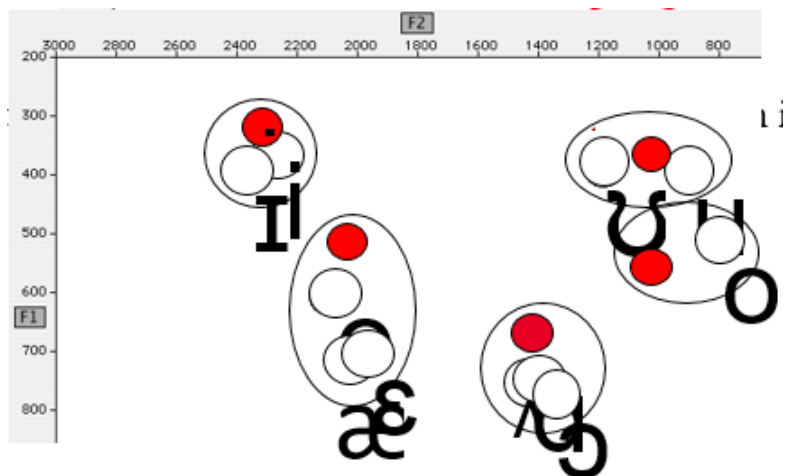
**FIGURE 5. THE AREA OF THE NORTHERN CITIES SHIFT (SHADED) AND THE AREA IN SOUTHERN MICHIGAN WITHIN IT WHERE THE WORK REPORTED HERE WAS DONE (ARROW)**



**FIGURE 6. PREDICTIONS OF THE SPANISH VOWEL SYSTEM'S INFLUENCE ON THE NCS SYSTEM**



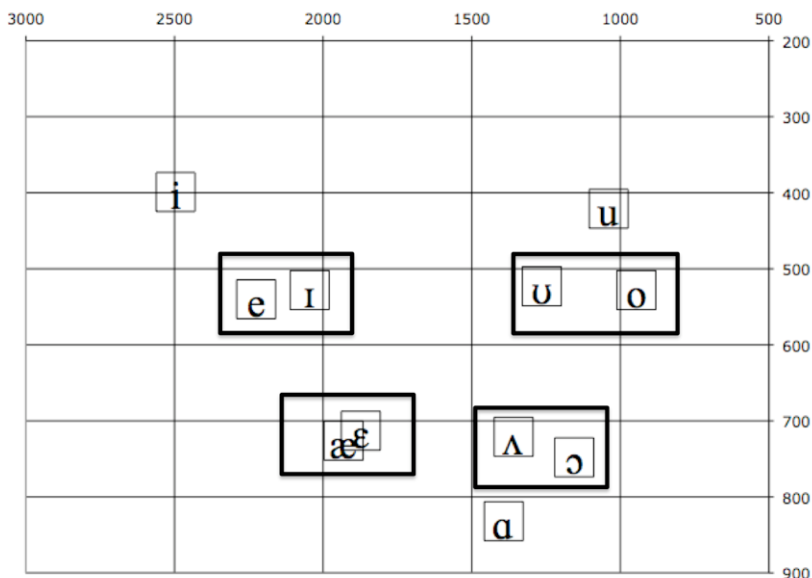
**FIGURE 7. AN ACTUAL SPANISH-DOMINANT SPEAKER (“DOMINGO”)  
BORN IN 1933 WHO LEARNED ENGLISH IN THE NCS AREA**



Source: Ocumpaugh, 2010.

Figure 4 shows an authentic vowel system of the area where the research reported in what follows was done – the area of southern Michigan within the “Northern Cities Shift” (NCS) (figure 5). It is a very different system from the imaginary one, and figure 6 represents how Spanish speakers might make sense of the NCS system rather than the phantom system of GA (figure 3). The organization of the vowels is very different, and figure 7 shows the system of a Spanish-dominant English speaker (“Domingo”) who learned English in this area. The influences of his original Spanish and the NCS system he encountered are both obvious, but I am less interested in him than the English of his grandchildren. Is there any trace of the Spanish and NCS systems in their speech?

FIGURE 8. THE “GRANDCHILDREN” OF “DOMINGO”



Source: Roeder, 2006.

Yes and no. Figure 8, a plot of the English vowels of young Mexican-American, English dominant speakers from Lansing, Michigan, suggests that it is not a system completely aligned to the five vowel spaces of Spanish, and the eleven vowels are not quite those of the NCS system either. Surely, however, this is the work of second language acquisition and heritage language specialists, not dialectologists. I believe such a limitation is historically and linguistically short-sighted.

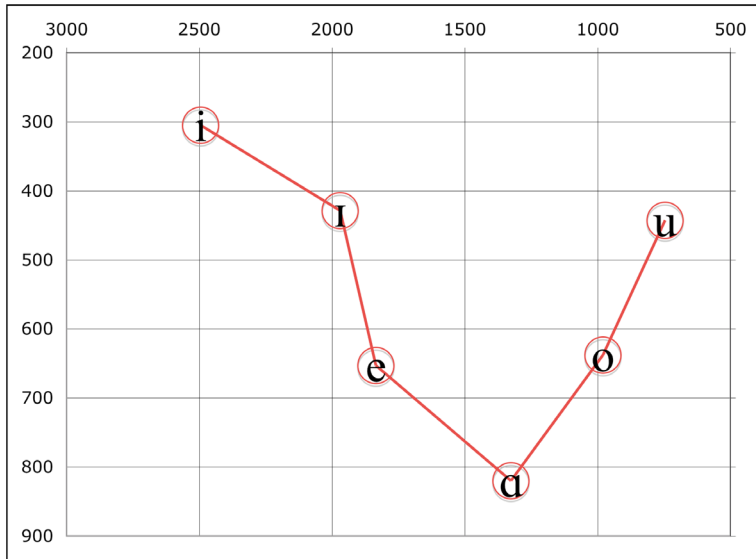
Look more closely at the system of Domingo’s grandchildren (Figure 8). This is an eleven-vowel system with distinctions between four pairs of tense and lax vowels in the high and low mid areas, both front and back. But the tense-lax pair /æ/ and /ɛ/ are surely the result of the NCS just as surely the Spanish system is responsible for the point vowels /i/, /a/, and /u/.

Neither the NCS nor Spanish, however, can be fully responsible for the other tense-lax pairs, front and back, nor for the impressive symmetry of this system. Like Spanish, it is a beautifully balanced system, perhaps even more impressive since it deals with eleven not just five vowels. But what is its source? To address this question, I turn to results collected in the same NCS regions of southern Michigan from speakers of other languages who moved to the same area and their grandchildren.

Figure 9 is the Polish system of six vowels (minus nasals), and figure 10 is the English vowel system of a first-generation respondent (“Jadwiga”) from Hamtramck, Michigan, who, like “Domingo”, has a system still showing the influence of her first

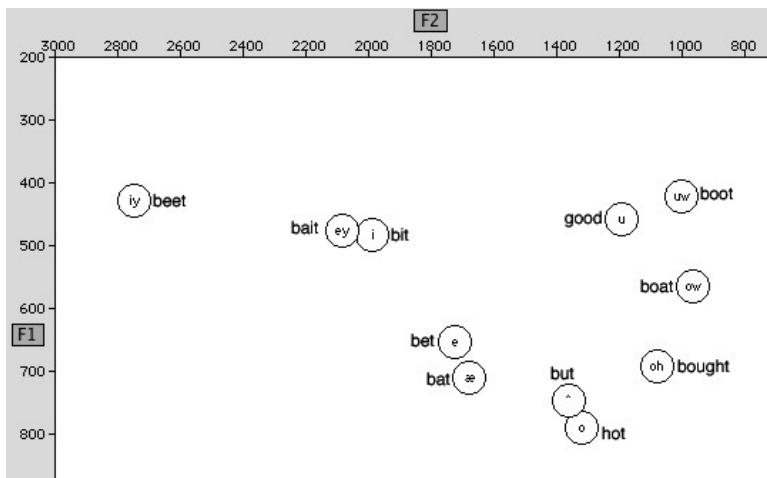
language but also influences of the NCS. Figure 11 shows “Jadwiga’s” grandchildren’s generation. They are not the offspring of Polish-American and Mexican-American intermarriage, but their English vowel systems look surprisingly similar to the latter.

**FIGURE 9. THE POLISH VOWEL SYSTEM, WITHOUT THE TWO NASAL VOWELS**



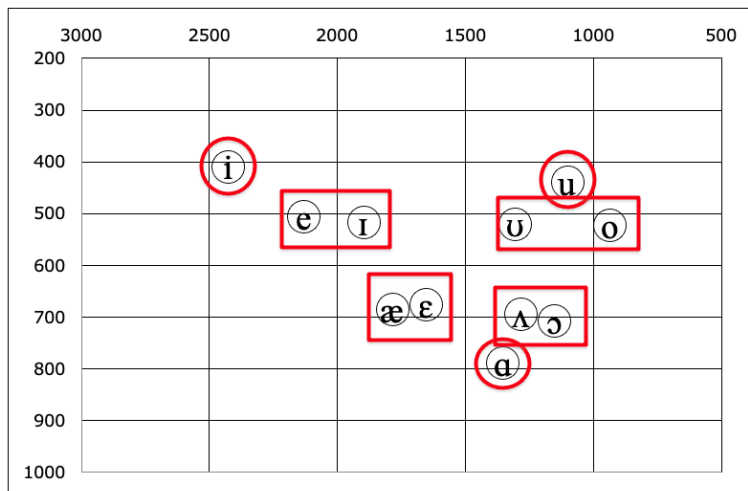
Source: Owsiany, 2001: 199.

**FIGURE 10. THE VOWEL SYSTEM OF “JADWIGA,” A FIRST-GENERATION POLISH AMERICAN**



Source: Firestone and Giese, 2008.

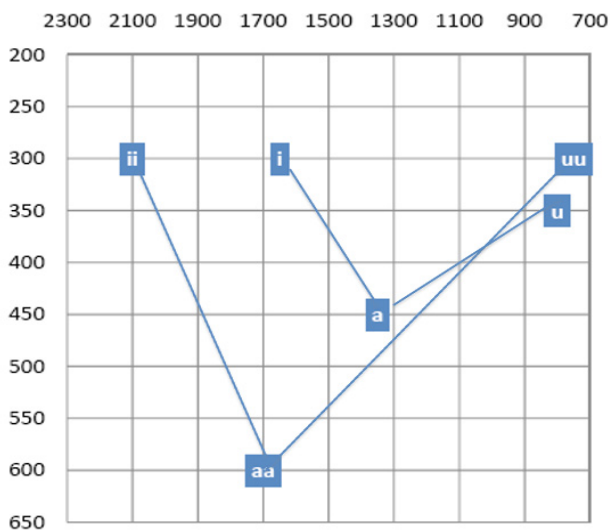
**FIGURE 11. THE VOWEL SYSTEM OF “JADWIGA’S” “GRANDCHILDREN”**



Source: Firestone and Giese, 2008.

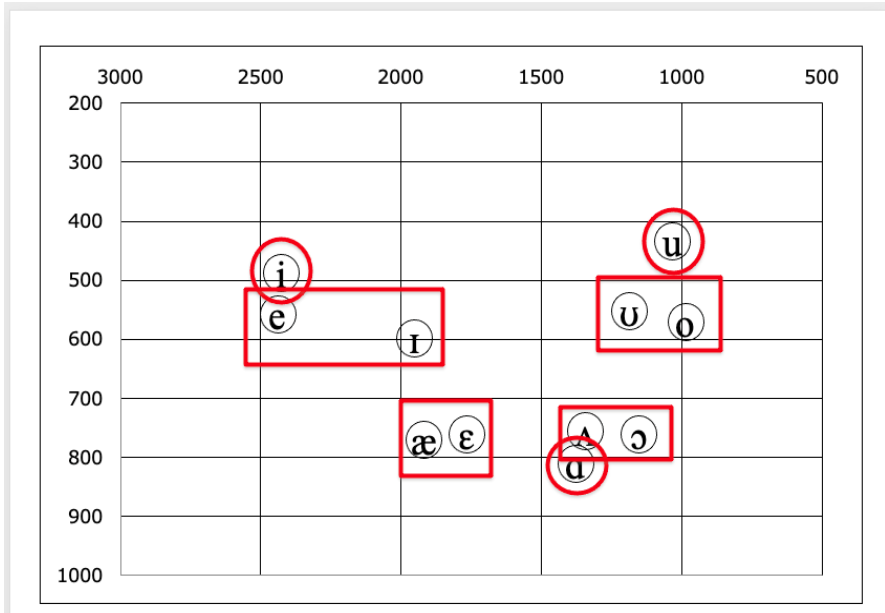
Figure 12 displays the very different system of Lebanese Arabic, and that of the grandchildren’s generation is shown in figure 13. It is very similar to the Spanish- and Polish-influenced systems, but there is one more that may be a bigger surprise than the similarities between these three.

**FIGURE 12. THE LEBANESE SIX-VOWEL SYSTEM**



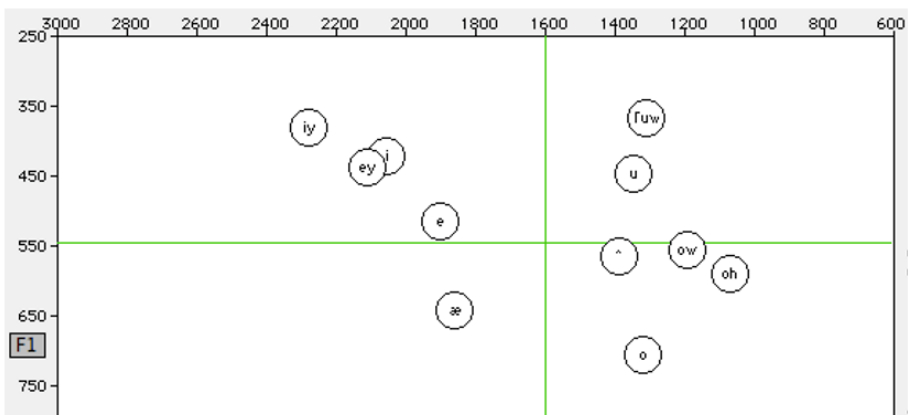
Source: Obrecht, 1968: 28.

**FIGURE 13. THE YOUNG LEBANESE-AMERICAN SYSTEM FROM DEARBORN, MICHIGAN**



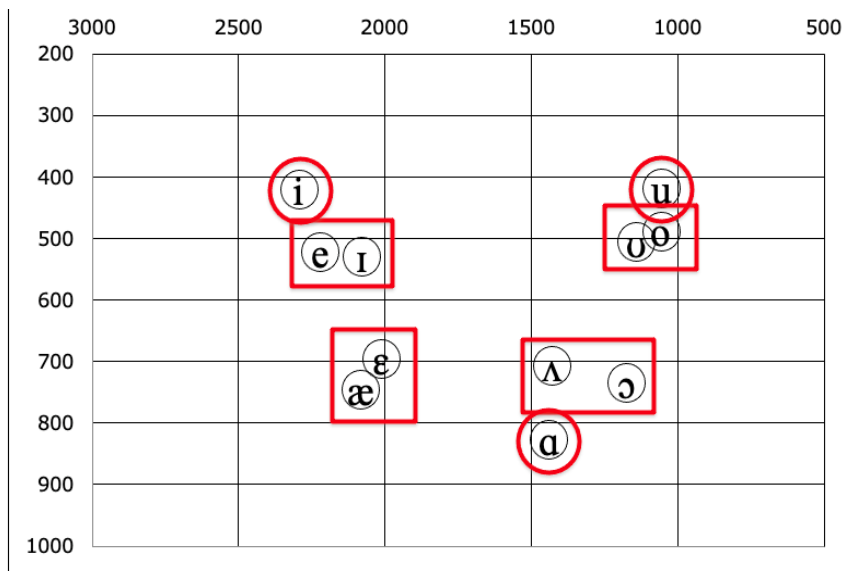
Source: Bakos, 2008.

**FIGURE 14. OLDER (“GRANDPARENTS”) ALABAMA AND MISSISSIPPI AFRICAN AMERICANS’ VOWEL SYSTEM, N=7, 3F, 4M, DERIVED FROM DICTIONARY OF AMERICAN ENGLISH (DARE) RECORDINGS**



Source: Jones, 2003.

**FIGURE 15. YOUNGER AFRICAN AMERICANS' VOWEL SYSTEM, EAST LANSING, MICHIGAN**



Source: Jones, 2003.

Figure 14 shows an African-American vowel system typical of many who came to the urban Midwestern North (the area of the NCS) from Mississippi and Alabama in what has been called the “Great Migration.” Figure 15 is the system of their “grandchildren,” one which should be familiar. Although the original system of the grandparents was also English, these younger speakers also found a way to modify it in a direction that creates the same symmetry as those who came from non-English language backgrounds. If the non-English systems were the only factors to be considered, symmetry might be the driving force inherited from the Spanish and Lebanese Arabic systems, but why do the asymmetric African-American and Polish systems end up with the same symmetry when they encounter the NCS?

The answer may lie in the fact that many studies have shown a decided preference for vowel symmetry (or balanced dispersion) from cognitive, theoretical, and empirical points of view. Maddieson (1984) shows this tendency in the examination of 192 world vowel systems from the UCLA Phonological Segment Inventory Database (UPSID); Honeybone (2010) provides a detailed history of the topic and Vaux and Samuels (2015) a more recent but still supporting view. It is surely reasonable to suggest that language and language variety learners, in their quest for simplifying strategies, will prefer symmetry, and that, having done so, may pass it on to future generations. It is also reasonable to assume that that system might not clash with surrounding local norms, and, in the long run, may even influence them. I believe

failure to consider local varieties that have emerged from contact and may show such tendencies as those seen here is an error in dialectology and sociolinguistics. How can such error be avoided?

In their conclusion to *Dialectology*, Chambers and Trudgill (1980) wrote that “Dialectology without sociolinguistics at its core is a relic”. It could be added that “Sociolinguistics and dialectology without linguistics at their core is a mistake”. What is the upshot of the fact that native speakers of English from other areas and backgrounds and descendants of some of the largest immigrant populations in southeastern Michigan (Polish, Mexican and Mexican American, and Lebanese) now all share a very similar system? One solution is to ignore them in studies of the English of the region. But there are thousands of them, and they are English-dominant speakers, many of them heritage or very weak speakers of the language or variety of their ancestors. Should they be excluded in work on English?

This is not the same problem Chambers and Trudgill identified. That was based on an overemphasis on NORMS (Nonmobile Older Rural Males), but the exclusion of large numbers of local English speakers from linguistic atlas activities suggests that another part of regional atlas errors rested in the mistake of ethnicity.

An interesting use of the word *ethnic* in US English is to refer to any person who is not a WASP (White Anglo-Saxon Protestant), and these so-called *ethnics* were systematically excluded from publication and even investigation in much work in the US. The *Handbook of the Linguistic Atlas of the Middle and South Atlantic States* (Kretzschmar *et al.*, 1993) explains how African American speakers from these areas were treated:

African-Americans were not included in the general grid of communities and informants. No African-American informants were interviewed in Kurath’s South Atlantic plan (11). ... (This) principle excluded African-Americans and other non-whites (17).

But many of these States contained (and still contain) a substantial African-American population, not to mention others who might have been classified as “non-whites.” This policy changed in the Linguistic Atlas of The Gulf States (LAGS), but only with regard to African-American respondents.

Studies of speakers of Spanish heritage were also excluded as “ethnic.” In a study of Texas vocabulary, here is what the author had to say concerning “border Spanish” influence on English and how he obtained the information:

My principal informants for Border Spanish were [a professor of folklore at the University of Texas, an assistant librarian at the University of Texas, a professor of anthropology at the University of Mexico City]. I have on some points questioned a number of graduate students from Central and South America, and the West Indies (Atwood, 1962: 124).

This is from a professor of linguistics and dialectology who insisted on the strictest attention to fieldwork but was apparently not interested in “ethnic” data nor how he obtained it. It is a mystery how the English of Texas, an area with a very large Hispanic population, can be studied without inclusion of those speakers.

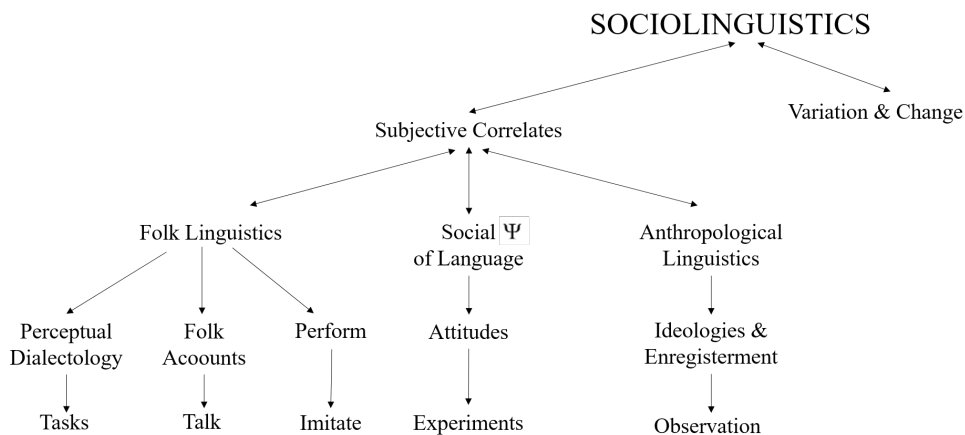
In an attempt to not only redress these wrongs but also configure a target population that takes into consideration the sorts of linguistic facts outlined above, the Linguistic Atlas Project (formerly the Linguistic Atlas of the United States and Canada) now takes sociolinguistics (and linguistics) seriously. A pilot project to identify and collect English data from heritage and bilingual speakers of Spanish in an area in North-Central Kentucky is underway. The aim is to establish informant selection protocols that include the wide range of English speakers in every area of the United States, to modify fieldwork training and materials appropriate to this new population, and to apply this new inclusiveness and methodology to all future projects of the LAP.

These last concerns reflect a move from the linguistic to the social in the considerations here, but it is one that uses linguistic evidence to argue for ethnic inclusiveness in dialectology and sociolinguistics. Spanish dialectology and sociolinguistics have generally done a much better job in this inclusiveness and will not be dwelled on here. This move to social concerns, however, allows a challenge to sociolinguistics in general and even the most sociolinguistically aware dialectology. Chambers and Trudgill’s warning may be paraphrased even again: “Sociolinguistics without social factors at its core is not living up to its name”.

Weinreich, Labov, and Herzog (1968) is the sacred book of US sociolinguistics and lists the principal “problems” of the field: Actuation, Transition, Constraints, Embedding, and Evaluation. Except for the last, these all seem to be primarily linguistic, but when the authors go on to describe each, the importance of social factors concern is everywhere. Such social facts always present an opportunity for the identification of what they call “subjective correlates”. Weinreich et al. state firmly that they cannot be derived from variable linguistic performance.

The theory of language change must establish empirically the subjective correlates of the several layers and variables in a heterogeneous structure. Such subjective correlates of evaluation cannot be deduced from the place of the variables within the linguistic structure (1968: 186).

FIGURE 16. THREE APPROACHES TO THE STUDY OF “SUBJECTIVE CORRELATES”



If these subjective correlates are so important, how can we find them? What means are available for determining the subjective correlates directly? Figure 16 shows the three principal approaches to methodologies for the independent determination of subjective correlates (or, perhaps more broadly, “language regard”, Preston, 2010; Evans *et al.*, 2018). Until recently the influence of the social psychology of language has predominated, and this domination by the study of language attitudes is related to the “observer’s paradox” (Labov, 1973), the fear that conscious awareness that language is being investigated will lead to inauthentic performance. Since that deals with language performance, however, we might ignore it; Weinreich *et al.* tell us that attitudes must be determined independently, not from language variation itself.

This fear is also related to the social psychological belief that direct questioning about language will lead to answers that do not reveal the authentic, nonconscious, automatic beliefs and attitudes of the respondent (Garrett, 2010), a fear that has led to the search for implicit responses in studies, from the “matched guise” technique of Lambert *et al.* (1960) to the Implicit Association Test (IAT, Campbell-Kibler, 2012) or even to ERP (event-related potential) neurological investigations (Loudermilk, 2015). These studies have provided important evidence of the subjective correlates, but they are often limited to specific topics that are already a part of the general sociolinguistic framework of investigation (e.g. the “-in” versus “-ing” variants of the English {-ING} morpheme).

More recently Eckert (2012) has appealed directly to the sociolinguistic content of variation and change by noting that more numerous uses of linguistic features used by certain social groups in topically and demographically specific interactions carry social meanings. This attempts to derive subjective correlates from performance, but it appeals to the results of social-psychological experiments (Campbell-

Kibler, 2007) or to the observational features of enregisterment located firmly in the anthropological tradition (Silverstein, 2003).

How has PRESEEA done in the investigation of these subjective correlates in the study of sociolinguistic variation in Spanish? It has a little sibling PRECAVES, which has as its stated aim the following:

[PRECAVES is] the project for the study of the varieties of Spanish in the 21st century ..., which aims to identify the beliefs and attitudes of different groups of individuals towards the normative varieties of Spanish (Cestero Mancera and Paredes Garcia, 2015: 3; translation mine).

This project characterizes its principal methodology as follows:

Linguistic beliefs and attitudes are undoubtedly a reflection of psychosocial perceptions and attitudes. Therefore, the first investigations on the subject were carried out within social psychology and its current theoretical framework (Cestero Mancera and Paredes Garcia, 2015: 1; translation mine).

This research effort has led to impressive results in various sites, and I do not want to criticize it. The founders of PRECAVES have already noted that the limitation of many of these studies to convenient university-student respondents and to the target of attitudes to standard language norms will be expanded on. More importantly, PRECAVES has also recognized that the methods suggested by social psychology are not the only ways to investigate subjective correlates, as I suggested earlier in figure 16 in the areas of Folk Linguistics and Anthropological Linguistics, the latter mentioned above in connection with Eckert's work on social meaning in what has come to be known as the "Third Wave" in sociolinguistic research (Eckert, 2012). I turn here, however, to the uncovering of subjective correlates in folk linguistics, an approach which has been underutilized in the field.

### 3. FOLK LINGUISTICS AND SUBJECTIVE CORRELATES

Folk linguistics has been encouraged by such scholars as Dell Hymes and has given rise to at least three approaches (figure 16). The first is perceptual dialectology, which, among other techniques, asks folk respondents to draw maps of their concepts of linguistic differences on a blank map of an area. Alfaraz and Mason (2019) is a good example of this work that takes the inclusion of Hispanic respondents seriously in the investigation of US English. They show that, although their respondents have many of the same perceptions as their European-American counterparts, those respondents do not focus on the same areas. In nearly all previous studies, upwards of 90% of the respondents identified an American South; in this study among Hispanics, the total fell to 59%, and the areas outlined were different as shown in the Hispanic respondent map in Figure 17 and the non-Hispanic map in Figure 18.

FIGURE 17. A HISPANIC HERITAGE SPEAKER'S HAND-DOWN MAP OF US DIALECTS

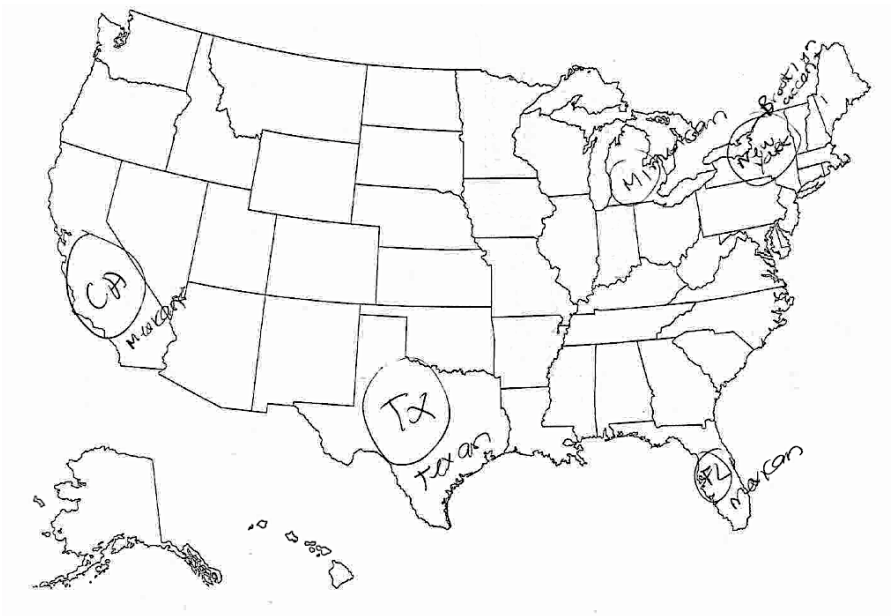


FIGURE 18. A NON-HISPANIC HERITAGE SPEAKER'S HAND-DOWN MAP OF US DIALECTS



It is also obvious that Hispanic-heritage respondents pay greater attention to those areas with large Spanish-heritage population.

Hispanic respondents	Non-Hispanic respondents
Texas 66% (1st)	29% (8th)
Florida 48% (4th)	34% (10th)
California 45% (5th)	23% (7th)

Such data as these, from English-dominant speakers of Hispanic heritage who constitute a significant population in Southeastern Michigan, cannot be excluded from a full study of the subjective correlates surrounding the speakers of English from this area.

The taxonomic chart of folk linguistics (figure 16) also identified the technique of eliciting variety imitations, a method that will serve us well if we want to know something about the salience of features in variety stereotyping and awareness, even if these imitations are inaccurate. Except for some by-chance occurrence of such imitations in folk linguistic interviews, research seeking variety imitations has rarely been done, but Evans (2002) shows that a nonprofessional imitator achieved low-level phonetic accuracy in the imitation of a prejudiced-against regional voice, one even identified as local by local respondents. Such results suggest that imitation can be a powerful tool in the identification and evaluation of subjective correlates, especially those that are linked to areas of linguistic performance that are beyond the specific control of the speaker (e.g. phonetic detail).

But prejudice against folk linguistics is apparent in this recent volume on perceptions of Spanish varieties:

[T]he gamut of different methodologies used in the studies in this volume... that seek to uncover the participants' linguistic perceptions, have moved away from "only" [emphasis in the original] relying on interpersonal interviews, where the participants could feel the need to offer the answer they believe to be politically correct (Ortiz-López and Suárez Büdenbender, 2022: 3).

This is not necessarily so. If people only asserted when they talked, there might be some reason to worry about their forthrightness. But all talk contains the very sorts of implicit materials that is so beloved by social psychologists. More than a little recent scholarship that suggests this is so. "As Silverstein (1979 and elsewhere) has suggested, the best place to look for language ideology may lie in the terms and presuppositions of metapragmatic discourse, not just in its assertions" (Irvine, 2001: 25)

What does all this mean? If you want to know, for example, how people do linguistic politeness, you should turn to the anthropological model of observation. If you want to see how people respond to cases of politeness and impoliteness, you should prepare a matched or implicit guise experiment. If you want to know what speakers think about linguistic politeness, why not just talk to them about it? Is this last a good idea? Why and why not?

First, although a researcher's observations in the culture may be long-standing and insightful, one may still miss the underlying beliefs, attitudes, and ideologies that talk about talk may reveal.

Second, in preparing an experiment, hasn't the researcher already concluded that this or that feature is salient? What if they miss those features that might more dramatically reveal the beliefs, attitudes, and ideologies of local respondents that talk about talk might provide?

Finally, if there is concern about political correctness, linguistic insecurity, or any other factor that may have caused a respondent to edit their responses, why not rely on the presupposed and implicit parts of talk which Irvine, citing Silverstein, called "the best place to find language ideology"?

How can we find such things in a principled (i.e. linguistically sophisticated) way? One way to do this is to apply the semantic and pragmatic tools that identify speaker and hearer intents and understandings of unasserted material. Such material is backgrounded in discourse and, as suggested above, a good parallel to the implicit sorts of judgments sought in many social psychologically-dominated attitude research models. This suggestion does not imply that there are not other linguistically sophisticated ways to interpret indirect meanings in talk about talk, especially perhaps in discourse analytic (Tannen *et al.*, 2015) and interactional (Gumperz, 1982) approaches. Here, however, the discussion is limited to semantic and pragmatic classifications of unasserted features as outlined in Potts (2015). His four subcategories are as follows:

1. Pragmatic presuppositions: Shared beliefs about the world; common ground. If one says "Iszol még kávét" (Do you have any coffee?), they pragmatically presuppose that the hearer understands Hungarian.
2. Pragmatic implicatures: The Gricean, Austinean, Searlean world of how conversational meanings, specifically how locutions (how utterances are linguistically structured) are understood as illocutions (what utterances accomplish). If one says "Do you have any coffee," although the locutionary structure is an interrogative, the pragmatic implication is that the illocutionary force of the utterance is a request.
3. Semantic presuppositions: Linguistic facts that trigger an additional, non-asserted meaning. If one asks, "Do you still drink coffee," *still* triggers a semantic presupposition that the speakers know that the hearer once drank coffee. At the syntactic level, the utterance "What Bill didn't lose was his billfold (in contrast to the simpler "Bill didn't lose his billfold") semantically presupposes that Bill lost something, and can even be triggered at the prosodic level by an utterance such as "Bill didn't lose his BILLfold" (with contrastive stress on *Bill-* and appropriate intonation).
4. Semantic implicature: Conventionalized linguistic cues that trigger additional, non-asserted meaning. The utterance "Bill drinks coffee but sleeps well" semantically implicates that there is an inconsistency between drinking coffee and sleeping well.

Using such tools, consider this edited conversation between professional airline pilots from the US State of Oklahoma. The group has just heard a seven-year old girl imitate “hick” speech, surely the source of S’s opening question. (“/” indicates short, nonterminal pauses; “=” indicates continued speech, and “[“ indicates the point at which overlapping speech begins.)

S: Is there a / is there a / an a / an a / an opinion or correlation or // about // intelligence / related to / how somebody speaks?

D: No / if [name] was who I fly with, / he sounds like the hickest of hicks / but that dude can do some crazy things with an airplane=

S: =But not-but not knowing that=

D: Yeah [

S: [ =If you were to meet somebody, / and they talk like // like the seven-year-old over here / talks like she’s from / the deep woo:ds./

D: You talking about perceptions,

S: Yeah / yeah.

D: Oh, perception you automatically assume they’re retarded

(Rodgers, 2017: 97-98)

When S asks D if he thinks language use correlates with intelligence, he pragmatically implicates that D can provide a yes or no answer.

D answers “no” and indicates that a man he flies with sounds like “the hickest of hicks.” This pragmatically presupposes that the person under discussion is unintelligent, a common ground association with “hick.” D adds quickly, however, “but that dude can do some crazy things with an airplane” in which *but* semantically implicates that that one who can “do crazy things with an airplane” is inconsistent with the identity of a “hick.” Finally, it is both conversationally implicated and pragmatically presupposed (at least among pilots) that one who has great flying skill cannot be unintelligent.

The rest of the conversation, which is not analyzed in such detail here, shows that D agrees that if one did not have personal knowledge about a speaker who sounds like they are from “the deep woods” (i.e., the sort of place where it is pragmatically presupposed that “hicks” reside), they would certainly think they were unintelligent, in fact “retarded”.

There are many things in even this short part of a conversation that are not asserted and are, therefore, more likely to be automatically delivered by the speaker and processed by the hearer, so long as they share subjective correlates. What are the more general linguistic matters involved?

- The relationship between language and intelligence may be questioned.
- Social and regional factors are candidates for consideration of that question.
- Rural and unsophisticated speakers use a variety of language that indicates a lack of intelligence.
- Flying airplanes well is an indication of intelligence– perhaps speech community restricted

- Speakers who exhibit intelligence in some skilled, professional way may overcome the stigma associated with their prejudiced against speech habits.
- Hearers who do not have contradictory personal information about a speaker will assign stereotypically presupposed interpretations (“perceptions”) to their language variety.

Is all this semanto-pragmatic detail necessary?

1. Sociolinguists already knew all this.

Maybe, but even if they did, the above details come from nonlinguists. This is surely confirming (and comforting) evidence for professionals.

2. Scholars don’t need to do such semanto-pragmatic work; if they had read this conversation carefully, they would have figured it out.

Maybe, but wouldn’t linguistic scholars rather have a formal, linguistic characterization of what they have understood, especially if they are a member of the same speech community as their respondents and share many of the same folk beliefs? They might make many of the points raised above easy to overlook.

3. Isn’t there a concern about the respondents’ concealing their actual beliefs and attitudes, especially when the interview involves talk about marginalized or stigmatized groups?

No. In folk linguistic interviews in general or in discussions of such folk linguistic tasks such as map-drawing and imitation, respondent enthusiasm for the topics of language in general and language variety in particular suggests that these are very successful methods for engaging respondents in natural talk. In the conversations studied in Folk Linguistics (Niedzielski y Preston, 2000), we observed excited, natural interactions on topics of regional and social varieties, language and gender, spelling, the uses of the passive construction, slang, etymologies, politeness, language acquisition and development, language education, language disorders, to name only a few. In many cases when more than one respondent was involved there were even open, natural disagreements. More importantly, even if respondents did disguise or avoid frankness in their responses, the fact that presupposed or implicated material is backgrounded in performance and not under the complete conscious control of the respondents does a great deal to mitigate any uneasiness about the authenticity and value of effective interview talk about talk in the sociolinguistic realm.

#### 4. CONCLUSIONS

Although linguistic matters are stronger than social and stylistic ones (remember Dominican [s] lenition), the deeper consideration of linguistic factors is inextricably bound

to a consideration of social topics necessary to the study of language variation and change (remember the NCS-influenced vowel systems). Linguistic factors are important in our reconsideration of who counts as a so-called native speaker in areal variationist work.

Linguistic facts need the explanatory value of subjective correlates. In the continuing work on worldwide Spanish sociolinguistics, it is important that PRECAVES is not taken as the little brother or sister of PRESEEA. Every one of the problems set by Weinreich et al. more than fifty years ago requires consideration of the subjective correlates necessary to the study of language variation and change.

Don't overlook the subjective correlates that can be derived from talk with respondents. They will tell you a lot, and a lot of what they say they presuppose and implicate, factors not really a part of their conscious plan. In fact, talk about talk is even one proven way to defeat the monster of the observer's paradox. Perhaps one might think about adding to the PRESEEA a list of suggested topics for the sociolinguistic interview, the topic of language itself. This would be an economical marriage of PRECAVES and PRESEEA.

Finally, the comprehensive study of international Spanish sociolinguistics will undoubtedly move forward for another twenty-five years, and, just to reassure younger scholars, another productive twenty-five will follow. The nice thing about the study of variation and change is that it never stops. We will always have a job to do and the linguistically sophisticated expansion of methods, tools, and approaches to the collection and analysis of data will make this an even more certain thing.

## REFERENCES

- Alafaraz, Gabriela G. y Alexander Mason (2019). "Ethnicity and Perceptual Dialectology: Latino Awareness of U.S. Regional Dialects". *American Speech*, 94(3): 352-379.
- Atwood, E. Bagby (1962). *The regional vocabulary of Texas*. Austin: The University of Texas Press.
- Babel, Anna M. (2009). "Dizque, evidentiality, and stance in Valley Spanish". *Language in Society*, 38(4): 487-511.
- Bakos, Jon (2008). *An examination of the adaptation of the Northern Cities Chain Shift by Lebanese immigrants in Dearborn, Michigan*. [MA Thesis]. East Lansing: Michigan State University.
- Bell, Alan (1984). "Language style as audience design". *Language in Society*, 13(2): 145-204.
- Bullock, Barbara E., Almeida Jacqueline Toribio, y Mark Amengual (2014). "The status of s in Dominican Spanish". *Lingua*, 143:20-35.
- Campbell-Kibler, Kathryn (2007). "Accent, (ING) and the social logic of listener perceptions". *American Speech*, 82: 32-64.
- Campbell-Kibler, Kathryn (2012). "The Implicit Association Test and sociolinguistic meaning". *Lingua*, 122: 753-763.

- Cestero Mancera, Ana M. y Florentino Paredes García (2015). “Metodología para el proyecto PRECAVES XXI”. Disponible en: <http://www.variedadesdespanol.es/metodologia.html>
- Chambers, Jack K. y Peter Trudgill (1980). *Dialectology*. Cambridge: Cambridge University Press.
- Eckert, Penelope (2012). “Three waves of variation study: The emergence of meaning in the study of variation”. *Annual Review of Anthropology*, 41: 87-100.
- Evans, Betsy E (2002). “An acoustic and perceptual analysis of imitation”. En D. Long and D. R. Preston (eds.), *Handbook of perceptual dialectology*, vol. 2. Amsterdam: Benjamins, pp. 95-112.
- Evans, Betsy, Erica Benson y James Stanford (eds.) (2018). *Language regard: Methods, variation, and change*. Cambridge: Cambridge University Press.
- Firestone, Megan, y Emma Giese (2008). “Jak jest Twoj angielski? The pronunciation of English in Hamtramck, Michigan”. *American Dialect Society meeting*, Chicago.
- Garrett, Peter (2010). *Attitudes to language*. Cambridge: Cambridge University Press.
- Gumperz, John J. (1982). *Discourse Strategies*. Cambridge: Cambridge University Press.
- Honeybone, Patrick (2010). “How Symmetrical are English Vowels?”. *Journal of Language and Language Behaviour*, 9: 33-63.
- Irvine, Judith (2001). “‘Style’ as distinctiveness: the culture and ideology of linguistic Differentiation”. En P. Eckert y J. R. Rickford (eds.), *Style and Sociolinguistic Variation*. Cambridge: Cambridge University Press, pp. 21-43.
- Jones, Jamila (2003). *African Americans in Lansing and the Northern Cities Vowel Shift*. [Tesis doctoral]. East Lansing: Michigan State University.
- Kretzschmar, William A, Jr., Virginia McDavid, Theodore Lerud, y Ellen Johnson (eds.) (1993). *Handbook of the Linguistic Atlas of the Middle and South Atlantic States*. Chicago: University of Chicago Press.
- Lambert, Wallace, R. C. Hodgson, Robert C. Gardner, y Samuel Fillenbaum (1960). “Evaluational reactions to spoken languages”. *Journal of Abnormal and Social Psychology*, 20(1): 44-51.
- Labov, William (1973). *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.
- Labov, William (1996). “The organization of dialect diversity in North America”. *Telsur Project*. Linguistics Laboratory: University of Pennsylvania.
- Loudermilk, Brandon C. (2015). “Implicit attitudes and the perception of sociolinguistic variation”. En A. Prikhodkine y D. R. Preston (eds.), *Responses to language varieties: Variability, processes and outcomes*. Amsterdam: Benjamins, pp. 137-156.
- Maddieson, Ian (1984). *Patterns of sounds*. Cambridge: Cambridge University Press.
- Niedzielski, Nancy A. y Dennis R. Preston (2000). *Folk Linguistics*. Berlín, Nueva York: De Gruyter Mouton. DOI: <https://doi.org/10.1515/9783110803389>

- Obrecht, Dean H. (1968). *Effects of the second formant on the perception of velarization consonants in Arabic*. The Hague: Mouton.
- Ocuppaugh, Jaclyn (2010). *Regional variation in Chicano English: Incipient dialect formation among L1 & L2 speakers in Benton Harbor, Michigan*. [Tesis doctoral]. East Lansing: Michigan State University.
- Ortiz-López, Luis Alfredo y Eva-Maria Suárez Büdenbender (2022). "Introduction: Linguistic perceptions among Spanish-speakers: on the pathway to a deeper understanding". En L. A. Ortiz-López y E. M. Suárez Büdenbender (eds.), *Topics in Spanish Linguistic Perceptions*. Londres: Routledge, pp. 1-10.
- Owsianny, Mariusz (2001). "Interaction between vocalic quality and fundamental frequency in the perception of Polish vowels". En S. Puppel y G. Demenko (eds.), *Prosody 2000*. Poznań: Faculty of modern Languages and Literatures, Adam Mickiewicz University, pp. 197-204.
- Pasquale, Michael (2001). *Patterns in Quechua-Spanish vowel alternation*. [Tesis doctoral]. East Lansing: Michigan State University.
- Peterson, Gordon E. and Harold L. Barney (1950). "Control methods used in a study of the vowels". *The Journal of the Acoustical Society of America*, 24(2): 175-84.
- Potts, Christopher (2015). "Presupposition and Implicature". En Shalom Lappin y Chris Fox (eds.), *The Handbook of Contemporary Semantic Theory*. Chichester, UK: John Wiley & Sons, Ltd., pp. 168-202. DOI: <https://doi.org/10.1002/9781118882139.ch6>.
- Preston, Dennis R. (1991). "Sorting out the variables in sociolinguistic theory". *American Speech*, 66(1): 3-26.
- Preston, Dennis R. (2010). "Variation in language regard". En P. Gilles, J. Scharloth y E. Zeigler (eds.), *Variatio delectat: Empirische Evidenzen und theoretische Passungen sprachlicher Variation* (für Klaus J. Mattheier zum 65. Geburtstag). Frankfurt am Main: Peter Lang, pp. 7-27.
- Rodgers, Elena A. (2017). *Discursive Constructions of Sociolinguistic Differences: A Rhetorical Analysis*. [Tesis doctoral]. Oklahoma State University.
- Roeder, Rebecca (2006). *Accommodation to the Northern Cities Shift among Mexican Americans in Lansing, Michigan*. [Tesis doctoral]. East Lansing: Michigan State University.
- Silva-Corvalán, Carmen (1994). *Language Contact and Change: Spanish in Los Angeles*. *Oxford Studies in Language Contact*. Oxford: Clarendon Press.
- Silverstein, Michael (2003). "Indexical order and the dialectics of sociolinguistic life". *Language & Communication*, 23: 193-229.
- Stevens, Kenneth N. (1998). *Acoustic phonetics*. Cambridge MA, USA/London, England: The MIT Press.
- Tannen, Deborah, Heidi E. Hamilton, y Deborah Schiffrin (eds.) (2015). *The handbook of discourse analysis* (2nd ed.). Hoboken NJ: John Wiley & Sons.



VARIACIÓN E INVARIACIÓN MORFOSINTÁCTICAS DEL ESPAÑOL.  
APUNTES DESDE PRESEEA

Carmen Silva Corvalán  
*Universidad del Sur de California*

## 1. INTRODUCCIÓN

En el *Congreso de La Lengua Española*, celebrado en Sevilla en 1992, apuntaba yo que los estudios de aspectos sociolingüísticos del español habían sido hasta entonces dispersos y habían empleado técnicas diversas de recolección y codificación de los datos. Todo esto hacía difícil la comparación de los resultados para establecer un cuadro cohesivo del valor sociolingüístico de las variables examinadas en diferentes comunidades hispanohablantes. PRESEEA ha respondido con el desarrollo de estudios sociolingüísticos coordinados del español gracias al trabajo de equipos de investigación en España y América. Investigadores de 16 países han realizado estudios de gran interés y han llevado a PRESEEA a convertirse en una organización lingüística vital que desde su inicio ha dado impulso a los más importantes estudios dialectales y sociolingüísticos del español contemporáneo.

Los estudios de variación sintáctica en el marco de PRESEEA han seguido la *teoría variacionista*, que observa que los sistemas lingüísticos no son homogéneos, sino que varían según *factores internos* al sistema de una lengua y según factores sociales y geográficos *externos* a ella. Esta teoría considera esencial la integración de aspectos lingüísticos y sociales en el estudio de las lenguas e incorpora un *componente cuantitativo* clave en el análisis de los factores que condicionan la variación.

Mucho se ha avanzado en variación fónica y en aspectos discursivos, pero en esta ocasión quiero enfocarme en dos variables morfosintácticas, la expresión del sujeto pronominal y *haber* impersonal, y una variable que considero más bien léxico-semántica, la variable *deber – deber de*.

Estos estudios justifican en parte el título de mi contribución: Variación e invariación morfosintácticas del español, ya que muestran diferenciación o variación, pero también unidad o invariación (morfo)sintáctica a través de las comunidades estudiadas.

Mi objetivo es explorar posibles áreas de convergencia y divergencia interdialectal. Para esto es necesario encontrar resultados comparables a través de los estudios realizados con datos de PRESEEA u otros orales, lo que me ha resultado algo difícil. Pero bien, aquí va lo que más o menos puedo contarles: empiezo con sujeto, luego *deber (de)* y finalmente *haber*.

Pero quisiera repasar algunas cuestiones teóricas antes de discutir estas variables.

Como es bien sabido, la *variable lingüística* se ha conceptualizado como dos o más realizaciones de una variable en el habla, las llamadas *variantes de una variable*. Por definición, las variantes son idénticas en cuanto a su valor referencial o de verdad, pero conllevan diferentes significados sociales, pragmáticos y/o estilísticos.

En el estudio de la variación (morfo)sintáctica se han tomado como punto de partida variantes cuya sinonimia lógica no es cuestionable y cuya estructura sintáctica varía solamente con respecto al fenómeno considerado variable. Fuera de todo contexto discursivo, esta variable (morfo)sintáctica no altera la equivalencia referencial. El análisis cualitativo de los contextos de uso de las variantes identificadas, sin embargo, tomando en cuenta su distribución en el discurso, ha permitido en algunos casos establecer posibles diferencias de significado semántico o pragmático. Así pues, los estudios variacionistas han ido más allá de lo distintivo en el nivel lingüístico estructural para poder identificar lo distintivo en los niveles semántico-pragmático, social y estilístico.

Gran parte de la variación morfosintáctica estudiada parece estar condicionada por factores internos a la lengua, incluso en algunos casos por factores fonológicos (como posiblemente el orden de los constituyentes oracionales). La sociolingüística integra, por tanto, los diferentes niveles analíticos que se han identificado tradicionalmente: la fonología, la morfología, la sintaxis y la semántica, agregando uno más: la pragmática (Silva-Corvalán y Enrique-Arias, 2017).

Por ejemplo, el uso no estándar de *de* antepuesto a *que* en complementos oracionales, fenómeno variable conocido como *dequeísmo*, parece responder a factores internos a la lengua relacionados con aspectos semánticos y pragmáticos. Así pues, se propone que el *dequeísmo* es favorecido por un sujeto de tercera persona y desfavorecido cuando el sujeto es primera persona. El tiempo pasado, en oposición a presente y futuro, también favorece el uso de *dequeísmo*. El ejemplo (1) ilustra la correlación entre *dequeísmo*, tiempo pasado y sujeto de tercera persona.

(1) Fue entonces que me dijo *de que* ya no le quedaba ni un peso.

La distribución del *dequeísmo* ha llevado a diferentes autores a proponer que el hablante *dequeísta* emplea este fenómeno con el propósito pragmático de comunicar

distanciamiento del contenido del enunciado y además, con el propósito pragmático-discursivo de marcar el complemento como información nueva o focal. La variable *dequeísmo*, entonces, conllevaría significados pragmáticos y no correspondería a «dos formas de decir la misma cosa», pero sí tendría el mismo significado referencial.

Hay variables sintácticas y morfológicas que no parecen conllevar diferencias de significado en ningún nivel de análisis lingüístico o *pragmalingüístico*. Por ejemplo, el doblado de clíticos en perífrasis verbales (ejemplo 2), posiblemente también algunos fenómenos de concordancia de número (ejemplo 3) e incluso de pérdida de marcación de caso en cláusulas relativas (ejemplo 4).

- (2) Yo (*la<sub>i</sub>*) quería ver*la<sub>i</sub>* pero llegué cuando se había marchado.
- (3) Un grupo de niños jugaban frente a su ventana
- (4) El hombre *que/al que* le entregué el mensaje ya se fue.

Estas son variables que se comportan, podríamos decir, como variables fonológicas. Su mayor o menor frecuencia de ocurrencia puede interpretarse como indicadora de significados sociales y/o estilísticos. Pero debo decir esto con una cierta cautela, sin embargo, ya que nuevas investigaciones podrían revelar significados semántico-pragmáticos. En verdad, la presencia variable de una copia pronominal en Cláusulas Relativas (CRs), como en el ejemplo (5), que había sido estudiada como un caso de variación solo estructural, ha resultado tener una motivación pragmática cuando la función sujeto se copia en una cláusula relativa:

- (5) Se puso un abrigo<sub>i</sub> que (*lo<sub>i</sub>*) había comprado el día antes.

Cerrón-Palomino (2015) examina esta variable en materiales de Lima (Perú) y argumenta de manera convincente que el sujeto reasuntivo tiene una función contrastiva: la propiedad introducida por la CR pertenece exclusivamente a su antecedente y descarta otros posibles candidatos. La función contrastiva del uso de estos pronombres se refuerza a veces con el adverbio *sí* o con acento prosódico en el pronombre (vean el ejemplo 6 a, de Cerrón Palomino, 2015- ej. 29).

- (6) a. Por eso se lo pregunto *a mi amigo*, *que él, sí* sabe de esto. Yo soy un ignorante.  
b. Por eso se lo pregunto *a mi amigo*, *que* sabe de esto.

Los factores lingüísticos internos que condicionan la variación han resultado ser de gran interés y han llevado a asignar un papel central a los aspectos semánticos, pragmáticos y cognitivos de las unidades lingüísticas como variables explicativas.

Por otra parte, no ha sido fácil asignar valor social o estilístico a todos los casos de variación morfosintáctica estudiados. Por ejemplo, algunos estudios de variedades del español en España concluyen que las variables sociales inciden en general poco (o nada) en la expresión variable de un sujeto pronominal. La escasa influencia de los factores sociales no es sorprendente si la expresión o no del sujeto conlleva significados diferentes en el contexto discursivo.

Veamos entonces lo que se ha demostrado en algunos enfoques relativamente recientes de variación morfosintáctica. Los estudios que he seleccionado presentan un panorama actual de los temas a los que se ha prestado mayor atención en sociosintaxis variacionista, varios de ellos incorporando en sus datos de análisis materiales de PRESEEA.

## 2. LA VARIABLE «EXPRESIÓN DE SUJETO»

La expresión *variable del sujeto gramatical* en español ha sido desde hace años motivo de gran interés. Las investigaciones han mostrado que la variación es controlada por múltiples factores, lo que hace su estudio enormemente complejo.

Los ejemplos (7) a (10) ilustran la variación según algunas de las variables explicativas. El sujeto no se requiere en (7); pero es en cambio obligatorio en (8) a (10).

(7) (a) Ø/yo toco bien el piano, pero la guitarra (b) Ø/yo la toco muy mal.

(8) *Contraste*

Mi hermana habla bien inglés pero yo \*(Ø) lo hablo muy quebrado.

(9) *Información focal*

A: ¿Quién trajo este diario?

B: Yo lo traje. / \*Ø lo traje.

(10) *Clarificación del referente del sujeto*

S: ... yo despertaba gritando y mi hermana tenía que levantarse a verme. Y (a) ella iba a mi lado (b) y yo \*(Ø) estaba temblando, que hasta los dientes se oían que pegaban.

Obviamente, si la información transmitida por el sujeto es nueva o focal (contrastiva), este debe expresarse. Es posible que entren a desempeñar un papel factores pragmáticos y semánticos: en el ejemplo (10), la expresión o no de «yo» en (10b) tiene consecuencias semánticas claras, ya que un sujeto cero implica que es «ella» la que «estaba temblando» y no «yo», que es lo que el hablante especifica al expresar «yo».

La correferencialidad ha resultado ser un factor estadísticamente significativo en casi todos los estudios. El ejemplo (7) la ilustra: dos formas verbales finitas en cláusulas contiguas con el mismo sujeto. En diversos dialectos del español, los sujetos correferenciales, como en (7b), se expresan en más o menos el 25 % de los casos y los sujetos que implican cambio de referente (8 y 10a-10b) se expresan en más del 40 % de los casos.

### *Sujetos*

*Correferenciales*

*Con cambio de referente*

### *Expresos*

más o menos el 25 %

más del 40 %

(Shin y Smith Cairns, 2012).

Es notable que las probabilidades asignadas al factor correferencialidad del referente del sujeto en estudios de diferentes dialectos sean exactamente iguales: Los Ángeles y Caracas: 0.34 con sujeto correferencial y 0.66 cuando no lo es. Valores

muy similares se obtienen en Barranquilla (0.35 y 0.67), en Castañer (0.34 y 0.70), en Lima (0.36 y 0.64) y en cuatro ciudades españolas (0.35 y 0.62 Norma Culta).

**TABLA 1. SUJETOS CORREFERENCIALES Y NO CORREFERENCIALES EN ALGUNAS CIUDADES HISPANOABLANTES**

Sujeto	Correferencial	NO correferencial	
Los Ángeles y Caracas	0.34	0.66	(Bentivoglio, 1987)
Barranquilla	0.35	0.67	(Orozco, 2015)
Castañer	0.34	0.70	(Holmquist, 2012)
Lima	0.36	0.64	(Cerrón-Palomino, 2014)
España*	0.35	0.62	(Samper Padilla <i>et al.</i> , 2006)

\* Norma Culta – Cuatro ciudades: Madrid, Sevilla, Salamanca, Las Palmas de Gran Canaria  
Fuente: Bentivoglio, 1987; Orozco, 2015; Holmquist, 2012; Cerrón-Palomino, 2014; Samper Padilla *et al.*, 2006.

El tipo de verbo ha resultado también significativo. Esta variable ha sido incluida, con algunas diferencias en cuanto a la clasificación de los verbos, en varios estudios que han refrendado la relación entre tipo de verbo y expresión del sujeto pronominal. El ranking de estos tipos de verbo no es el mismo, pero en todos los estudios citados los verbos de percepción (*oler*, *ver*) y los cognitivos (*pensar*, *saber*, *creer*) (unidos a veces en el grupo «verbos mentales») son los que más favorecen la expresión de un sujeto.

Por ejemplo, factores determinantes en la expresión del sujeto pronominal en el corpus PRESEEA de Granada han resultado ser el tipo de verbo y la especificidad del sujeto.

Como muestra la Tabla 2, tal como en otras comunidades, los verbos que significan procesos mentales del tipo *creer* (69 %), *saber* (27 %) o *ver* (26 %), presentan una alta tendencia a expresarse con el sujeto pronominal. En datos PRESEEA de Sevilla, *creer* y *recordar* aparecen con el porcentaje más alto de sujetos pronominales.

**TABLA 2. VERBOS MENTALES CON SUJETO EXPRESO**

	<i>creer</i>	<i>saber</i>	<i>recordar</i>	
Granada	69%	27%	543	(Manjón-Cabeza Cruz, <i>et al.</i> , 2016)
Sevilla	60%		62%	(Repede, 2019)

Fuente: Manjón-Cabeza Cruz, *et al.*, 2016; Repede, 2019.

La persona gramatical es un factor determinante en la expresión de un sujeto. Además del pronombre *uno* indefinido, que presenta generalmente la probabilidad más alta de empleo, es común que los sujetos singulares se expresen con mucha mayor frecuencia que los plurales, como se observa en la Tabla 3, en los pesos probabilísticos y porcenta-

jes obtenidos en el análisis del corpus de Granada, columnas «A» y en los porcentajes obtenidos en 24 entrevistas del sociolecto alto en el corpus de Sevilla, columnas «B».

**TABLA 3. EXPRESIÓN DEL SUJETO SEGÚN PERSONA GRAMATICAL**

Factor persona	Peso probabilístico	GRANADA		SEVILLA	
		%	N/total	%	N expresos
	A	A	A	B	B
1s	0.646	25	424/1715	30	170
2s	0.496	15	68/452	24	22
3s	0.516	16	64/397	29	15
1p	0.254	6	25/432	6	8
3p	0.200	4	18/419	11	3

Fuente: Los resultados en las columnas «A» son tomados de Manjón-Cabeza Cruz et al., 2016 y en «B» de Repede, 2019.

En general los factores sociales no tienen gran influencia en la expresión variable del sujeto. Algunos estudios en datos de PRESEEA, sin embargo, han indicado que el sexo y la edad parecen incidir en la variación (e.g., Manjón-Cabeza Cruz *et al.*, 2016; Repede, 2019; también en Orozco, 2015 en Barranquilla). Las mujeres son las que más uso hacen del sujeto pronominal. En los datos de Sevilla, por ejemplo, las mujeres casi duplican el empleo de *yo* frente a los hombres: 61 % versus 39 %.

Uso de *yo* – Sevilla (Repede, 2019)

Mujeres 61 %

Hombres 39 %

Por otra parte, hay contradicción con respecto a la edad: en Sevilla, es la tercera generación (> 55) del sociolecto alto la que menos uso hace del sujeto pronominal (18 %), mientras que en Granada son los jóvenes (< 35) de los niveles educativos altos los que menos uso hacen del sujeto. Lastra y Martín-Butragueño (2015) muestran la misma tendencia en Ciudad de México (en datos orales, pero no de PRESEEA); como en Granada, los jóvenes (de 20 a 34 años) expresan el sujeto con mucha menor frecuencia que los mayores de 55 años. Se preguntan si se tratará de un cambio en marcha y ofrecen una forma de respuesta: «se necesita más investigación, sin embargo, para descubrir la causa de este efecto etario» (pp. 53-54). No podría yo estar más de acuerdo con el llamado a más investigación.

En resumen, es evidente que se ha demostrado que hay más semejanzas que diferencias interdialectales en cuanto a la expresión del sujeto. Aunque los porcentajes de frecuencia de uso de pronombres sujeto no son iguales en los dialectos estudiados, estos porcentajes y los análisis multivariantes muestran probabilidades similares que validan la consideración de una gramática con variables explicativas y tendencias comunes con respecto a la correferencia, los verbos mentales y la persona gramatical. Lo mismo se observa en los datos de la norma culta estudiados por Samper *et al.* (2006)

en cuatro comunidades urbanas (Sevilla, Madrid, Las Palmas de Gran Canaria, Salamanca). Estos investigadores confirman la significatividad de las variables lingüísticas y encuentran que la diferenciación diatópica es «la que ofrece una significación menos importante»; en cuanto a la expresión del sujeto, predominan los rasgos coincidentes. Existe variación e invariación en este dominio de la lengua.

### 3. *DEBER Y DEBER DE*

Ahora quisiera añadir algunas palabras sobre la variación entre *deber* y *deber de*. De acuerdo con las gramáticas prescriptivas, los usos de estos verbos modales corresponden a diferentes significados. *Deber* se usa para indicar obligación o necesidad, modalidad deóntica, (por ejemplo, «Juan debe presentarse al trabajo mañana»), mientras que *deber de* indica suposición o probabilidad, modalidad epistémica (por ejemplo, «Juan debe de tener unos 30 años»).

Obligación:	<i>Juan debe presentarse al trabajo mañana</i>	(Modalidad deóntica)
Probabilidad:	<i>Juan debe de tener unos 30 años</i>	(Modalidad epistémica)

Sin embargo, los trabajos sobre esta variable concuerdan en notar que los hablantes de español a menudo usan *deber de* y *deber* para expresar tanto obligación como posibilidad (Aracil, 2001; Cuervo, 1994). La «confusión entre las dos formas» parece ser típica en todo el mundo hispanohablante y no una característica de una variedad particular. Se ha documentado en Houston (Texas), Madrid, Valencia y otras ciudades de España, en varias capitales americanas: México D.F., San José, Bogotá, Lima, Santiago, Buenos Aires, La Paz, Caracas, San Juan.

La «confusión» documentada en:

- EEUU: Houston y Texas (Fairclough, 2000).
- En varias capitales americanas: México D.F., San José, Bogotá, Lima, Santiago, Buenos Aires, La Paz, Caracas y San Juan.
- En ciudades de España (Samper *et al.*, 1998): Valencia (Gómez Molina, 2013a) y Madrid (Ruiz Gurillo, 2002).

Los ejemplos (1) a (4) ilustran el uso de *deber de* y *deber* para expresar tanto obligación como posibilidad:

#### *Obligación*

- (1) Todos *debemos* obedecer las leyes del tránsito.
- (2) *Debes de* ir a la clínica si tienes fiebre.

#### *Probabilidad*

- (3) Ese hombre alto *debe* ser el atleta.
- (4) *Debe de* estar enfermo, yo creo.

En un estudio que hice con un colega de la universidad Brigham Young (en Utah) (Eddington y Silva-Corvalán, 2011) examinamos *deber* y *deber de* en el corpus Arthus (Archivo de textos hispánicos de la Universidad de Santiago de Compostela) y los sometimos a análisis de regresión logística con el fin de explorar una serie de factores que juegan un papel en la elección del verbo modal con o sin *de*. Arthus incluye materiales orales y escritos pertenecientes a diferentes etapas de la historia del español. Nosotros estudiamos los textos contemporáneos incluidos en el corpus hasta 1995. Los datos no son comparables con los obtenidos en PRESEEA, pero dan una idea de la extensión de la variación *deber* - *deber de* en principalmente materiales escritos, que constituyen el 85% del total de los datos que estudiamos.

Identificamos 568 casos de *deber* y 106 casos de *deber de* en el corpus (para un total de 674 casos). De los 106 casos de *deber de*, solo 12 ocurren en datos de Hispanoamérica y 94 en España.

674 total casos de *deber* (*de*) - corpus Arthus (Archivo de textos hispánicos de la Universidad de Santiago de Compostela)

*deber* - 568 casos: 115 Hispanoamérica [25%]  
453 España.

*deber de* - 106 casos: 12 Hispanoamérica [13%]  
94 España

En cuanto a *deber*, la Tabla 4 muestra similitud entre Hispanoamérica y España: 99% y 94%, respectivamente, de casos que implican una expresión de obligación y se expresan con *deber*.

**TABLA 4. ELECCIÓN DE *DEBER* (VERSUS *DEBER DE*) EN HISPANOAMÉRICA Y EN ESPAÑA (85% MATERIALES ESCRITOS)**

Tipo de modalidad	Probabilidad estadística	%	N modal/ situaciones
<i>Obligación</i>			
Hispanoamérica*	.637	99	167/169
España	.389	94	320/340
<i>Probabilidad</i>			
Hispanoamérica*	16	<b>81</b>	42/52
España	6	<b>35</b>	39/113
Total			568/674

Fuente: Datos de Argentina (46%), Colombia (11%) y México (43%), principalmente escritos.

*Deber de* es muy poco frecuente en el español hispanoamericano; sin embargo, 10 de los 12 ejemplos que encontramos siguen el uso prescrito y comunican el valor contextual de probabilidad. En los datos españoles, en los que *deber de* es más numeroso: 74 casos de *deber de*, el 65% comunica probabilidad (74 de 113).

La diferencia entre Hispanoamérica y España se refleja en el uso extendido de *deber* a contextos de probabilidad. En estos contextos encontramos *deber* con una alta frecuencia de 81%: en los datos de Argentina (46%), Colombia (11%) y México (43%), mientras que en España encontramos solo el 35% de *deber* en estos contextos. El uso de *deber* para obligación y *deber de* para probabilidad es respetado con más frecuencia en el español peninsular, donde en todo caso también existe variación.

El aspecto verbal incide en la interpretación del verbo modal. Verbos dinámicos y expresiones de obligación son más comunes con *deber*. Verbos estativos y expresiones de posibilidad ocurren con mayor frecuencia con *deber de*.

Sin embargo, en Hispanoamérica la clase de verbo principal no es estadísticamente significativa en la elección de *deber* o *deber de*, pero la interpretación de un enunciado es diferente según el aspecto del verbo modalizado.

Nótense las diferentes interpretaciones de los ejemplos (5) y (6) con el modal *deber*. Sin más contexto, con un verbo estativo, (5) se interpreta como probabilidad, y (6) con un verbo dinámico se interpreta como obligación. Estas interpretaciones no son sorprendentes ya que el tipo de evento que las personas se sienten obligadas a realizar, o intentan obligar a otros a realizar, es normalmente dinámico, como en (6). Por otro lado, las suposiciones acerca de un posible estado de cosas tienden a involucrar situaciones estáticas más que dinámicas, como en (5).

(5) Pepe *debe* tener 30 años. - probabilidad

(6) Pepe *debe* visitar a su madre. - obligación

La semántica del verbo auxiliado es también examinada por Gómez Molina en Valencia. Sus resultados para verbos dinámicos y de estado coinciden con los nuestros en cuanto a que en Valencia los dinámicos aparecen con *deber* en el 67% de los casos y los estativos con menor frecuencia: 53%. De manera opuesta, los verbos estativos favorecen el uso de *deber de* y los dinámicos desfavorecen *deber de*, 47% versus 33%, respectivamente.

	<i>Deber</i>		<i>Deber de</i>	
Verbo modalizado				
dinámico	49/73	67%	24/73	33%
estativo	148/279	53%	131/279	47%

Fuente: Gómez-Molina, 2013a, Tabla 4.

Queda claro que los dos modales se usan para expresar tanto modalidad deóntica como epistémica, al menos en España, como indica la Tabla 5, que recoge resultados en corpus orales de España y Caracas. Nuevamente aquí hay que considerar los porcentajes con cautela pues la cantidad de datos de Granada es pequeña; Caracas, Castellón y Madrid examinan datos orales *no* de PRESEEA; y Valencia incluye también datos de Valesco, aunque estos *sí* son orales y componen menos del 40% del total de los datos estudiados por Gómez Molina.

TABLA 5. VALORES MODALES EN DIFERENTES COMUNIDADES DE HABLA

Valor	<i>deber</i> + infinitivo					<i>Deber de</i> + infinitivo				
	Deóntico		Epistémico		Total	Deóntico		Epistémico		Total
	N	%	N	%		N	%	N	%	
Caracas	55	<b>51</b>	52	<b>49</b>	107	12	<b>92</b>	1	8	13
Granada	12	<b>60</b>	8	<b>40</b>	20	31	<b>97</b>	1	3	32
Valencia	261	<b>74</b>	92	<b>26</b>	353	129	<b>60</b>	86	40	215
Castellón	195	<b>61</b>	125	<b>39</b>	320	50	<b>54</b>	42	46	92
Madrid	54	<b>55</b>	45	<b>45</b>	99	25	<b>53</b>	22	46	47

Fuente: Gómez Molina, 2013a, Tabla 2 (Madrid, Caracas, Castellón, Valencia). Manjón-Cabeza Cruz, 2017, Granada. Nota: He redondeado los porcentajes que aparecen en los trabajos originales.

Los usos de *deber* con valor deóntico son altos en estas cinco ciudades, pero solo en Valencia hay una diferencia notable con su uso en contextos epistémicos, 74% versus 26%. Quizás exceptuando Valencia, creo que podemos considerar el uso de *deber* un ejemplo de invariación interdialectal.

*Deber de*, por otro lado, mantiene una cierta robustez en contextos epistémicos en Valencia, Castellón y Madrid, pero incluso en estas tres comunidades el uso de *deber de* con sentido deóntico, el más criticado por la llamada norma culta, sobrepasa el 50% de los casos de este modal, es decir, es más frecuente que su uso epistémico.

Consideremos ahora la distribución de estos modales en la Tabla 6, según sociolecto y nivel educativo, que me imagino son más o menos comparables.

TABLA 6. DISTRIBUCIÓN DE *DEBER* Y *DEBER DE* EN VALENCIA Y GRANADA

VALENCIA										
Valor	<i>deber</i>					<i>deber de</i>				
	Deóntico		Epistémico		Total	Deóntico		Epistémico		Total
	N	%	N	%		N	%	N	%	
Soc. alto*	123	<b>70</b>	52	<b>30</b>	175	45	<b>50</b>	45	<b>50</b>	90
Soc. medio	95	<b>77</b>	29	<b>23</b>	124	48	<b>66</b>	23	<b>34</b>	73
Soc. bajo	43	<b>80</b>	11	<b>20</b>	54	36	<b>69</b>	16	<b>31</b>	52
Total					<b>353</b>					<b>215</b>
GRANADA										
Valor	<i>deber</i>					<i>deber de</i>				
	Deóntico		Epistémico		Total	Deóntico		Epistémico		Total
	N	%	N	%		N	%	N	%	
Tercer**	8	<b>57</b>	6	<b>43</b>	14	12	<b>100</b>	0		12
Segundo	3	<b>75</b>	1	<b>25</b>	4	6	<b>86</b>	1	<b>14</b>	7
Primer	1	<b>50</b>	1	<b>50</b>	2	13	<b>100</b>	0		13
Total					<b>20</b>					<b>32</b>

Fuente: Gómez Molina, 2013a; Manjón-Cabeza Cruz, 2017. \*Soc. = nivel sociocultural. \*\*Nivel educativo: Tercer = universitario; Segundo = secundario; Primer = primario.

El mayor número de datos de Valencia nos permite decir con cierta confianza que el uso deóntico de *deber de* que ya habíamos visto en el Cuadro 4 anterior se constata en los tres sociolectos, con un porcentaje menor en el sociolecto alto (50%) y más elevado en el sociolecto bajo (69%). Estos y otros resultados probabilísticos llevan a Gómez Molina (2013a: 106) a sugerir que deber «incrementará su uso para expresar probabilidad y que la presencia de la preposición en deber de actúa a favor del énfasis expresivo de la obligación» (Gómez Molina, 2013a: 106). Si así fuera, Valencia se acercaría a lo observado en Granada donde, como apunta Manjón-Cabeza Cruz (2017: 10), los datos muestran que se está produciendo «una clara tendencia a usar la variante *deber* + inf. para expresar la probabilidad», y cito también su ejemplo (11):

- (11) Yo entré a trabajar en el mes de julio del noventa / eso *debió ser* en mayo o junio del año noventa.

Observamos, entonces, por ahora, diferenciación en datos de Valencia y Granada; diferenciación en cuanto a la casi absoluta desaparición del uso epistémico de *deber de* en Granada.

Para terminar esta sección sobre *deber* y *deber de*, es posible sugerir además que, quizá con excepción de Valencia, la extensión de *deber* a contextos epistémicos no establece gran diferencia entre dialectos del español y el uso de *deber de* en contextos deónticos establece una clara diferencia entre el español hispanoamericano y el peninsular (exceptuando Granada por la escasez de los datos, Tabla 5).

#### 4. PLURALIZACIÓN DE HABER IMPERSONAL

Finalmente, quiero detenerme unos minutos en lo que he aprendido sobre *haber*. La pluralización de *haber* es una variable sintáctica que ha recibido bastante atención en los últimos decenios. Su estudio ha sido privilegiado por algunos equipos de PRESEEA. Con metodologías comunes, se espera determinar con más precisión tanto su difusión geográfica y social como los factores lingüísticos que condicionan la variación.

La variante plural, como en el ejemplo (1a)

- (1.) a) *Han* habido muchas protestas      b) *Ha* habido muchas protestas

en el que la frase nominal (FN) «muchas protestas» concuerda en número con el verbo, alterna con la variante normativa en singular, «Ha habido muchas protestas» en (1b).

*Haber* es un verbo presentativo, es decir, introduce en el discurso el referente de una FN considerada el objeto directo, usualmente en posición postverbal, que parece reinterpretarse como el sujeto de la oración y se hace concordar con el verbo.

La variante concordada de *haber* impersonal en la norma culta de varias ciudades de España e Hispanoamérica había sido constatada ya en los años noventa. No se

habían encontrado casos de pluralización en las normas cultas de Madrid y Sevilla, pero en 2016 se comprueba la existencia de *haber* plural en los corpus PRESEEA de Madrid en todos los estratos sociales, aunque en un porcentaje bajo.

La ocurrencia de *haber* plural en todos los estratos sociales es compartida por estudios realizados en países hispanoamericanos y en otras ciudades de España, como se ve en la Tabla 7, con porcentajes que van de 2,1% en Madrid y 8,4 % en Ciudad de México a 82 % en Táchira, Venezuela. Estos porcentajes deben considerarse con precaución, pues no todos los estudios están basados en las mismas metodologías, pero al menos dan una idea de la extensión de este fenómeno, que parece ir en aumento en muchas variedades del español.

Las formas pluralizadas *habemos* y *hubieron*, en cambio, no son frecuentes y parecen estar estigmatizadas en el nivel educativo alto en algunos dialectos.

**TABLA 7. PORCENTAJES DE PLURALIZACIÓN DE *HABER* EN COMUNIDADES HISPANOHABLANTES**

	Comunidad de habla	%
Madrid		<b>2,1</b>
Ciudad de México		<b>8,4</b>
Las Palmas de Gran Canaria		<b>36,2</b>
Valencia, España		<b>46,2</b>
Santiago de Chile		<b>50</b>
Castellón		<b>52</b>
Mérida (México)		53
San Juan de Puerto Rico		54,3
Caracas		55,3
La Habana		60,6
Santo Domingo		61,2
Mérida (Venezuela)		70,7
San Salvador		79,6
Táchira (Venezuela)		82

Fuente: Adaptado de información en Bentivoglio y Sedano, 1989; de los Cuadros 3 y 16 de Hernández Cabrera, 2016: 148, 156 y de Tabla 3 de Paredes García, 2016: 220-221.

La pluralización de *haber* y de los verbos auxiliares usados con haber (‘poder’, ‘deber’, ‘soler’, etc.), como en (2), se manifiesta en todos los países hispanohablantes.

(2) *Pueden/Deben/Suelen haber* unas 50 personas en el gimnasio.

En España, se ha sugerido que los porcentajes más altos constatados en Valencia y Castellón son resultado del contacto con el catalán (Gómez Molina, 2013b; Blas Arroyo, 2018 y 2019). En Castellón, Blas Arroyo (2018) observa que las formas

plurales de *haver-hi* son producidas con frecuencia en catalán: 52% de pluralización en catalán, similar al porcentaje de formas plurales en el español de Castellón (56%).

Castellón: 52% de pluralización en catalán (verbo *haver-hi*)  
5 6% de pluralización en español

Además, el análisis de regresión múltiple comparativo muestra variables explicativas similares en las dos lenguas, apoyando una posible situación de influencia recíproca.

En Hispanoamérica, al menos dos variantes que favorecen la pluralización son comunes a casi todos los estudios realizados: el rasgo [+ humano] de la FN y la idea de pluralidad (como en los ejemplos 3 y 4).

- (3) *Habían* unos niños en el parque. FN [+ humana]  
(4) *Habían* más de quinientos libros. FN [+ plural]

Otras variables son significativas en unos estudios y no en otros, por ejemplo, el tiempo del verbo y variables sociales como el sexo y el nivel socioeconómico.

El rasgo [+ humano] de la FN y el tiempo verbal resultan significativos en Puerto Rico, en datos orales que no forman parte de PRESEEA. Los investigadores, Rivas y Brown (2012), analizan 27 horas de conversación con 33 hablantes nativos de edades entre los 24 y los 90 años y encuentran 58% de haber pluralizado.

La Tabla 8 muestra que *Goldvarb* determina que son altamente favorables la FN [+ humana] (con un peso probabilístico de 0.81) y el imperfecto de indicativo (con 0.58).

**TABLA 8. PORCENTAJES VARIABLES LINGÜÍSTICAS QUE FAVORECEN LA PLURALIZACIÓN DE *HABER* EN PUERTO RICO. TOTAL DE CASOS CUANTIFICADOS: 98**

	% <i>haber</i> plural	Peso probabilístico	N casos
FN			
Humana	<b>84</b>	<b>0.81</b>	19
no humana	52	0.41	79
Tiempo verbal			
Imperfecto	<b>65</b>	<b>0.58</b>	61
Otros	47	0.36	37
<i>Hubieron</i> : 11/19 casos de pretérito (58 %) - Ej.: <i>Hubieron</i> muchos desórdenes			

Fuente: Rivas y Brown, 2012.

El pretérito *hubieron* se constata en 11 de 19 casos de pretérito (*Hubieron muchos desórdenes*), mientras que el imperfecto alcanza el 65% de pluralización.

A diferencia de Puerto Rico, en la comunidad urbana de Valencia no es significativa ninguna de las variables lingüísticas según Gómez Molina (2013). Gómez Mo-

lina (2013b) establece diez variables lingüísticas, entre estas el tiempo del verbo y el rasgo [ $\pm$  humano] de la FN; tres estilísticas y cuatro sociológicas. Siguiendo PRE-SEEA, incluye 36 mujeres y 36 hombres distribuidos uniformemente en tres niveles socioculturales (alto, medio, bajo) y tres grupos etarios (mayores de 55; 35-55 y 20-34 años). De todas las variables independientes consideradas, *Goldvarb* identificó solo dos significativas: la edad y el grupo sociocultural. Los resultados, adaptados de Gómez Molina (2013b: 275), se presentan en la Tabla 9.

**TABLA 9. DOS VARIABLES INDEPENDIENTES SELECCIONADAS POR *GOLDVARB* EN LA PLURALIZACIÓN DE *HABER* EN VALENCIA (ESPAÑA)**

Variables independientes	Peso probabilístico	Nº de Casos	%
<i>Edad</i> (rango 0,20)			
20-34	<b>0.59</b>	35/63	<b>55,6</b>
35-55	0.52	62/127	48,8
Más de 55	0.39	30/85	35,3
<i>Nivel sociocultural</i> (rango 0,16)			
Medio	<b>0.56</b>	43/81	<b>53,1</b>
Bajo	0.53	47/94	50
Alto	0.40	37/100	37

Fuente: Tabla 3 de Gómez Molina, 2013b: 275.

Basado en el uso mayor de *haber* plural entre los hablantes jóvenes y en el sociolecto medio, Gómez Molina sugiere que la forma plural parece estar en expansión en Valencia. Esta conclusión encuentra apoyo en el relativamente alto índice de aceptación de las formas personales de *haber* en esta comunidad (37 %). La prueba de aceptabilidad incluía ocho formas personales de *haber*. Por el contrario, aunque el nivel de aceptabilidad hacia la pluralización es incluso más alto en Las Palmas de Gran Canaria, Hernández Cabrera (2016: 158) sugiere «la regresión de la personalización de *haber*». Esta sugerencia se deriva del porcentaje de aceptación más bajo en el grupo etario más joven (50 %) que en el de los mayores de 55 años (65 %), y del uso menos frecuente de *haber* concordado entre los jóvenes de nivel educativo alto (21 %).

Aceptación de *haber* personalizado en:

Valencia: 37 %

Las Palmas de Gran Canaria: grupo etario más joven: 50 %

mayores de 55 años: 65 %

Uso de *haber* concordado: jóvenes de nivel educativo alto: 21 %

Además de los estudios de *haber* de Gómez Molina y Hernández Cabrera, otro estudio realizado siguiendo los criterios de PRESEEA es el de Lastra y Martín Bu-

tragueño (2016) en Ciudad de México. La Tabla 10 presenta resultados; los guiones indican que *Goldvarb* no ha seleccionado este factor como estadísticamente significativo.

Hernández Cabrera y Lastra y Martín Butragueño incluyen los mismos factores que Gómez Molina (2013b), pero a diferencia de este, determinan que una variable lingüística, diferente en cada estudio, es significativa: el rasgo semántico de la FN en Las Palmas ([+humana] 0,59; y la definitud de la FN en Ciudad de México (definida 0,69). Dos variables sociales son también significativas: la edad y el nivel de estudios. El comportamiento según edad es similar en Valencia y México y bastante similar según nivel de estudios en Valencia y Las Palmas. La variable sexo solo resulta significativa en México, donde las mujeres favorecen la pluralización, mientras que en Valencia y Las Palmas esta variable no es estadísticamente significativa.

**TABLA 10. PLURALIZACIÓN DE *HABER*. VARIABLES EXPLICATIVAS ESTADÍSTICAMENTE SIGNIFICATIVAS**

	Peso probabilístico			
	Valencia (N= 275)	Las Palmas (N= 365)		Ciudad de México (N= 561)
<i>Variables independientes</i>				
<i>Lingüísticas:</i>				
Frase Nominal [-humana]		--	0,41	--
[-humana]		--	<b>0,59</b>	--
Frase Nominal [-definida]		--	--	0,46
[-definida]		--	--	<b>0,69</b>
<i>Sociológicas:</i>				
Edad				
20-34 años		<b>0,59</b>	0,44	<b>0,56</b>
35-55 años		<b>0,52</b>	0,44	<b>0,68</b>
más de 55		0,39	<b>0,62</b>	0,33
Nivel de estudios				
Alto		0,40	0,31	0,36
Medio		<b>0,56</b>	<b>0,64</b>	0,39
Bajo		<b>0,53</b>	<b>0,56</b>	<b>0,76</b>
Sexo				
Femenino		--	--	<b>0,61</b>
Masculino		--	--	0,42

Fuentes: Hernández Cabrera, 2016; Gómez Molina, 2013b; y Lastra y Martín Butragueño, 2016. Elaboración propia.

Los resultados dejan claro que estas tres comunidades comparten dos variables externas como condicionantes de la variación *haber* singular~plural: la edad y el nivel de estudios o nivel sociocultural de los hablantes. Es otro ejemplo de invaria-

bilidad entre dialectos. Los niveles medio y bajo potencian la pluralización y el nivel alto la restringe en las tres localidades. Así pues, el análisis probabilístico subraya las semejanzas en sus gramáticas y también las diferencias entre Ciudad de México y las dos comunidades peninsulares.

## 5. CONCLUSIÓN

Y ahora ya concluyo, con unas pocas notas optimistas.

No podemos dejar de reconocer los avances logrados en el estudio de la variación sintáctica, aunque todavía quede mucho por hacer. En el continuo proceso de investigación de los condicionamientos de la variación, el sociolingüismo seguirá haciendo aportes importantes a la sintaxis, semántica y pragmática del español. La variación se ha considerado sistemáticamente y se han buscado explicaciones tanto dentro como fuera del sistema lingüístico propiamente tal. El análisis probabilístico ha permitido hacer predicciones confiables sobre competencia lingüística y equivalencias o divergencias interdialectales.

Es de esperar que estudios coordinados continúen dando resultados significativos para el avance de nuestro conocimiento de la vida real de la lengua española en sus principales variedades y para el esclarecimiento de ciertos aspectos fundamentales de la teoría sociolingüística.

Un resultado de crucial relevancia que nos permita evaluar el futuro de las variedades del español hablado será la identificación de fenómenos de variación que correspondan a cambios en marcha, ya sea hacia patrones más conservadores o, por el contrario, alejándose de ellos y por tanto presagiando quizá mayores divergencias interdialectales. Una comparación de materiales recogidos en dos períodos diferentes podría arrojar resultados reveladores.

Últimamente, en programas televisivos de Chile, por ejemplo, he notado que el imperfecto de subjuntivo en *-ra* ha dado paso, variablemente, a la forma en *-se*, acercándose así, creo yo, al uso peninsular.

TVN Chile, «24 Horas Central»

- Imagínese el daño inmenso que, que va a hacer todo esto si todo esto *llegase* a bajar- (3/15/23)
- Una decisión tan compleja que en la mayoría de los casos no se *hubiese* tomado. (3/15/23)

Hace solo unos 50 años quizá en Chile habrían dicho «si todo esto *llegara* a bajar-» y «no se *hubiera* tomado (esa decisión)», que es lo que yo (¿y otros?) todavía digo.

H: y sigo el tema y lo converso y todo, como si fuera normal

En la televisión española parecen favorecer el imperfecto en *-se*, pero se escucha también de vez en cuando la forma en *-ra*.

TVE, «Telediario 1ª. Edición: 2/2023

- No han dejado de luchar para que este caso no *cayera* en el olvido.  
[interesante: los subtítulos dicen «no *cayese* en el»]  
3/2023
- Otros ex jugadores tampoco creen que el Barça *comprara* árbitros. (14/03/23)
- Hace unos días la juez le pidió que *entregase* el contrato firmado con la Federación para que se *disputase* en Arabia. (3/14/23)

Otro fenómeno que también me llama la atención es la marca variable de objeto directo con la llamada «a personal», que ha sido objeto de numerosísimos estudios, principalmente generativistas, bajo la denominación de *marcado diferencial de objeto directo* o, en inglés, (DOM) ‘differential object marking’.

Me llama la atención porque en mi variedad de español creo que en general no marcaría con *a* objetos directos definidos o indefinidos que designan objetos inanimados:

Puso *EL LIBRO* en la mesa (no *al libro*)  
(Diccionario Panhispánico de Dudas)

Excepto si produjera confusión: «Desvía la corriente el arenal», pero he observado este uso en la televisión española:

TVE: Su discurso [de Erdogan] nacionalista y su propuesta de un estado islamista ha vencido *a* las esperanzas de cambio y acercamiento a occidente.

TVE: El gobierno está inspeccionando *a* los supermercados.

TVE: No hemos visto *a* turistas haciéndose la típica foto.

TVE: Es imposible ya pasear por Valencia sin ver *a* ningún ninot.

Con algunos verbos de percepción como *ver* y *conocer*, puede aparecer o no la preposición, pero en mi variedad, ¿o quizá mi idiolecto?, no la usaría.

Mi intuición me dice que DOM podría representar un cambio en marcha: ¿Se están re-interpretando los ODs y siendo marcados como OIs? Nótese que el leísmo es un caso de OD marcado como OI: *le* vi a Juan ayer. En todo caso, es necesario examinar empíricamente esta escurridiza *a*.

Como he dicho, queda tanto por hacer. El futuro de la investigación sobre la variación sintáctica en español es promisorio. La labor parece titánica pues los estudios variacionistas son complejos y requieren el esfuerzo de muchos. Afortunadamente, sin embargo, hay ya grupos de lingüistas que independientemente o como parte de PRESEEA continúan haciendo aportes valiosísimos a nuestro conocimiento de la lengua española y al avance de la sociolingüística en general.

Así pues, como bien lo escribió Fran Carriscondo en la primera estrofa de su poesía para PRESEEA en 2006:

Preparados para investigar  
a la lengua en sociedad  
no hay variable, ni comunidad  
que dejemos sin analizar.

## REFERENCIAS

- Aracil, Nadia (2001). «Deber/deber de», En Y. Macchi, *Colloque de linguistique hispanique*, Lille: Université Charles de Gaulle, pp. 195-205.
- Bentivoglio, Paola (1987). *Los sujetos pronominales de primera persona en el habla de Caracas*. Caracas: Universidad Central de Venezuela.
- Bentivoglio, Paola y Mercedes Sedano (1989). «Haber: ¿un verbo impersonal? Un estudio sobre el español de Caracas». En *Estudios sobre español de América y Lingüística afroamericana: ponencias presentadas en el 45 Congreso Internacional de Americanistas*, Bogotá: Instituto Caro y Cuervo: 59-81.
- Blas Arroyo, José Luis (2018). «At the Boundaries of Linguistic Convergence: Variation in Presentational haber / haver-hi. A Sociolinguistic Comparative Analysis of Spanish and Catalan grammars». *Languages in Contrast*, 18(1): 35-68.
- Blas Arroyo, José Luis (2019). «La sociolingüística del español en España». En E. Ridruejo (ed.), *Manual de lingüística española*, Berlin, Boston: De Gruyter, pp. 582-613.
- Cerrón-Palomino, Álvaro (2014). «Ser más pro o menos pro: variación en la expresión de sujeto pronominal en el castellano limeño». *Lingüística*, 30(1): 61-83.
- Cerrón-Palomino, Álvaro (2015). «Resumption or Contrast? Non-standard Subject Pronouns in Spanish Relative Clauses. *Spanish in Context*, 12(3): 349-372.
- Cuervo, Rufino (1994). *1886-1893. Diccionario de construcción y régimen de la lengua castellana*, Bogotá: Instituto Caro y Cuervo.
- Eddington, David y Carmen Silva-Corvalán (2011). «Variation in the use of deber and deber de in written and oral materials from Latin America and Spain». *Spanish in Context* 8(2): 257-271.
- Fairclough, Norman (2000). «Discourse, social theory, and social research: The discourse of welfare reform». *Journal of sociolinguistics*, 4(2): 163-195.
- Gómez Molina, José Ramón (2013a). «Norma y usos de las perífrasis ‘deber + infinitivo’ / ‘deber de + infinitivo’», En J. R. Gómez Molina (coord.), *El español de Valencia. Estudio sociolingüístico*, Berna: Peter Lang, pp. 71-108.
- Gómez Molina, José Ramón (2013b). «Pluralización de haber impersonal en el español de Valencia». *Verba*, 40: 253-284.
- Hernández Cabrera, Clara Eugenia (2016). «Variación de haber impersonal en el español de Las Palmas de Gran Canaria». *ELUA*, 30: 141-162.

- Holmquist, Jonathan (2012). «Frequency Rates and Constraints on Subject Personal Pronoun Expression: Findings from the Puerto Rican Highlands». *Language Variation and Change*, 24(2): 203-220.
- Lastra, Yolanda y Pedro Martín Butragueño (2015). «Subject pronoun expression in oral Mexican Spanish», En A. M. Carvalho, R. Orozco y N. Lapidus Shin (eds.), *Subject pronoun expression in Spanish: A cross-dialectal perspective*, Washington, D.C.: Georgetown University Press, pp. 39-57.
- Lastra, Yolanda y Pedro Martín Butragueño (2016). «La Concordancia de haber Existencial en la Ciudad de México». *Boletín de Filología*, 51(2): 121-145.
- Manjón-Cabeza Cruz, Antonio, Francisca Pose Furest, y Francisco José Sánchez García (2016). «Factores determinantes en la expresión del sujeto pronominal en el corpus PRESEEA de Granada». *Boletín de Filología* 51(2):181-207.
- Manjón-Cabeza Cruz, Antonio (2017). «Deber (de) + infinitivo en el corpus PRESEEA de Granada». *LingRed, Lingüística en la Red*, 15: 4.
- Orozco, Rafael (2015). «Pronoun variation in Colombian Costeño Spanish». En A. M. Carvalho, R. Orozco, y N. L. Shin (eds.) *Subject pronoun expression in Spanish: A cross-dialectal perspective*, Washington, D.C.: Georgetown University Press, pp. 17-37.
- Paredes García, Florentino (2016). «La pluralización del verbo haber existencial en Madrid: ¿etapas iniciales de un cambio lingüístico?», *Boletín de Filología*, 51(2): 209-234.
- PRESEEA. *Proyecto de Estudio Sociolingüístico del Español de España y América*, <http://preseea.linguas.net/Inicio.aspx>
- Repede, Doina (2019). «Condicionantes sociolingüísticos de la expresión del sujeto pronominal en el sociolecto alto del corpus Preseea-Sevilla». *Revista de Investigación Lingüística*, Universidad de Murcia, 22: 397-423.
- Rivas, Javier y Esther Brown (2012). «Concordancia variable con haber en español puertorriqueño». *Boletín de Lingüística*, 24(37-38): 102-118.
- Ruiz Gurillo, Leonor (2002). «Algo más sobre el empleo de deber y deber de». *Español Actual*, 77: 157-159.
- Samper Padilla, José Antonio, Clara Eugenia Hernández y Magnolia Troya (1998). «El uso de deber + infinitivo y deber de + infinitivo en la norma lingüística culta de América y España». *Boletín de Filología*, 37(2): 1059-1083.
- Samper Padilla, José Antonio, Clara Eugenia Hernández y Magnolia Troya (2006). «Presencia / ausencia del sujeto pronominal de primera persona singular en la norma culta de España». En C. Company y P. Martín Butragueño (eds.) *El español en América. Diatopía, diacronía e historiografía*. México: UNAM, pp. 87-109.
- Shin, Naomi Lapidus y Helen Smith Cairns (2012). «The Development of NP Selection in School-age Children: Reference and Spanish Subject Pronouns». *Language Acquisition: A Journal of Developmental Linguistics*, 19(1): 3-38.
- Silva-Corvalán, Carmen y Andrés Enrique-Arias (2017). *Sociolingüística y Pragmática del Español*. Washington, D.C.: Georgetown University Press.



## **PRESEA EN AMÉRICA**



## CONSTRUCCIÓN DEL PRESEEA-BUENOS AIRES Y ANÁLISIS DE DATOS

Claudia Borzi

*Consejo Nacional de Investigaciones Científicas y Técnicas  
Universidad de Buenos Aires*

### 1. INTRODUCCIÓN

En este trabajo se describe cómo se formó la ciudad de Buenos Aires a lo largo de su historia, su desarrollo económico, político y social. Luego se presentan características de la muestra y finalmente se da una reseña de los estudios, presentaciones en reuniones científicas, artículos y capítulos publicados hasta el momento sobre los materiales recolectados del PRESEEA Buenos Aires y se resumen brevemente algunos resultados<sup>1</sup>.

Cabe inicialmente un agradecimiento a las Profesoras Paola Bentivoglio y Mercedes Sedano, que en el XV Congreso ALFAL de Montevideo, en agosto de 2008, me invitaron a participar de este valioso proyecto en representación de Buenos Aires. Ya en Buenos Aires, con Norma Carricaburo como co-coordinadora comenzamos a trabajar en la recolección y transcripción de los datos, tarea en la que cooperó un valeroso grupo integrado en distintos momentos por Sofía Blasco López, María Soledad Funes, Carolina Gattei, Rocío Martínez, Julián Martínez Vázquez, Silvina Peri, María Sol Portaluppi y Mónica Trípodí quienes constituyeron un activo equipo desde el inicio de la creación del grupo en 2009. Como ocurre muchas veces en los proyectos de largo alcance, el trabajo contó con poca financiación específica,

---

<sup>1</sup> Vaya mi agradecimiento, no solo a todo el equipo PRESEEA que estuvo siempre presente acompañando, sino en esta oportunidad especialmente a Isabel Molina Martos y a Pedro Martín Butragueño que con su paciencia y comprensión me permitieron llegar a participar de este volumen a pesar de mis continuos retrasos.

solamente durante los dos primeros años de recolección. Con Carricaburo como directora y Borzi como integrante ganamos un Proyecto de Investigación Plurianual (PIP)<sup>2</sup> «Proyecto para el estudio sociolingüístico del español hablado de Buenos Aires» financiado por el Consejo Nacional de Investigaciones Científicas y Técnicas de Argentina.

El proyecto fue descrito y presentado en diversas reuniones científicas, la primera oportunidad fue la lectura de «Metodología y Cuerpos de datos: el PRESEEA Buenos Aires» que presentó Borzi en el marco del VI Simposio de la Asociación Argentina de Lingüística Cognitiva (AALiCo), que tuvo lugar el 5 y 6 de noviembre de 2012, en la Universidad de Buenos Aires, con el objetivo de convocar a los especialistas participantes a trabajar sobre este cuerpo de datos. El cognitivismo prototipalista aboga por la realización de estudios sobre datos auténticos de situaciones reales de uso y, en ese contexto, la oralidad es la fuente más confiable y representativa del uso de una lengua. Hasta ese momento, del habla de Buenos Aires contábamos solamente con los registros llevados adelante en el marco de la Norma Culta hispánica en las décadas de 1960/70, 40 años antes. Si bien en 2004 Ana María Barrenechea, directora de la Norma Culta de Buenos Aires, había dejado en manos de Borzi las cintas originales que contienen las 400 horas de grabación del proyecto, para que las sintetizara y salvaguardara como patrimonio intangible de la ciudad, en ese momento, el rescate estaba en proceso y teníamos solamente, desde 1987, 24 horas y 16 minutos publicados. Actualmente, la digitalización ya está concluida (cf. Borzi, 2021a) y estoy desarrollando el estudio etnográfico. Estas fueron las motivaciones, que originalmente llevaron a Borzi a interesarse por las recolecciones de datos del habla de su ciudad y a promoverlos.

En este marco, los materiales científicamente recolectados del PRESEEA eran no solo muy novedosos, porque aportaban nuevos criterios y estrategias de recolección, sino además muy necesarios. Diez años después, en 2022, Borzi organizó una mesa, para el Congreso Internacional de Estudios Lingüísticos, convocado por la Universidad Nacional de Moreno y la Sede ALFAL de Argentina, que tuvo lugar del 30 de noviembre al 2 de diciembre de 2022. Dicha mesa se denominó «Proyecto 1 de la ALFAL: La Norma Culta Hispánica ‘J.M. Lope Blanch’ y Proyecto 5: PRESEEA. Núcleo BUENOS AIRES» y reunió a investigadores de distintas universidades que presentaron trabajos sobre los datos recolectados en PRESEEA y, en varios casos, llevando adelante comparaciones con la norma culta que permitieron sacar conclusiones diacrónicas. Expusieron de la Universidad Nacional de San Juan, Verónica Orellano, sobre pausas internas en nominales que llamó «demorados» y Adriana Collado, quien se refirió al uso de «pero bueno» como marcador argumentativo; de la Universidad de Buenos Aires Vanina Barbeito, que expuso sobre estudio de la concordancia en esquemas verbales ditransitivos; Ma. Soledad Funes, quien trató el aporte de la preposición

---

<sup>2</sup> PIP Conicet N° 11220090100089CO – 2010-2012.

«por» al marcador discursivo «por lo menos», y de la Universidad del Museo Social Argentino, Sofía Gutiérrez Böhmer, quien se refirió a la didáctica en la enseñanza de tres usos de «se» para su aplicación en español como lengua extranjera.

## 2. BUENOS AIRES: ELEMENTOS DE SU HISTORIA, SU ECONOMÍA, SU GENTE

En el territorio que poblaban tehuelches y querandíes, Pedro de Mendoza desembarcó desde España en el río que había descubierto Solís diez años antes buscando una mítica sierra de plata. Fundó la ciudad de Buenos Aires en la margen derecha del Río de la Plata en 1536 y la nombró Puerto de Santa María del Buen Ayre. Por la resistencia indígena, la falta de alimentos y enfermedades, los pobladores se desplazaron hacia el norte y fundaron Asunción en la actual República del Paraguay. Desde allí partió Juan de Garay para abrir un puerto que rompiera el aislamiento de Asunción, frenara el avance portugués y fuera una alternativa de salida para los metales del Alto Perú. El 11 de junio de 1580 refundó la Ciudad de la Santísima Trinidad y Puerto de Nuestra Señora de los Buenos Ayres, acompañado por españoles que ya llevaban más 40 años en el territorio, criollos e indígenas.

Pronto se desarrolló el valor estratégico de esa ciudad abierta directamente a España, que mejoraba el tránsito hacia el Tucumán, el Alto Perú y Chile y ofrecía la posibilidad de comerciar, legal y también ilegalmente, en especial con el contrabando, que, oculto en barcos de comercio de esclavos que circulaban al margen de los navíos de flota, dejaba pingües ganancias en el puerto. Atentos a esta posibilidad, los gobernadores prefirieron ya a fines del 1500 radicarse en Buenos Aires y pugnaron por la división de la Gobernación del Paraguay, a la que pertenecía Buenos Aires, obteniendo en 1617 la creación de la Gobernación del Río de la Plata, dependiente todavía del Virreinato del Perú. En 1776 se lograría la creación del Virreinato del Río de la Plata que satisfizo el deseo de independencia del puerto y vino a reforzar la presencia española en el comercio, frente a la británica, en el Atlántico sur. Buenos Aires, como su capital, fue el centro político y cultural de una población local que, sin intervención de España, repelió las invasiones inglesas en 1806 y en 1807 y que ya desde inicios del siglo XIX quería tener participación no solo en las decisiones comerciales sino también en las políticas, que, hasta ese momento, descansaban en representantes directos de la corona española. Debilitada España ante la presión francesa y la caída de la Junta de Sevilla, el 25 de mayo de 1810, Buenos Aires es el escenario de la Revolución independentista que da inicio a las Provincias Unidas de Sur América y en 1826 será capital de la República Argentina.

Sin embargo, luchas fratricidas combinadas con incursiones de los nativos desde el Gran Chaco y el llamado «desierto» que estaba poblado de indígenas, hicieron que Buenos Aires se separara de las otras provincias del territorio hasta 1861. Se reincorpora entonces, mantiene las rentas de su aduana, es sede del gobierno presidencial y luego, tras las luchas por su federalización, logra ser la capital del país sin depender

de ninguna provincia. De 1862 a 1885 se resuelven las guerras civiles y se finaliza el sometimiento de indígenas de la zona pampeana, de sus culturas, de sus lenguas y se ocupan sus tierras. Se desarrolla el proyecto nacional liberal agroexportador que descansa en un intenso período inmigratorio y de distribución de tierras, con afluencia especialmente de españoles, italianos y franceses (en mucha menor cantidad). Estos inmigrantes, a medida que dejaron de recibir parcelas para labrar, no abandonaron la ciudad, dejaron su indeleble huella especialmente en la lengua y fueron constructores de una Buenos Aires que resultó el centro del armado de la nación y del estado, el espacio cultural donde se dieron la modernización jurídica privada, penal, comercial, comunicacional y política, y desde donde se desarrolló la educación común obligatoria y gratuita para todo el territorio, que integró en la escuela pública al hijo del inmigrante y al del criollo. Mientras tanto, el enfrentamiento del gobierno nacional con los indígenas de la zona del Gran Chaco se extendió hasta 1917, para resolverse recién en 1994 en el reconocimiento de los derechos indígenas y de sus lenguas. En el siglo XX el país mostró períodos alternados de crecimiento económico con desigual distribución de ganancias y de aumento poblacional con nuevas migraciones europeas y desde los países limítrofes<sup>3</sup>, que en los censos de 2001 y 2010 superaron ampliamente a las de otros países. La afluencia de latinoamericanos de Paraguay, y Bolivia especialmente, duplicó (en 2001) y triplicó (en 2010) la migración de otras procedencias, entre las que siguen prevaleciendo los italianos y los españoles.

A lo largo de todo ese siglo XX, siempre ligada a los beneficios del puerto, Buenos Aires capital federal creció como centro administrativo, comercial, financiero y cultural, mostrando la mayor tasa de actividad del país en el mercado laboral, el menor nivel de pobreza, la menor mortalidad infantil, aunque con alta desigualdad en la distribución del ingreso. En 1994, con la reforma constitucional, pasó a llamarse Ciudad Autónoma de Buenos Aires, nivel que le garantiza independencia administrativa y jurídica, y beneficios económicos en la coparticipación en la recaudación nacional. Geográficamente ubicada sobre el Río de la Plata, es el río su límite este, limitando al sur, oeste y norte con la provincia de Buenos Aires, de la que la separan el Riachuelo y la autopista de circunvalación, la Avenida General Paz. Su superficie es de 203 km<sup>2</sup>; desde 2008, a nivel político y administrativo, la integran 15 comunas y es parte, junto con 24 partidos del Gran Buenos Aires, de la Región Metropolitana. Según el Censo nacional de población, hogares y viviendas de 2010, su población total<sup>4</sup> ascendía en ese momento a 2.890.151 habitantes, con una distribución de 14.450,8 por km<sup>2</sup>. Ese censo registró 32.294 hogares donde una o más personas se reconocieron como descendientes de pueblos indígenas u originarios, que represen-

---

<sup>3</sup> En 1869, 12.1% de la población había nacido en el extranjero; con fuerte diferencia en 1895 al constituir el 25.4%, que ascendió, en 1914, al 29.9%. En 1947 se mostró un claro descenso al 15.3% y siguió descendiendo. De esos porcentajes totales de migrantes extranjeros, de un 2% a un 3% en cada período mencionado previamente corresponde a migrantes de países limítrofes.

<sup>4</sup> La población total incluye a las personas viviendo en situación de calle.

taban solamente un 2.80% del total de hogares censados en la ciudad, 1.150.134. A su vez, hubo 8.056 hogares (0.7%) donde una o más personas declararon tener antepasados de origen afrodescendiente o africano en relación con la totalidad de hogares censados. El número de extranjeros ascendía a 381.778 (13.21%); 297.325 (10.21%) provenientes de otros países de América (límitrofes, especialmente Paraguay y Bolivia; y no límitrofes, en particular de Perú) y 66.083 (2.28%) provenientes de Europa, especialmente de España e Italia.

En relación con el sexo declarado, el censo de 2010 registró sobre una población total de 2.890.151 personas, 1.329.681 varones y 1.560.470 mujeres; de los que, en la franja etaria de 20 a 34 años, se registraron 351.697 varones y 372.091 mujeres; de la franja etaria de 35 a 54 años se registraron 345.870 varones y 392.979 mujeres y de la franja etaria 55 o más (que en el censo abarcó hasta 100 o más) 308.568 varones y 478.754 mujeres. En conclusión, en Buenos Aires se registran más mujeres que hombres: 54% de la población es femenina mientras que solamente un 46% es masculina.

La población argentina en general y la de Buenos Aires en particular muestra disminución de la tasa de natalidad. De acuerdo con el censo 2010 la proporción de adultos de más de 65 años ya supera el 10.2% a nivel total país, siendo la capital la que concentra más personas mayores, con un 16.4%. La edad media del total de la población es aproximadamente 29 años, lo que hace de Argentina el tercer país más envejecido de la región, aunque muy lejos de Europa cuya edad media es 43 años.

En cuanto a las características de su lengua, Buenos Aires integra la Región lingüística Bonaerense<sup>5</sup>, que forma parte de la Macrorregión Litoral (junto con la Región Litoraleña y Patagónica); la distinguen de las otras macrorregiones, el voseo pronominal y verbal con paradigma único, el yeísmo rehilado y la /r/ no asibilada. La norma de la ciudad orienta la norma general del país.

### 3. LA MUESTRA

Como se dijo, según el Censo Nacional de 2010, en la Ciudad Autónoma de Buenos Aires había 2.890.151 pobladores, por lo que para el proyecto se realizaron 108 encuestas, 12 de cada grupo social y banda etaria, a 6 hombres y a 6 mujeres respectivamente. La tasa de urbanización de Buenos Aires es del 86.3% según la metodología general (Moreno Fernández, 2021). En Argentina, de acuerdo con la legislación, la educación primaria es común, gratuita y obligatoria a partir de la Ley 1420 publicada el 8 de julio de 1884 y la educación secundaria, a partir de la Ley de Educación Nacional 26.206 de 2006, cumpliéndose en total 12 años de instrucción común gratuita y obligatoria. Así, el sistema educativo está organizado en 4 niveles: Educación Inicial, Educación Pri-

---

<sup>5</sup> Borzi, C. (2021c).

maria, Educación Secundaria y Educación Superior (que incluye educación terciaria y universitaria). A su vez, la educación terciaria y universitaria son también gratuitas. En este marco, en la ciudad de Buenos Aires especialmente, la población sin escolaridad es muy reducida, por lo que no se entrevistaron personas sin estudios<sup>6</sup>. Así, la variable educativa se organizó en 3 niveles: enseñanza primaria (incompleta/completa); secundaria (incompleta con al menos 1 año de estudio/completa); y terciaria (incompleta al menos con 1 año de estudio/completa) que abarcó personas con estudios terciarios y/o universitarios. En relación con la variable etaria, conforme la metodología general, se categorizaron 3 niveles: de 20<sup>7</sup> a 34; de 35 a 54 y de 55 o más.

Comenzada la recolección se privilegió el grupo de estudios terciarios y/o universitarios con el objetivo de comparar el PRESEEA con los datos de la Norma Culta Hispánica<sup>8</sup>.

#### 4. ESTUDIOS SOBRE LOS DATOS DEL PRESEEA BUENOS AIRES

A medida que se fueron recolectando materiales se comenzaron a hacer estudios sobre los datos. Con la dirección de Borzi, se llevaron a cabo 2 tesis doctorales sobre estos datos, defendidas en la Facultad de Filosofía y Letras de la Universidad de Buenos Aires. Por un lado, Funes (2015) analizó el uso de la preposición «de». Comparó datos de la Norma Culta de Buenos Aires con datos del PRESEEA\_BUE. Se concluyó del análisis que «de» presenta como significado básico, prototípico, estable en cada ocurrencia analizada, y central en la categoría, la posesión, que constituye un *continuum* entre atribución y propiedad. La descripción semántica de los esquemas nominales con «de» permite proponer una estructura categorial radial. A su vez, se propuso que, dentro del discurso, «de» entabla relaciones sintácticas de centro-periferia y de coordinación.

Por otro lado, con la dirección de Borzi y a partir de una de sus hipótesis, y la codirección de Jorge Gurlekián, en su Tesis de doctorado, Trípodí (2019) analizó la

---

<sup>6</sup> Según el Censo Nacional de 2001 (cuando se comenzó la reunión de datos no estaban publicados todavía los resultados del Censo 2010), sobre 2.307.177 habitantes en la Ciudad de Buenos Aires, se detectó solamente un 1.8% de personas sin instrucción; 4.2% con primaria incompleta; 18.2% con primaria completa; 16.9% con secundario incompleto; 24% con secundario completo; 15% con estudios superiores incompletos y un 19.9% con estudios superiores completos. (Fuente: Instituto Nacional de Estadísticas y Censos (INDEC). Dirección Nacional de Estadísticas Sociales y de Población. Dirección de Estadísticas Sectoriales en base a procesamientos especiales del Censo Nacional de Población, Hogares y Viviendas 2001). En 2012 se publicó la versión definitiva del Censo 2010 que también contemplamos, mostrando que de una población total de 3 años o más, solamente un 1% nunca asistió a la escuela; un 70.9% asistió y un 28.1% asistía.

<sup>7</sup> La Ley 26.579 «Mayoría de edad» del Código Civil de la Nación de 2009 modificó la mayoría de edad de 21 a 18 años, de manera que los entrevistados eran todos mayores de edad al momento de la toma de datos.

<sup>8</sup> En 2012, Borzi se integró como Miembro de la Comisión Permanente al *Proyecto de la Norma Culta* (ALFAL y UNAMéxico) bajo la presidencia de la Dra. Alba Valencia.

prosodia de oraciones con cláusulas condicionales, en posición antepuesta, introducidas por «si». Se compararon datos del PRESEEA\_BUE con otras entrevistas hechas por la doctoranda y con un corpus leído construido. Se concluyó que se registra pausa entre prótasis antepuesta y su apódosis, como algunos autores habían propuesto en sus ejemplos escritos. Sin embargo, a diferencia de estas propuestas que marcaban pausa extensa con prótasis modificadoras de la modalidad que tienen menor ligazón semántico-pragmática con la apódosis, se encontró, tras un análisis acústico cualitativo y cuantitativo con pruebas de percepción incluidas, que existe una correlación entre el grado de coherencia entre las prótasis y la apódosis y la duración de la pausa. Es decir que a mayor coherencia entre los conectados la pausa es más extensa (“Si llueve; me quedo en casa»); mientras que a menor coherencia entre los conectados la pausa es más breve (“Si llueve voy al parque»).

El primer trabajo que menciona el PRESEEA\_Buenos Aires es Carricaburo (2010), «La variación acentual en el español de Buenos Aires» donde se atendió al uso de «por ahí/por ahí» (con y sin diptongo), inicialmente locativos, y se sugirió que la primera generación culta tiende a una especialización que mantiene «por ahí» con valor locativo, mientras que hay una reformulación en «por ahí» como índice de modalidad de duda o posibilidad con el significado «quizás» (“Por ahí, te hacés millonario»). Miembros del Equipo de la Cátedra Borzi<sup>9</sup> retomaron la problemática sobre la muestra de nivel culto. Funes (2019) concluye que «por ahí» no funciona solamente como locativo, sino también como marcador discursivo (MD) al igual que «porái»; y que presenta mayor frecuencia que el uso sin hiato. Ambos casos recurren en tres contextos: dubitativo, contrastivo y atenuante. Después de análisis acústicos, Trípodí en varios trabajos en colaboración (Funes *et al.*, 2021; Trípodí y Funes 2022 y Funes *et al.*, 2022), concluyeron que se distinguen con la localización de pausas, tres usos: 1) «por ahí» locativo, con pausa posterior, en frase entonativa con el discurso que lo precede; 2) «por ahí» MD con ausencia de pausa a ambos lados de la construcción, integrado a la frase entonativa que lo contiene discursivamente; 3) «porái» MD con pausa previa en frase entonativa con el discurso que le sigue.

A su vez, otros miembros de la Cátedra, siguiendo el Enfoque cognitivo-prototípico en la línea de Lakoff (1987) y Langacker (1987 y 1991), estudiaron verbos, marcadores discursivos, partículas, preposiciones, y en cláusulas relativas, formas de conexión y pronombres. El estudio de los verbos copulativos y semicopulativos lo desarrolló J. Martínez Vázquez (2012) y M. Morón Usandivaras y Borzi (Morón Usandivaras y Borzi, 2017; Borzi, 2018a) se ocuparon de verbos ditransitivos. Se concluyó para los copulativos y semicopulativos aspectuales de cambio que, si el hablante percibe ese atributo como parte de un estado del actante (es decir, dentro de un marco de cambio) seleccionará *estar*, *ponerse*, *quedarse*; si lo percibe, por

---

<sup>9</sup> Cátedra «C» de Gramática del Departamento de Letras de la Facultad de Filosofía y Letras de la Universidad de Buenos Aires.

el contrario, como cualidad intrínseca (o que se ha vuelto intrínseca) elegirá *ser*, *hacerse* o *volverse*. Luego se confrontaron dichas expresiones con las que emplean verbos adjetivales (*entristecerse* en vez de *ponerse triste*) viéndose que en Madrid predomina *ponerse enfermo*, y en Buenos Aires *enfermarse*. En cuanto a los esquemas ditransitivos, Morón Usandivaras y Borzi clasificaron usos de «dar» con distinto grado de prototipicidad. Borzi (2017a, 2017b y 2019a) comparó el uso de «dale» MD en Buenos Aires con las distintas ciudades PRESEEA subidas a la red en ese momento: Montevideo, Caracas, Medellín, La Habana, Santiago de Chile, Lima, Madrid, Alcalá y Valencia (España). Lo registró como MD en: Buenos Aires, Montevideo, Caracas, La Habana, Santiago de Chile y Alcalá. Identificó 3 contextos: 1. para instar o animar a continuar una acción o a repetirla, registrado en Buenos Aires, Montevideo, Caracas, La Habana y Santiago, «Ya estoy verde de tanto mate, pero *dale* igual»; 2. Para dar aceptación o acuerdo «E.: Bueno, vamos a hablar un poco de tu familia, ¿quierés?/I.: *Dale*», muy frecuente en Buenos Aires, Montevideo, Caracas y Santiago; y 3. un uso exclusivo de La Habana con valor «y ya está», «yo siempre le echo aceite para que para no se me para que no se me pegue ¿no? y ya lo demás le corresponde a la técnica ¿no? con esto yo pongo y *dale* que ella [la olla arrocera] más o menos ya ella misma me avisa cuando está» interpretado gracias a la comprometida respuesta de nuestra querida colega Ana María González Mafud.

Se describieron también los marcadores discursivos, «después» (Borzi, 2021b) y el más identitario del español en Argentina, «che» (Borzi, 2015, 2016 y 2017a). «Después» MD se estudió comparativamente con el uso de «luego» en la Norma culta y en el PRESEEA. En el corpus de 1960/70 se registraron 6 casos de uso de «luego» marcador discursivo («vinieron unas expertas cuba... eh... peruanas, vino una colombiana, en fin--- gente valiosa en el... en los diferentes campos sociales o educativos que pueden hacer o desarrollar cursos para ir formando a las dirigentes locales. Bueno, y *luego* en... eh... la Feria Fantástica de la Solidaridad, que es otra coordinación»), pero no como MD en ese grupo social en el PRESEEA, donde «después» MD es muy frecuente («los mundiales de fútbol sí me gustan, me gustan eh... bueno, ahora ya más de grande me gusta mucho ver, eh... la reacción de la gente ¿no? como fenómeno eh, pero *después* los corsos no, a mis hijos le gusta, a mí la verdad que nunca me gustaron los corsos»). Por su parte, se identificaron 3 contextos de uso de «che» organizados según presencia o ausencia previa de atención del interlocutor: el vocativo o apelativo que da la instrucción de guiar la atención hacia la fuente de la apelación, «*Che*, loco, ¿cuándo pagan acá?»; el Enfocador de la Alteridad orientado a disminuir o aumentar la distancia entre el hablante y su interlocutor «*Che*, tratame de vos/»No, *che*, nene» y el Marcador de modalidad que orienta al interlocutor hacia la afectividad del mensaje, «y yo me dije ¡*Che!* Pero qué raro».

Otro par opositivo estudiado fue «entonces» y «así que» (Borzi, 2011a, 2012a, 2013a, 2013b, 2014a), Los estudios sobre el PRESEEA permitieron agregar a estudios anteriores otro atributo saliente en la elección: el tipo discursivo del fragmento, argumentativo o narrativo. En comparación con los resultados de Borzi (1997) sobre

la Norma culta de Buenos Aires (Barrenechea, 1987) se confirmó que «entonces» se especializa en contextos narrativos de causa/consecuencia-resultado y se acumula con «y» sin desplazarse. En contexto argumentativo, el PRESEEA permitió agregar que «entonces» es acompañado por tiempos pretéritos, verbos en 3<sup>era</sup> persona, marcas de causalidad en el primer conectado (“porque» y «como») y que en este contexto sí se desplaza dentro del segundo conectado señalizando el cambio de función retórica/superestructural, o la información más nueva o importante -desplazamiento que había sido registrado en texto escrito (Borzi, 2000). Borzi, García Jurado y Trípodí (2016), Trípodí y García Jurado (2018) y Trípodí y Borzi (2019), aportaron argumentos acústicos a la descripción posicional. Por el contrario, «así que», que prevalece en contexto argumentativo y de razón/conclusión, registra tiempo presente, 1<sup>era</sup> persona y verbos modales o de lengua; mantiene posición fija entre los conectados y, precedido por pausa (Borzi y García Jurado, 2018), señala la conclusión; rechaza la convivencia no solamente con «y» sino con toda marca de causalidad en el conectado precedente. Se confirma su frecuencia en discurso oral (Borzi, 2014b). Se propuso que, en contexto consecutivo, «entonces» activa no solamente su valor mostrativo temporal anafórico, sino también mostración hacia adelante. «Así que», por el contrario, constituye un límite y se especializa en focalizar la conclusión explotando su capacidad catafórica.

También se estudió sobre esos datos el uso de la forma «ahora» MD (Borzi, 2012b y 2019b) y se establecieron comparaciones con los resultados obtenidos en Borzi (2004) sobre la Norma culta. A diferencia de lo que plantean Martín Zorraquino y Portolés Lázaro (1999: §63.3.4.12 y 63.6.1 nota 97) y la NGLÉ (RAE/ASALE, 2009: §30.12h) el PRESEEA mostró que la forma no solamente funciona en contextos de contraste, sino que es un marcador conversacional de inicio y toma de turno que se emparenta con el fático o que cambia de tema (Silva Corvalán, 2001: 5.4.2.2.). A diferencia de Fuentes Rodríguez (1998: 56) PRESEEA mostró que «ahora» puede marcar énfasis. Estos contextos conversacionales actualizan su significado adverbial deíctico y aspectual puntual. Se aportó así más evidencia a que la conversación tiene marcadores propios.

En cuanto a preposiciones y partículas, tras su tesis doctoral, Funes (2017a, 2017b, 2017c, 2020 y 2021) analizó el uso de la preposición «por» y luego buscó explicar la elección de «por» en lugar de «de» en complemento agente. Planteó un ítem polisémico para la preposición «por» de estructura radial, cuyo significado básico es el locativo, que denota un camino, un trayecto de extensión grande, pero también imprecisa. La variedad rioplatense actual, muestra que la imprecisión enlaza todos los significados hallados de «por», y que los usos más frecuentes son los de locativo (el básico) y el de causa (más abstracto, derivado de la metáfora espacial de tiempo) que da lugar al valor agentivo. Los pocos casos de significado ‘en favor de’ derivan de estos últimos. La imprecisión del nominal que acompaña lleva hacia adelante en el discurso, persiste; funcionando «por» como focalizador. Participa así de la relación sintáctica de centro-periferia, marcando el centro. En otros contextos como en

«por lo menos» manifiesta un continuum de significados, del más concreto (cantidad mínima esperable) al más abstracto (atenuación), pasando por un significado intermedio que combina cantidad y atenuación.

Finalmente, otros temas que se investigaron sobre estos datos, fueron la conexión y la posición del sujeto en las cláusulas relativas (Borzi, 2011b, 2018b) ofreciéndose una explicación unitaria para unos fenómenos descritos de manera desarticulada entre sí y tangencialmente en las gramáticas: la elección de «que» por preposición + artículo + que («La noche que/en que (en la que) velaron a Teodelina»); el uso de «que» galicado en lugar de preposición + artículo + que («Fue ese día que (en el que) los invitó») y el agregado de un pronombre Reasuntivo («Una decisión que ella no *la* tomó»). El hablante elige estas estrategias para lograr un texto menos complejo cuando el antecedente está indeterminado y alejado por incisos, para aclararlo reconstruye en la relativa la distribución informativa oracional de información conocida a nueva (SVO), usa «que», antepone el sujeto e introduce un pronombre reasuntivo.

Más allá de los encuentros en el marco de los congresos ALFAL y de ALFAL Argentina, los resultados y los materiales PRESEEA se difundieron en reuniones científicas dentro del país y en universidades o congresos en universidades de Brasil, Chile, Uruguay, Estonia, Suecia, Alemania, Francia y España. Las actividades de extensión, cursos sobre redacción académica y normativa dictados en la Universidad de Buenos Aires desde 2005, y las muy numerosas capacitaciones docentes dictadas en diferentes instituciones educativas, así como los talleres de las sucesivas Jornadas de didáctica de la lengua y la literatura, que anualmente organiza la cátedra de Borzi desde 2018, permitieron el acceso de un público más amplio a los resultados de los estudios sobre PRESEEA y al proyecto en sí, como fuente de datos científicamente organizada y muy confiable, y como instrumento para el tratamiento de la oralidad en el aula escolar.

## REFERENCIAS

- Barrenechea, Ana María (coord.) (1987). *El Habla Culta de la Ciudad de Buenos Aires. Materiales para su Estudio, Tomos I y II*. Buenos Aires: Universidad de Buenos Aires.
- Borzi, Claudia (1997). *Syntax, Semantik und Pragmatik der Konnektoren «entonces», «así que» und «(tal/...)que»*. München: Ludwig-Maximilians-Universität Profidruck.
- Borzi, Claudia (2000). «El entonces focalizador y el anafórico temporal». Comunicación leída en el *VIII Congreso de la Sociedad Argentina de Lingüística (SAL)*, Universidad Nacional de Mar del Plata, 20 al 23 de Setiembre de 2000. Mar del Plata: Argentina.
- Borzi, Claudia (2004). «Los usos de *ahora* en el Habla Culta de Buenos Aires». En B. Guaycochea, Walter Solar y S. Betti (eds.), *Actas de las 3<sup>ras</sup>. Jornadas 'La Creatividad Lingüística'*. San Luis, Argentina: Universidad Nacional de San

- Luis, pp. 1-9.
- Borzi, Claudia (2011a). «La variación ‘así que/entonces’ en el PRESEEA Buenos Aires». Comunicación leída en el *II Coloquio internacional Marcadores discursivos en las lenguas románicas: un enfoque contrastivo*. Universidad de Buenos Aires, 1<sup>er</sup>o al 3 de diciembre de 2011, Buenos Aires, Argentina.
- Borzi, Claudia (2011b). «Preposición + (artículo) + ‘que’ en cláusulas relativas (adjetivas). Evaluación de una hipótesis». Comunicación leída en el *XVI Congreso Internacional de la ALFAL*, Universidad de Alcalá de Henares, 6 al 9 de junio de 2011, Alcalá de Henares, España.
- Borzi, Claudia (2012a). «La posición de ‘entonces’ en contexto argumentativo en el PRESEEA Buenos Aires». Comunicación leída en el *I Congreso de la Delegación Argentina de la ALFAL*, Universidad Nacional de La Plata, 21 al 23 de marzo de 2012, La Plata, Argentina.
- Borzi, Claudia (2012b). «‘Ahora’ marcador conversacional: su uso en el PRESEEA Buenos Aires». Comunicación leída en el *IV Simpósio nacional de estudos filológicos e linguísticos*, Universidade Federal Fluminense, 2 a 4 de abril de 2012, Niterói, Brasil.
- Borzi, Claudia (2013a). Conferencia «Marqueurs et connecteurs: L’usage des connecteurs ‘entonces’ et ‘así que’ dans un contexte narratif». *Groupe d’études et de recherches en Linguistique Hispanique*, Université de la Sorbonne Nouvelle, Paris 3, 25 de mayo de 2013, Paris, France.
- Borzi, Claudia (2013b). «El uso de ‘entonces/ así que’ en contexto narrativo sobre datos del Preseea-Buenos Aires». *Estudios de Lingüística de la Universidad de Alicante*, 27: 61-88.
- Borzi, Claudia (2014a). «‘Entonces’ en alternancia con ‘así que’ en contexto argumentativo en el PRESEEA Buenos Aires». En M. M. García Negroni (ed.), *Marcadores del Discurso. Perspectivas y contrastes*. Buenos Aires: Santiago Arcos Editor, pp. 309-336.
- Borzi, Claudia (2014b). «Marcadores discursivos en la oralidad culta de Buenos Aires». Comunicación leída en el *XVII Congreso Internacional de la ALFAL*, Universidad Federal da Paraíba, 14 al 19 de Julio de 2014, Joao Pessoa, Brasil.
- Borzi, Claudia (2015). «¿Por qué una misma forma puede ser útil como vocativo, interjección y marcador discursivo? Una propuesta para el caso de ‘che’». Comunicación leída en el *VII Simposio de la Asociación Argentina de Lingüística Cognitiva (AALiCo)*, Universidad Nacional de San Luis, 4 y 6 de noviembre de 2015, San Luis, Argentina.
- Borzi, Claudia (2016). «El ‘che’ argentino: sus contextos de uso y su significado». *Philologia Hispalensis. Sección monográfica*, 30/ 1: 9-32.
- Borzi, Claudia (2017a). Conferencia «El español coloquial de Buenos Aires: ‘che’ y ‘¡dale!’. Identidad e innovación». Departamento de estudios románicos y clásicos de la Facultad de Humanidades de la Universidad de Estocolmo, 3 de febrero de 2017, Estocolmo, Suecia

- Borzi, Claudia (2017b). «Variación y uso: contextos de ‘¡dale!’ en distintas ciudades del PRESEEA». Comunicación leída en el *XVIII Congreso Internacional de la ALFAL. Proyecto N°17: Variación y Pragmática Regional*, Universidad Nacional de Colombia, 24 al 28 de julio de 2017, Bogotá, Colombia.
- Borzi, Claudia (2018a). «‘Dar’ en el discurso». Comunicación leída en el *XVI Congreso de la Sociedad Argentina de Estudios Lingüísticos*. Universidad Nacional de San Martín, 3 al 6 de Julio de 2018, San Martín, Argentina.
- Borzi, Claudia (2018b). «La conceptualización del antecedente como explicación que unifica cuatro fenómenos marginados por los gramáticos». *Cuadernos de la ALFAL*, 10: 24-44.
- Borzi, Claudia (2019a). «Variación regional en el uso de ‘dale’ marcador discursivo. Estudio del contexto de aceptación». *Revista Internacional de Lingüística Iberoamericana*, 17.1(33): 15-35.
- Borzi, Claudia (2019b). Conferencia «La mostración de lo evidente para capturar la atención del interlocutor. ‘Ahora’ marcador discursivo». Comunicación leída en el *Seminario Internacional Las investigaciones lingüísticas en el Mundo Hispánico*, Universidad de La Habana, 28 de enero al 1<sup>er</sup>o febrero de 2019, La Habana, Cuba.
- Borzi, Claudia (2021a). «La Norma Culta de Buenos Aires dirigida por Ana María Barrenechea. Informe sobre el inventario y la digitalización de las grabaciones originales». *Anuario 30 de la Academia Boliviana de la Lengua*, 129-153.
- Borzi, Claudia (2021b). «El uso de ‘después’ en el PRESEEA Buenos Aires». Comunicación leída en el *XIX Congreso Internacional de la ALFAL*, Asociación de Lingüística y Filología de América Latina, 8 al 13 de agosto de 2021, La Paz, Bolivia.
- Borzi, C. (2021c). «El español en Argentina». En F. Moreno Fernández y R. Caravedo (eds.). *Dialectología hispánica / The Routledge Handbook of Spanish Dialectology*. Londres: Routledge, pp. 189-201.
- Borzi, Claudia, M.<sup>a</sup> Amalia García Jurado y Mónica Trípodí (2016). «Análisis preliminar de las configuraciones prosódicas de ‘entonces’ en el discurso oral de Buenos Aires». En Gabriela Leighton (ed.), *Actas de las IV Jornadas Internacionales y I Congreso de Didáctica de la Fonética de las Lenguas Extranjeras*. Universidad Nacional de San Martín: San Martín, Argentina, Publicación en CD.
- Borzi, Claudia y M.<sup>a</sup> Amalia García Jurado (2018). «Descripción preliminar del entorno prosódico de ‘así que’ en contexto argumentativo». Comunicación leída en el *V Congreso Internacional: Didáctica de la fonética y de la fonología de las lenguas extranjeras*, Universidad Nacional de San Martín, 24 al 25 de agosto de 2018, San Martín, Argentina.
- Carricaburo, Norma (2010). «La variación acentual en el español de Buenos Aires». *Boletín de la Academia Argentina de Letras*, LXXV, mayo-agosto: 363-372.
- Fuentes Rodríguez, Catalina (1998). *Las Construcciones Adversativas*. Madrid: Arco/Libros.

- Funes, M.<sup>a</sup> Soledad (2015). *Estudio del uso de la preposición «de» en el español de Buenos Aires. Un acercamiento desde el Enfoque Cognitivo-Prototípico*. Múnich: LINCOM Europa.
- Funes, M.<sup>a</sup> Soledad (2017a). «The preposition ‘de’ as a polysemic item in Buenos Aires Spanish». Comunicación presentada en la *14. International Cognitive Linguistics Conference*, Universidad de Tartu, 10 al 14 de julio de 2017, Tartú, Estonia.
- Funes, M.<sup>a</sup> Soledad (2017b). «Las construcciones nominales con preposición ‘de’ en español. Un estudio desde la Lingüística cognitiva». *Cuadernos AISPI. Estudios de lenguas y literaturas hispánicas*, 10: 39-62.
- Funes, M.<sup>a</sup> Soledad (2017c). «La posesión inalienable en los esquemas nominales con preposición ‘de’». En F. Carranza, N. Giollo, C. Muñoz Pérez y F. Soriano (coords.), *Actas de las II Jornadas de Jóvenes Lingüistas. Ciudad Autónoma de Buenos Aires*. Buenos Aires: Editorial de la Facultad de Filosofía y Letras, Universidad de Buenos Aires. Libro digital, pp. 134-142.
- Funes, M.<sup>a</sup> Soledad (2019). «Por ahí es un marcador discursivo: una propuesta cognitiva para el español de Buenos Aires». *Oralia*, 22(1): 153-166.
- Funes, M.<sup>a</sup> Soledad (2020). «Del espacio a la causa: el ítem polisémico de la preposición ‘por’». *Textos en proceso*, 6(1): 27-54.
- Funes, M.<sup>a</sup> Soledad (2021). «La preposición ‘por’ como focalizador: una propuesta cognitiva». Comunicación presentada en el *XVII Congreso de la Sociedad Argentina de Estudios Lingüísticos*, Universidad Nacional de Tucumán, 17 al 22 de mayo de 2021, San Miguel de Tucumán, Argentina.
- Funes, M.<sup>a</sup> Soledad, Mónica Trípodí, M.<sup>a</sup> Florencia Tenaglia y Muriel Troncoso (2021). «Presencia o ausencia de pausa en el marcador discursivo ‘por ahí’: un estudio prosódico y discursivo». Comunicación presentada en el *III Encuentro Iberoamericano de Retórica*. Universidad de Buenos Aires y Asociación Argentina de Retórica, 10 al 12 de noviembre 2021, Buenos Aires, Argentina.
- Funes, M.<sup>a</sup> Soledad, Mónica Trípodí, Muriel Troncoso y M.<sup>a</sup> Florencia Tenaglia (2022). «¿Porái o por ahí? La variación diptongo versus hiato en el marcador discursivo ‘por ahí’ en el español de Buenos Aires». En G. Leighton, R. Basso y A. Renato (eds.). *Pedagogía de la Fonética y la Fonología de las Lenguas Extranjeras*. Universidad Nacional de San Martín: San Martín, Argentina, Libro digital, pp. 33-52
- Instituto Nacional de Estadísticas y Censos (2012). *Censo nacional de población, hogares y viviendas 2010. Censo del bicentenario. Resultados definitivos*. Serie B N° 2, Tomos 1 y 2, Buenos Aires: Instituto Nacional de Estadísticas y Censos.
- Lakoff, George (1987). *Women, Fire and dangerous Things*, Chicago: The University of Chicago Press.
- Langacker, Ronald (1987). *Foundations of Cognitive Grammar. Theoretical Prerequisites*. Stanford: Stanford University Press.
- Langacker, Ronald (1991). *Foundations of Cognitive grammar. Descriptive application*. Stanford: Stanford University Press.

- Martín Zorraquino, M.<sup>a</sup> Antonia y José Portolés Lázaro (1999). «Los marcadores del discurso». En I. Bosque y V. Demonte (dirs.), *Gramática Descriptiva de la Lengua Española*. Madrid: Espasa Calpe, pp. 4051-4213.
- Martínez Vázquez, Julián (2012). «Esquemas Pseudocopulativos Aspectuales de Cambio: Estudio sobre datos del PRESEEA Buenos Aires». Comunicación leída en el *I Congreso de la Delegación Argentina de la Asociación de Lingüística y Filología de América Latina*, Universidad Nacional de La Plata, 21 al 23 de marzo de 2012, La Plata, Argentina.
- Morón Usandivaras, Mariana y Claudia Borzi (2017). «Concepción de eventos y cláusulas con ‘dar’ en el PRESEEA Buenos Aires». Comunicación leída en el *XVIII Congreso Internacional de la ALFAL*, Universidad Nacional de Colombia, 24 al 28 de julio 2017, Bogotá, Colombia.
- Moreno Fernández, Francisco (2021). *Metodología del «Proyecto para el estudio sociolingüístico del español de España y de América» (PRESEEA)*. Alcalá de Henares: Universidad de Alcalá.
- Real Academia Española y Asociación de Academias de la Lengua Española (2009). *Nueva gramática de la lengua española*. Madrid: Espasa.
- Silva Corvalán, Carmen (2001). *Sociolingüística y pragmática del español*. Washington: Georgetown University Press.
- Trípodí, Mónica Noemí (2019). *Estudio de la prosodia de las cláusulas condicionales con «si» inicial: Fonética y Gramática*. Editorial Académica Española: Saarbrücken.
- Trípodí, Mónica Noemí, Claudia Beatriz Borzi y Jorge Gurlekián (2012). *La prosodia y la sintaxis en las condicionales introducidas por «si»*. Editorial Académica Española: Saarbrücken.
- Trípodí, Mónica y M.<sup>a</sup> Amalia García Jurado (2018). «Trama sintáctico-discursiva y prosodia en el uso narrativo y argumentativo de ‘entonces’». Ponencia leída en el Coloquio Estudios fonético-fonológicos desde diversas perspectivas en el marco del *XVI Congreso de la Sociedad Argentina de Estudios Lingüísticos*, Universidad Nacional de San Martín, 3 al 6 de julio de 2018, San Martín, Argentina.
- Trípodí Mónica y Claudia Borzi (2019). «Marcador ‘entonces’. Segmentación del habla coloquial: el método del grupo Val.Es.Co y distinciones prosódicas». En *Actas del I Congreso Internacional de Ciencias Humanas «Humanidades entre Pasado y Futuro»*, Universidad Nacional de San Martín, San Martín, Argentina: 1-14. Disponible en: <https://www.aacademica.org/1.congreso.internacional.de.ciencias.humanas/1580>
- Trípodí, Mónica y M.<sup>a</sup> Soledad Funes (2022). «Variación acentual en el marcador ‘por ahí/porái’: un estudio prosódico y discursivo». Comunicación leída en las *VII Jornadas Internacionales de Fonética y Fonología*, Universidad Nacional del Comahue, 19 al 21 octubre de 2022, Río Negro, Argentina.

## PRESEEA-LA HABANA: BREVE HISTORIA Y RESULTADOS

Ana María González Mafud  
*Universidad de La Habana*

### 1. INTRODUCCIÓN

La cuestión del español de Cuba fue siempre un problema que, más allá de su específico carácter lingüístico, se proyectaba necesariamente sobre el más amplio y candente tema de la identidad cultural cubana<sup>1</sup>. Luego de un largo proceso de transculturación y particulares interinfluencias de lenguas diferentes, el español se convirtió en nuestra lengua nacional. Y desde que comenzaron a perfilarse las características de un hablar propio, distintivo, se gestó la conciencia de investigar –a veces de forma rudimentaria y otras con todo el conocimiento que la época y las circunstancias de Cuba permitían– esos rasgos comunes y diferenciadores que nos identifican y distinguen.

Desde entonces las investigaciones lingüísticas en Cuba han transitado por varias etapas con resultados de diferente alcance, con la ampliación de áreas de estudio, determinadas por el desarrollo que ha alcanzado la lingüística en el siglo XX y los aportes de disciplinas como la sociolingüística, la pragmática y el análisis del discurso. En cada uno de esos momentos, sin embargo, el interés fundamental ha estado centrado en la descripción y caracterización de nuestra variedad, en el entendido de que la forma en la que nos expresamos, la manera en que hablamos en registros más o menos formales, es parte inseparable de nuestro modo de ser, de nuestra idiosincrasia, de nuestra identidad cultural.

---

<sup>1</sup> Cfr., por ejemplo, el estudio «Orígenes [*sic*] del lenguaje criollo» de Juan Ignacio de Armas (1882), reproducido en *Antología de la lingüística cubana* de Gladys Alonso y Ángel Luis Fernández (1977).

El panorama de las investigaciones y los estudios lingüísticos hoy en Cuba es amplio y diverso, de madurez y consolidación en el trabajo teórico, pero sobre todo en la aplicación a la caracterización y descripción de nuestra variedad. Se recogen los resultados de muchos años dedicados, por una parte, a la formación de lingüistas a lo largo y ancho del país y, por otra, a la creación y el desarrollo de instituciones dedicadas a estos fines, así como el intercambio con profesores e investigadores de diversos países y la publicación de materiales, textos, trabajos de culminación de grado, máster y tesis doctorales, ensayos y artículos especializados.

En el ámbito de las investigaciones sociolingüísticas y la construcción de corpus, el verdadero *boom* comienza en Cuba en la década del 90 del pasado siglo, cuando se inserta en algunos de los principales proyectos internacionales. Por razones de diversa índole, Cuba siempre ha llegado tarde a esos proyectos, pero es justo señalar, con toda modestia, que, gracias a la colaboración de magníficos equipos de profesores y estudiantes, hicimos bien los deberes o la tarea, como diríamos nosotros.

Es el caso, por ejemplo, del «Estudio coordinado de la norma lingüística culta ‘Juan M. Lope Blanch’» (NLC), el decano o el de más solera como le gustaba decir a José Antonio Samper. Habiendo sido diseñado en 1964, Cuba, en su segunda época, es decir, registrando el habla y el léxico de los habaneros que viven en La Habana, se incorpora en 1993. Sucedió también con el «Proyecto para el estudio sociolingüístico del español de España y América» (PRESEEA) al que nos sumamos en 2009, luego de casi 10 años de haberse iniciado.

A todos esos proyectos hemos podido aportar los resultados más importantes que se espera de ellos: las *Muestras del habla culta de La Habana, el Léxico culto de La Habana* (González Mafud *et al.*, 2010); las 18 muestras que exige PRESEEA, las 108 muestras de La Habana, que, aunque no han sido publicadas, sí están recogidas y transcritas, y que han sido utilizadas por investigadores cubanos y de otros países. Y se ha podido establecer igualmente el primer corpus de conversaciones coloquiales en Cuba (Hab.Es.Co), lo que ha permitido introducir y visibilizar algunos de los fenómenos pragmlingüísticos del estudio del español coloquial contemporáneo en Cuba. Hab. Es.Co abarca el período 2012-2018 y está conformado por 82 grabaciones, equivalentes a 21 horas de conversación, 274 participantes, 186 mujeres y 88 hombres, distribuidos en tres grupos de edades y niveles alto, medio y bajo de instrucción.

Y todo ello ha tenido una gran significación en cuanto a que no solo ha contribuido a poner de relieve la labor de los lingüistas cubanos en esos ámbitos específicos, sino también porque ha favorecido el trabajo con corpus que confirman, legitiman o reconocen diferencias y procesos de evolución en los rasgos de la variedad cubana del español.

## 2. PRESEEA-LA HABANA: EL INICIO

En particular, en el caso de PRESEEA, recuerdo vívidamente que hace 13 años fui a ver a Paco Moreno al Instituto Cervantes. En esa época yo no lo conocía perso-

nalmente. Quería comentarle dos temas: y uno de ellos era mi interés en que Cuba, La Habana, formara parte de PRESEEA. Él se puso contentísimo y me dio todas las facilidades posibles, así que le agradezco aquí su inestimable apoyo.

El reto para los investigadores cubanos, sin embargo, era grande pues apenas nos separaban 10 meses de la fecha de cierre acordada para la realización de las entrevistas.

En cuanto regresé a La Habana, creamos un grupo de trabajo con profesores jóvenes y estudiantes del último año de la carrera de Letras, quienes acogieron el proyecto con gran interés y motivación y fueron los que hicieron posible, en buena medida, los resultados que hoy podemos mostrar. Desde ese entonces me acompaña en las tareas de coordinación mi colega Marisela Pérez Rodríguez.

El equipo de trabajo inicial quedó conformado por un grupo de investigadores solo de la Facultad de Artes y Letras de la Universidad de La Habana; con posterioridad hubo algunas bajas y se añadieron colegas de otras facultades de la Universidad de La Habana, de otras instituciones y otras universidades. Hoy forman parte del equipo las siguientes instituciones:

- Universidad de La Habana.
- Universidad Central «Marta Abreu» de las Villas.
- Instituto Superior de Relaciones Internacionales.

También contamos con investigadores externos de las Universidades de Zaragoza, Oviedo, Querétaro y de la Real Academia Española.

En el breve tiempo de que disponíamos, logramos realizar las 108 entrevistas que nos correspondían, teniendo en cuenta la población de la capital cubana. Ya hoy, aunque no están publicadas, están transcritas, etiquetadas y revisadas.

Hubo que hacer algunas adaptaciones, en particular, en el grado de instrucción, pues resultaba muy difícil encontrar en La Habana personas de los distintos grupos etarios que tuvieran un nivel escolar de sexto grado, por lo que PRESEEA-La Habana comienza en noveno grado, según se observa en la tabla 1.

**TABLA 1. ESTRATIFICACIÓN DE LAS ENTREVISTAS DE PRESEEA-LA HABANA**

	<i>Generación I (20 a 30 años)</i>		<i>Generación II (35 a 54 años)</i>		<i>Generación III (más de 55 años)</i>	
	H	M	H	M	H	M
Grado de instrucción I (hasta 9º grado)	6	6	6	6	6	6
Grado de instrucción II (hasta 12º grado o técnico medio)	6	6	6	6	6	6
Grado de instrucción III (universitarios)	6	6	6	6	6	6

Fuente: PRESEEA-La Habana. Población: 2,13 millones de habitantes. 108 informantes distribuidos uniformemente por edad, sexo, y grado de escolaridad.

Otro problema que se nos presentaba, dada la peculiar realidad económica y social de Cuba, era la determinación de las variables de posestratificación. Ingresos económicos, condiciones de alojamiento y profesión eran variables sin ningún tipo de correspondencia entre sí y en relación con los estándares internacionales. No obstante, hicimos adecuaciones y las trabajamos. Sí tuvimos que prescindir de la determinación del modo de vida que ya la metodología señalaba su no obligatoriedad.

Asimismo, desde 2010 se encuentran disponibles en la página de PRESEEA las 18 muestras exigidas por el proyecto.

### 3. PRESEEA-LA HABANA: ACTUALIDAD

Pero ¿cuál es la realidad de PRESEEA-La Habana hoy?, ¿cuáles han sido los temas que más hemos investigado?, ¿qué resultados hemos alcanzado?

Con investigaciones de diferente alcance y profundidad, desde tesis de grado, de maestría hasta trabajos de doctorado o proyectos de investigación, el estudio del corpus PRESEEA-La Habana ha sido objeto de análisis en las siguientes líneas:

#### *De morfología y sintaxis:*

- Expresión del pronombre sujeto
- *Haber* impersonal
- *Queísmo-dequeísmo*
- El futuro morfológico y la perífrasis *ir a* + infinitivo
- Análisis de los adverbios espacio-temporales
- Diminutivos

#### *De pragmática:*

- Recursos de atenuación
- Recursos para expresar desacuerdo y actos disidentes.

#### *De lexicografía:*

- Variación fraseológica

#### *De aspectos fónicos y prosódicos:*

- Realización de la /d/ intervocálica
- Análisis sociolingüístico de la entonación en enunciados declarativos del español de Cuba.
- Estudio sociolingüístico de la entonación de Cuba.
- La entonación de los enunciados interrogativos de eco en el español de Cuba.

#### *Del discurso:*

- Descripción de partículas discursivas en hablantes cultos habaneros.

En sentido general, se ha trabajado con las guías metodológicas preparadas por los colegas del proyecto. Asimismo, en algunas ocasiones estos estudios han sido

realizados, cuando metodológicamente ha sido posible, en comparación con otros corpus como el de La Norma Culta de La Habana, cuyas muestras se encuentran publicadas desde 2010 (González Mafud *et al.*, 2010), o con otros de diferentes provincias del país, lo que ha permitido, en determinados temas, un examen diacrónico que ha aportado importantes valoraciones en cuanto a la evolución o posibles cambios lingüísticos en proceso de ciertos fenómenos y también en relación con la distribución dialectal de los temas de análisis.

Entre los años 2010 y 2014 se defendieron varias tesis de licenciatura (grado) en las que se trabajó ya con los materiales de PRESEEA-La Habana. Es el caso de los estudios sobre la variación del fonema /d/ (Sánchez, 2014), los usos de *haber* impersonal (Darlin Valdés, 2014), la variación fraseológica, con el análisis de las locuciones contextualizadas en una muestra de 36 informantes (Cruz Barrios, 2014) o las estructuras del desacuerdo y los actos disentivos (Blanco, 2014). Desafortunadamente, muchos de esos temas, que hoy forman parte de los importantes estudios coordinados, no han tenido continuidad en las investigaciones del PRESEEA-La Habana, a lo que sin dudas debemos atender con prontitud<sup>2</sup>.

Es así que han recibido particular atención en los últimos años el examen de aspectos fónicos y prosódicos, lo que ha dado lugar a publicaciones y participaciones en eventos que han contribuido no solo a describir patrones de la entonación de enunciados declarativos e interrogativos y a contrastarlos con otros similares de España, por ejemplo, sino a destacar la importancia y utilidad del proyecto. En particular, en cuanto al carácter del consonantismo, a las evidencias del carácter relajado del consonantismo posnuclear (asimilación de /l/ y /r/ distensivos en Santana Cepero (2013), se añaden ahora datos en muestras de PRESEEA de la propensión de la postdental sonora /d/ en posición intervocálica a su realización como una aproximante debilitada y a la elisión, en particular en estructuras participiales. Vinculado este comportamiento con variables sociolingüísticas, se advierte que favorecen su realización las situaciones informales para hablantes de cualquier edad, y específicamente para los de nivel escolar bajo. A pesar de que esta ha sido una tendencia constante en la historia del español, se comprueba que sigue siendo estigmatizada en los estratos educacionales más altos.

Sobre la entonación de la variedad cubana se han realizado recientes acercamientos a la descripción y análisis de enunciados declarativos de la muestra del PRESEEA, teniendo en cuenta variables sociolingüísticas (sexo y edad). Allí se comprueban dos configuraciones prosódicas del tonema en relación con su extensión y carácter paroxítono u oxítono, sin que se aprecien valores diferentes en relación con el español general (Pedrosa Ramírez *et al.*, 2019).

En la línea de los estudios sobre cuestiones morfológicas ha sido decisiva la contribución de la Dra. Marisela Pérez Rodríguez, bien como investigadora, bien como

---

<sup>2</sup> Estas tesis de licenciatura forman parte del repositorio de la Facultad de Artes y Letras de la Universidad de La Habana.

tutora de otros trabajos. Los resultados, en particular los relacionados con la expresión del pronombre sujeto, *haber* impersonal o los diminutivos, han permitido identificar e ilustrar de manera actualizada la presencia de algunos rasgos tenidos como esenciales en la variedad cubana del español.

Así, estudiados ciertos sufijos para comprobar si existía alguna peculiaridad (de presencia, ausencia o frecuencia) en muestras del PRESEEA, se aprecia en los colectivos que la selección *-al / -ar* se advierte influida por la disimilación motivada por la presencia de otra líquida en la base (*cañaverál, melonar, pedregal*). Se documenta el empleo de despectivos singulares *-ada* (*negrada*), *-erío* (*mujerío, chiquillerío*), así como se corrobora la productividad de *-ero* como indicador de ‘gran cantidad’ (*pulguero, churrero*) y *-ear* para la formación de verbos (*chapistear, cabecear*), con lo cual, incluso, se forman pares con distintos semas sobre la misma base (*relajar / relajear* ‘burlarse de alguien, diciéndole algo en tono de broma o chiste’ (‘alguien relaja a alguien’) [Cu: *bachatear, bonchear, changuear, guarachear, usar jaranas, vacilar*] (Haensch *et al.*, 2000; Pérez Rodríguez, 2022).

Según Pérez Rodríguez (2022), llama la atención que un verbo como *tutorear* sea, en el español de Cuba, la forma más extendida, sobre *tutorar* y *tutorizar*, ambas, sobre todo la última, prácticamente inexistentes en la isla. Sigue el patrón de *golpear, pestañear* y otros, en los que el sufijo está asociado a la repetición. No está, sin embargo, reconocida en el *Diccionario de la lengua española* (DLE) (Real Academia Española (RAE)), aunque sí lo está el verbo *apuñalear*, que se usa en varios países de América, entre ellos, Cuba, con el significado de ‘apuñalar’, y que también aparece en el *Atlas Lingüístico de Cuba* (ALCu) (Montero *et al.*, 2018) como opción preferente (86 % en el habla urbana y 85 % en la rural). Tanto *tutorear* como *apuñalear* están recogidos en el *Diccionario del español de Cuba* (Cárdenas Molina *et al.*, 2000). El primero de ellos significa ‘asesorar un profesor o investigador a un alumno universitario o asesorar en una tesis de doctorado un profesor o investigador con una categoría científica o docente superior a un profesor o investigador de menor categoría’ y el segundo tiene como acepciones ‘dar puñaladas a alguien’ o, en la forma reflexiva, *apuñalearse*, ‘quedarse una persona con una parte del dinero que debía repartir o devolver’, respectivamente. Además, aparece el adjetivo participial formado con el mismo patrón: *tutoreado*. Llama la atención este ejemplo, que constituye una reflexión metalingüística, evidencia de que pudiera ser interesante el análisis de las creencias y actitudes, sobre todo en hablantes cultos, hacia las formas con *-ear*.

Insiste Pérez Rodríguez (2022) en que al consultar el DLE, advertimos que la palabra *tutorear* no existe en nuestro idioma, por lo tanto, tampoco existe su participio *tutoreado*. Y que este fenómeno, aunque tiene una representatividad escasa en los corpus, caracteriza al español de Cuba. No es, sin embargo, exclusivo de la isla. Aparece documentado en otros países hispanoamericanos: en Venezuela (en PRESEEA) y en el CE/WD; en Nicaragua, en el CE/WD; en México, en el Corpus de referencia del español actual (CREA) y, curiosamente, se ha encontrado un ejemplo en prensa de Miami, Estados Unidos, que pertenece al propio CREA.

De otra parte, habida cuenta de que en los imaginarios sociales americanos se describe al hablante cubano como propenso al empleo frecuente del diminutivo, y en particular en las mujeres, es de interés observar ahora en muestras del PRESEEA (Díaz González, 2016), que la presencia del diminutivo –mayoritariamente con morfema *-it(o, a)* salvo que la base incluya una dental /t/ en el ataque de la sílaba anterior al sufijo diminutivo (*patica*)– se condiciona a los temas familiares y afectivos y a la interacción cercana, sin que se pueda afirmar categóricamente el predominio femenino de su uso.

Como resultado de la metodología conjunta para el análisis de la pluralización del verbo *haber* (Gómez Molina y Hernández Cabrera, 2021) en PRESEEA, existe una tesis de maestría que emplea esas muestras para compararlas con la Ciudad de México (Arteaga, 2017). La autora encontró que, en las muestras de la capital cubana, en cerca de la mitad (47,2 %) de las construcciones de *haber* seguido de grupo nominal plural o colectivo, este verbo aparece pluralizado y las variables lingüísticas que resultaron estadísticamente significativas fueron la colectividad y la animacidad de la forma nominal, mientras que ningún factor social lo fue. Desde el punto de vista morfológico, como en el caso de las muestras del habla culta y de otras variedades del español, el tiempo copretérito o pretérito imperfecto es el más productivo en la pluralización de *haber*.

Desde la perspectiva sintáctica, la preferencia por el futuro perifrástico sobre el sintético es también un tópico en la descripción del español de Cuba, que los datos del PRESEEA confirman, aunque dependiendo de tipos de verbo y de texto. Se observa que, mientras el perifrástico es la forma prácticamente categórica del futuro, los remanentes del sintético, con cierta presencia en lo escrito –de hablantes, sobre todo de estrato alto– adquieren valores estilísticos de modalización.

En el análisis del discurso, otro tema que ha dado lugar a diversos trabajos ha sido los marcadores discursivos (MD), refrendado en diversas publicaciones. Marialys Perdomo es, sin dudas, una de nuestras mejores especialistas en este ámbito y sus aportes han sido decisivos en los diferentes trabajos y publicaciones. Perdomo (2020) caracteriza, en su investigación doctoral<sup>3</sup>, el funcionamiento de cinco marcadores discursivos de uso frecuente en el habla de La Habana, identificados en las indagaciones previas (González y Perdomo, 2014 y 2015): *bueno*, *por ejemplo*, *¿no?*, *es decir* y *eh*. La muestra examinada en esta oportunidad pertenece al corpus del PRESEEA-La Habana y consta de 36 informantes. Desde un punto de vista cualitativo, se describe la entidad morfológica de los marcadores del discurso, sus propiedades sintagmáticas y sus funciones, y desde una perspectiva cuantitativa, se analiza la distribución del uso de los MD en los factores lingüísticos (la posición

---

<sup>3</sup> La tesis titulada *Contribución al estudio de los marcadores discursivos en un corpus oral en el español actual de Cuba* fue codirigida por las catedráticas María Antonia Martín Zorraquino y Ana María González Mafud, y defendida en la Universidad de Zaragoza en enero de 2020.

discursiva), sociales (la edad, el sexo y el grado de instrucción) y estilísticos (secuencias discursivas, módulos temáticos y fase de la entrevista). La autora concluye que las formas estudiadas se ajustan al estatuto de marcador discursivo, si bien algunas propiedades parecen ser más estables que otras. Determina que los más frecuentes son los que proceden de una categoría de base ya polisémica *-bueno-* y los que presentan un significado estrictamente procedimental *-eh-*, que pueden aparecer en una mayor diversidad de contextos lingüísticos; en cambio, los que tienen un menor número de ocurrencias manifiestan usos más específicos, relacionados con el interlocutor. El análisis cuantitativo refleja que la posición lingüística es el factor que más claramente contribuye al comportamiento funcional de los marcadores discursivos, a diferencia de los factores sociales y estilísticos, que no muestran una relación significativa con el uso variable de los marcadores estudiados.

De este modo los materiales orales del PRESEEA, sostiene Perdomo (2022), permitieron corroborar que estos elementos discursivos tienen vitalidad en la conversación, fundamentalmente como marcadores del discurso, pues su uso en el plano discursivo supera, en número, su empleo como categorías gramaticales dentro de la estructura proposicional. En las entrevistas, *bueno* se utiliza también como adjetivo, aunque en menor medida; ¿no? tiene escasas ocurrencias como enunciado interrogativo y *por ejemplo, es decir* y *eh* se manifiestan exclusivamente como marcadores discursivos, pues no se documenta su uso como sintagma preposicional, ni como construcción verbal no fijada, ni como interjección, respectivamente. Los signos más frecuentes son *eh* y *bueno*, y *es decir* ha sido el menos empleado por los hablantes habaneros. Estos resultados se deben al tipo significativo de los dos primeros –la cierta indeterminación semántica de la interjección *eh* y la polifuncionalidad del adjetivo de base *bueno* y la desarrollada también como marcador discursivo– y a la progresiva sustitución de *es decir* por una variante menos formal (*o sea*).

A pesar de que en esta oportunidad, los marcadores investigados en la muestra del PRESEEA-La Habana pertenecen al sistema de la lengua española y son de uso general (en estos materiales, por la modalidad de la entrevista parecen aflorar menos variantes dialectales), la descripción de las propiedades gramaticales, semánticas y del comportamiento funcional de estos elementos corrobora que se ajustan a la caracterización general de los marcadores del discurso. Ahora bien, en la conversación algunas propiedades parecen más estables que otras y dependen, sobre todo, del tipo significativo y del estatuto gramatical de las clases de palabras a partir de las cuales se origina cada marcador. Los cinco signos analizados presentan un significado de procesamiento, con instrucciones que reflejan rasgos del significado conceptual de la palabra de base.

Así, se establece un camino en la investigación de estos elementos, que podrá ser completado con análisis contrastivos con los resultados que han obtenido los investigadores en otras comunidades de habla. De momento, hemos iniciado el estudio de este tema en La Habana, en los materiales de PRESEEA, en los que, por ejemplo, Santiago de Chile tiene larga data, con los resultados obtenidos fundamentalmente por Abelardo San Martín (2011), para los reformuladores, marcadores del control del contacto, entre otros.

#### 4. CONCLUSIONES

Pero quizás la conclusión más importante que podamos derivar del valor de PRESEEA-La Habana es su contribución a la descripción del español de Cuba, en particular de La Habana y la realización de estudios comparativos con otras variedades del español. Así lo demuestran, por ejemplo, por una parte, la reciente publicación del libro: *El español en Cuba: aportes a su descripción en el siglo XXI* (González Mafud y Perdomo Carmona, 2022) donde aparecen siete artículos, incluidos, entre otros, amplios panoramas de los estudios fónicos, morfológicos y sintácticos de la variedad cubana del español en el siglo XXI, en los que se puede identificar el trabajo con el corpus y los datos del PRESEEA-La Habana. Asimismo, la solicitud del corpus PRESEEA-La Habana, o de algunas muestras de este por colegas de Chile, México y la propia España para realizar trabajos contrastivos, pone de relieve también la importancia del trabajo realizado.

La caracterización de variables contribuye a la descripción de ciudades. Es tarea pendiente reunir todas las investigaciones sobre PRESEEA-La Habana, y estoy segura de que podríamos tener un hermoso libro sobre la realidad y el imaginario lingüísticos de la capital cubana.

#### 5. CONSIDERACIONES FINALES

Y para los que se hayan quedado intrigados con el otro tema que le comenté a Paco Moreno, era solo la posibilidad de visitar la Universidad de Alcalá y establecer contactos con sus profesores. Esa visita al Cervantes y la probada generosidad de Paco Moreno, me han permitido disfrutar de entrañables relaciones humanas y de trabajo con Ana María Cestero, Isabel Molina, Florentino Paredes y Manuel Martí, a quienes hoy también les agradezco públicamente el privilegio de su amistad y la voluntad de colaboración y apoyo al desarrollo y difusión de los estudios lingüísticos en Cuba.

#### REFERENCIAS

- Alonso, Gladys y Ángel Luis Fernández (1977). *Antología de lingüística cubana*. La Habana: Editorial de Ciencias Sociales.
- Arteaga Santos, Sandra (2017). *La (im)personalización del verbo haber en La Habana y en la Ciudad de México* [Tesis de maestría]. Universidad Autónoma de Querétaro.
- Blanco, Yenisleydis (2014). *Desacuerdo y actos disentivos en las muestras de PRESEEA-La Habana*, [Tesis de licenciatura], Universidad de La Habana.
- Cárdenas Molina, Gisela, Antonia María Tristán Pérez y Reinhold Werner (2000). *Diccionario del español de Cuba*. Madrid: Gredos.

- Cruz Barrios, Jessica (2014). *Variación fraseológica en PRESEEA. Estudio de las locuciones contextualizadas en una muestra de 36 informantes*, [Tesis de licenciatura], Universidad de La Habana.
- de Armas, Juan Ignacio (1882). *Orijenes [sic] del lenguaje criollo*. Imprenta de la Viuda de Soler.
- Díaz González, Ailín (2016). *Algunas noticas de pragmagramática: diminutivos en el español habanero (Análisis de una muestra de PREESEA- La Habana)*, [Tesis de licenciatura], Universidad de La Habana.
- González Mafud, Ana María, Marialys Perdomo Carmona, Marisela Pérez Rodríguez y Gretel Gutiérrez Fuentes (2010). *Muestras del habla culta de La Habana*, La Habana: Editorial UH.
- González Mafud, Ana María y Marialys Perdomo Carmona (2014). «Marcadores del discurso de La Habana», En A. Valencia (coord.), *Marcadores discursivos en la norma culta hispánica: 1964-2014*. Cuaderno ALFAL, 5: 107-139.
- González Mafud, Ana María y Marialys Perdomo Carmona (2015). «Marcadores discursivos de La Habana», En A. Valencia y A. Vigueras (coord.), *Más sobre marcadores hispánicos: Usos de España y América en el Corpus de estudio de la norma culta*, Universidad Nacional Autónoma de México, pp. 123-150.
- González Mafud, Ana María y Marialys Perdomo Carmona (coords.) (2022). *El español en Cuba: aportes a su descripción en el siglo XXI*, Santiago de Compostela: Servizo de publicacións e intercambio.
- Haensch, Günther; Reinhold Werner, Gisela Cárdenas Molina y Antonia María Tristán Pérez. (2000). *Diccionario del español de Cuba: español de Cuba-español de España*. Madrid: Gredos.
- Gómez Molina, José Ramón y Clara E. Hernández Cabrera. (2021). *Guía PRESEEA de estudio del verbo impersonal haber*. Disponible en: [https://preseea.uah.es/sites/default/files/2022-02/Gu%C3%ADa%20PRESEEA%20de%20estudio%20del%20verbo%20impersonal%20haber\\_G%C3%B3mez%20Molina%20y%20Hern%C3%A1ndez%20Cabrera%20%282021%29.pdf](https://preseea.uah.es/sites/default/files/2022-02/Gu%C3%ADa%20PRESEEA%20de%20estudio%20del%20verbo%20impersonal%20haber_G%C3%B3mez%20Molina%20y%20Hern%C3%A1ndez%20Cabrera%20%282021%29.pdf)
- Montero Bernal, Lourdes E., et al. (2018). «Atlas Lingüístico de Cuba». *Anales de la Academia de Ciencias de Cuba*, 8(1): 1-14. Disponible en: <https://revistaccuba.sld.cu/index.php/revacc/article/view/413>
- Perdomo Carmona, Marialys (2020). Contribución al estudio de los marcadores discursivos en un corpus del español actual de Cuba [Tesis doctoral], Zaragoza: Universidad de Zaragoza.
- Perdomo Carmona, Marialys (2022). «Los estudios sobre los marcadores discursivos en el español de Cuba», En A. M. González Mafud y M. Perdomo Carmona (coords.). *El español en Cuba. Aportes a su descripción en el siglo XXI*. Universidad de Santiago de Compostela: Servizo de publicacións e intercambio, pp. 321-334.
- Pedrosa Ramírez, Adriana; Madeleyne Bermúdez Sánchez y Raquel García Riverón. (2019). «Análisis sociolingüístico de la entonación en enunciados declarativos

- del español de Cuba y España: primer acercamiento al corpus PRESEEA», *Cuaderno de Letras*, (33): 105-122.
- Pérez Rodríguez, Marisela. (2022). Morfología del español de Cuba, En A. M. González Mafud y M. Perdomo Carmona (coords.). *El español en Cuba. Aportes a su descripción en el siglo XXI*. Universidad de Santiago de Compostela: Servizo de publicacións e intercambio, pp.77-98.
- Sánchez, Claudia. (2014). *Variación del fonema /d/ en posición intervocálica en las muestras de PRESEEA-La Habana*, [Tesis de licenciatura], Universidad de La Habana.
- San Martín Núñez, Abelardo. (2011). «Los marcadores interrogativos de control de contacto en el corpus PRESEEA de Santiago de Chile», *Boletín de filología*, 46(2): 135-166.
- Santana Cepero, Elizabeth. (2013). «Patrones fónicos de la asimilación consonántica. Análisis instrumental de /l/ y /r/ distensivos en el español de La Habana», *Congreso Internacional Lingüística 2013*. La Habana: Instituto de Literatura y Lingüística.
- Valdés, Darlin. (2014). *Usos de 'haber' impersonal en las muestras de PRESEEA-La Habana*, [Tesis de licenciatura], Universidad de La Habana.



## ELISIÓN DE /d/ INTERVOCÁLICA: ASPECTOS DE UNA TENDENCIA HISPÁNICA

María Claudia González-Rátiva  
*Universidad de Antioquia*

### 1. INTRODUCCIÓN

El interés por la convivencia de rasgos lingüísticos, la alternancia de formas en el habla y la combinación de posibilidades discursivas inter e intradialectales devino en estudios cuantitativos y probabilísticos que en conjunto forman el trabajo sociolingüístico de la variación cambiante en la concreción de la lengua. La corriente hispanista cuenta con una tradición de este tipo de estudios que buscan frecuencias relativas, probabilidades de aparición y significación social para diferentes segmentos lingüísticos de la integración discursiva.

Los aspectos que condicionan a una variable lingüística se han buscado en factores internos y externos, aquellos que van desde los aspectos propiamente estructurales y lingüísticos hasta los idiolectales, cognitivos, estilísticos, sociales, diafásicos, contextuales y de interacción. De hecho, el habla es susceptible en extremo a la variación y al cambio; la lengua está en cambio constante y en la variación está su riqueza. Por lo tanto, cuestiones o considerandos disímiles pueden favorecer en algún sentido la variabilidad de los fenómenos lingüísticos.

Entendemos, entonces, que la variación sociolingüística es el resultado de la expresividad y la identidad de los seres humanos y de los grupos en la historia. El plano fonológico de la lengua es tal vez el más accesible a las investigaciones variacionistas, dada la facilidad de la aparición de las formas en el plano superficial del habla. En dicho plano, la variable hispana /d/ es uno de los elementos más investigados, dada su frecuencia de aparición en la realidad fónica, su distribución o las diversas posiciones que ocupa en los lexemas, su integración a los morfemas lingüísticos y las variantes y sus posibilidades de afectación. Una de las variantes que aparece con

mayor recurrencia en las variedades lingüísticas es la aproximante [ɸ]; otra de sus variantes, que la sucede en el habla, es el cero fonético o la elisión del sonido; las tres variantes en cuestión se ven afectadas por la posición intervocálica y se encuentran estratificadas extralingüísticamente de diferente manera para las variedades estudiadas. Además, las tres fonías [oclusiva, aproximante y elidida] son tratadas en muchos de los textos revisados como las opciones de mantenimiento y de supresión, las dos primeras se relacionan con el conservadurismo y la tercera con las hablas innovadoras (Cedergren, 1973; López Morales, 1983; Samper, 1990; Molina y Paredes, 2014; Malaver y Samper, 2016; Nitola, 2018).

Este texto presenta un breve recorrido por el fenómeno de la elisión como representante de la innovación y de la espirantización total de /d/ intervocálica, incluidos aquellos del proyecto PRESEEA, con el fin de ejemplificar este cambio en marcha estable y su evidente dispersión en el ámbito hispanohablante. Pretende, además, exponer que los estudios sociolingüísticos como los aquí presentados pueden enmarcarse bien dentro de una gramática basada en el uso, cuyos resultados fundamentan la idea de que la similitud de las formas y su frecuencia de aparición incrementan las probabilidades de aparición de los elementos innovadores en el habla.

## 2. SOBRE LA APROXIMANTE [ɸ]

El fonema /d/ en el sistema fonológico del español es oclusivo dental sonoro y presenta dos alófonos, una variante oclusiva [d̪] y una variante aproximante [ɸ]. En la mayoría de los casos (V-V), la distinción se establece entre la oclusiva y la aproximante; por ello, el español pertenece a las lenguas en las cuales el contraste sonoro se da con la aproximante y no con la oclusiva; en este sentido, el español pertenece al grupo con la menor proporción, 6.7% (Maddieson, 2013).

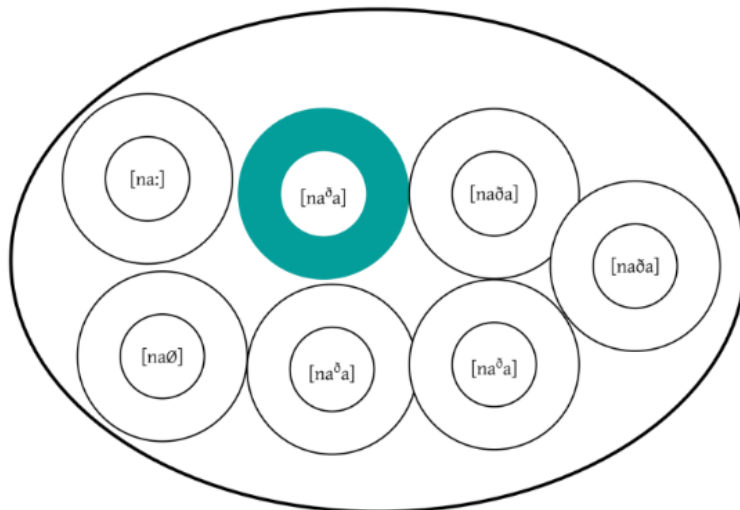
La distribución complementaria estructuralista entre [d̪] y [ɸ] en el sistema hispánico se perfila a partir de un proceso fonológico de espirantización en el habla. Harris (1969) arguye que las obstruyentes oclusivas sordas se transforman en continuas mediante la regla de espirantización, una alteración contextual del sonido; desde el punto de vista histórico, un tipo de lenición (Trask, 1996: 332). El fonema oclusivo dental sonoro se transforma contextualmente en un sonido aproximante. Para la mayoría de variedades de habla, este sonido en contexto intervocálico, principalmente, se modifica para concretizarse como alófono aproximante.

La similitud acústica y espectrográfica con las vocales dio lugar a la noción de aproximante. Ladefoged (1975) describió estos sonidos como aquellos generados por aproximación y carentes de corriente de aire turbulento. En el plano hispánico, Martínez Celdrán (1984, 1996, 1998) agregó a estos fonos características de mayor cierre que una vocal y más apertura que una nasal, lateral o semiconsonante. Serán aproximantes en español las obstruyentes no fricativas [β, ɸ, γ], con esa transcripción

del Alfabeto Fonético Internacional<sup>1</sup> detallada y, desde el punto de vista acústico, ausentes de ruido, con armonicidad en forma de estrías y un decremento de intensidad comparado con las vocales adyacentes, así como el menor ennegrecimiento de la estructura formántica. Ya en 2013, Martínez Celdrán las definió como: «segmentos que, poseyendo cierto grado de constricción, carecen de la precisión articulatoria requerida para producir una corriente turbulenta de aire, bien sea por la falta de tensión en los órganos articuladores, bien sea porque el tracto vocal no está suficientemente constreñido o ambas cosas conjuntamente» (15).

Martínez Celdrán clasifica a los sonidos aproximantes como sonidos cerrados, abiertos y vocálicos. La naturaleza acústica del sonido aproximante es una gradación que va desde el decremento de su intensidad y el asomo de resonancias armónicas, hasta la completa aparición de estructura formántica con menor intensidad que en sus acompañantes vocálicas adyacentes. El sonido aproximante se muestra gradual y como un proceso subfonémico (Díaz-Campos, Geeslin y Gurzynski-Weiss, 2021); es decir, existirían diversos grados de debilitamiento que se pueden explicar fonética y sociolingüísticamente en cada variedad, grupo o corpus, tal como se muestra en la imagen 1. Allí se contrasta con los fenómenos de alargamiento y elisión. Nótese que la aproximante vocalizada se transcribe como superíndice en este texto, sin el diacrítico referido en menciones anteriores en este y las demás variantes.

**IMAGEN 1. EJEMPLO DE LA CATEGORIZACIÓN PARA LAS VARIANTES APROXIMANTE Y ELIDIDA EN UN GRUPO ESPECÍFICO**



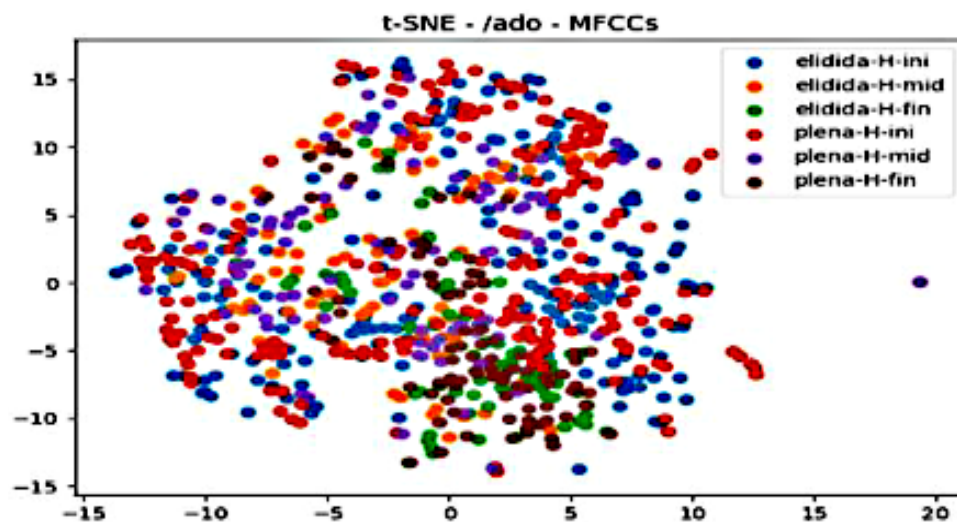
Fuente: Díaz-Campos, Geeslin y Gurzynski-Weiss, 2021: 105.

<sup>1</sup> Consultable en <https://www.internationalphoneticassociation.org/>

Aparece en el habla el fenómeno de la elisión como una variante posible al nivel subfonémico y para el cual se ha establecido su favorabilidad desde el plano léxico y el plano morfológico, especialmente este último al aparecer de manera redundante en la estructura oracional. Desde el punto de vista gradacional, se amplía la posibilidad de un abanico de fonías que van desde la oclusividad hasta el cero fonético, pasando por las clasificaciones obstruyentes no fricativas o aproximantes, cerrada, abierta y vocálica.

Para la ciudad de Medellín, el estudio de González-Rátiva, Orozco-Arroyave, Ospina, y Chaparro (2019) se realizó a partir de una prueba experimental de 108 audios segmentados en *Praat* (Boersma y Weenink, 2009) de palabras terminadas en *-ado*, llevados a automatización con nueve datos acústicos por segmento y presentados con la técnica t-SNE12 (Van der Maaten, 2019), para visualizar el resultado bidimensionalmente, como se presenta en la Imagen 2. Allí se estableció, con una prueba experimental en palabras terminadas en *-ado*, que la naturaleza acústica de la /d/ era mayoritariamente una aproximante vocalizada, dado que los resultados arrojaron poca diferenciación entre los sonidos vocálicos adyacentes y los segmentos analizados para la /d/, lo que mostró que no hubo tendencia de agrupamiento hacia las vocales adyacentes *a* y *o*.

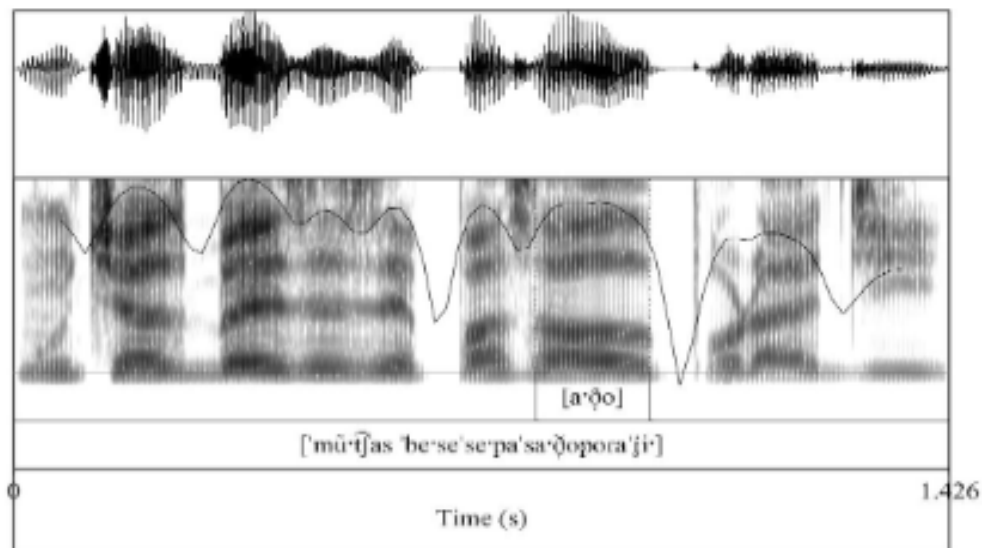
IMAGEN 2



Fuente: González-Rátiva *et al.*, 2019: 527. Prueba masculina para el habla de Medellín.

Es decir, para el corpus analizado, hasta en los casos de pronunciación cuidada, ni el descenso de la intensidad relativa permite determinar con precisión los límites de adyacencia con las vocales del segmento *-ado*, tal como se muestra en la Imagen 3. Por ello, se estaría ante la realización de una variante aproximante vocalizada; un mantenimiento de difícil segmentación.

IMAGEN 3



Fuente: González-Rátiva *et al.*, 2019: 527. Aproximante vocálica masculina para el habla de Medellín.

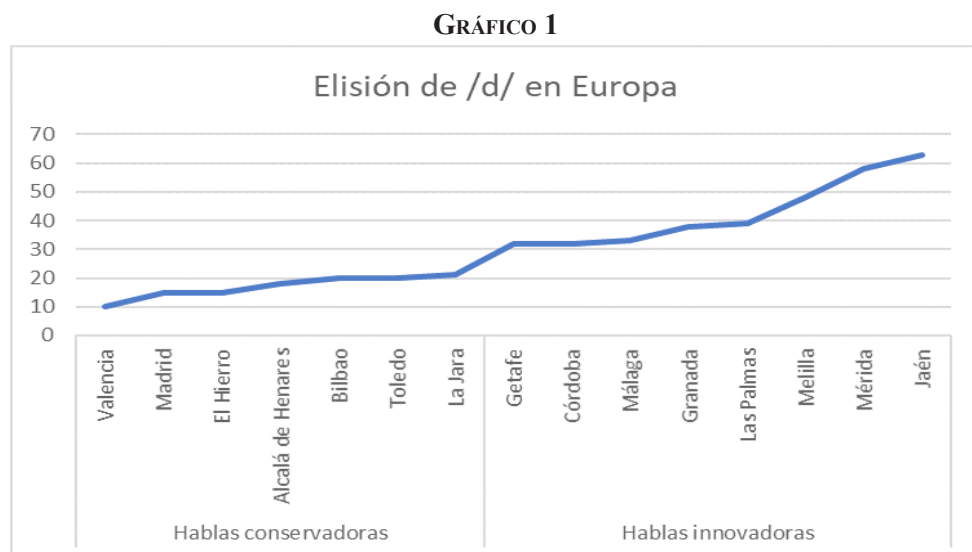
### 3. SOBRE LA ELISIÓN DE /d/. [d̪] > [∅]

El proceso fonológico de espirantización permite dar cuenta de la disminución progresiva de la energía vocálica y articuladora del sonido hasta su completa eliminación o elisión fónica. La presencia del cero fonético para el sistema fonológico español ha sido reportada desde la historia de la lengua, un aporte por demás valioso para los estudios sincrónicos. Así, Frago (1999: 258-259), registra la caída de la /d/ intervocálica en varios tipos de textos. Uno de ellos, específico a un exvoto de 1700 cordobés en la forma [ˈao] por *-ado*, con la que, junto a otros fenómenos fonéticos, explica que la manera de documentar dichas alteraciones regionalmente con un ejemplario vasto para el andaluz no cambia durante los siglos XVIII y XIX, «una situación dialectal ya consolidada y sin sustanciales cambios durante esos doscientos años» (Frago, 1999: 258-259) y que está en constante difusión a ambos lados del Atlántico.

Reviste de gran importancia este registro ya que nos indica que el fenómeno de espirantización de /d/, especialmente intervocálico, no es un fenómeno americano, ya que aparece el documento en territorio ibérico; no es un fenómeno específico a una región, puesto que demuestra que se encuentra en varios dialectos de las hablas andaluzas; y no es un fenómeno reciente ni efímero en la historia de la lengua, pues su aparición, de acuerdo con dicha fuente se puede documentar ya en el XVIII.

Desde el punto de vista de los estudios sociolingüísticos con corpus oral, se muestra lo atinente a las consideraciones hechas para la variante en cuestión y los resultados. En esta parte, no se incluyen los datos del PRESEEA. De acuerdo con Blas Arroyo (2008: 48) la influencia de los factores estructurales puede diferir no solo entre comunidades de habla diferentes, sino también entre sociolectos diversos dentro de una misma comunidad.

El debilitamiento de la /d/ intervocálica, su elisión, está reportado para un buen número de comunidades de habla (Blas Arroyo, 2008: 55). La elisión de la /d/ está estudiada principalmente en la terminación *-ado* (Gómez Molina y Gómez Devís, 2010), contexto que estructural y frecuentemente es más susceptible de debilitamiento, dada la redundancia de significado gramatical a la que apunta la presencia del fonema. Algunas de las hablas hispánicas europeas donde se reporta debilitamiento intervocálico de /d/ se muestran en el siguiente Gráfico 1<sup>1</sup>:

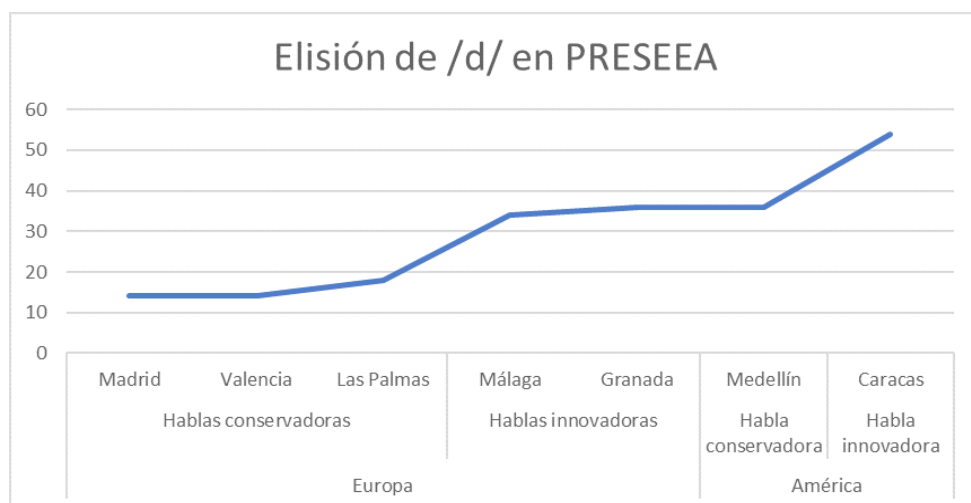


Fuente: elaboración propia con base en Fernández de Molina, 2019.

Para América se muestra una distribución diversa y no tan acuciosa en términos de datos en RAE-ASALE (2011, 4.7e: 146), ya que presenta diferentes soluciones en varios lugares, «según los segmentos fónicos que entren en contacto». Son clásicos los estudios realizados para Panamá (Cedergren, 1973), Caracas por D’Introno y Sosa (1986), San Juan (López Morales, 1983); y Lima (Caravedo, 1990), solo por mencionar algunos. También Martín Butragueño (2016: 19) sustenta la variabilidad a partir de trabajos realizados para México y propone en su reflexión teórica el estudio minucioso de los segmentos fónicos a partir de conceptos como la indización, precisamente para no caer en la «supuesta estabilidad».

El surgimiento del Proyecto para el estudio sociolingüístico del español de España y América (PRESEEA, 2014) construyó y desarrolló un proyecto coordinado que permitió mejorar la presentación y el contraste de los datos. Surgen así varios trabajos que visibilizan los resultados del debilitamiento de la /d/. Coordinados por Samper Padilla, un grupo de trabajo realizó la metodología para el estudio de la /d/ en el ámbito hispánico, una propuesta variacionista que abarcaba todos los factores que podía implicar el fenómeno. Poco a poco las ciudades se fueron decantando de acuerdo con los resultados parciales, hasta dar cuenta de la tendencia del debilitamiento de /d/. Los resultados de estos trabajos se exponen en el Gráfico 2:

GRÁFICO 2



Fuente: Elaboración propia a partir de Malaver, I. y J.A. Samper Padilla, 2016: 330 y González-Rátiva, 2023.

Estos resultados nos permiten observar que la elisión avanza de los dialectos conservadores (Madrid, Valencia y Las Palmas) a los innovadores (Málaga y Granada) en España y que el contraste con América, del habla conservadora de Medellín al habla innovadora de Caracas, no es en términos de datos puntuales sino de consideraciones igualmente amplias y con un gran foco. Caracas y Medellín muestran un avance enorme frente a España, pero la innovación caraqueña alcanza la mayor frecuencia de debilitamiento, como se espera de una ciudad del espacio caribeño. Hablamos aquí del *continuum* de gradación previsible variación cambio, expuesto por Chela-Flores (2022).

Esta relación muestra un comportamiento similar para ambos lados del atlántico; tanto en el espacio ibérico, como en el espacio americano, se reproduce una tendencia que va en gradación desde las variantes que muestran un mayor comportamiento

de retención y con avances de elisión, hasta el polo en que se muestran las variantes que reproducen una situación de mayor elisión sobre las fonías aproximantes o de retención. Es importante destacar que la tendencia es convergente hacia las formas innovadoras, lo que apunta a una posible perspectiva diacrónica o la tendencia diacrónica, la del espacio románico occidental, hipótesis que valdría la pena indagar en profundidad.

Otro aspecto de gran interés en el fenómeno de la elisión es su tímida tendencia hacia la lenición. Hasta ahora, para la mayoría de las ciudades en estudio, la variación de la elisión tiene explicaciones sociolingüísticas, pero no de significado sociolingüístico como los casos del español meridional «cantado» vs. «canta». Este caso demuestra la pervivencia del fenómeno y la búsqueda de esas significaciones en América, donde el fenómeno de elisión alcanza altos niveles de avance.

En este sentido, estamos de acuerdo con Bybee (2004) en que, al haber un aumento de la frecuencia de uso, se demuestra que ella en sí coadyuva al resultado del proceso, ya que «la repetición tiene ciertos efectos en las representaciones neuromotoras y cognitivas» (253). Además, se arguye también que «la repetición de los actos comunicativos lleva a la automatización y la reducción de forma, a la habituación y generalización del significado y también al convencionalismo de las inferencias pragmáticas (255-6). Sin embargo, su interpretación deberá estar relacionada con las otras rutas lingüísticas para la determinación de su proceso radical/conservador o innovador.

#### 4. NOTAS FINALES

Sigue siendo para la lingüística y para la sociolingüística el establecimiento de modelos de análisis fonológicos claros y relevantes, ya que estos constituyen un modelo a seguir en la disciplina. Con ello, es importante aclarar que las unidades de análisis, en este caso los fonemas y el sistema fonológico del español, sigue en la búsqueda de determinar la clase de sonidos y las relaciones que se pueden establecer entre ellos. ¿Es alófono un sonido aproximante –sonante– de un sonido oclusivo –obstruyente–? ¿La ausencia del fono esperado en el habla es una variante más de un fonema o es una sustitución fonológica de otra categoría? ¿Si las variantes del habla superan a las alofónicas, estamos ya en el debate de la trascendencia de categorías contrastivas a las categorías gradientes y de valor sociolingüístico? Son inquietudes que se desprenden de la aproximación al ejemplario sobre la /d/.

La metodología coordinada PRESEEA (2014) generó un avance en la comparación de fenómenos lingüísticos del habla y de las entrevistas dirigidas. Sin embargo, aún hay datos que no se pueden contrastar, ya que no se toman en cuenta la totalidad de los factores o solo se establecen dos de tres variantes, por ejemplo. Estudios que den cuenta del sentido social de las variantes deben seguir realizándose, establecer la diferencia diatópica y la distancia perceptual de las variantes. Por ello, se deben

tomar estos contrastes con la relativización necesaria para una mejor comprensión del fenómeno continuo y el interés que despierta en la geografía y la sociedad.

Fernández de Molina (2019: 42) arguye que «el desarrollo del fenómeno ha ido moderándose a lo largo de los últimos siglos debido a factores lingüísticos, sociales y estilísticos». Parece más bien que la gradación de este cambio en marcha constante es netamente diferenciado y gradual de acuerdo a las tendencias conservadoras o innovadoras de las comunidades; es decir, está relacionado con el grado de innovación consonántica generalizada en la comunidad.

Para seguir adelantando estudios sobre el fenómeno de espirantización y elisión es importante, entre otros aspectos: (1) indagar por la frecuencia léxica del corpus. La palabra en la que ocurre un sonido tiene un papel importante debido a que los fenómenos de variación y cambio parecen ser más comunes en las palabras de alta frecuencia donde se observa la automatización de los gestos articulatorios. (2) basar el estudio en una aproximación acústica que permita determinar cortes específicos para la intensidad del fono en el caso de la espirantización y ausencia del decremento de intensidad para la corroboración de la elisión. Y (3) llevar los resultados a los aspectos teóricos de mayor abstracción para la determinación de la tendencia hispánica, como el concepto de indización, esbozado en Martín Butragueño (2016).

## REFERENCIAS

- Blas Arroyo, José L. (2008). *Sociolingüística del español. Desarrollos y perspectivas en el estudio de la lengua española en contexto social*. Madrid: Cátedra.
- Bybee, Joan (2004). «Los mecanismos de cambio como universales lingüísticos». En R. Mairal y J. Gil (eds.), *En torno a los universales lingüísticos*. España: Ediciones AKAL, pp. 245-264.
- Boersma, Paul y David J. M. Weenink (2009). *Praat. Doing phonetics by computer (Versión 6.0.17)* [Programa para computadora]. Disponible en <http://www.fon.hum.uva.nl/praat/>.
- Caravedo, Rocío. (1990). *Sociolingüística del español de Lima*. Lima: Pontificia Universidad Católica del Perú
- Cedergren, Henrieta (1973). «La elisión de la /d/: un ensayo de comparación dialectal». *Boletín de la Academia Puertorriqueña de la Lengua Española*, 7: 12-29.
- Chela-Flores, Godsuno (2022). «Reinterpretación del continuum variación-cambio y conceptualización de la situación especial del universo hispánico». En A. M. Ávila Muñoz, M. Vida Castro y J. L. Jiménez Ruiz (eds.), *Monográfico, Estudios de variación lingüística: homenaje a Juan Andrés Villena Ponsoda*. Linred, XIX. DOI: <https://doi.org/10.37536/linred.2022.XIX.1873>
- D'Introno, Francesco y Juan M. Sosa (1986). «Elisión de la /d/ en el español de Caracas: Aspectos sociolingüísticos e implicaciones teóricas». En R. Núñez

- Cedeño, I. Páez Urdaneta y J. Guitart (eds.), *Estudios sobre la fonología del español del Caribe*. Caracas: La Casa de Bello, pp. 135-165
- Díaz-Campos, Manuel, Kimberly L. Geeslin y Laura Gurzynski-Weiss (2021). *Introducción y aplicaciones contextualizadas a la lingüística hispánica*. Nueva Jersey: John Wiley & Sons.
- Fernández de Molina Ortés, Elena (2019). «Factores lingüísticos y extralingüísticos influyentes en la elisión de la /d/ en posición intervocálica en la comarca de Mérida (España)». *Philologica Canariensis*, 25: 40-65.
- Frago García, Juan Antonio (1999). *Historia del Español de América*. Madrid: Gredos.
- Gómez Molina, José Ramón y María Begoña Gómez Devís (2010). «Mantenimiento y elisión de la /d/ intervocálica en el español de Valencia». *Verba. Anuario Galego de Filoloxía.*, 37: 89-122.
- González-Rátiva, María Claudia, Juan Rafael Orozco-Arroyave, Liliana Estefanía Ospina Grialdo y Erika Patricia Chaparro Manrique (2019). «Aproximante [ð̣] en contexto -ado en el habla de Medellín: prueba experimental para la identificación automática de variantes alofónicas y su caracterización acústica». *Lenguaje*, 47(2S): 514-536. DOI: <https://doi.org/10.25100/lenguaje.v47i3.7560>
- González-Rátiva, María Claudia (2023). «Estudio sociolingüístico de la /d/ en el contexto -ado en el Corpus PRESEEA-Medellín». En: M. Dieck, J.M. Molina y M.C. González-Rátiva (eds.), *Estudios sociolingüísticos en Colombia. Investigaciones a partir del Corpus PRESEEA-Medellín y sobre lenguas minoritarias. Cap. I*. Medellín: Editorial FOCO, Facultad de Comunicaciones y Filología, Universidad de Antioquia, pp. 15-42.
- Harris, James W. (1969). *Spanish phonology*. M.I.T. Press. Disponible en: <https://mitpress.mit.edu/books/spanish-phonology>.
- Ladefoged, Peter (1975). *A Course in Phonetics*. Nueva York: Harcourt Brace Jovanovich.
- López Morales, Humberto (1983). *Estratificación social del español de San Juan de Puerto Rico*. México: UNAM.
- Malaver, Irania y José Antonio Samper Padilla (2016). «Estudio de la /d/ intervocálica en los corpus PRESEEA». *Boletín de Filología*, LI (2): 325-345.
- Martín Butragueño, Pedro (2016). «El Español visto desde México». En R. Barriga Villanueva, y P. Martín Butragueño (dirs.). *Historia sociolingüística de México, vol. 5: Nuevas visitas al pasado y al presente*. México: El Colegio de México, pp. 3133-3178. Consultado en: (99+) «El español visto desde México» (draft 3). | Pedro Martín Butragueño - Academia.edu
- Martínez Celdrán, Eugenio (1984). «Cantidad e intensidad en los sonidos obstruyentes del castellano: hacia una caracterización acústica de los sonidos aproximantes». *Estudios de Fonética Experimental*, 1: 71-129. Disponible en: <https://www.raco.cat/index.php/EFE/article/view/144191>.
- Martínez Celdrán, Eugenio (1996). *El sonido en la comunicación humana. Introducción a la fonética*. Barcelona: Octaedro.

- Martínez Celdrán, Eugenio (1998). *Análisis espectrográfico de los sonidos del habla*. Barcelona: Ariel.
- Martínez Celdrán, Eugenio (2013). «Caracterización acústica de las aproximantes espirantes en español». *Estudios de Fonética Experimental*, XXII: 11-35.
- Maddieson, Ian (2013). «Voicing in Plosives and Fricatives». En M. S. Dryer y M. Haspelmath (eds.) *WALS Online* (v2020.3). [Data set]. DOI: <https://doi.org/10.5281/zenodo.7385533> Disponible en: <http://wals.info/chapter/4>
- Molina Martos, Isabel y Florentino Paredes García (2014). «Sociolingüística de la elisión de la dental –/d/- en Madrid (distrito de Salamanca)». *Cuadernos de Lingüística*, 2: 55-114.
- Nitola Torres, Jholman Fabián (2018). *Variación sociofonética en la alternancia [d̥]~[ø] intervocálica del habla de Tunja*. [Trabajo de grado]. Colombia: UPTC. Disponible en: [https://repositorio.uptc.edu.co/bitstream/001/2349/1/TGT\\_999.pdf](https://repositorio.uptc.edu.co/bitstream/001/2349/1/TGT_999.pdf)
- PRESEEA (2014-): Corpus del Proyecto para el estudio sociolingüístico del español de España y de América. Alcalá de Henares: Universidad de Alcalá. [<http://preseea.uah.es>]. Consultado: [12 de marzo de 2023]
- Trask, Robert Lawrence (1996). *A Dictionary of Phonetics and Phonology*. Nueva York: Routledge.
- Samper Padilla, José Antonio (1990). *Estudio sociolingüístico del español de Las Palmas de Gran Canaria*. Las Palmas de Gran Canaria: La Caja de Canarias.
- RAE-ASALE (2011) *Nueva Gramática de la Lengua Española*. Madrid: Espasa.
- Van der Maaten, Laurens (2019). t-SNE. Recuperado de <https://lvdmaaten.github.io/tsne/>.



## HACIA UN MODELO DE LA CIUDAD LINGÜÍSTICA: ESTRUCTURACIÓN SOCIOLINGÜÍSTICA BÁSICA DE CIUDAD DE MÉXICO A INICIOS DEL S. XXI

Pedro Martín Butragueño  
*El Colegio de México*

### 1. INTRODUCCIÓN: EL CONCEPTO DE CIUDAD LINGÜÍSTICA

La estructura sociolingüística de una ciudad no puede entenderse estudiando variables lingüísticas aisladas<sup>1</sup>. Un conjunto de variables, en cambio, confirma –o no– los posibles patrones, al tiempo que permite perfilar el carácter de un grupo sociolingüístico específico. Por otra parte, buena parte de los enclaves urbanos han crecido de tal manera en las últimas décadas que es difícil sostener que la Ciudad de México que tenemos ante los ojos, en un momento dado, sea la misma que hace treinta, cincuenta o setenta años (para este problema, véanse las reflexiones de Labov, 1994, §1.2: 18-25, especialmente pp. 20-21), lo que cimbra mucho de lo que suponemos o asumimos metodológicamente al trabajar con el tiempo aparente (Labov, 1994, cap. 3: 43-72; Cukor-Avila y Bailey, 2018).

Al hacer referencia a la *ciudad lingüística* (Villena Ponsoda 1994, Martín Butragueño, 2010b: 1018-1055; 2013) se está pensando en un cierto tipo de conglomerado de hablas urbanas y, sobre todo, macro-urbanas, caracterizadas por la complejidad social, espacial, demográfica y étnica, por la multiplicidad lingüística y dialectal y, muchas veces, por el crecimiento abrupto asociado a la inmigración. En la macro-

---

<sup>1</sup> Un avance de este trabajo se presentó en la mesa «PRESEEA en América», dentro del *Simposio Internacional PRESEEA. 25 años de sociolingüística hispánica* (Alcalá de Henares, Universidad de Alcalá, 12-13 de junio de 2023). Agradezco la generosa invitación que me hicieron en aquel momento los organizadores del simposio –Ana M. Cestero Mancera, Isabel Molina Martos, Francisco Moreno Fernández y Florentino Paredes García. Y doy las gracias a Isabel Molina Martos por su cuidadosa lectura del texto.

ciudad, los flujos migratorios rompen en algún momento el crecimiento más sosegado de etapas pretéritas, lo que no impide que parte de las antiguas estructuras, sin embargo, se sostengan, a veces semiocultas, entre los habitantes actuales: es esta la *vieja ciudad lingüística*, una comunidad dentro de otra comunidad, con otras reglas y otros valores lingüísticos (Martín Butragueño, 2019: 52). Los habitantes de estas ciudades-conglomerado construyen espacios lingüísticos (Blommaert, 2010: 5-6; Caravedo, 2009, 2014: 257-264; Martín Butragueño, 2021: 156-157; 2023) y adoptan *perspectivas* sociolingüísticas que enfatizan ciertos saberes y sus potenciales (Martín Butragueño, 2023), al hilo de planteamientos que se remontan, cuando menos, al Ortega y Gasset de *El tema de nuestro tiempo*, publicado en 1923 (y consultado ahora por la edición de 2022; véase especialmente el cap. X: 98-107).

El concepto de ciudad lingüística tiene semejanzas y diferencias con el concepto de comunidad lingüística. Ambos conceptos son coincidentes en cuanto refieren a conglomerados humanos en los que se adensan las relaciones, pero disímiles en la medida en que una ciudad no tiene por qué ser una sola comunidad, sino que puede incluir varias, algunas de ellas extraordinariamente opuestas entre sí. Que puedan existir relaciones potenciales entre los asentados en un lugar y los inmigrantes no necesariamente significa que los segundos estén integrados o que puedan completar su proceso de integración. La marginación y la minorización pueden ser tales que ni siquiera haya espacios de convivencia física. Las experiencias vitales pueden ser enormemente diferentes según donde se viva. En las macro-ciudades es común no conocer ni visitar jamás amplias zonas de la ciudad ajenas al circuito vital de los habitantes de una zona físico-espacial, asociada a su vez a diversos atributos sociales. De ahí la importancia de disponer de una unidad de trabajo diferente a la de comunidad lingüística.

Los estudios estratificacionales tienen cierto potencial para revelar algunos, no todos, de los viejos y los nuevos significados sociales (Silverstein, 1976; Blommaert, 2010, especialmente §2.3-§2.4: 37-41; Eckert, 2018). La complejidad de Ciudad de México se subraya por su posición, en el centro de la megalópolis cuya lengua general puede, hasta cierto punto, identificarse con el español mexicano central (Martín Butragueño, 2020a, 2021: 156-157; 2023). Deben reconocerse, en cualquier caso, los límites de un proyecto estratificacional como CSCM-PRESEEA<sup>2</sup> (Martín Butragueño y Lastra, 2011-2015), aun cuando el *Corpus sociolingüístico de la Ciudad de México*, CSCM, recogió, en su conjunto, una muestra bastante mayor, que acogió diversos perfiles sociales, entre ellos varios grupos de inmigrantes, más allá del módulo dedicado a PRESEEA dentro del CSCM.

Es quizá el momento de intentar síntesis explicativas sobre la multitud de materiales y análisis disponibles, con el objetivo de formular un modelo de ciudad lingüística. Cabe preguntarse si es factible, en realidad, generar tal modelo sociolin-

---

<sup>2</sup> PRESEEA es el «Proyecto para el estudio del español sociolingüístico del español de España y de América» (cf. Moreno Fernández, 1996; Cestero Mancera, 2012).

güístico unitario para Ciudad de México. Dado que la respuesta final será, casi con seguridad, que *no totalmente*, o que *solo parcialmente*, es probable que sea mejor conformarse con atisbar algunos fragmentos de cómo podría ser ese modelo.

Los requisitos mínimos para ese objetivo son emplear diversas variables lingüísticas en correlación con un catálogo básico de categorías sociales. La hipótesis de trabajo pasa por sugerir que la ciudad lingüística es un constructo social con consecuencias específicas sobre la estructura de la variación de rasgos lingüísticos disímiles.

A continuación, se describen algunas decisiones y problemas metodológicos, para después presentar los datos que se comparan de una forma homogénea. Una primera discusión ofrece una interpretación general de los hechos, mientras que una segunda discusión establece los límites de esa misma interpretación. Se termina con unas breves conclusiones.

## 2. METODOLOGÍA

Al esbozar un modelo de ciudad lingüística para la capital mexicana, conviene entonces empezar por integrar en un solo componente una serie de hechos lingüísticos de naturaleza bastante variada entre sí. Se toman aquí, con ese propósito, las siguientes variables:

- (1) a. El uso de apreciativos en *-ito*.
- b. La concordancia entre *haber* y un argumento plural.
- c. La expresión de sujeto pronominal.
- d. La distribución de los indoamericanismos léxicos.
- e. La circunflexión entonativa de enunciados aseverativos.

Como se verá más adelante, todos estos problemas han sido ya estudiados<sup>3</sup> en materiales de CSCM-PRESEEA (Martín Butragueño y Lastra, 2011-2015), lo que allana el camino. Se trata, además, en la mayoría de los casos, de cuestiones que también han sido analizadas en muchos otros lugares, incluidas otras ciudades asociadas a PRESEEA, lo que habrá de facilitar, en el futuro, la comparación entre ciudades lingüísticas diversas. Como es también claro, (1a) es una variable en principio morfológica, pero con numerosas ramificaciones pragmáticas. (1b) y (1c) son, en principio, variables sintácticas, pero también incardinadas en el discurso, sobre todo (1c). Por su parte, (1d) refiere a la dimensión léxica. Y (1e) apela a la variación prosódica, relacionada, a su vez, con la estructura sintáctica y la estructura informativa y con diversos aspectos pragmáticos. En su variedad, pues tocan muy distintos aspectos lingüísticos, la expectativa es que contribuyan a caracterizar la ciudad lingüística desde diversos ángulos.

---

<sup>3</sup> Las referencias concretas se anotan *infra*.

Podrían compararse estas cinco variables a la luz de los predictores lingüísticos que las configuran, pero la ponderación más sencilla y evidente ahora es la que toca a los factores sociales, en particular, a la edad, el nivel de instrucción y el sexo. La base de datos de CSCM-PRESEEA recoge mucha más información social y reticular sobre cada colaborador, pero parece prudente por el momento limitarse a estos tres predictores sociales, ordenados como se anota en (2), siguiendo los lineamientos básicos generales del proyecto PRESEEA (cf. Moreno Fernández, 1996):

- (2) a. Edad: 20-34 años, 35-54, 55 en adelante.
- b. Nivel educativo: bajo (primaria o menos), medio (estudios medios), alto (estudios universitarios).
- c. Sexo: hombres y mujeres.

El principal problema metodológico es cómo comparar de manera coherente los estudios ya realizados. Si bien todos los trabajos previos reanalizados ahora proceden de CSCM-PRESEEA, a veces se ha trabajado con las 108 entrevistas, pero a veces no, y los cálculos se han realizado de diversas maneras: a veces con modelos de efectos fijos, a veces con modelos mixtos (Johnson, 2009; Díaz-Campos y Dickinson, 2019). Hay también ocasiones en que no se han generado verdaderas variables, sino que se muestran las tasas de uso para diversos grupos y circunstancias lingüísticas y discursivas. La solución ha sido repetir un cálculo de ANOVA de un factor, cuando existía, o llevarlo a cabo, cuando no, como ha sido en la mayor parte de los casos. Este cálculo permite, de una manera sencilla, observar la significación de los factores y ponderar la desviación estándar de los niveles de cada factor, por lo que se vuelve idóneo para establecer el papel social de diversas variables lingüísticas en un modelo básico de comportamiento ciudadano.

### 3. ALGUNOS DATOS SOCIOLINGÜÍSTICOS DE CIUDAD DE MÉXICO

A continuación, se presentan los datos relevantes para las cinco variables lingüísticas mencionadas en (1), a la luz de los predictores sociales anotados en (2). En cada caso, se ofrece primero un cuadro con el resumen de resultados de cada ocasión, y luego un cálculo de ANOVA de un factor, que permite la comparación transversal que se hará más adelante, al discutir los datos.

#### 3.1. Los apreciativos en «-ito»

Los datos de PRESEEA han permitido ofrecer un análisis actualizado de la distribución social y geográfica de los sufijos apreciativos, a partir de materiales que son razonablemente comparables entre sí; puede verse Patiño Agreda (2023) para un análisis comparativo de seis ciudades. Para Ciudad de México, en Martín Butra-

gueño (2021) se consideraron los casos en *-ito* producidos por los 108 colaboradores entrevistados en el módulo PRESEEA del *Corpus sociolingüístico de la Ciudad de México*, CSCM (Martín Butragueño y Lastra, 2011-2015). Al considerar el número de apreciativos producidos por cada uno de 108 individuos, se analizó una amplia matriz de factores sociales en un cálculo de ANOVA de un factor (Martín Butragueño, 2021, cuadro 15: 202-203). Los datos en juego son del tipo de los de (3).

(3) *Cuellito, molito, lanita, abuelita, Virgencita, afuerita.*

Fuente: Martín Butragueño, 2021, ítem 8: 192.

La tabla (1a) presenta los datos frecuenciales considerados ahora. Como puede observarse, las personas de más edad están 15 puntos porcentuales por encima de los jóvenes, mientras que los adultos, es decir, la generación intermedia, muestra una puntuación casi equidistante de los otros grupos: a más edad, más apreciativos en *-ito*. Puede verse asimismo que las personas de estudios bajos casi duplican en casos a las personas de estudios altos, y que las mujeres están prácticamente 23 puntos porcentuales por arriba de los hombres.

**TABLA 1A. APRECIATIVOS EN *-ITO* NO LEXICALIZADOS, SEGÚN EDAD, ESTUDIOS Y SEXO. N=3 242**

<i>Factores</i>	<i>Niveles</i>	<i>Frec</i>	<i>% (vert)</i>
Edad	Jóvenes	825	25.4
	Adultos	1059	32.7
	Mayores	1358	41.9
Estudios	Bajos	1388	42.8
	Medios	1083	33.4
	Altos	771	23.8
Sexo	Hombres	1250	38.6
	Mujeres	1992	61.4

Fuente: Martín Butragueño, 2021, cuadro 15: 202-203; calculado con SPSS 2012.

El cuadro 1b, por su parte, muestra que los tres factores considerados son estadísticamente significativos, como exhibe la ANOVA de un factor, cuya significación está siempre por abajo del umbral de 0.05. Llama la atención también que, siendo relativamente abultadas las desviaciones estándar anotadas para todos los niveles de los factores, estas no superan los promedios establecidos para cada nivel, a diferencia de lo que se verá *infra* para algunos otros casos.

**TABLA 1b. ANOVA DE UN FACTOR PARA APRECIATIVOS EN –ITO NO LEXICALIZADOS, SEGÚN EDAD, ESTUDIOS Y SEXO. N=108**

<i>Factores</i>	<i>Niveles</i>	<i>Media</i>	<i>Desv. Est.</i>	<i>F</i>	<i>Sign</i>
Edad	Jóvenes	22.9	12.6	3.86	0.0241
	Adultos	29.4	17.8		
	Mayores	37.7	32.6		
Estudios	Bajos	38.6	30.4	5.28	0.0066
	Medios	30.1	17.1		
	Altos	21.4	17.0		
Sexo	Hombres	23.1	14.8	10.23	0.0018
	Mujeres	36.9	27.9		

Fuente: Martín Butragueño, 2021, cuadro 15: 202-203; calculado originalmente con *SPPS* 2012 y recalculado ahora con *Stata/MP 17* = StataCorp 1985-2021.

La significación de los tres factores y la relativa moderación de la desviación estándar en (1b) no impide, sin embargo, que los individuos como tales –o las entrevistas en que se documentan– carezcan de un cierto papel, como se estudia en mayor detalle en el trabajo original<sup>4</sup>. No se trata de una cuestión menor, pues la comparación entre variables en una ciudad debería confrontar también la homogeneidad o no de los comportamientos individuales.

### 3.2. Los casos de «habían» y de «yo comí»

En lo que toca a la pluralización de *haber* con argumentos plurales, como en (4), y a partir de un cálculo de regresión logística escalonada efectuado con *Goldvarb Lion* (Sankoff, Tagliamonte y Smith, 2012), el trabajo de Lastra y Martín Butragueño (2016, cuadro 3: 131-132) reveló para Ciudad de México la significación de tres predictores sociales y una variable lingüística: el nivel de estudios, la edad, la definitud del argumento de *haber* y el sexo del hablante, en ese mismo orden de importancia.

- (4) a. Había muchos niños.  
b. Habían muchos niños

Ahora, en este trabajo, se parte de la misma base de datos empleada en el análisis de Lastra y Martín Butragueño (2016), pero los materiales se analizan estadísticamente desde otro ángulo. En parte, como se ha dicho, para hacerlos más directamente comparables con las otras variables consideradas en este informe; en parte, para discernir más claramente la dispersión de los resultados dentro de los grupos sociales. Se analizan solo los mismos tres aspectos sociales –edad, estudios y sexo–, sin considerar variables lingüísticas, dejando

<sup>4</sup> Véase particularmente la figura 9, p. 222, y el ítem §5.4, pp. 224-231, de Martín Butragueño (2021).

fuera 14 colaboradores que no produjeron ningún contexto variable relevante de *haber* impersonal, lo que da una N de 94 entrevistas. Para el cálculo de ANOVA, cada entrevista se puntúa con la frecuencia relativa correspondiente, en una escala de 0 a 1, al dividir el número de casos plurales entre el número total de casos de *haber* pertinentes.

El cuadro (2a) ofrece el resumen de resultados y el cuadro (2b) el análisis de varianza. Los resultados frecuenciales del primero de estos dos cuadros muestran porcentajes superiores para los adultos, en lo que toca a la edad; para las personas de estudios bajos, en relación al nivel de instrucción; y a las mujeres, en la distribución por sexos.

**TABLA 2A. CONCORDANCIA DE *HABER* CON SU ARGUMENTO PLURAL, SEGÚN EDAD, ESTUDIOS Y SEXO**

<i>Factores</i>	<i>Niveles</i>	<i>Frec/N</i>	<i>% (hor.)</i>
Edad	Jóvenes	9/116	7.8
	Adultos	25/199	12.6
	Mayores	13/246	5.3
Estudios	Bajos	26/177	14.7
	Medios	10/174	5.7
	Altos	11/210	5.2
Sexo	Hombres	26/331	7.9
	Mujeres	21/230	9.1

Fuente: Lastra y Martín Butragueño, 2016, cuadro 2: 128; calculado con *Goldvarb Lion*, de Sankoff, Tagliamonte y Smith, 2012.

Una diferencia importante entre los resultados de Lastra y Martín Butragueño (2016, cuadro 3: 131-132) y los presentados ahora en el cuadro (2b) es que el predictor social sexo no es significativo en el nuevo análisis. Debe considerarse que ya en el estudio original era el predictor significativo menos relevante, además del hecho de que ahora no se consideran factores lingüísticos y se han descartado las entrevistas sin casos útiles.

**TABLA 2B. ANOVA DE UN FACTOR PARA CONCORDANCIA DE *HABER* CON SU ARGUMENTO PLURAL, SEGÚN EDAD, ESTUDIOS Y SEXO. N= 94**

<i>Factores</i>	<i>Niveles</i>	<i>Media</i>	<i>Desv. Est.</i>	<i>F</i>	<i>Sign</i>
Edad	Jóvenes	0.04	0.15	3.01	0.0544
	Adultos	0.15	0.24		
	Mayores	0.06	0.14		
Estudios	Bajos	0.16	0.27	3.48	0.0349
	Medios	0.05	0.11		
	Altos	0.06	0.14		
Sexo	Hombres	0.07	0.15	1.08	0.3021
	Mujeres	0.11	0.22		

Fuente: Datos procedentes de Lastra y Martín Butragueño, 2016, recalculados ahora con *Stata/MP 17 = StataCorp 1985-2021*.

También ahora son las personas de estudios bajos quienes acumulan más pluralizaciones del verbo, y lo mismo puede decirse del grupo intermedio de edad, los adultos. Estudios y edad son significativos, pero no el sexo. Y la desviación estándar en cada nivel de todos los factores es siempre superior al valor alcanzado por el promedio, lo que sugiere una dosis elevada de variación individual. Los patrones de la pluralización de *haber* son complejos, como muestran diversos trabajos (véase Lastra y Martín Butragueño, 2016), entre ellos el reciente de Morales del Valle (2022, cap. 4: 267-320).

También es relevante ahora la expresión del sujeto pronominal, como en (5). El trabajo de Lastra y Martín Butragueño (2015, cuadro 3.3: 49) encontró, a partir de un cálculo de regresión logística realizado con *Goldvarb* (Sankoff *et al.*, 2012), que la edad es significativa, con tasas más elevadas cuanto mayor es la persona. Se hacen ahora algunos ajustes menores sobre los datos (cuadro 3a) y se recalculan con ANOVA de un factor por medio de *Stata/MP 17* (StataCorp, 1985-2021) para mejorar la comparabilidad (cuadro 3b). La ANOVA se calcula a partir de un índice por entrevista, consistente en la frecuencia relativa producto de dividir el número de presencias de sujeto pronominal entre el total de casos relevantes.

- (5) a. Yo soy aficionado al ciclismo.  
b. Soy aficionado al ciclismo.

Como puede observarse en el cuadro (3a), los mayores presentan tasas de uso bastante más elevadas que sus contrapartes, prácticamente el doble que las de los jóvenes. Las personas de estudios bajos y las mujeres destacan en sus factores, pero de modo no tan marcado como ocurre con la edad.

**TABLA 3A. PRESENCIA DEL SUJETO PRONOMINAL, SEGÚN EDAD, ESTUDIOS Y SEXO. N=2044**

<i>Factores</i>	<i>Niveles</i>	<i>Frec/N</i>	<i>% (hor.)</i>
Edad	Jóvenes	122/750	16.3
	Adultos	131/651	20.1
	Mayores	191/643	29.7
Estudios	Bajos	161/667	24.1
	Medios	158/754	21.0
	Altos	125/623	20.1
Sexo	Hombres	201/1014	19.8
	Mujeres	243/1030	23.6

Fuente: Lastra y Martín Butragueño, 2015, especialmente pp. 49-50; recalculado ahora con *Microsoft Excel para Mac* = Microsoft 2023.

El reanálisis del cuadro (3b) confirma lo ya establecido en Lastra y Martín Butragueño (2015), el hecho de que solo la edad es significativa en el modelo

social. Debe observarse también que la desviación estándar es pequeña en los niveles de cada factor, en comparación con lo que ocurre para otras variables lingüísticas.

**TABLA 3B. ANOVA DE UN FACTOR PARA PRESENCIA DEL SUJETO PRONOMINAL, SEGÚN EDAD, ESTUDIOS Y SEXO. N= 18**

<i>Factores</i>	<i>Niveles</i>	<i>Media</i>	<i>Desv. Est.</i>	<i>F</i>	<i>Sign</i>
Edad	Jóvenes	0.16	0.03	5.26	0.0186
	Adultos	0.20	0.05		
	Mayores	0.31	0.13		
Estudios	Bajos	0.26	0.13	0.35	0.7135
	Medios	0.21	0.09		
	Altos	0.21	0.08		
Sexo	Hombres	0.22	0.12	0.19	0.6692
	Mujeres	0.24	0.08		

Fuente: Datos procedentes de Lastra y Martín Butragueño, 2015, especialmente pp. 49-50; recalculados ahora con *Stata/MP 17* = StataCorp 1985-2021.

### 3.3. El caso de los indoamericanismos léxicos

El cálculo presentado ahora toma como punto de partida los lemas indomexicanos comunes anotados en Lastra y Martín Butragueño (en prensa; véase allí, en especial, los cuadros 1, 2a y 2b)<sup>5</sup>. Estas voces comunes son las que quedan después de excluir topónimos, antropónimos (y afines) y gentilicios. Algunas de ellas son las de (6):

- (6) a. *Chamaco, cuate, metate, quelite, etc.*

Fuente: Lastra y Martín Butragueño (en prensa, cuadro 1).

El número total de lemas indomexicanos comunes producidos por los entrevistados en CSCM-PRESEEA es de 149, repartidos en 1 097 ocurrencias. El cuadro (4a) resume el perfil social de su uso, al mostrar la distribución según tres grupos de edad y de estudios, y dos grupos de sexo.

<sup>5</sup> En Lastra y Martín y Butragueño (en prensa, n. 9; véase en especial cuadro 2a para el número de ocurrencias) la unidad de trabajo es todo el corpus, analizado con un enfoque semiautomático, lo que arroja 1 336 ocurrencias para 155 indomexicanismos comunes. La diferencia en las cifras con el análisis manual actual se debe a que ahora se toman en cuenta exclusivamente las producciones de los entrevistados que eran objeto de cada entrevista.

**TABLA 4A. INDOMEXICANISMOS COMUNES, SEGÚN EDAD, ESTUDIOS Y SEXO. N=1 097**

<i>Factores</i>	<i>Niveles</i>	<i>F</i>	<i>% (vert.)</i>
Edad	Jóvenes	289	26.3
	Adultos	350	31.9
	Mayores	458	41.8
Estudios	Bajos	528	48.1
	Medios	321	29.3
	Altos	248	22.6
Sexo	Hombres	457	41.7
	Mujeres	640	58.3

Los datos muestran una documentación más abundante de casos cuanto mayor es la edad del grupo considerado, habiendo unos 15 puntos porcentuales de diferencia. Algo parecido ocurre con el nivel de estudios, pues las personas de estudios bajos suman casi la mitad de las ocurrencias, seguidas, a bastante distancia, por los colaboradores de estudios medios y altos. Y no menos llamativo es el patrón exhibido por hombres y mujeres, pues las segundas rozan los 17 puntos de diferencia. Parecería, pues, a primera vista, que el empleo de lemas indomexicanos muestra una rica matriz social con respecto a los tres factores analizados.

Las cosas, sin embargo, no son exactamente así. Dado que en la base de datos se observan discrepancias notables entre los individuos, el cuadro (4b) estudia, como para las otras variables en juego, y tras establecer el número de casos para cada individuo, la ANOVA de un factor para edad, estudios y sexo.

**TABLA 4B. ANOVA DE UN FACTOR PARA INDOMEXICANISMOS COMUNES, SEGÚN EDAD, ESTUDIOS Y SEXO. N= 108**

<i>Factores</i>	<i>Niveles</i>	<i>Media</i>	<i>Desv. Est.</i>	<i>F</i>	<i>Sign</i>
Edad	Jóvenes	8.0	11.7	1.05	0.3527
	Adultos	9.7	11.8		
	Mayores	12.7	17.4		
Estudios	Bajos	14.7	16.5	3.15	0.0469
	Medios	8.9	13.0		
	Altos	6.9	10.8		
Sexo	Hombres	8.5	10.8	1.61	0.2070
	Mujeres	11.9	16.4		

Fuente: Datos procedentes de Lastra y Martín Butragueño, en prensa, especialmente cuadro 1, 2a y 2b; analizados ahora con *Stata/MP 17* = StataCorp 1985-2021.

Los datos son ahora 108, es decir, cada una de las cantidades de uso alcanzadas por cada individuo. Las medias de los factores confirman lo ya visto en el cuadro (4a), pues

los mayores, las personas de estudios bajos y las mujeres son quienes ostentan las medias más elevadas. Obsérvese, sin embargo, que las desviaciones estándar son bastante altas, siempre mayores a las medias<sup>6</sup>. La interpretación más simple de este hecho es que las divergencias entre los individuos que constituyen cada subgrupo son fuertes, lo suficiente al menos para impedir que las variables de edad y de sexo sean significativas en el cálculo de ANOVA, aunque sí lo sea el nivel de estudios.

Al no haber estratificación por edad, con estos datos ni siquiera puede sugerirse que haya cambio en curso. Esto refuerza la interpretación de Lastra y Martín Butragueño (en prensa), que sugieren que las diferencias entre sus datos y los de Lope Blanch (1979 [1969]) tienen que ver más con el corpus que con el tiempo real (frente a lo que exponen Bogomilova Lozanova, 2000: 69-71, y Torres Sánchez, 2014: 392-394)<sup>7</sup>.

Si bien el léxico es un excelente candidato para mostrar la evolución comunitaria, como ha mostrado Serrano (2014, especialmente el cap. 5: 191-311) para Ciudad de México, el léxico indomexicano –visto al menos como conjunto–, no delimita en estos datos la ciudad actual respecto de la vieja ciudad. Hay, sin embargo, individuos específicos que se comportan de manera sobresaliente en sus entrevistas, aspecto que, junto a la diferenciación educativa, merece ser examinado pausadamente en otro lugar.

### 3.4. Ascensos en núcleos de aseveraciones

Por fin, la última variable considerada atiende el ascenso tonal en las sílabas nucleares –las últimas tónicas– de aseveraciones no marcadas terminadas en palabra grave, tal como se estudió en Martín Butragueño (2011). Los datos que se ofrecen ahora oponen aquellos casos en que el ascenso tonal fue de +3.0 st o más, frente a los casos en que no ocurrió tal ascenso. El cuadro (5a) muestra cierto potencial sociolingüístico. Así, las personas de más edad muestran más ascensos notables que los adultos, y estos que los jóvenes; por su parte, los hombres puntúan más alto que las mujeres. Menos claro es el patrón por nivel de estudios, pues las personas de estudios altos muestran tasas parecidas a las de estudios bajos, a diferencia de las de estudios medios, que reducen notoriamente los casos de ascenso. En el trabajo original, se sugerían cruces relevantes entre la edad, de un lado, y los estudios y el sexo, de otro, de forma que la retracción generacional sería especialmente pronun-

---

<sup>6</sup> Considérese lo expuesto *supra* para los apreciativos en *-ito* no lexicalizados (cuadro 1b), para las mismas entrevistas y con cálculos realizados de modo semejante: si bien en los apreciativos también las desviaciones estándar alcanzan magnitudes notorias, nunca superan los promedios y la significación de ANOVA se extiende a todos los factores ahora relevantes, es decir, edad, estudios y sexo.

<sup>7</sup> Quedaría por hacer un análisis basado no en la frecuencia absoluta de los préstamos en cada entrevista, sino en la densidad de aparición según el número de palabras. Otro ángulo que deberá explorarse es la estratificación social no por el número de ocurrencias de indomexicanismos, sino por el empleo de lemas diferentes. Ambas perspectivas podrían cambiar o matizar el enfoque que ahora se ofrece.

ciada en las personas de estudios medios y en los hombres (Martín Butragueño, 2011: 115-118).

**TABLA 5A. PRESENCIA DE ASCENSOS NUCLEARES DE +3.0 ST Y MÁS, SEGÚN EDAD, ESTUDIOS Y SEXO. N= 540**

<i>Factores</i>	<i>Niveles</i>	<i>Frec/N</i>	<i>% (hor.)</i>
Edad	Jóvenes	59/180	32.8
	Adultos	68/180	37.8
	Mayores	82/180	45.6
Estudios	Bajos	80/180	44.4
	Medios	51/180	28.3
	Altos	78/180	43.3
Sexo	Hombres	118/270	43.7
	Mujeres	91/270	33.7

Fuente: Martín Butragueño, 2011, véase especialmente las tablas 10 y 11, pp. 113-114 y comentarios adyacentes en pp. 113-115 y, más en general, las tablas y comentarios de las pp. 104-113; recalculado ahora con *Microsoft Excel para Mac* = Microsoft 2023.

Para lo que interesa ahora, el cuadro (5b) muestra que, como factores independientes, las diferencias por edad no son significativas. Sí es relevante, en cambio, el nivel de estudios, y el sexo está muy cercano al umbral de significación (0.0579), por lo que también se mencionará en los cuadros comparativos posteriores. Por lo pronto, el ascenso tonal notorio es más recurrente en las personas de estudios bajos y altos y en los hombres.

**TABLA 5B. ANOVA DE UN FACTOR PARA PRESENCIA DE ASCENSOS NUCLEARES DE +3.0 ST Y MÁS, SEGÚN EDAD, ESTUDIOS Y SEXO. N= 54**

<i>Factores</i>	<i>Niveles</i>	<i>Media</i>	<i>Desv. Est.</i>	<i>F</i>	<i>Sign</i>
Edad	Jóvenes	0.33	0.18	2.06	0.1385
	Adultos	0.38	0.20		
	Mayores	0.46	0.19		
Estudios	Bajos	0.44	0.11	4.35	0.0180
	Medios	0.28	0.22		
	Altos	0.43	0.20		
Sexo	Hombres	0.44	0.18	3.76	0.0579
	Mujeres	0.34	0.20		

Fuente: Datos procedentes de Martín Butragueño, 2011; recalculados ahora con *Stata/MP 17* = StataCorp 1985-2021.

#### 4. DISCUSIÓN 1: HACIA UNA INTERPRETACIÓN GENERAL

Para entender mejor los hechos sociolingüísticos que afectan a Ciudad de México, es necesario cruzar los datos vistos hasta ahora de manera aislada. Para ello, se siguen tres rutas: la densidad de rasgos sociales significativos (tabla 6), la significación general de las variables lingüísticas y predictores sociales (tabla 7) y la dispersión entre individuos para las mismas variables y predictores (tabla 8).

**TABLA 6. SÍNTESIS COMPARATIVA DE CINCO VARIABLES LINGÜÍSTICAS Y TRES PREDICTORES SOCIALES SIGNIFICATIVOS (SEGÚN LAS MEDIAS DE LOS CUADROS B PARA SIGN. < 0.05)<sup>8</sup>**

<i>Variables</i>	<i>Edad</i>	<i>Estudios</i>	<i>Sexo</i>
Apreciativos en <i>-ito</i>	Mayores > adultos > jóvenes	Bajos > medios > altos	Mujeres > hombres
Concordancia de <i>haber</i>	Adultos > mayores > jóvenes	Bajos > altos > medios	n.s.
Presencia del sujeto pronominal	Mayores > adultos > jóvenes	n.s.	n.s.
Indomexicanismos comunes	n.s.	Bajos > medios > altos	n.s.
Ascensos nucleares de +3.0 st y más	n.s.	Bajos > altos > medios	(Hombres > mujeres)

( ) = significación < 0.1; n. s. = no significativo.

La tabla 6 resume de forma transversal la significación social de las cinco variables lingüísticas consideradas. El empleo de los apreciativos en *-ito* muestra una matriz social especialmente compleja, con los tres predictores sociales —edad, estudios y sexo— ejerciendo un efecto significativo. En un rango medio, con dos predictores sociales activos, se encuentra la concordancia de *haber* con su argumento plural y, hasta cierto punto, la presencia de ascensos nucleares de +3.0 st o más, pues el sexo está muy cerca de ser significativo al 0.05. Por fin, la variable más refractaria a los efectos sociales es el empleo de indomexicanismos comunes, correlacionados de modo significativo solo con el nivel de estudios.

Cabe realizar también una lectura vertical de la tabla 6. El nivel de estudios es el predictor más efectivo, con significación en cuatro de las cinco variables, siempre con las personas de estudios bajos en las gradas de uso más elevadas. En segundo lugar, aparece la edad, en tres de los cinco casos, siempre con los jóvenes asociados

<sup>8</sup> Al ofrecer el cuadro 6 según las medias de grupos, surgen a veces pequeñas diferencias respecto a los estudios originales.

a las tasas más bajas. Por fin, el sexo surge únicamente como relevante en dos de las cinco variables, y eso suponiendo que se acepte como interesante la significación  $< 0.1$  del predictor en el caso de los ascensos nucleares.

Vistas así las cosas, los efectos sociales más consistentes en la variación lingüística de Ciudad de México son el tener estudios bajos y el ser joven, aunque con efectos opuestos. Ambas características se vuelven fundamentales a la hora de entender la estructuración de la variación y de los posibles cambios lingüísticos.

**TABLA 7. JERARQUIZACIÓN DE LAS VARIABLES Y LOS PREDICTORES SEGÚN LA SIGNIFICACIÓN ALCANZADA EN LA ANOVA DE UN FACTOR**

<i>Variable lingüística</i>	<i>Predictor social</i>	<i>Significación</i>
Apreciativos en <i>-ito</i>	Sexo	0.0018
Apreciativos en <i>-ito</i>	Estudios	0.0066
Ascensos nucleares de +3.0st y más	Estudios	0.0180
Presencia del sujeto pronominal	Edad	0.0186
Apreciativos en <i>-ito</i>	Edad	0.0241
Concordancia de <i>haber</i>	Estudios	0.0349
Indomexicanismos comunes	Estudios	0.0469
Concordancia de <i>haber</i>	Edad	0.0544
Ascensos nucleares de +3.0st y más	Sexo	0.0579

( ) = significación  $< 0.1$ ; n. s. = no significativo.

La tabla 7 traza otro de los ángulos de la comparación entre variables y predictores, al jerarquizar la significación en el conjunto de los datos, y no únicamente dentro de una sola variable. Se pretende con ello ofrecer una imagen más dinámica y global de la relación entre las variables lingüísticas y los predictores sociales. Los dos binomios más significativos quedan asociados al manejo de los apreciativos en *-ito*, en relación al sexo y al nivel de estudios. Un segundo bloque está formado por los binomios cuya significación se mueve en el nivel de las centésimas, que alberga parejas muy variadas asociadas a la edad y al nivel de estudios. Por fin, los dos últimos pares se encuentran en el borde del umbral de 0.05 de significación.

No parece trazarse un dibujo muy evidente, lo que sugiere la complejidad de los patrones sociales. En todo caso, los apreciativos en *-ito* muestran una sólida significación, en contraste, por ejemplo, con la concordancia de *haber*. El nivel de estudios reafirma la importancia ya vista en la tabla 6, mientras que la edad, y sobre todo el sexo, resultan tener una significación más aleatoria.

**TABLA 8. DISPERSIÓN DE LAS VARIABLES SEGÚN EL PROMEDIO DE LAS DIFERENCIAS PORCENTUALES ENTRE LAS DESVIACIONES ESTÁNDAR Y LOS MEDIAS POR FACTOR EN CADA PREDICTOR SOCIAL**

<i>Variable lingüística</i>	<i>Predictor social</i>	<i>Dispersión</i>
Presencia del sujeto pronominal	Edad	28.6
Ascensos nucleares de +3.0 st y más	Sexo	49.9
Ascensos nucleares de +3.0 st y más	Estudios	50.0
Apreciativos en <i>-ito</i>	Edad	67.3
Apreciativos en <i>-ito</i>	Sexo	69.8
Apreciativos en <i>-ito</i>	Estudios	71.7
Indomexicanismos comunes	Estudios	138.3
Concordancia de <i>haber</i>	Edad	202.8
Concordancia de <i>haber</i>	Estudios	207.4

Por su parte, la tabla 8 proyecta un retrato diferente, en la medida en que compara las dispersiones en los binomios que están siendo considerados. Para valorar tal dispersión, se ha calculado, en primer término, a qué porcentaje del promedio corresponde la desviación estándar en cada nivel; después, se ha sacado el promedio de los porcentajes de los niveles de un factor (o predictor). Así, la presencia del sujeto pronominal tiene, con respecto a la edad, una dispersión promedio de apenas 28.6%. En el otro extremo, la concordancia de *haber* con su argumento plural presenta, en lo que toca a los estudios, una dispersión de 207.4%, es decir, una dispersión más de 7 veces mayor que en el caso que se acaba de mencionar.

El dato puede ser relevante, si se considera que es probable que cuanto mayor sea la dispersión, mayor podrá ser el papel de los individuos en un modelo mixto. Así, se sugiere que en los apreciativos en *-ito* pesan más los factores individuales que cuando se trata de la presencia del sujeto pronominal o de los ascensos nucleares, pero menos que con el uso de los indigenismos comunes o en la concordancia con *haber*.

De forma incipiente, parece relevante que en la comparación entre variables, en el camino para producir estructuras sociolingüísticas de una ciudad, debería entrar en juego la consideración del tipo de modelo, en términos de los agregados fijos y aleatorios que le corresponden como mixtura. Estas valoraciones pueden ser, a su vez, fundamentales para entender la estabilidad o no de la variación y, en consecuencia, de su grado de involucramiento en situaciones de cambio lingüístico.

Los elementos expuestos en los cuadros 6, 7 y 8 deberían dar la pauta primera para una interpretación cuantitativa general.

## 5. DISCUSIÓN 2: LOS LÍMITES DE LA INTERPRETACIÓN GENERAL

Puede pensarse que la integración de los perfiles sociales de un grupo de variables lingüísticas es un paso útil en la dirección de construir modelos sociolingüísticos del comportamiento urbano. Poder comparar con unas mismas herramientas los correlatos sociales, establecer la jerarquía de importancia de los binomios variable lingüística / predictor social y esbozar la dispersión del comportamiento individual parecen tareas pertinentes en la dirección trazada.

Por supuesto, no es esta toda la historia que hay que contar. Para empezar, es necesario formular una hipótesis rectora que dé un sentido general al conjunto de los acontecimientos. La pugna entre lo igual y lo diferente en cada intercambio tiene paralelo en el conflicto entre el arraigo y la migración (*cf.* Kerswill, 1996, 2018; Moreno Fernández 2019), entre la acomodación y la desacomodación (*cf.* Trudgill, 1986), entre lo identitario y lo oponible (Eckert, 2018). La gestión de las necesidades comunicativas (Martín Butragueño, 2010a: 46-60) al servicio de la imagen social (Félix-Brasdefer, 2019, en especial §6.3: 155-156) nos posiciona en gentil equilibrio en el espacio sociolingüístico del que formamos parte y que construimos a cada paso (*cf.* Caravedo, 2009, 2014: 257-264). El perspectivismo (Ortega y Gasset, 2022 [1923]) desde el que interpretamos tal espacio nos acompaña en cada gesto y en cada acción social (Martín Butragueño, 2023: 304-306, 308-309 y 314-315).

Más en concreto, los datos actuales dejan ver, particularmente, dos grandes efectos, derivados del nivel de estudios y de la edad de la población. Como se aprecia en la tabla 6 para los hechos significativos, las personas con el nivel más bajo de estudios son quienes puntúan más alto en las variables consideradas. Al tiempo, los jóvenes son quienes puntúan más bajo en los datos de esa misma tabla.

Saliéndonos de estos datos, sabemos que el número de años de instrucción viene aumentando en México a lo largo del tiempo (INEGI, 2023). También es bien conocido el papel de los jóvenes y, en general, de la edad en el cambio lingüístico (Kerswill, 1996, Chambers, 2018, especialmente §8: 310-312 y §10: 314-317). Sabemos también que existen procesos de desdialéctalización de los migrantes en un contexto de contacto dialectal, no abordado aquí, pero sí en otros trabajos (véase Martín Butragueño, 2013, 2020b, Soler Arechalde, 2022). A la vista de todo esto, la hipótesis rectora para establecer el papel de la ciudad lingüística, tomando como caso específico el de Ciudad de México, adopta la forma de una predicción normalizadora, expuesta en (7).

- (7) *Hipótesis o predicción normalizadora.* Tanto la variación asociada a estratificación como la asociada a contacto dialectal en Ciudad de México confluyen en la misma dirección general, más allá de posibles incidencias particulares. Existe un contexto normalizador y, más allá de las intuiciones, este se demuestra en el comportamiento de diversas variables lingüísticas.

Por supuesto, la predicción de (7) no es particularmente novedosa, pues puede alegarse que funciona para muchas otras ciudades y, además, que es lo esperado o

lo supuesto en multitud de trabajos de dialectología y sociolingüística. Lo relevante ahora es contribuir al contenido empírico de la hipótesis, con datos precisos que muestren los trayectos de cambio y la estructuración social de la variación.

En segundo lugar, debería contarse con un catálogo de variables y de predictores mucho mayor al aquí resumido. Esto sí es posible para Ciudad de México, pues a lo largo de los años se han realizado diversos estudios sobre tales materiales, al tiempo que se dispone de información social bastante detallada sobre cada colaborador en la muestra. Reprocesar y uniformar los análisis podría ser, sin embargo, bastante laborioso.

Un tercer aspecto es que no todos los rasgos sociolingüísticos son interesantes para el conjunto de la misma ciudad lingüística. Estos *rasgos puntuales* pueden ser muy variados: la pronunciación de un sonido, digamos /j/ o /x/ en ciertos grupos de inmigrantes; ciertas soluciones morfosintácticas en personas de edad (Martín Butragueño, 2019: 47-54; 2021); la presencia de un léxico muy característico en redes y grupos particulares. Por supuesto, con lo que se topa aquí es con el hecho de que una ciudad lingüística compleja está formada por un mosaico de subcomunidades, donde lo local, lo glocal y lo global (para Ciudad de México, véase Martín Butragueño, 2020a y bibliografía allí citada) adquieren diversas valoraciones. Difícilmente podría decirse que Ecatepec, en la zona conurbada (*cf.* Martín Butragueño, 2016), Xochimilco (Martín Butragueño, 2019: 45) o la Condesa tengan características fácilmente equiparables.

Un modelo sociolingüístico de Ciudad de México –y quizá de cualquier otra gran ciudad– no puede reducirse a suponer un consenso generalizado (*cf.* Labov, 1972, cap. 6: 143-159; 2006 [1966], cap. 11: 265-299), mucho menos en una sociedad caracterizada por las profundas brechas económicas y por los conflictos de todo tipo (Barriga Villanueva y Martín Butragueño, 2010-2022).

Podría, con todo, intentarse un modelo de ciudad lingüística:

- (8) Una *ciudad lingüística* (o al menos *esta* ciudad lingüística) es un mecanismo social de normalización lingüística, con efectos muy variados, especialmente visibles a través del nivel de estudios y de la edad de las personas.

## 6. CONCLUSIONES BÁSICAS: HACIA LA CIUDAD LINGÜÍSTICA

Una primera conclusión, de orden general, es la importancia de estudiar la ciudad lingüística como artefacto sociolingüístico (Martín Butragueño, 2013). Si bien esta idea subyace a la razón de ser de la sociolingüística urbana, seguramente no esté de más enfatizar las características de las comunidades ciudadinas y macro-ciudadinas, cuyas dinámicas son distintas a las de otros tipos de unidades poblacionales.

Una segunda conclusión es la necesidad de considerar variables lingüísticas de muy variada naturaleza para comprender de forma adecuada el efecto de las ciudades lingüísticas (*cf.* Martín Butragueño, 2020a). Si bien la perspectiva variacionista que

subraya las diferencias dialectológicas en una sola variable es perfectamente legítima, es más factible entender la dimensión sociolingüística de una ciudad a través del reflejo en diferentes fenómenos.

Un tercer punto es la necesidad de comparar con unas mismas herramientas esas variables lingüísticas. Esto no es necesariamente sencillo por diversas razones. Para empezar, los aspectos lingüísticos relevantes pueden aludir a dimensiones muy diversas entre sí. En lo social, el catálogo de rasgos comparables puede ser limitado. Y si se están reanalizando estudios previos, es probable que las aproximaciones cuantitativas no sean exactamente iguales.

En cuarto lugar, debe resaltarse la importancia de comparar las variables lingüísticas en términos de los agregados fijos y aleatorios que contienen. Estos últimos, en especial, deben verse en lo que toca al papel de los hablantes individuales, pues si la variación individual es elevada, cabe pensar en etapas difusas de cambio lingüístico (Le Page, 1978, Le Page y Tabouret-Keller, 1985), en las que el mecanismo citadino no ha conseguido la normalización de las soluciones lingüísticas.

Más en concreto, como una quinta cuestión, debe señalarse la relevancia del nivel de estudios y de la edad en el comportamiento de las variables lingüísticas en Ciudad de México.

Por fin, cabe apuntalar empíricamente el papel de la ciudad como agente normalizador, en el sentido de un manejo endógeno de la variación y de una disminución de esta, es decir, en el sentido de la focalización (de nuevo Le Page, 1978; Le Page y Tabouret-Keller, 1985).

## REFERENCIAS

- Barriga Villanueva, Rebeca y Pedro Martín Butragueño (dirs.) (2010-2022). *Historia sociolingüística de México*, Ciudad de México: El Colegio de México, 5 vols.
- Blommaert, Jan (2010). *The Sociolinguistics of Globalization*. Cambridge: Cambridge University Press.
- Bogomilova Lozanova, Elena (2000). «Notas sobre la vitalidad del léxico indígena en el español contemporáneo de la Ciudad de México». En P. Martín Butragueño (ed.), *Estructuras en contexto: estudios de variación lingüística*. Ciudad de México: El Colegio de México, pp. 61-79.
- Caravedo, Rocío (2009). «Percepción, espacios mentales y variedades lingüísticas en contacto». *Neue Romania*, 39: 171-195.
- Caravedo, Rocío (2014). *Percepción y variación lingüística. Enfoque sociocognitivo*. Madrid: Iberoamericana – Frankfurt: Vervuert.
- Cestero Mancera, Ana M. (2012). «El proyecto para el estudio sociolingüístico del español de España y América (PRESEEA)». *Español Actual*, 98: 227-234.
- Chambers, Jack (2018). «Patterns of variation including change», En J.K Chambers y N. Schilling-Estes (eds.), *The Handbook of Language Variation and Change*. Oxford: Wiley-Blackwell, pp. 349-372.

- Cukor-Avila, Patricia, y Guy Bailey (2018). «Real time and apparent time», En J.K Chambers y N. Schilling-Estes (eds.), *The Handbook of Language Variation and Change*. Oxford: Wiley-Blackwell, pp.239-262.
- Díaz-Campos, Manuel y Stephanie Dickinson (2019). «Using statistics as a tool in the analysis of sociolinguistic variation: A comparison of current and traditional methods». En G. Rei-Doval y F. Tejedo-Herrero, *Lusophone, Galician and Spanish Linguistics: Bridging Frames and Traditions*. Londres: Routledge, pp. 205-226.
- Eckert, Penelope (2018). *Meaning and Linguistic Variation. The Third Wave in Sociolinguistics*. Cambridge: Cambridge University Press.
- Félix-Brasdefer, J. César (2019). *Pragmática del español: contexto, uso y variación*. Londres - Nueva York: Routledge.
- INEGI (2023). Instituto Nacional de Estadística y Geografía. «México en cifras... Características educativas de la población». México: INEGI. <https://www.inegi.org.mx/app/areasgeograficas/?ag=0700#tabMCcollapse-Indicadores> [consultado el 26 de julio de 2023].
- Johnson, Daniel Ezra (2009). «Getting off the Goldvarb standard: introducing Rbrul for mixed-effects variable rule analysis». *Language and Linguistics Compass*, 3(1): 359-383.
- Kerswill, Paul (1996). «Children, adolescents, and language change». *Language Variation and Change*, 8: 177-202.
- Kerswill, Paul (2018). «Koineization». En J.K Chambers y N. Schilling-Estes (eds.), *The Handbook of Language Variation and Change*. Oxford: Wiley-Blackwell, pp. 519-536.
- Labov, William (1972). *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.
- Labov, William (1994). *Principles of Linguistic Change*. Vol. 1: *Internal Factors*. Blackwell: Oxford.
- Labov, William (2006 [1ª. ed.,1966]). *The Social Stratification of English in New York City*. 2ª. ed. Cambridge: Cambridge University Press.
- Lastra, Yolanda, y Pedro Martín Butragueño (2015). «Subject pronoun expression in oral Mexican Spanish». En R. Orozco, N. Shin y A. M. Carvalho, *Subject Pronoun Expression in Spanish: A Cross-dialectal Perspective*. Washington: Georgetown University Press, pp. 39-57.
- Lastra, Yolanda, y Pedro Martín Butragueño (2016). «La concordancia de haber existencial en la Ciudad de México». *Boletín de Filología de la Universidad de Chile*, 51(2): 121-145.
- Lastra, Yolanda, y Pedro Martín Butragueño (en prensa). «Aproximación al estudio de los indoamericanismos léxicos presentes en el Corpus sociolingüístico de la Ciudad de México». *Homenaje a Everardo Mendoza* [título provisional].
- Le Page, Robert (1978). «Projection, focussing, diffusion: or, steps towards a sociolinguistic theory of language, illustrated from the sociolinguistic survey

- of Multilingual communities. Stages I: cayo district, Belize (formerly British Honduras), and II: St. Lucia». Society for *Caribbean Linguistics Occasional Paper*, 9. [También en *York Papers in Linguistics*, 9: 7-32].
- Le Page, Robert y Andrée Tabouret-Keller (1985). *Acts of Identity: Creole-based Approaches to Language and Ethnicity*. Cambridge: Cambridge University Press.
- Lope Blanch, Juan Miguel (1979 [1ª. ed., 1969]). *Léxico indígena en el español de México*. 2ª. ed. Ciudad de México: El Colegio de México.
- Martín Butragueño, Pedro (2010a). «Perspectiva sociolingüística de la historia». En R. Barriga y P. Martín Butragueño (dirs.), *Historia sociolingüística de México. Vol. 1: México prehispánico y colonial*. Ciudad de México: El Colegio de México, pp. 41-96.
- Martín Butragueño, Pedro (2010b). «El proceso de urbanización: consecuencias lingüísticas». En R. Barriga y P. Martín Butragueño (dirs.), *Historia sociolingüística de México. Vol. 2: México contemporáneo*. Ciudad de México: El Colegio de México, pp. 997-1093.
- Martín Butragueño, Pedro (2011). «Estratificación sociolingüística de la entonación circunfleja mexicana». En P. Martín Butragueño (ed.), *Primer encuentro de cambio y variación lingüística: Realismo en el análisis de corpus orales*. Ciudad de México: El Colegio de México, pp. 93-121.
- Martín Butragueño, Pedro (2013). «Ciudades lingüísticas». Presentación en el Seminario *Relación entre factores lingüísticos y factores sociales: mutuas influencias*. Ciudad de México, Universidad Nacional Autónoma de México, 12 de junio.
- Martín Butragueño, Pedro (2016). «Inmigración y reconstrucción de la identidad lingüística: el caso de Ecatepec». *Cuadernos AISPI. Estudios de Lenguas y Literaturas Hispánicas. Revista Semestral de la Associazione Ispanisti Italiani*, 8: 145-170; <http://dx.doi.org/10.14672/8.2016.1099>.
- Martín Butragueño, Pedro (2019). *El presente como puerta del pasado: la documentación del español hablado en México*. Ciudad de México: Academia Mexicana de la Lengua - Universidad Nacional Autónoma de México.
- Martín Butragueño, Pedro (2020a). «Building the Megalopolis: dialectal leveling and language contact in Mexico City». En A. Lynch (ed.), *The Routledge Handbook of Spanish in the Global City*. Londres - Nueva York: Routledge, pp. 234-274.
- Martín Butragueño, Pedro (2020b). «Historia de dos medidas: contacto entonativo en la Ciudad de México». En A. Soler Arechalde y J. Serrano (eds.), *Contacto lingüístico y contexto social. Estudios de variación y cambio*. Ciudad de México: Universidad Nacional Autónoma de México, pp. 403-460.
- Martín Butragueño, Pedro (2021). «La aportación de los apreciativos en -ito/a a la estructuración sociolingüística de Ciudad de México». *Memorias de la Academia Mexicana de la Lengua*, 47: 155-248.
- Martín Butragueño, Pedro (2023). «El español en México», En R. Caravedo y F. Moreno Fernández (eds.), *Dialectología hispánica / The Routledge Handbook of Spanish Dialectology*. Londres - Nueva York: Routledge, pp. 304-318.

- Martín Butragueño, Pedro, y Yolanda Lastra (2011-2015). *Corpus sociolingüístico de la Ciudad de México*. Ciudad de México: El Colegio de México, 3 vols.
- Microsoft (2023). *Microsoft Excel para Mac*. Versión 16.72. Microsoft 365. <https://www.microsoft.com/es-mx>.
- Morales del Valle, Hugo Heriberto (2022). *El lugar de Guanajuato en la dialectología mexicana: tres niveles de variación y cambio*. [Tesis doctoral]. Ciudad de México: El Colegio de México.
- Moreno Fernández, Francisco (1996). «Metodología del «Proyecto para el estudio sociolingüístico del Español de España y de América» (PRESEEA)», *Lingüística*, 8: 257-287.
- Moreno Fernández, Francisco (ed.) (2019). «El español, lengua migratoria = Archiletras Científica». *Revista de Investigación de Lengua y Letras*, 2.
- Ortega y Gasset, José (2022 [1ª. ed., 1923]). *El tema de nuestro tiempo y otros ensayos*. Madrid: Alianza.
- Patiño Agreda, Claudia (2023). *Aspectos semánticos, pragmáticos y sociales en el uso de la morfología apreciativa del español*. [Tesis doctoral]. Ciudad de México: El Colegio de México.
- Sankoff, David, Sali Tagliamonte, y Enric Smith (2012). *Goldvarb Lion. A Multivariate Analysis Application*. Toronto: University of Toronto - Ottawa: University of Ottawa. <http://individual.utoronto.ca/tagliamonte/goldvarb.html>.
- Serrano, Julio (2014). *Procesos sociolingüísticos en español de la Ciudad de México. Estudio en tiempo real*. [Tesis doctoral]. Ciudad de México: El Colegio de México.
- Silverstein, Michael (1976). «Shifters, linguistic categories, and cultural description». En K. H. Basso y H. A. Selby (eds.), *Meaning in Anthropology*. Albuquerque: University of New Mexico Press, pp. 11-55.
- Soler Arechalde, María Ángeles (2022). «Contacto de dialectos del español en México». En R. Barriga-Villanueva y P. Martín Butragueño (dirs.), *Historia sociolingüística de México*. Vol. 5: *Nievas visitas al pasado y al presente*. Ciudad de México: El Colegio de México, pp. 3017-3064.
- SPSS (2012). *IBM SPSS Statistics [Statistical Package for the Social Sciences]*. Versión 21. IBM Corp.
- StataCorp (1985-2021) *Stata/MP 17.0 for Mac*. College Station, TX: StataCorp LLC. <https://www.stata.com>.
- Torres Sánchez, Nadiezdha (2014). «Léxico indígena en la ciudad de Guadalajara». En P. Martín Butragueño y L. Orozco (eds.), *Argumentos cuantitativos y cualitativos en sociolingüística*. Ciudad de México: El Colegio de México, pp. 371-396.
- Trudgill, Peter (1986). *Dialects in Contact*. Oxford – Nueva York: Basil Blackwell.
- Villena Ponsoda, Juan Andrés (1994). *La ciudad lingüística: fundamentos críticos de la sociolingüística urbana*. Universidad de Granada.



## **EL EQUIPO DEL PRESEEA DE SANTIAGO DE CHILE: SITUACIÓN ACTUAL Y DESAFÍOS A FUTURO**

Abelardo San Martín Núñez  
*Universidad de Chile*

### **1. INTRODUCCIÓN**

En este capítulo, nos hemos propuesto un doble objetivo. Por una parte, informar brevemente del estado actual de la investigación realizada por el equipo del PRESEEA de Santiago de Chile (en adelante, PRESEEA-SCHI) y, por otra, reflexionar sobre el estudio de la variación social y dialectal de los marcadores del discurso con los materiales del mencionado proyecto. Consecuentemente con estos objetivos, dividiremos nuestra participación en dos secciones principales, luego de este sucinto párrafo introductorio, para finalizar con una también compendiosa sección de síntesis.

### **2. INFORME SOBRE EL ESTADO ACTUAL DEL PRESEEA-SCHI**

La incorporación del equipo PRESEEA-SCHI se inició, formalmente, en 2007 con la firma del acuerdo correspondiente. La recogida del corpus se realizó entre 2007 y 2009, el envío de materiales etiquetados, en 2013, y la publicación de la muestra de 18 entrevistas (transcripciones y audios) en la plataforma del proyecto (<https://preseea.uah.es/>) en 2015. Por lo tanto, podemos afirmar que el trabajo del equipo ha sido continuo desde 2007 hasta el presente, lo que se refleja, sobre todo, en la publicación de estudios por miembros del equipo que comenzó ya en 2011, poco después de haber recopilado el corpus. A este respecto, es importante también destacar que desde 2016 el trabajo se dividió en dos sub-equipos de investigadores:

uno a cargo de la Dra. Silvana Guerrero y otro a cargo del Dr. Abelardo San Martín, autor de esta apretada reseña.

Con respecto a una caracterización de las variables analizadas por el equipo del PRESEEA-SCHI, basándonos en sus publicaciones, podemos señalar que, hasta ahora, principalmente han sido discursivas y morfosintácticas, es decir, no se han realizado estudios de variables fonéticas, ni léxicas. En específico, los principales fenómenos abordados, desde un punto de vista sociolingüístico, han sido la estructura de las narraciones de experiencia personal y los recursos empleados en su elaboración, el funcionamiento y la estratificación social de los marcadores del discurso, los mecanismos de introducción del discurso referido, los recursos y las estrategias de atenuación, la redundancia de clíticos, el comportamiento del voseo y del sujeto pronominal, entre otros objetos de estudio. Para mayor detalle de las referencias correspondientes puede consultarse el sitio del equipo del PRESEEA-SCHI en la plataforma del proyecto (<https://preseea.uah.es/equipos/santiago-de-chile>). Por supuesto, la causa de estas preferencias o tendencias analíticas, como es lógico, radica en las líneas de especialización e intereses investigativos particulares de los miembros del equipo.

Por supuesto, dentro del equipo ha habido un decidido interés en colaborar en la elaboración de las guías de estudios coordinados del PRESEEA, en específico, sobre la expresión del sujeto pronominal (Guerrero y Manjón-Cabeza Cruz), el presente narrativo (Guerrero y Arriagada) y los marcadores discursivos de reformulación (San Martín, Rojas, Ruiz González y Córdova Abundis). También ha habido una activa colaboración desde 2013 con proyectos relacionados con el PRESEEA como el PRECAVES sobre actitudes lingüísticas hacia las variedades del español (Cestero Mancera y Paredes García, 2018). Nos parece oportuno, en este sentido, destacar el interés entre los miembros del equipo del PRESEEA-SCHI por establecer redes de colaboración al interior del proyecto, en general.

Un hito importante para el desarrollo y la consolidación del equipo del PRESEEA-SCHI fue la edición del número monográfico «Estudios sobre la lengua española hablada en el mundo hispánico en su variedad geográfica y social con materiales del PRESEEA» en el *Boletín de Filología* (San Martín y Guerrero, 2016b). Concebido, en un inicio, como una «sección» monográfica, la cantidad y la calidad de los trabajos presentados determinó que los editores lo transformaran en un volumen independiente. En este monográfico se publicaron 13 estudios, entre artículos y notas, sobre variables fonético-fonológicas (/d/ intervocálica y final de palabra, yeísmo y alternancia entre seseo y ceceo), sintácticas (*haber* existencial y su pluralización, expresión de sujeto pronominal y selección modal entre subjuntivo e indicativo) y discursivas (marcadores *ahora sí que*, *por ser* y *onda*) en comunidades de habla incluidas en el proyecto (en específico, Caracas, Granada, Madrid, Málaga, Medellín, México D.F., Santiago de Chile y Sevilla).

Por lo que refiere a la proyección de nuevas metas del equipo del PRESEEA-SCHI, quisiéramos plantear algunas reflexiones, a título personal. Por un lado, en nuestra opinión, sería óptimo ampliar el espectro de variables analizadas, es decir,

abordar también el análisis de fenómenos fonéticos y léxicos. Para ello sería oportuno profundizar los vínculos de colaboración con otros especialistas dentro y fuera del proyecto, a efecto de diversificar el tipo de variables analizadas. Por otro lado, creemos que es necesario ampliar la representación de Chile en el PRESEEA, incluyendo en este macro-proyecto corpus sociolingüísticos de otras ciudades chilenas importantes, desde un punto de vista demográfico. A pesar de la innegable importancia del habla capitalina para el dialecto chileno, su figuración en esta plataforma investigativa de alcance internacional no puede reducirse a la muestra de Santiago. En este sentido, un avance para este cometido sería concretar la recogida de materiales del PRESEEA Valparaíso y de otras comunidades del norte y del sur de Chile. Afortunadamente, ya hemos establecido un acuerdo inicial con el Dr. Siegfried Muñoz de la Universidad de Playa Ancha de Valparaíso y la Dra. Marcela Rivadeneira de la Universidad Arturo Prat de Iquique para el logro de este importante desafío.

### 3. LA INVESTIGACIÓN DE LOS MARCADORES DEL DISCURSO EN EL PRESEEA-SCHI

Como ya adelantamos, a continuación, presentaremos algunas reflexiones sobre el estudio de la variación social y dialectal de los marcadores del discurso con los materiales del PRESEEA-SCHI. La realización de investigaciones monográficas en esta línea de trabajo nos ha permitido llegar a algunas conclusiones, en relación con algunos marcadores relevantes, que quisiéramos destacar aquí.

En un primer trabajo (San Martín, 2011) analizamos el uso de los marcadores interrogativos de control de contacto, también conocidos como «apéndices comprobativos», en el corpus PRESEEA-SCHI. Aunque identificamos un repertorio amplio de este tipo de partículas, el más frecuente –por mucho– fue *¿cachái?*, cuyo comportamiento ilustramos con el siguiente ejemplo:

E: hacer carrete piola en la casa/ un asado

I: sí po/ su asado *¿cachái?*/ pero más así como otra onda/ no/ *¿cachái?*/ porque salí/ igual/ salí cuando tení plata no más po/ pero a veces no hay las posibilidades de salir/ no todos los amigos po *¿cachái?*/ o sea que esas fecha// suben todos los precios po huevón [...] (SCHI\_H12\_039).

Como se ve *¿cachái?* es una partícula de valor apelativo y metadiscursivo que permite al hablante corroborar la comprensión de su interlocutor y que ocupa una posición final en el enunciado. *¿Cachái?* se emplea de modo muy frecuente en el español de Chile, al punto de que se ha convertido en un estereotipo del habla chilena, en términos de una marca identitaria de esta variedad dialectal. La asociación de variables sociales con el uso de este marcador resultó pertinente para el factor etario, ya que es más frecuente en el discurso de los sujetos jóvenes.

En un segundo estudio (San Martín y Guerrero, 2016a) abordamos el empleo de los marcadores discursivos que cumplen la función de reformulación (“reformula-

dores») en los materiales recogidos por el equipo. Uno de nuestros resultados fue la polifuncionalidad que asume *o sea* dentro de la reformulación como función general, debido a que revelamos casos en los que cumplía las subfunciones de explicación, rectificación y recapitulación. La única subfunción dentro de la reformulación que no desempeña este marcador es la de reformulador de distanciamiento. Los siguientes ejemplos ilustran los valores de reformulación de *o sea*:

1) Explicativo:

los amigos se hacen más rápido/ y// la relación se/ forma más se se forma de forma más intensa digamos/ y todo en menos/ en menos tiempo ¿cachái?/ *o sea* tú podí tener un mejor amigo en un día/ y puede ser tu amigo para siempre/ o puede no ser tu mejor amigo para siempre puede ser tu mejor amigo por años ¿cachái?» (SCHI\_H13\_074).

2) Rectificativo:

no es como antes/ que antes nuestros papás huevón/ se casaron cuando teníamos/ *o sea* cuando ellos tenían/ quince a veinte años quizás po/ y ahora se casan como a los treinta (SCHI\_H11\_001).

3) Recapitulativo:

yo creo que los fundamentos teóricos son son los correctos/ pero la aplicación en términos de definir los recorridos y la frecuencia/ y los/ y el número de transbordos/ ee/ hay un error de implementación po/ *o sea* ee/ hay/ tienen que revisar más la malla de/ de recorridos/ y lo otro que un aspecto negativo// que yo no entiendo cómo antes los empresarios microbuseros ganaban mucha plata y hoy día el/ el gobierno está perdiendo tanta plata/ *o sea* hay algo que está mal/ mal hecho (SCHI\_H33\_098).

Como se aprecia, *o sea* sirve para ampliar un enunciado anterior en otros términos a modo de aclaración (ejemplo 1), corregir una formulación previa que el hablante considera imprecisa (ejemplo 2) y sintetizar un conjunto de antecedentes en un enunciado de cierre (ejemplo 3). A nuestro juicio, esta polifuncionalidad deriva de la prototipicidad de *o sea* dentro de la reformulación, ya que se trata del reformulador más característico y frecuente, a nivel panhispánico, esto es, en el conjunto de variedades de habla española.

Asimismo, en San Martín y Guerrero (2016a) determinamos la relevancia de la partícula *igual* como reformulador de distanciamiento, que ilustramos con el siguiente ejemplo:

no po/ *igual* de repente es que de verdad no es por nada pero es que *igual* ya me acostumbré tanto a andar en el auto que de repente ando en micro y es así como que huevón/ y más con guagua que es súper incómodo ¿cachái?/ que tení que estar con el pedazo de bolso la leche que los pañales y la huevada y/ *igual* mala (SCHI\_M11\_010).

Como vemos, con esta función *igual* es conmutable por marcadores como *de todas maneras*, *de todas formas*, *o en todo caso*, es decir, se emplea para indicar que

el segmento discursivo que sigue anula las inferencias que se pueden desprender del segmento anterior. Este funcionamiento es bastante usual en el español de Argentina y México y, por supuesto, también en Chile muestra una frecuencia destacable. Se trata de un caso interesante porque consiste en un valor pragmático diferente al que se da en el español peninsular en el que *igual*, a nivel pragmático, funciona como un operador epistémico de posibilidad, conmutable por *quizás* o *tal vez* (García Negroni y Marcovecchio, 2014). También se trata de un marcador que es más frecuente en el discurso de los hablantes más jóvenes en Santiago.

Igualmente, nuestra indagación sobre los reformuladores en el corpus del PRE-SEEA-SCHI estableció la preponderancia de *al final* como reformulador de recapitulación, tal como ejemplificamos, a continuación:

por ahí tengo/ como la/ esa/ ese/ el deseo// que creo que todo el mundo debe tener pero// pero no vivo mi vida en base a eso/ o sea no ando jugando esas/ encuentro absurdo andar gastando plata todas las semanas para jugar eso si *al final* las expectativas de que uno se lo gane son mínimas (SCHI\_M23\_095).

En específico, acá la hablante cierra su intervención ofreciendo una síntesis comprensiva del modo en que lo haría también mediante marcadores como *en fin* o *en definitiva*. Se trata de un caso no tan estudiado como marcador discursivo, pero que se usa también en otras comunidades de habla española.

En un tercer artículo (San Martín, Rojas y Guerrero, 2016) estudiamos la polifuncionalidad general de los marcadores *por ser* y *onda* en el corpus PRESEEA-SCHI, que desempeñan las siguientes funciones:

#### 1) Reformulador explicativo:

E.: y/ ¿las características de las personas/ de la clase alta?/ o mejor dicho/ ¿cuáles son las características de los cuicos?

I.: eeh/ la misma/ la misma entrada que tienen/ *por ser* tienen más/ debe ser más educación/ más desplante/ es otro tipo de educación que tiene/ son más notorios/ son como más cancheros/ eso (SCHI\_H21\_015).

I.: sí fumo verde/ pero cuando no hay verde igual fumo prensada no tengo/ lo tengo que decir/ pero hoy día pero ahora como es abril cogollos mil la he pasado súper bien porque me han traído hartos cogollitos y buenos buenos buenos/ unos cogollos buenos buenos buenos así *onda* dos con dos aspiradas y quedáis volada al tiro así no tenía para que seguir fumando (SCHI\_M11\_007).

#### 2) Operador de concreción:

I.: eso claro y de mujeres/ a ver/ estaba la señora CS// em de también de los antiguos *por ser* estaba la señora esta locutora eh la de los tangos/ ¿cómo se llamaba ésta? eeh AC (SCHI\_H32\_066).

I.: de la V fue su arpillera eso fue lo y y más sus no sé sus décimas yo creo que también así *onda* yo creo que hay una que me gusta mucho que dice eh «mira cómo se visten cabos y sargentos para teñir de sangre los pavimentos» (SCHI\_M11\_007).

3) Operador enunciativo con valor continuativo o aproximativo:

I.: claro porque a ver/ compromiso no sería porque entonces ya no sería amistad po/ porque resulta que cuando uno es amigo es amigo/ pero es que ahora está tan difícil la cosa que uno a veces/ y como yo/ tal como en mi situación/ como yo/ donde vivo no no tengo eem/ eh *por ser* amistad con los vecinos/ entonces digo yo/ de amigos no/ ahora ya mayor prácticamente no me preocupo de eso (SCHI\_H32\_066).

I.: en mis años/ por ejemplo cuando yo pololeaba/ era/ nos juntábamos un lote/ fumábamos/ pero cigarros/ éramos *onda*/ éramos sanos po/ no éramos lo/ locos de esos que fumaban marihuana/ no/ nos gustaba escuchar música/ los C/ eeh/ puro rock clásico/ pero la juventud de antes era mucho mejor/ ahora está muy/ está muy acelerada la juventud/ no me gusta la juventud de ahora (SCHI\_M21\_019).

4) Introdutor de discurso directo:

y// y y bueno y se entiende que uno también tiene que saber corresponderle en ese sen/ en ese sentido/ una persona que no/ que no ande pidiendo favores *por ser* eeh/ «oye yo te hago este este este favor a ti pero tú me tienes que hacer/ uno a mí» (SCHI\_H32\_066) que tenga buena voluntad con uno// que tenga muy buena voluntad para cualquier cosa/ no que sea como ahí flojo y que esté echado y que/ uno le diga *onda*/ «¿me puede acompañar?» o/ no/ alguien que sea atento y que/ que te ayude y que te sirva/ que te sirva/ obvio (SCHI\_M12\_043).

Como se aprecia, ambas partículas se usan para aclarar un segmento discursivo anterior (ejemplos de 1), introducir un ejemplo o un caso concreto (ejemplos de 2), continuar el discurso de modo aproximativo (ejemplos de 3) e introducir una cita en estilo directo (ejemplos de 4). Resulta destacable, además, que se trate de marcadores usados en Chile, pero no en otras variedades dialectales del español. Con respecto a su asociación con factores sociales, *por ser* y *onda* muestran un comportamiento inverso en el habla santiaguina: el primero es más frecuente en hablantes mayores y del grupo de instrucción bajo, mientras que el segundo, lo es en sujetos jóvenes del grupo alto.

Los avances en la investigación de los marcadores del discurso en el corpus PRESEEA-SCHI nos estimuló a la formulación, adjudicación y ejecución del Proyecto FONDECYT N°119019, «Los marcadores del discurso en el español hablado en las dos orillas: análisis de su variación pragmática regional en España y América» (2019-2022), que codirigimos con Cristian Rojas. El objetivo de este proyecto fue contrastar el funcionamiento de los marcadores del discurso más frecuentes en diferentes comunidades de habla del español peninsular y americano. Para este cometido se analizó la muestra disponible del PRESEEA en su plataforma electrónica. En síntesis, realizamos un estudio contrastivo de los marcadores en diferentes variedades dialectales, abordando su variación pragmática en las dimensiones espacial y social para determinar la existencia de marcadores de uso general o más panhispánico y de marcadores característicos o de uso más idiosincrásico de cada comunidad de habla.

A este respecto, como señalamos en San Martín (2020a: 220), el análisis del empleo de estas unidades con valor pragmático en diferentes variedades de la lengua española nos permitiría identificar marcadores discursivos que:

- 1) existen en algunas regiones y no existen en otras;
- 2) asumen funciones diferentes en algunas regiones;
- 3) se emplean con las mismas funciones, pero con diferentes frecuencias de uso en algunas regiones respecto de otras;
- 4) se emplean con las mismas funciones, pero con una diferente distribución de uso de acuerdo con los factores sociodemográficos de los hablantes;
- 5) se emplean con las mismas funciones, pero con una diferente distribución de uso de acuerdo con las situaciones comunicativas en el continuo formal-informal.

Solo por citar uno de los artículos producidos en el marco del mencionado proyecto, en San Martín (2020b) analizamos el uso de *por último* como operador argumentativo escalar, que ilustramos con el siguiente ejemplo:

I.: [...] yo creo que no debería existir el aborto porque/ ya/ si puta no querías no quieres o no puedes en muchos casos o en el caso sois sois/ bueno/ siendo mujer ultrajada y de eso nace un hijo/ al fin y al cabo es una vida y si ya no la quieres tener/ *por último* dala en adopción/ hay personas que no pueden tener hijos/ qué más le gustaría tener un hijo una guagüita con su pareja (SCHI\_H11\_005).

Se trata, como puede desprenderse del caso precedente, de un valor consolatorio semejante al de *al menos* o *por lo menos*, es decir, donde el marcador introduce el argumento más bajo en una escala, pero que resulta suficiente para apoyar una tesis. Nuestro trabajo nos llevó a concluir que este comportamiento de *por último* solo se da en Chile, ya que no lo identificamos en los materiales de alguna otra variedad dialectal incluida en el PRESEEA. Asimismo, otro producto asociado con la ejecución del mencionado proyecto fue la elaboración y la publicación de la Guía PRESEEA de estudio de los marcadores discursivos de reformulación (San Martín, Rojas, Ruiz González y Córdova Abundis, 2022). De hecho, próximamente, esperamos organizar una sección monográfica en la que se publiquen los resultados del trabajo colaborativo de este grupo de estudio, en alguna revista de corriente principal.

A partir de la investigación realizada sobre los marcadores discursivos en los materiales del PRESEEA, quisiéramos plantear algunas reflexiones. En primer lugar, podríamos señalar como limitación el que las particularidades del instrumento (entrevista sociolingüística semiestructurada) pudieran restringir la aparición de marcadores de índole más conversacional, esto es, que son más frecuentes en interacciones libres. Sin embargo, si se cumple el criterio de lograr que el o la informante se relaje y obtener, así -dentro de lo posible- su estilo de habla más natural o espontánea, esa limitación podría mitigarse, en cierto grado. Por lo demás, es muy probable que en

interacciones conversacionales otros tipos de marcadores, como los estructuradores de la información y los conectores, no tengan una aparición tan abundante.

Por otro lado, es pertinente recordar que, además de la Guía PRESEEA de reformuladores antes mencionada, existe una Guía PRESEEA de estudio de los apéndices interrogativos de control de contacto (Cestero Mancera, 2022). En otras palabras, la tendencia en el proyecto hasta ahora ha sido coordinar estudios según tipos de función de los marcadores, es decir, indagaciones con una orientación onomasiológica (de la función hacia sus marcadores). No obstante, podría también proponerse una guía sobre los marcadores discursivos más general, esto es, que permita explorar las diferentes funciones de un mismo marcador, a saber, con una orientación semasiológica (del marcador hacia sus funciones). En esa guía general más básica podrían considerarse funciones en dos niveles: principales y secundarias, así como factores internos como la posición, la secuencia discursiva y la combinación de los marcadores con otras partículas, junto con los factores externos (sexo, edad y nivel de instrucción).

Asimismo, nos parece importante destacar que existe una asociación entre el uso de algunos marcadores discursivos y la variedad dialectal de los hablantes, en términos identitarios. Por ejemplo, las partículas vocativas de tipo apelativo como *che* y *güey* están vinculadas con las identidades de los hablantes rioplatense y mexicano, respectivamente. De modo semejante, los apéndices comprobativos *¿cachái?* y *¿sabes?* particularizan las identidades de los hablantes chilenos y españoles, correspondientemente. En este sentido, a nuestro juicio, la perspectiva de análisis de la identidad dialectal y social relacionada con el empleo de los marcadores del discurso es una línea de trabajo interesante de explorar. A este propósito, resulta fundamental y muy estimulante la concepción de la identidad lingüística desarrollada por Moreno Fernández (2020), a la que podríamos añadir la formulación de la sociolingüística de la identidad de Eckert (2018). Se trata de una perspectiva que se puede aplicar con provecho, ya que, desde nuestro punto de vista, los marcadores discursivos cumplen una doble indexicalidad: discursiva y social. En otras palabras, por una parte, indican u orientan la interpretación del discurso de los hablantes y, por otra, marcan o indican la identidad social (grupal o dialectal) de los individuos.

#### 4. CONCLUSIÓN

Inevitablemente, son muchos los aspectos que -por motivos de espacio- han quedado fuera de esta apretada recapitulación de la andadura del equipo del PRESEEA-SCHI. En lo fundamental, informamos el estado actual de la investigación realizada por este equipo y reflexionamos, con mayor detención, sobre el estudio de la variación socio-dialectal de los marcadores del discurso en el corpus recogido en Santiago. De esta manera, no hemos pretendido más que dar cuenta de algunas de las dimensiones de la estrecha relación entre la sociolingüística y la dialectología que pudiera ser de provecho indagar en los materiales del proyecto PRESEEA. Asi-

mismo, nos ha parecido relevante identificar algunas de las metas que podrían ser importantes de lograr a futuro, en cualquier caso, solo a modo de orientación, en términos de motivaciones e inquietudes personales.

## REFERENCIAS

- Cestero Mancera, Ana María (2022). *Guía PRESEEA de estudio de apéndices interrogativos de control de contacto*. Alcalá de Henares: Editorial Universidad de Alcalá.
- Cestero Mancera, Ana María y Florentino Paredes (2018). «Creencias y actitudes hacia las variedades cultas del español actual: el proyecto PRECAVES XXI». *Boletín de Filología*, LIII(2): 11-43.
- Eckert, Penelope (2018). *Meaning and Linguistic variation. The Third Wave in Sociolinguistics*. Cambridge: Cambridge University Press.
- García Negroni, María Marta y Ana María Marcovecchio (2014). «Igual a un lado y otro del Atlántico. Un origen común para dos valores argumentativos». En M. M. García Negroni (ed.), *Marcadores del discurso: Perspectivas y contrastes*. Buenos Aires: Santiago Arcos, pp. 141-157.
- Moreno Fernández, Francisco (2020). *La lengua y el sueño de la identidad*. Canterano: Aracne Editrice.
- San Martín Núñez, Abelardo (2011). «Los marcadores interrogativos de control de contacto en el corpus PRESEEA de Santiago de Chile». *Boletín de Filología*, XLVI(2): 135-166.
- San Martín Núñez, Abelardo (2020a). «Marcadores discursivos y variedades del español: el español de América». *Archiletras Científica*, IV: 217-229.
- San Martín Núñez, Abelardo (2020b). «Por último como operador argumentativo en español: función pragmática y variación regional con datos del corpus PRESEEA». *Revista de Lingüística Teórica y Aplicada*, 58(2): 93-116.
- San Martín Núñez, Abelardo y Silvana Guerrero González (2016a). «Los marcadores de reformulación en el corpus PRESEEA de Santiago de Chile». *Forma y Función*, 29(2): 15-38.
- San Martín Núñez, Abelardo y Silvana Guerrero González (eds.) (2016b). «Número monográfico ‘Estudios sobre la lengua española hablada en el mundo hispánico en su variedad geográfica y social con materiales del PRESEEA’». *Boletín de Filología*, LI(2).
- San Martín Núñez, Abelardo, Cristian Rojas Inostroza y Silvana Guerrero González (2016). «La función discursiva y la distribución social de los marcadores *por ser* y *onda* en el corpus del PRESEEA de Santiago de Chile». *Boletín de Filología*, LI(2): 235-254.
- San Martín Núñez, Abelardo, Cristian Rojas Inostroza, Natalia Ruiz González y Patricia Córdova Abundis (2022). *Guía PRESEEA de estudio de los marcadores discursivos de reformulación*. Alcalá de Henares: Editorial Universidad de Alcalá.



# **PRESEA EN ESPAÑA**



## PRESEEA-PAMPLONA. PRESENTACIÓN Y PRIMEROS ESTUDIOS<sup>1</sup>

María José Gallucci  
*Universidad de Navarra*

Carmela Pérez-Salazar  
*Universidad de Navarra*

### 1. NAVARRA Y SU RIQUEZA LINGÜÍSTICA A LO LARGO DE LA HISTORIA

Las peculiaridades históricas de Navarra, encrucijada geográfica y humana durante siglos, han determinado la singularidad de sus rasgos lingüísticos. Al igual que en otros territorios peninsulares, la romanización hizo desaparecer buena parte de la heterogeneidad lingüística anterior, pero en el caso de Navarra iba a dar lugar a una nueva realidad también diversa en la que coexisten, primero, vascuence y latín; después, vascuence y romance medieval (González Ollé, 1970a: 37-50); finalmente, y hasta el momento actual, vascuence y español<sup>2</sup>.

En el espacio físico que ocupaba el antiguo Reino de Navarra y ocupa hoy la Comunidad Foral, «territorio nuclear de los antiguos vascones» (Gorrochategui, 2006: 113)<sup>3</sup>, se distinguen tradicionalmente tres zonas de norte a sur (González Ollé, 2004:

---

<sup>1</sup> Este trabajo se adscribe al Grupo de Investigación sobre Oralidad (UNoral) de la Universidad de Navarra y cuenta con el apoyo de la Red Temática sobre comunicación conflictiva y mediación: interacción, vínculos relacionales y cohesión social (CoCoMInt; ref. RED2022-134123-T) financiada por el Ministerio de Ciencia e Innovación.

<sup>2</sup> Erize (2023: 763) señala que, en la actualidad, los hablantes de vasco son una minoría social, «como resultado de complejos procesos sociales e históricos de mantenimiento, de retroceso o de recuperación del idioma».

<sup>3</sup> Para la distinción entre vascones y vascófonos, véanse Wulff (2009: 47-56) y González Ollé (1970a: 31-37 y 2016).

230-231): la de la Montaña (*Saltus Vasconum*), la más septentrional, de clima húmedo y escasa densidad de población; la Zona Media (*Ager Vasconum*), de clima más continental, donde el multilingüismo se manifiesta con especial claridad en la toponimia (González Ollé, 2016: 148), y que comprende la capital, Pamplona; y la Ribera, zona meridional, con núcleos de población más densos. La división es oportuna porque condiciona desde antiguo los desplazamientos y asentamientos de la población, a veces determinantes para la historia lingüística<sup>4</sup>, y porque influye en las formas de vida y los caracteres. No obstante, las diferencias entre el norte y el sur son compatibles, hoy por hoy, con un sentimiento compartido de pertenencia a una comunidad (González Ollé, 2016: 231; Nausia *et al.*, 2023: 17)<sup>5</sup>.

Invasada por el latín y la cultura romana<sup>6</sup>, la Navarra medieval conoce, además de su *lingua navarrorum* –así se denomina en documentos medievales al euskera o vascuence (González Ollé, 1970a: 50)–, dos modalidades neolatinas: el dialecto o romance navarro, que comenzó a configurarse hacia el siglo X en tierras del oriente navarro y fue cobrando importancia en ámbitos eclesiásticos, monásticos y regios (González Ollé, 1970b); y el occitano, llegado de ultrapuertos con las poblaciones asentadas en distintos enclaves del territorio (González Ollé, 1969). En romance navarro, al que los documentos oficiales se refieren como *ydimate navarre terre*, se redactaron el *Fuero General de Navarra* y las Actas de Coronación de Carlos II y Carlos III de Navarra (González Ollé, 2004: 232). Situada casi en el extremo norte de la Zona Media, la capital representa singularmente la convivencia de lenguas de la Edad Media: la población es mayoritariamente vascohablante (González Ollé, 1970a: 45), pero también emplea y difunde el romance autóctono<sup>7</sup>.

De la convivencia euskera-latín y euskera-romance (o hablado-escrito) se accede, al finalizar la Edad Media, a otras circunstancias lingüísticas. En el siglo XV el romance navarro confluye, esencialmente, con el castellano (González Ollé, 1991: 60-61). El euskera continúa siendo lengua hablada mayoritaria en la capital (Jimeno Jurío, 1995b: 58) y en buena parte del reino (González Ollé, 1970a: 51), pero se anuncian tiempos de cambio con la incorporación de Navarra a Castilla en 1515 y la llegada de personajes castellanohablantes, virreyes, obispos, miembros de los tribunales de Corte, Consejo y Comptos, de la curia eclesiástica y la guarnición militar del Castillo nuevo o Ciudadela, desconocedores del vascuence, pese a la ley o *regula de idimate*, por la que ningún vicario ni regente de parroquia podía serlo si no com-

---

<sup>4</sup> González Ollé (2004: 231) se refiere al transporte fluvial de la madera de norte a sur y a la trashumancia ganadera, circunstancias que determinaron la necesidad de comunicación y, en concreto, el aprendizaje del romance o del español por parte de los vascohablantes.

<sup>5</sup> La división de norte a sur es compatible con la idea de la existencia de un área geolingüística horizontal que recorre, de este a oeste, parte de Aragón, Navarra y la Rioja (González Ollé, 1991: 64-68).

<sup>6</sup> Sobre la latinización de los vascos y las vías de penetración del latín, véase Echenique (1987: 64-68).

<sup>7</sup> “La difusión del romance navarro arranca desde la corte, se encuentre donde se encuentre», explica González Ollé (1970b: 68).

prendía o hablaba la lengua del país de destino (Jimeno Jurío, 1995b: 58-59). Ajena a este bilingüismo, una parte de la zona septentrional se mantuvo como vascohablante hasta la Edad Moderna (Saralegui, 2011: 79-80).

A lo largo de los siglos XVI y XVII el castellano fue adquiriendo prestigio paulatinamente por su vinculación con las situaciones comunicativas oficiales en las que se utilizaba: «sesiones de regimiento, en bandos y pregones, en juntas de feligreses, de mayoresales de barrios y cofradías; en la predicación ordinaria en parroquias y conventos (...) en la enseñanza escolar (...). En castellano está escrita casi toda la documentación de los protocolos notariales, de los tribunales civiles y eclesiásticos, de ayuntamientos, de parroquias y conventos; en ellas se escribían gran parte de los libros» (Jimeno Jurío, 1995b: 60). En el siglo XVII, el panorama lingüístico del antiguo reino se caracteriza por la convivencia desigual de las dos lenguas: había hablantes monolingües vascófonos, castellanohablantes monolingües y hablantes bilingües que empleaban uno y otro idioma según la situación comunicativa (Jimeno Jurío, 1995a: 166-180)<sup>8</sup>. En cuanto a Pamplona, en el siglo XVII la población era mayoritariamente euskaldún (un 75 %), pero este porcentaje desciende notablemente en los siglos posteriores (Mikelarena, 2003: 190).

El retroceso paulatino del vascuence a partir de la Edad Moderna<sup>9</sup> se estima provocado por dos factores (Jimeno Jurío, 1994): las medidas de restricción impuestas (el castellano se hace obligatorio en la enseñanza y se adoptan medidas punitivas para quienes emplearan la lengua vasca), y la renuncia, por parte de los clérigos, a predicar, confesar y enseñar el catecismo en euskera. En cualquier caso, el retroceso no afecta a la parte septentrional de Navarra, la Navarra norteña o Montaña, que no conoció el romance autóctono o patrimonial y adoptó, a partir de la Edad Moderna, el castellano (González Ollé, 1991: 61-62), o más precisamente, el español regional de Navarra (González Ollé, 1991: 62). Es incluso posible establecer diferencias entre la zona nororiental, que adoptó el castellano (y abandonó el euskera) por contacto natural entre vecinos con los habitantes de la Navarra media y de Aragón a lo largo del siglo XIX, y la zona noroccidental, que recibe más tardíamente el español estándar por la industrialización, la llegada de hablantes monolingües castellanos y la enseñanza, impartida en lengua española (Saralegui, 2011: 79-80). Avanzado el siglo XX, el norte de la Comunidad Foral todavía tiene vascohablantes monolingües.

---

<sup>8</sup> Jimeno Jurío (1995b: 60) recoge las palabras de Miguel de Itúrbide, pamplonés y caballero de Santiago, en el siglo XVII: «La lengua matriz de este reino (...) y en particular en esta Ciudad de Pamplona, es la basconica o bascongada, como lo dicen muchos ystoridores»; «la jente ordinaria no ablan otra lengua, y si saben algunas voces o frases en lengua castellana, no la saben con tal espedición que puedan confesar sus culpas con todas sus circunstancias».

<sup>9</sup> González Ollé (1970a: 51-59), de acuerdo con Caro Baroja, explica que a partir del siglo XVII la presencia del vascuence se fue reduciendo horizontal y verticalmente. Los límites son bien conocidos en el siglo XVI gracias a un documento que distingue las localidades vascohablantes de las no vascohablantes en la diócesis de Pamplona. Después de esta centuria y hasta el siglo XIX, los textos revelan los límites cambiantes entre euskera y castellano.

## 2. LA NAVARRA ACTUAL. PECULIARIDADES LINGÜÍSTICAS

Heredera de los avatares histórico-lingüísticos que se han descrito hasta aquí, en la Comunidad Foral de Navarra conviven hoy español y vascuence. La presencia del euskera es diversa según la zona (Nausia *et al.*, 2023: 14) y está regulada legalmente: existe una *Ley Foral del Vascuence*, aprobada en 1986 (BON 54, 17/12/86), que reconoce las dos lenguas como propias de Navarra y el derecho a conocerlas y utilizarlas. La Ley distingue tres zonas en el territorio: una zona vascófona (en la que el euskera es cooficial), una zona mixta (en la que se fomenta y se intenta recuperar el euskera) y una zona no vascófona. Esta zonificación, que pretende regular una realidad lingüística compleja, suscita debate hasta la actualidad.

En las últimas décadas se han producido cambios importantes en la Comunidad Foral, cambios lingüísticos, sociolingüísticos y psicolingüísticos, sobre todo desde el final de la dictadura franquista, que persiguió, como en otros lugares de España, el uso de lenguas distintas del español. Si hasta aquí se ha descrito la deseuskeringización social o geográfica, acaecida por distintos motivos durante siglos, y política, es obligado constatar, con la llegada de la democracia, la implantación de *ikastolas* y de modelos de enseñanza en euskera y, en consecuencia, la recuperación de la lengua, ahora normalizada (el *euskera batúa*), en el intercambio oral y escrito (Vilches, 2002, cap. 2.2.).

Los años posteriores a la dictadura han dejado también muestras de politización lingüística. Poco han ayudado en este sentido la defensa del vascuence abanderada por el terrorismo y su entorno y las reacciones de quienes asumen esa identificación y proponen un españolismo excluyente. Hoy puede decirse que perdura, en parte, en Navarra, la actitud negativa hacia el español o hacia el vascuence (Nausia *et al.*, 2023: 12)<sup>10</sup>. La implantación del euskera en la enseñanza ha hecho aumentar el número de hablantes bilingües<sup>11</sup>. En concreto, en la zona mixta, que incluye la capital, la preferencia por la enseñanza bilingüe se ha instalado en las nuevas generaciones: un 41,5 % de los habitantes entre 16 y 24 años conoce el euskera (Nausia *et al.*, 2023: 13).

El español hablado hoy en Navarra es heredero del romance autóctono y del castellano, y producto también de la convivencia multiseccular de ambos con el euskera. La convergencia del navarro con el castellano, o, en su caso, la adopción del castellano (en lo que respecta al área norteña –Saralegui, 2011–), no impide que pueda hablarse, en pleno siglo XXI, de un español de Navarra, con peculiaridades en todos los ámbitos lingüísticos. La modalidad hablada en Navarra pertenece al área dialectal del denominado *español castellano* (Moreno Fernández, 2014: 120-121) o

---

<sup>10</sup> En los capítulos 4.1. y 4.2. de esta obra se muestran datos sobre la percepción del vascuence y sobre la politización y la ideologización de las lenguas por parte del nacionalismo vasco y el nacionalismo español.

<sup>11</sup> Véase Erize (2023: 764-767) para cifras de matriculación, en 2021, en los modelos lingüísticos que ofrecen euskera y otros indicadores de la evolución social de esta lengua. Para este autor, la promoción del euskera se ve dificultada por la falta de diálogo entre las instituciones, entidades y personas que se ocupan de ella y la población que habla castellano.

*español oriental* (Fernández Ordóñez, 2012: 31), pero incluye, pese a lo reducido de su espacio geográfico, dos modalidades: el *castellano de áreas bilingües* (castellano vasco) y el *castellano norteño oriental*, entre las que no existen límites nítidos ni «aduanas lingüísticas» (Moreno Fernández, 2014: 134).

### 3. ESTUDIOS SOBRE EL HABLA ACTUAL DE NAVARRA. EL INTERÉS DE PAMPLONA

La tradición de los estudios histórico-dialectológicos iniciada en la Universidad de Navarra (González Ollé, 1970a y 1970b<sup>12</sup>), que ha tenido continuidad hasta el momento actual (Pérez-Salazar, 1995; Martínez Pasamar y Taberero, 2012); las investigaciones sobre oralidad desde fuentes escritas en el norte peninsular (Pérez-Salazar, 2002, 2007 y 2012; Taberero *et al.*, 2020; Pérez-Salazar, 2020) y los estudios realizados sobre el español actual de Navarra (Saralegui, 1985; Rioja Arano, 1997; Saralegui y Taberero, 2002) son la base sobre la que se debe asentar la andadura de la sociolingüística variacionista en la comunidad de habla de Pamplona (Blas Arroyo, 2019, cap. 1).

La singularidad de la Comunidad Foral ha dado lugar a iniciativas que se han ocupado de recopilar el español hablado. Navarra cuenta desde mediados del siglo XX con un repertorio léxico excepcional, el *Vocabulario navarro* de José María Iribarren (Iribarren, 1984). Más recientemente, el proyecto denominado *Patrimonio Cultural Inmaterial*, promovido y desarrollado por la Universidad Pública de Navarra, ha creado el *Archivo del Patrimonio inmaterial de Navarra* ([www.navarchivo.com](http://www.navarchivo.com)), una plataforma interactiva con material oral, multimodal y no verbal que cuenta, entre sus objetivos, con la realización de estudios lingüísticos. Pese a la calidad del archivo y de las publicaciones que han surgido de este proyecto (Romera y Medina, 2014; Aznárez y Romera, 2016; entre otros), en el ámbito lingüístico, el corpus no se ha realizado con metodología sociolingüística y, por tanto, no colisiona con los fines del PRESEEA-Pamplona.

Son también dignas de mención aportaciones excepcionales que se han ocupado del habla de Pamplona, en el aspecto léxico, con metodología sociolingüística: valga citar la tesis doctoral de Itziar Aragües, *Estudio sociolingüístico del léxico del habla de Pamplona* (Aragües, 2000<sup>13</sup>) y los trabajos surgidos del proyecto de disponibilidad léxica (Taberero, 2008; Areta, 2009<sup>14</sup>).

Por último, es preciso mencionar aquí dos trabajos recientes, ya citados en esta aportación, que abordan la cuestión de la identidad y la conciencia lingüística en Na-

<sup>12</sup> La investigación iniciada por González Ollé en 1970 sobre el romance de Navarra y continuada por Saralegui en 1977 ha conocido un desarrollo notable durante varias décadas. Razones de espacio nos impiden presentar aquí una bibliografía representativa del alcance de los estudios dialectológicos sobre Navarra.

<sup>13</sup> Existe una publicación parcial de esta tesis doctoral: Aragües, Itziar, *Y tú ¿cómo dices? Vocabulario actual de Pamplona*. Ayuntamiento de Pamplona, 2003.

<sup>14</sup> Tesis doctoral inédita.

varra. El primero, un artículo especializado (Erize, 2023) en el que el autor describe las características sociolingüísticas de Navarra en lo que respecta a la convivencia de castellano y euskera y propone, como expresa el título, un «diálogo con la población castellanohablante navarra para el fomento de la lengua vasca». El segundo es un libro que acaba de ver la luz y que recoge los resultados de un estudio realizado a lo largo de cuatro años sobre la polarización social y lingüística de la Comunidad Foral (Nausia *et al.*, 2023). Cualquier indagación sobre estas cuestiones en distintos sectores de la población contribuirá a completar el conocimiento sociolingüístico de Navarra y, más precisamente, de la capital (Moreno Fernández, 2015, cap. 9).

Pese a todo, las investigaciones sociolingüísticas sobre Navarra, y más precisamente sobre Pamplona, se encuentran muy por detrás del alcance que han permitido desarrollar los corpus orales en otros lugares del mundo hispanohablante (Díaz-Campos, 2011; Hernández Campoy, 2011; Blas Arroyo, 2019; Moreno Fernández, 2021a y 2021b; Cortés Rodríguez, 2022).

#### 4. EL CORPUS PRESEEA-PAMPLONA

La incorporación formal de Pamplona al PRESEEA empieza a finales del año 2020, con la firma del acuerdo de participación entre el coordinador general del proyecto, el profesor Francisco Moreno Fernández, y las profesoras María José Gallucci y Carmela Pérez-Salazar, responsables del equipo de investigación sociolingüística de la ciudad de Pamplona.

El siguiente paso consistió en organizar un equipo de trabajo que apoyara las actividades inherentes a un proyecto de este tipo. A las coordinadoras se sumaron tres investigadores (Inés Olza Moreno, Dámaso Izquierdo Alegría y Nekane Celayeta Gil); dos colaboradoras (Ana Zúñiga Lacruz y Virginia Marín Marín) y dos alumnas internas del Departamento de Filología (Nadia Aldunate Martín y Emilia Ferran Gusseva). Los investigadores y las colaboradoras están adscritos a distintas dependencias de la Universidad de Navarra (Instituto Cultura y Sociedad y Facultad de Educación y Psicología).

Paralelamente se empezaron a tomar algunas decisiones metodológicas previas a la recolección y transcripción de las muestras de habla; siempre en el marco establecido por el PRESEEA (Moreno Fernández, 2021a, 2021b): la descripción de Pamplona como comunidad de habla, el número de muestras, su distribución en función de los barrios de la ciudad, los módulos temáticos de las entrevistas, entre otros aspectos que se detallan en los puntos §4.1 y §4.2 de este trabajo.

##### 4.1. Pamplona: la comunidad de estudio

Pamplona es la capital de la Comunidad Foral de Navarra. Además del contexto histórico y de las particularidades lingüísticas que ya se han apuntado, en ella se

concentran entes gubernamentales, poderes públicos y dos universidades; además, en esta ciudad funcionan y se editan distintos medios de comunicación (periódicos, radios y televisoras).

La ciudad tiene una extensión de 23,55 kilómetros y una población de más de 207.000 personas (INE, 2018). Teniendo en cuenta esta última cifra, y acorde con las pautas metodológicas del PRESEEA sobre la densidad poblacional de las comunidades de habla que integran el macrocorpus, el número total de informantes establecido en el subcorpus de Pamplona es de 54.

Tal como se aprecia en la tabla 1, las 54 muestras se agrupan en cuotas de afiliación uniforme (3 por casilla) en función de las variables edad, sexo y grado de instrucción establecidas por el proyecto y que son comunes a todos los equipos que participan en él.

**TABLA 1. DISTRIBUCIÓN POR CUOTAS DE LOS HABLANTES**

<i>Nivel de instrucción</i>	<i>Grupo generacional 1 20-34 años</i>		<i>Grupo generacional 2 35-54 años</i>		<i>Grupo generacional 3 55 o más</i>		<i>Total</i>
	H	M	H	M	H	M	
1: Primaria	3	3	3	3	3	3	18
2: Media	3	3	3	3	3	3	18
3: Superior	3	3	3	3	3	3	18
	9	9	9	9	9	9	
Total	18		18		18		54

El número de informantes establecido constituye una muestra exhaustiva y suficiente que hará posible uno de los objetivos generales del PRESEEA: la comparación de los resultados obtenidos en distintas comunidades de habla, tanto de España como de América, que han seguido los mismos lineamientos de estratificación (Moreno Fernández, 2005, 2021a).

En la tabla 2 se presentan las muestras del habla de Pamplona con los códigos por hablante que garantizan su identificación anónima. El código de los informantes debe leerse, de izquierda a derecha, de la siguiente forma: i) PAMP\_ = Pamplona; ii) sexo o género: H = hombres, M = mujeres; iii) edad: 1 = 20 a 34 años; iv) grado de instrucción: 3 = enseñanza superior; v) identificación del hablante dentro del corpus: 01, 02, 03, etc. Por ejemplo, con el código PAMP\_H13\_37 se identifica a un hombre pamplonés, de 20 a 34 años, con instrucción superior, al que le corresponde el número 37 en el cuadro general del corpus.

**TABLA 2. IDENTIFICACIÓN DE LOS HABLANTES POR CÓDIGOS**

Nivel de instrucción	Grupo generacional 1 20-34 años		Grupo generacional 2 35-54 años		Grupo generacional 3 55 o más	
	H	M	H	M	H	M
1: Primaria	PAMP_H11_01	PAMP_M11_04	PAMP_H21_07	PAMP_M21_10	PAMP_H31_13	PAMP_M31_16
	PAMP_H11_02	PAMP_M11_05	PAMP_H21_08	PAMP_M21_11	PAMP_H31_14	PAMP_M31_17
	PAMP_H11_03	PAMP_M11_06	PAMP_H21_09	PAMP_M21_12	PAMP_H31_15	PAMP_M31_18
2: Media	PAMP_H12_19	PAMP_M12_22	PAMP_H22_25	PAMP_M22_28	PAMP_H32_31	PAMP_M32_34
	PAMP_H12_20	PAMP_M12_23	PAMP_H22_26	PAMP_M22_29	PAMP_H32_32	PAMP_M32_35
	PAMP_H12_21	PAMP_M12_24	PAMP_H12_27	PAMP_M12_30	PAMP_H32_33	PAMP_M32_36
3: Superior	PAMP_H13_37	PAMP_M13_40	PAMP_H23_43	PAMP_M23_46	PAMP_H33_49	PAMP_M33_52
	PAMP_H13_38	PAMP_M13_41	PAMP_H23_44	PAMP_M23_47	PAMP_H33_50	PAMP_M33_53
	PAMP_H13_39	PAMP_M13_42	PAMP_H23_45	PAMP_M23_48	PAMP_H33_51	PAMP_M33_54

Aunque no existe un número de entrevistados universal ni mínimo que garantice la representatividad total de la comunidad (Molina *et al.*, 2014: 15), a fin de asegurar una distribución espacial equilibrada que permita contar con hablantes de los 14 barrios de la ciudad y, en consecuencia, garantizar que se reflejen las peculiaridades del español oral de Pamplona, estos barrios se han agrupado en tres grandes zonas, tal como se observa en la tabla 3 y en la figura 1.

**TABLA 3. ZONAS, BARRIOS Y NÚMERO DE HABITANTES**

Zona 1		Zona 2		Zona 3	
Barrio	Nº habitantes	Barrio	Nº habitantes	Barrio	Nº habitantes
Casco Viejo	11.496	Azpilagaña	7.352	Buztintxuri	9.005
Ensanches	21.111	Iturrama	22.718	Chantrea	19.439
Milagrosa	15.005	Mendebaldea	16.804	Rochapea	26.621
Lezkairu	6.255	Echavacoiz	5.317	San Jorge	12.318
Mendillorri	13.951	San Juan	19.970	—	—
Total	67.818		72.161		67.383

Como se puede apreciar, la zona 1 incluye los barrios del Casco Viejo, Ensanches, Milagrosa, Lezkairu, Mendillorri; la zona 2 está constituida por Azpilagaña, Iturrama, Mendebaldea-Ermitagaña, Echavacoiz y San Juan; mientras que la zona 3 está formada por Buztintxuri, Chantrea, Rochapea y San Jorge.

Aunque de cada zona se grabarán 18 hablantes en total (3 hombres y 3 mujeres por grado de instrucción) (ver tabla 4), en la distribución se tendrán en cuenta, a su vez, los barrios con mayor población. Por ejemplo, la zona 1, que cuenta con un total de 67.818 habitantes, tendrá mayor representación de Ensanches, con 21.111 habitantes, que de Lezkairu, pues este último agrupa 6.255.

FIGURA 1. BARRIOS DE PAMPLONA AGRUPADOS POR ZONAS

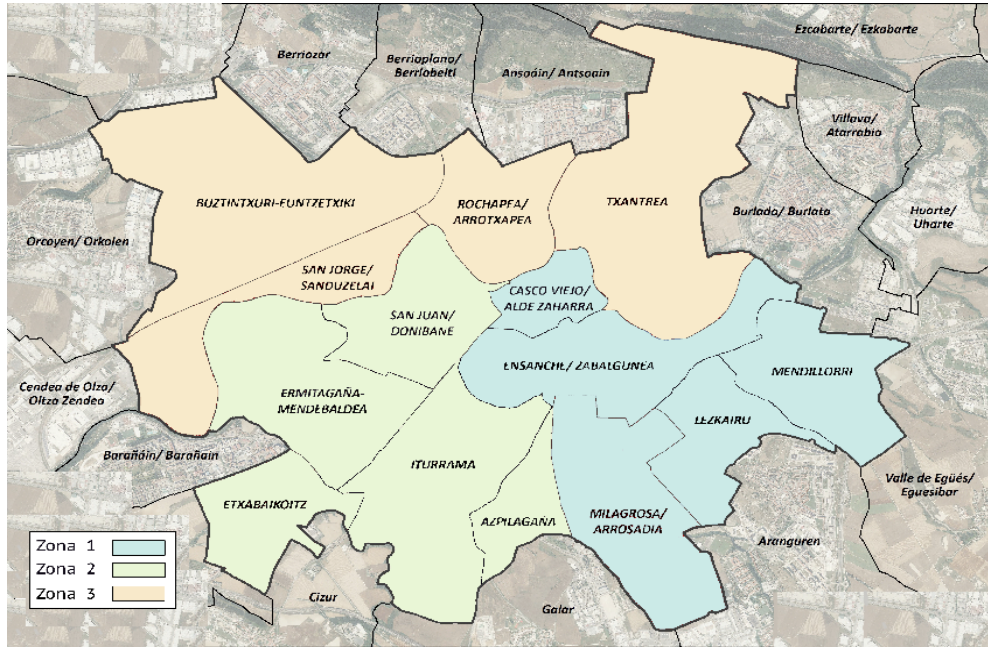


TABLA 4. DISTRIBUCIÓN GENERAL DE LOS HABLANTES POR ZONAS

Nivel de instrucción	Zona 1		Zona 2		Zona 3		Total
	H	M	H	M	H	M	
1: Primaria	3	3	3	3	3	3	18
2: Media	3	3	3	3	3	3	18
3: Superior	3	3	3	3	3	3	18
	9	9	9	9	9	9	
Total	18	18	18	54			

Fuente: Datos procedentes de Lastra y Martín Butragueño, 2016, recalculados ahora con *Stata/MP 17* = StataCorp, 1985-2021.

#### 4.2. Grabaciones y transcripciones

La recogida de los materiales del PRESEEA en Pamplona empezó en diciembre de 2022 y se está llevando a cabo siguiendo los lineamientos generales del proyecto: conversaciones semidirigidas –con consentimiento informado<sup>15</sup> entre un entrevista-

<sup>15</sup> El proyecto general cuenta con la certificación del Comité de ética de Investigación de la Universidad de Alcalá (código CEI/HU/2015/20), que garantiza el cumplimiento de principios éticos en los procedimientos, recopilación y análisis de materiales.

dor y un entrevistado, de al menos 45 minutos de duración, con grabadora a la vista, en contextos propios de cada una de las comunidades de habla (Moreno Fernández, 2005, 2015, 2021a).

Las entrevistas de este tipo buscan conseguir, en lugar de respuestas breves, muestras de habla extensas que, en la medida de lo posible, reproduzcan el estilo utilizado espontáneamente por los hablantes en situaciones familiares en las que la presión de la norma de prestigio es mínima (Recalde *et al.*, 2008: 3). Dentro del grupo de las técnicas de encuestas la entrevista de este tipo supone menor estructuración que las dirigidas y, por su propia naturaleza, más flexibilidad en la obtención de los datos que reflejan la actuación lingüística. Labov (1972, 1991) ha sido una figura esencial en este sentido, pues el autor, interesado en dar cuenta del habla vernácula (habla espontánea que emplean los sujetos en situaciones en las que el autocontrol lingüístico es mínimo), propone, bajo la premisa de la implicación emocional, una serie de módulos temáticos (infancia, familia, peligro de muerte, etc.) para la entrevista que se traducen en un conjunto de preguntas relacionadas bajo un tronco temático común (Moreno Fernández, 2021a).

A partir de todo lo anterior, las entrevistas del PRESEEA-Pamplona se han estructurado teniendo en cuenta los nueve módulos temáticos del proyecto, a saber: saludos, tiempo, lugar donde vive, familia y amistad, costumbres, peligro de muerte, anécdotas importantes en la vida, deseo de mejora económica y final.

Asimismo, se siguen las orientaciones para el tratamiento de cada módulo propuestas por Moreno Fernández (2021a: 18-23), aunque cada uno puede variar según las circunstancias de cada entrevista (anécdotas de interés para el informante, etc.). La intención última es, como señala el autor, que el conjunto de materiales de una comunidad ofrezca muestras proporcionadas de cada uno de los módulos.

Tomando en cuenta la situación particular de Navarra, comunidad en la que, como se ha visto, el castellano y el euskera son lenguas oficiales, en el último módulo, que funciona a manera de cierre de la entrevista, se formulan algunas preguntas para conocer la percepción lingüística que tienen los entrevistados sobre el euskera y sobre el español. Esta información puede ser de gran utilidad en materia de política y planificación lingüísticas, y es de esperar que permita observar cambios según la edad, los antecedentes familiares o incluso el barrio.

Según datos oficiales,<sup>16</sup> un 23,2 % de la población navarra de 16 años o más tiene conocimientos de euskera: un 12,9 % lo habla bien y un 10,3 % tiene algún conocimiento, aunque sin llegar a dominarlo. Asimismo, la población joven de 16 a 24 años presenta la tasa más alta de conocimiento de euskera: el 38 %. Un 25,8 % lo domina y un 12,2 % tiene algún conocimiento sin llegar a hablarlo bien. En función de esta realidad lingüística, hemos propuesto las siguientes preguntas orientativas que serán más o menos productivas o pertinentes en función de las características de cada informante:

---

<sup>16</sup> <https://lenguas.navarra.es/es/lenguas-oficiales-de-navarra>

¿Sabes euskera? ¿Con quién sueles hablar en euskera? ¿Te gustaría aprenderlo? ¿Por qué? ¿Debería ser obligatorio en la enseñanza? ¿Prefieres hablar en euskera o en español? ¿Qué piensas del español en Navarra en relación con el euskera?

En cuanto a las transcripciones, se llevan a cabo a partir de un modelo adaptado del *Standard Generalized Markup language* (SGML), siguiendo las convenciones internacionales de la TEI (*Text Encoding Initiative*) y en dos versiones: con etiquetas y sin etiquetas (Moreno Fernández, 2021b).

## 5. PRIMEROS ESTUDIOS

Las entrevistas del PRESEEA-Pamplona se empezaron a recoger en diciembre de 2022 en distintos espacios (salas de reuniones de la universidad, domicilio del entrevistador, etc.). Todas son producto de la interacción entre un informante y un único entrevistador que comparte con el primero la misma variedad dialectal.

Como suele ser habitual, empezamos a grabar hablantes del grado de instrucción 3. Por esta razón, de momento, las entrevistadoras del equipo y los entrevistados han compartido grupo generacional y nivel de instrucción (Aldunate Martín, 2023). Aunque este hecho podría facilitar una recopilación de resultados de menor artificialidad y más próximos a la realidad social, hemos partido de la premisa de que los interlocutores no se encontraban en una situación conversacional natural y, en consecuencia, a lo largo de la entrevista se ha procurado rebajar la distancia entre la entrevistadora y los entrevistados.

A partir de las primeras grabaciones y transcripciones (hablantes de grado de instrucción 3 de los grupos generacionales 1 y 3), parte del equipo ha empezado a estudiar dos fenómenos lingüísticos: la sufijación apreciativa (Pérez-Salazar, 2023) y las citas en estilo directo como manifestación del discurso referido (Aldunate Martín, 2023).

Asimismo, el equipo está trabajando en dos cuestionarios: uno sobre fórmulas de tratamiento y otro sobre percepciones lingüísticas del español y el euskera en Navarra. El primero ya lo han rellenado los hablantes grabados hasta ahora; el segundo, en proceso, busca completar la información referida en el último módulo de las entrevistas. En ambos casos, la idea no es que se limiten únicamente a los informantes del PRESEEA-Pamplona, sino que a través de ellos se recojan datos en distintas comunidades.

## 6. REFLEXIÓN FINAL

Los estudios sobre el habla como variedad funcional o diafásica del lenguaje a partir de muestras orales auténticas y representativas, creadas, en este caso, a partir de entrevistas semidirigidas (López Morales, 1989; Silva-Corvalán y Enrique-Arias, 2017), garantizarán, sin duda, la representación de Pamplona en la comunidad internacional de habla española que forma parte del PRESEEA.

Asimismo, aunque todavía queda mucho camino por recorrer, la puesta en marcha del PRESEEA-Pamplona facilitará el conocimiento y la difusión de los fenómenos de variación y los procesos de cambio lingüístico en el español de Navarra, comunidad rica y compleja desde el punto de vista histórico-lingüístico, pero también por los cambios acaecidos en los últimos años con la incorporación de hablantes de distintas naciones hispanohablantes y de otras lenguas cuya integración en el territorio es muy desigual.

## REFERENCIAS

- Aldunate Martín, Nadia. (2023). *Contadores de historias: el estilo directo como estrategia narrativa en el discurso oral de jóvenes de Pamplona*, [Trabajo de fin de grado], Navarra: Universidad de Navarra.
- Aragües, Itziar (2000). *Estudio sociolingüístico del léxico del habla de Pamplona*, [Tesis doctoral], Navarra: Universidad de Navarra.
- Aragüez, Itziar (2003). *Y tú ¿cómo dices? Vocabulario actual de Pamplona*, Pamplona: Ayuntamiento de Pamplona.
- Areta, María. (2009). *El léxico disponible de los estudiantes preuniversitarios navarros, Estudio sociolingüístico*, [Tesis doctoral], Navarra: Universidad de Navarra.
- Aznárez, Mónica y Magdalena Romera (eds.). (2016). *Estudio y transmisión del Patrimonio Cultural Inmaterial: aportaciones desde la lingüística, la literatura y la didáctica*. Pamplona: UPNA.
- Blas Arroyo, José Luis. (2019). «La sociolingüística del español en España», En E. Ridruejo Alonso, *Manual de lingüística española*, Berlin / Boston: De Gruyter, pp. 582-613. <https://doi.org/10.1515/9783110362084>.
- Cortés Rodríguez, Luis. (2022). *Los estudios del español hablado entre 1950 y 1999*. Anejo Oralía. Madrid: Arco/Libros.
- Díaz-Campos, Manuel. (ed.). (2011). *The Handbook of Hispanic Sociolinguistics*. Oxford: Wiley-Blackwell.
- Echenique Elizondo, María Teresa. (1987). *Historia lingüística vasco-románica*. Madrid: Paraninfo.
- Erize, Xabier. (2023). «Diálogo con la población castellano hablante navarra en el fomento de la lengua vasca», *Príncipe de Viana* 284: 761-787. DOI: <https://doi.org/10.35462/pv.284.9>
- Fernández Ordóñez, Inés. (2012). «El norte peninsular y su papel en la historia de la lengua española», En S. Gómez Seibane y K. Sinner (eds.), *Estudios sobre tiempo y espacio en el español norteño*, San Millán de la Cogolla: Cilengua, pp. 23-68.
- González Ollé, Fernando. (1969). «La lengua occitana en Navarra». *RDTP*, 25: 285-300.
- González Ollé, Fernando. (1970a). «Vascuence y romance en la historia lingüística de Navarra». *BRAE* 50(189): 31-76.

- González Ollé, Fernando. (1970b). «El romance navarro». *RFE*, 53: 1-4, 45-93.
- González Ollé, Fernando. (1991). «La posición de Navarra en el dominio lingüístico navarro-aragonés». En *Actas del I Congreso de Lingüistas Aragoneses*, Zaragoza, Diputación General de Aragón, pp. 55-68.
- González Ollé, Fernando. (2004). «Navarra, Romania Emersa y ¿Romania Submersa? Aemilianense», *Revista internacional sobre la génesis y los orígenes históricos de las lenguas romances*, 1: 225-270.
- González Ollé, Fernando. (2016). «Vascones y vascuence. Historia (para romanistas) de una relación». *Revista Iberoamericana de Lingüística*, 11: 35-173.
- Gorrochategui, Joaquín. (2006). «Onomástica vascónica y aquitana: elementos para el conocimiento de la historia antigua de Navarra», En *Navarra en la antigüedad. Propuesta de actualización*. Pamplona: Príncipe de Viana, pp. 111-134.
- Hernández Campoy, Juan Manuel. (2011). «Variation and identity in Spain», En J. M. Díaz-Campos (ed.), *Handbook of Hispanic Sociolinguistics*, Oxford: Wiley-Blackwell, pp. 704-727.
- Iribarren, José María. (1984 [1952]). *Vocabulario Navarro. Nueva edición preparada y ampliada por Ricardo Ollaquindia*. Pamplona: Institución Príncipe de Viana.
- Jimeno Jurío, José María. (1994). «El vascuence a las puertas de Pamplona (1790)». *Fontes Linguae Vasconum*, 26(66): 327-334.
- Jimeno Jurío, José María. (1995a). *Historia de Pamplona y de sus lenguas*. Pamplona: Txalaparta.
- Jimeno Jurío, José María. (1995b). «Las lenguas habladas y escritas en Pamplona». *Fontes Linguae Vasconum*, 27 (68): 51-68.
- Labov, William. (1972). *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.
- Labov, William. (1991). «The Intersection of Sex and Social Class in the Course of Linguistic Change». *Language Variation and Change*, 2 (2): 205-254.
- López Morales, Humberto. (1989). *Sociolingüística*. Madrid: Gredos.
- Martínez Pasamar, Concepción y Cristina Taberero, (2012). *Hablar en Navarra. Las lenguas de un reino (1212-1512)*, Pamplona: Gobierno de Navarra-Diario de Navarra.
- Mikelarena, Fernando. (2003). «La evolución demográfica de la población vascoarablante en Navarra entre 1553 y 1936». *Fontes linguae vasconum: Studia et documenta*, 35(92): 183-197.
- Molina, Isabel; Florentino Paredes y Ana María Cestero (2014). *La lengua hablada en Madrid. Corpus PRESEEA-Madrid (distrito de Salamanca). Vol. II. Hablantes de instrucción media*. Alcalá de Henares: Universidad de Alcalá.
- Moreno Fernández, Francisco. (1997). «Metodología del «Proyecto para el Estudio Sociolingüístico del Español de España y América», En *Trabajos de sociolingüística hispánica*, Alcalá de Henares: Universidad de Alcalá, pp. 137-161.
- Moreno Fernández, Francisco. (2005). «Corpus para el estudio del español en su variación geográfica y social. El corpus «PRESEEA»». *Oralia*, 8: 123-139.

- Moreno Fernández, Francisco (2014). *La lengua española en su geografía*, Madrid: Arco/Libros.
- Moreno Fernández, Francisco (2015). *Principios de sociolingüística y sociología del lenguaje*, Barcelona: Ariel.
- Moreno Fernández, Francisco (2021a). «Metodología del «Proyecto para el Estudio Sociolingüístico del Español de España y América» (PRESEEA)», *Documentos PRESEEA de investigación, Documentos de trabajo 1*. DOI: 10.37536/PRESEEA.2021.doc1.
- Moreno Fernández, Francisco (2021b). «Marcas y etiquetas mínimas obligatorias para materiales de PRESEEA», *Documentos PRESEEA de investigación, Documentos de trabajo 2*. DOI: 10.37536/PRESEEA.2021.doc2.
- Nausia Pimoulier, Amaia, Julen Zabalo Bilbao y Txoli Mateos González (2023). *Diversidad y convivencia en Navarra. El euskera y las identidades nacionales*, Pamplona: Eusko Ikaskuntza.
- Pérez-Salazar, Carmela (1995). *El romance navarro en documentos reales del siglo XIV (1322-1349)*, Pamplona: Institución Príncipe de Viana.
- Pérez-Salazar, Carmela (2002). «Manifestaciones del habla coloquial en un tipo de escrito: las cartas de amor (siglos XVI-XVIII)», En M. T. Echenique y J. P. Sánchez (coords.), *Actas del V Congreso Internacional de Historia de la Lengua Española*, Madrid: Gredos, pp. 877-890.
- Pérez-Salazar, Carmela (2007). «La expresión de la existencialidad en cartas privadas (siglos XVII-XVIII): creatividad y evolución», En L. M. Cortés (coord.), *Discurso y oralidad. Homenaje al profesor J. J. de Bustos Tovar*, Madrid: Arco/Libros, pp. 401-416.
- Pérez-Salazar, Carmela (2012). «La sufijación apreciativa en el norte de la Península entre los siglos XVI y XVIII», En J. M. Martín, T. B. Candón y M. R. Zancarrón (eds.), *Actas del IX Congreso Internacional de Historia de la Lengua Española*, Madrid: Newbook, pp.1975-1992.
- Pérez-Salazar, Carmela (2020). «Entre trabajillos, mocetes, añitos y momenticos. Aproximación a la forma y significado del diminutivo en Navarra», *Revista de Filología, Universidad de la Laguna*, 40: 181-203.
- Pérez-Salazar, Carmela (2023). «La sufijación apreciativa en el habla de Pamplona. Preliminares». Comunicación presentada en el *Simposio Internacional PRESEEA. 25 años de sociolingüística hispánica*, Alcalá de Henares: Universidad de Alcalá, 12 al 13 de junio de 2023.
- Recalde, Montserrat, Victoria Vázquez Rosas, Alba Fernández Sanmartín y Marcos García Salido (2008) «Reflexiones metodológicas sobre la técnica de entrevista semidirigida», En *Actas del XV Congreso Internacional de la ALFAL*, Montevideo: Universidad de La República.
- Rioja Arano, Catalina (1997). «Rasgos lingüísticos de la Ribera de Navarra en las obras de José María Iribarren». *Príncipe de Viana*, 58(211): 445-474.

- Romera, Magdalena y Nayim Medina (2014). «Castellano y euskera en contacto en Navarra. El seseo vasco en los datos del Archivo del Patrimonio Inmaterial de Navarra». En P. Salaberri (coord.), *El Patrimonio Cultural Inmaterial: ámbito de la tradición oral de las particularidades lingüísticas*, Pamplona: Cátedra Archivo del Patrimonio Inmaterial de Navarra-UPNA, pp. 35-52.
- Silva-Corvalán, Carmen y Andrés Enrique-Arias (2017). *Sociolingüística y pragmática del español*, Washington, D.C.: Georgetown University Press.
- Saralegui, Carmen (1997). *El dialecto navarro en los documentos del Monasterio de Irache (958-1397)*. Pamplona: Institución Príncipe de Viana.
- Saralegui, Carmen (1985). «Caracterización lingüística de las hablas ribereñas a través de un escrito de José María Iribarren». *RILCE*, 1(1): 113-136.
- Saralegui, Carmen (2011). «Sobre geografía lingüística de Navarra: de nuevo el norte y el sur». *AFA*, 67: 75-112.
- Saralegui, Carmen y Cristina Taberero (2002). *Navarrismos en el Diccionario de la Real Academia Española*. Pamplona: Gobierno de Navarra, 2ª ed. aumentada.
- Silva-Corvalán, Carmen (2001). *Sociolingüística y pragmática del español*. Washington D.C.: Georgetown University Press.
- Taberero, Cristina (2008). «Disponibilidad léxica y contacto de lenguas». *Oihenart*, 23: 545-565.
- Taberero, Cristina; Pichel, Ricardo y Andrés Enrique-Arias (2020). «El español en contacto con otras lenguas peninsulares: los corpus diacrónicos de Galicia, Navarra y Mallorca». *Estudios de lingüística del español*, 41: 177-217.
- Vilches Plaza, Carlos (2002). *Evolución de las actitudes ante el euskera de los universitarios de Navarra: 1992-2002* [Tesis doctoral], Universidad Pública de Navarra.
- Wulff Alonso, Fernando (2009). «Vascones, autoctonía, continuidad, lengua. Entre la historia y la historiografía», En J. Andreu, (ed.). *Los vascones en las fuentes antiguas. En torno a una etnia en la Antigüedad peninsular*, Barcelona, Edicions Universitat Barcelona, pp. 23-56.



## EL PRESEEA EN VALENCIA. HISTORIA Y RESULTADOS

José Ramón Gómez Molina  
*Universidad de Valencia*

### 1. INTRODUCCIÓN

Un hito importante para el impulso de la sociolingüística hispánica fue la propuesta de Silva-Corvalán en el *Congreso de la Lengua Española* celebrado en Sevilla en 1992 sobre la conveniencia de ampliar el estudio de la norma culta del mundo hispánico, proyecto iniciado por Lope Blanch (1968, 1977, 1986), y la prioridad de un estudio sociolingüístico coordinado del español hablado en las principales ciudades de España y América. Silva Corvalán (1994: 400-402) argumentó la necesidad e importancia vital de este proyecto porque nos permitiría el esclarecimiento de ciertos aspectos controvertidos de la lengua española, así como el avance de la teoría sociolingüística en cuestiones relacionadas con la determinación de la variación y/o del cambio lingüístico.

En ese marco, el objetivo de este capítulo es describir y valorar, de forma sucinta, nuestra aportación al Proyecto para el estudio del español de España y América (PRESEEA), durante más de dos décadas tanto en la elaboración del macrocorpus de español hablado como en el desarrollo metodológico y análisis de los datos sociolingüísticos. La comunidad de habla objeto de estudio es Valencia y su área metropolitana (1.300.000 habitantes en el año 2000), y la tarea ha sido realizada por el equipo PRESEVAL (Proyecto para el estudio del español en Valencia).

### 2. EL PRESEEA EN VALENCIA

La labor iniciada en 1990 por el grupo Val.Es.Co. (Valencia Español Coloquial), surgido en el Departamento de Filología Española de la Universidad de Valencia

y dirigido por el profesor Briz Gómez (1995), se amplió y complementó en 1996 puesto que algunos profesores nos integramos, además, en el proyecto panhispánico PRESEEA. Asistimos a su puesta en marcha en el XI Congreso Internacional de la Asociación de Lingüística y Filología de la América Latina (ALFAL), celebrado en Las Palmas de Gran Canaria, y participamos en las distintas reuniones celebradas en la Universidad de Alcalá los dos años siguientes. En estos encuentros se concretaron los objetivos y la metodología para investigar los procesos de variación y/o de cambio lingüístico en el español del ámbito hispánico.

El equipo PRESEVAL, coordinado por el profesor José Ramón Gómez, ha aportado un corpus de español hablado obtenido de acuerdo con las pautas metodológicas comúnmente establecidas para todos los equipos del proyecto y que contiene muestras de habla representativas de esta área urbana con una cantidad próxima al medio millón de palabras. Por otra parte, la utilidad y aplicaciones de los dos corpus de lengua oral, Val.Es.Co. y PRESEVAL, permiten conocer la estructura y el uso de la lengua española en esta comunidad de habla bilingüe castellano-catalán en dos registros comunicativos: coloquial y semiformal. Nuestra tarea investigadora ha contado en todo momento con ayudas y subvenciones de diferentes entidades: Universitat de València, Generalitat Valenciana, Ministerio de Ciencia y Tecnología, así como Fondos FEDER.

Hoy, más de veinte años después y con una población de 1.581.057 habitantes en el área metropolitana, se están recogiendo nuevos materiales que conformarán un nuevo corpus y que permitirá comparar las distintas comunidades de habla bilingüe castellano-catalán (Valencia, Barcelona y Palma de Mallorca). Estas nuevas entrevistas se irán transcribiendo y etiquetando según los requisitos establecidos por el PRESEEA y se publicarán para su uso por otros investigadores.

Centrándonos en la tarea realizada por nuestro equipo como aportación al proyecto coordinado panhispánico, conviene destacar de forma descriptiva:

- La recopilación del corpus. Dado el número de habitantes de esta comunidad de habla bilingüe se obtuvo un corpus de 72 entrevistas semidirigidas recogidas entre 2000 y 2006: sociolecto alto (24), medio (24) y bajo (24) combinando las variables sociales de género, edad, nivel de estudios y lengua habitual (Gómez Molina 2001, 2005, 2007). Además, contamos con 24 grabaciones de español coloquial, registradas entre hablantes de distintos niveles socioculturales, que nos permiten establecer correspondencias entre ambos registros.
- El etiquetado de dicho corpus según las normas acordadas en el proyecto coordinado y adaptadas a las convenciones TEI/SGML (*Text Encoding Initiative/Standard Generalized Markup Language*).
- La digitalización de las grabaciones de todas las muestras de habla y su almacenamiento en Creative y mp3.
- Para su difusión se enviaron todas las publicaciones y material digitalizado a los demás equipos del proyecto coordinado (Alcalá, Granada, Málaga, Las

Palmas, Lérída, Sevilla) y a otros equipos del PRESEEA (Barranquilla, Oviedo, Monterrey, Caracas).

- La elaboración y actualización de la página electrónica de PRESEVAL (<http://www.uv.es/preseval>).
- Los fenómenos sociolingüísticos estudiados han sido tanto los establecidos en las reuniones periódicas de coordinación científica del PRESEEA como aquellos relevantes para nuestra comunidad de habla.
- Se ha procedido a la difusión y transferencia de los resultados a la comunidad científica mediante la participación en múltiples congresos internacionales y la publicación de los mismos.
- Hemos elaborado las fichas de codificación provisional de cinco variables sociolingüísticas para su utilización por diferentes equipos del proyecto coordinado y su posible estudio comparado: /d/ intervocálica, fonemas palatales sonoros lateral /ʎ/ y central /j/, presencia-ausencia de la preposición *de*/Ø, usos de *deber(de)* y usos impersonales del verbo *haber*.
- Hemos colaborado en la guía de codificación definitiva de dos fenómenos sociolingüísticos: ‘estudio del verbo impersonal *haber*’ y estudio del *(de)queísmo*’.
- Y en esta línea de cooperación cabe señalar, junto a las numerosas y periódicas reuniones de trabajo durante estos años con los equipos de otras universidades, que nuestro equipo realizó la coordinación científica de todos los equipos durante el periodo 2004-2010.

Por su parte, la tarea investigadora realizada a partir del corpus lingüístico obtenido ha sido constante y centrada en el establecimiento de patrones sociolingüísticos tanto en el análisis de variables interesantes para nuestra comunidad de habla como el análisis de variables comunes para los distintos equipos del PRESEEA, hecho que permite el estudio comparado de los resultados obtenidos en diferentes ciudades.

Enumeramos a continuación los distintos fenómenos sociolingüísticos analizados en sus diferentes niveles:

1. En la variación fónica se han estudiado dos variables, consideradas prioritarias por el proyecto coordinado:
  - ‘elisión de la /d/ en posición intervocálica’ (Gómez Molina y Gómez Devís, 2010; Gómez Molina, 2012, 2013b; Gómez Molina, Molina y Paredes, 2012); y
  - ‘el *yeísmo*, realización del fonema lateral palatal /ʎ/ como fricativo palatal [y]’ (Gómez Molina y Gómez Devís, 2016).
2. En el nivel morfosintáctico se ha establecido cierta cantidad de variables sociolingüísticas en función de su utilidad para la elaboración de una gramática del español hablado:
  - ‘verbos semiauxiliares con especial atención a las construcciones perifrásticas modales de obligación y de posibilidad (*deber, deber de*)’ (Gómez

- Molina, 2006, 2008b, 2013e);
- ‘paradigma verbal de futuro: futuro morfológico, futuro perifrástico y presente pro futuro’ (Buzón García, 2013; Gómez Molina y Buzón García, 2015);
  - ‘concordancia del verbo *haber* con usos impersonales’ (Gómez Molina, 2013a, 2013d; Gómez Molina y Gómez Devís, 2020);
  - ‘construcciones con ‘*de* /  $\emptyset$ ’ en cláusulas integradas o completivas *-(de) queísmo-*’ (Gómez Molina, 2011a, 2011b, 2013c);
  - ‘laísmo, leísmo y loísmo’ (Roselló Verdeguer, 2017a);
  - ‘clíticos: posición y duplicación’ (Roselló Verdeguer, 2018); y
  - ‘expresión y posición del sujeto’ (Roselló Verdeguer, 2021).
3. En el nivel léxico-semántico se ha realizado un estudio sobre ‘frecuencias léxicas y su estratificación’ (Terrádez Gurrea, 2013).
  4. En el nivel pragmático se han considerado prioritarios los estudios sobre patrones de interacción:
    - ‘las formas de tratamiento’ (Roselló Verdeguer, 2013, 2017b); así como
    - ‘la atenuación y la intensificación’ (Albelda Marco, 2004, 2012, 2013; Estellés y Cabedo, 2017).
  5. Finalmente y dado que se trata de una comunidad de habla bilingüe, tan importante como la identificación de las correlaciones sociolingüísticas ha sido la investigación de los fenómenos de transferencias morfosintácticas entre castellano y catalán, en su variedad valenciano, para comprobar el grado de permeabilidad de ambos sistemas (Gómez Molina, 2004; Gómez Devís, 2013) las transferencias léxicas castellano  $\leftarrow \rightarrow$  catalán (Gómez Molina, 2008a) y la tipología de los cambios de código (Gómez Molina, 2004). En este contexto y relacionada con los usos lingüísticos se han investigado las actitudes lingüísticas hacia las dos lenguas en contacto (Gómez Molina, 1998; Gómez Molina, Buzón García y Gómez Devís, 2017).

Tras esta exposición descriptiva, presentamos a continuación un breve comentario valorativo sobre algunas de las variables lingüísticas investigadas, el establecimiento de sus patrones sociolingüísticos y la consideración de si se trata de un fenómeno estable o en proceso de cambio lingüístico.

### 3. EL ANÁLISIS DEL HABLA DE VALENCIA

#### 3.1. Nivel fónico. Variable ‘/d/ intervocálica’

La elisión de la /d/ intervocálica en nuestra comunidad de habla es un fenómeno reducido y variable en el entorno general (14.05%); pero ese proceso se intensifica cuando precede inmediatamente una vocal tónica (18.1%) y tiende a generalizarse de

forma bastante automática en el entorno /-ado/ (48.2%). Una vez que se inicia este proceso de cambio, la tendencia a conservar las formas subyacentes de las palabras (restricción de fidelidad) evita su avance frenándolo en los lexemas (6.8%), especialmente en lexías marcadas (pares mínimos, nombres propios, términos marcados).

La relación jerarquizada de las variables independientes seleccionadas y los rangos obtenidos en el análisis multivariante confirma que los factores que más inciden en la elisión de la /d/ intervocálica son lingüísticos o internos: vocal anterior [a-, o-], categoría gramatical (participio, determinante, adjetivo), vocal posterior [-o], límite morfemático (morfema derivativo y gramatical), estructura acentual de la palabra (paroxítona), homofonía (participio-adjetivo-nombre) y entorno consonántico secuencial (hay [d] elididas en el entorno); ello nos indica que la elisión se ve favorecida fundamentalmente por factores lingüísticos.

Al mismo tiempo, el modelo probabilístico nos indica la evolución del proceso: la generalización de la elisión avanzará en el contexto fónico vocálico [ádo → á<sup>o</sup>o → áØo] y en menor medida en los contextos [ódo → ó<sup>o</sup>o → óØo] [óða → ó<sup>o</sup>a → óØa]. Se trataría, pues, de un proceso de cambio fónico condicionado léxica y gramaticalmente, tendente a facilitar y simplificar la pronunciación natural y no marcada de la dental, restringido a contextos gramaticales de gran frecuencia de uso.

En cuanto al patrón sociolingüístico observamos tres factores extralingüísticos relevantes: cuando el hablante pertenece a la misma red social que sus interlocutores y tiene una edad aproximada, si es un individuo de la segunda generación (35-55 años) y si es castellanohablante (monolingüe). La tendencia general hacia la elisión queda favorecida por el sociolecto medio y frenada por el sociolecto alto, mientras que la variable sexo no ejerce especial incidencia. Se trata de un fenómeno de variación estable (*input* general, 0,140).

### 3.2. Nivel fónico. Variable ‘yeísmo’, pérdida de la distinción fonológica de los fonemas palatales sonoros lateral /ʎ/ y central /j/

El avance del yeísmo es una opinión compartida por varios autores. Para el análisis sociolingüístico de la variable dependiente (fonema /ʎ/) se han establecido dos variantes: lateral o canónica [ʎ] y no lateral, variante que agrupa los alófonos [j]+[dʒ]+[j] en el castellano de esta comunidad de habla<sup>1</sup>.

Se trata de un hecho que afecta al esquema fonológico de nuestra lengua y, si bien es cierto que esta confluencia de alófonos tiene un origen bajomedieval, parece legítimo plantear la hipótesis de la existencia de un estado de latencia del fenómeno yeísta. Pero el rápido desarrollo de este proceso de deslateralización a favor de la palatal central durante el siglo XX comporta que «su presencia en amplias zonas, así

<sup>1</sup> Aproximante fricativa [j], sonido africado [dʒ] y semiconsonántico [j].

como su creciente expansión, hacen del yeísmo un fenómeno aceptado en la norma culta» (RAE, 2005: 682).

Los resultados del análisis de regresión multivariante nos permiten señalar que el estado de latencia de este fenómeno está dando paso a un cambio lingüístico también en nuestra comunidad de habla. Señalamos algunos aspectos relevantes:

- El proceso de desfonologización queda favorecido, mayoritariamente, por los factores siguientes: estrato generacional joven (20-34 años,  $p=0.809$ ), cuando se habla sobre temas cotidianos ( $p=0.791$ ), si residen en la ciudad de Valencia ( $p=0.703$ ) frente a los pueblos del área metropolitana, y si pertenecen al género femenino ( $p=0.719$ ). Por su parte, mantienen la pronunciación lateral de [ʎ] o retrasan el proceso de deslateralización los mayores de 55 años ( $p=0.209$ ), cuando se tratan temas técnicos ( $p=0.209$ ) y si pertenecen al sexo masculino ( $p=0.281$ ).
- La incidencia de los factores lingüísticos queda debilitada en este proceso, pues solo la vocal anterior /i-/ facilita el yeísmo.
- Otras variables sociales que muestran cierta incidencia en el proceso de desfonologización son el nivel sociocultural (rango 0.250) y la lengua habitual de los hablantes (rango 0.218). No obstante, los individuos de nivel sociocultural bajo y los castellanohablantes ofrecen pesos probabilísticos con valores explicativo y predictivo importante.

Si estudiamos el patrón sociolingüístico, los análisis de regresión logística binomial corroboran que son los factores sociales y estilísticos los de mayor incidencia en la variación sociolingüística del yeísmo. Estamos ante un patrón sociolingüístico en el que intervienen esencialmente variables sociales (edad, rango 0.600; residencia, rango 0.479; sexo, rango 0.438) y factores estilísticos (campo o tema de conversación, tenor-edad, tenor-estatus y tenor-proximidad). Por su parte, si nos centramos en el nivel sociocultural (rango 0.248) para establecer el patrón sociolingüístico, nos hallamos ante un patrón de estratificación lineal (sociolectos bajo, peso 0.648; medio, 0.449; alto, 0.400).

### 3.3. Nivel morfosintáctico. Variable ‘uso de las perífrasis modales ‘deber’ y ‘deber de’

El estudio cuantitativo de uso de las construcciones perifrásticas de modalidad deóntica en el corpus PRESEVAL nos muestra las frecuencias según sociolectos en la tabla siguiente. Conviene tener presente que en catalán la construcción paralela *deure+infinitiu* expresa probabilidad y la perífrasis *haver de+infinitiu* indica obligación.

TABLA 1. FRECUENCIAS DE USO SEGÚN SOCIOLECTOS

Estructuras	Sociolecto alto	Sociolecto medio	Sociolecto bajo	Global
<i>tener que + inf</i>	408 (54.6%)	496 (67.6%)	314 (74.1%)	1218 (63.9%)
<i>deber + inf</i>	175 (23.4%)	124 (16.9%)	54 (12.7%)	353 (18.6%)
<i>deber de + inf</i>	90 (12.1%)	73 (9.9%)	52 (12.3%)	215 (11.3%)
<i>haber de + inf</i>	74 (9.9%)	41 (5.6%)	4 (0.9%)	119 (6.2%)
<b>TOTAL</b>	747 (100.0%)	734 (100.0%)	424 (100.0%)	1905

Por su parte, si comparamos los usos de los auxiliares perifrásticos ‘*deber*’ para expresar modalidad deóntica y probabilística, y ‘*deber de*’ para manifestar modalidad epistémica, se comprueba que la diferenciación establecida por la norma en el segundo caso no queda refrendada plenamente por la conciencia lingüística de los hablantes en el español oral.

TABLA 2. USOS MODALES DE DEBER(DE) + INFINITIVO

VALOR CONTEXTUAL	<i>deber</i> + infinitivo	<i>deber de</i> + infinitivo
obligación + necesidad	66.9%	33.1%
probabilidad	51.7%	48.3%

Los patrones de estratificación etaria indican que el mejor uso canónico de ambas perífrasis corresponde a los mayores de 55 años: ‘*deber* + infinitivo’ (85.4%) y ‘*deber de* + infinitivo’ (41.7%). El análisis de regresión logística nos permite conocer que, de todos los factores estudiados, las variables que favorecen o propician dicha elección son: tres de carácter lingüístico (tipo sintáctico de oración, tipo semántico de verbo auxiliado y tipo de sujeto), dos de carácter estilístico (registro y valor contextual-discursivo) y dos de índole sociológica (lengua habitual y nivel sociocultural).

Mostramos un ejemplo del índice de variabilidad (*input*), influencia de la variable social en la variación conjunta (rango) y pesos probabilísticos del nivel sociocultural (*p*):

- Modalidad deóntica. Variante: ‘*deber* + infinitivo’. *Input*: 0.669. Total: Nivel sociocultural (rango 0.353): Alto ( $p = 0,721$ ), Medio ( $p = 0,491$ ), Bajo ( $p = 0,368$ )
- Modalidad epistémica. Variante: ‘*deber* + infinitivo’. *Input*: 0.517. Nivel sociocultural (rango 0.130): Medio ( $p = 0,520$ ), Alto ( $p = 0,519$ ), Bajo ( $p = 0,392$ )

Globalmente, el patrón sociolingüístico obtenido indica que estamos ante un caso de variación sociolingüística estable -estratificación continua-, donde los análisis realizados muestran que la perífrasis sin preposición ‘*deber*’ incrementará su uso

para expresar probabilidad y que la presencia de la preposición ‘*deber de*’ actúa a favor del énfasis expresivo de la obligación, probablemente por influencia de la perífrasis del catalán ‘*haver de*’.

### 3.4. Nivel morfosintáctico. Variable ‘Expresión de la futuridad: usos del futuro perifrástico, futuro simple y presente con valor de futuro’

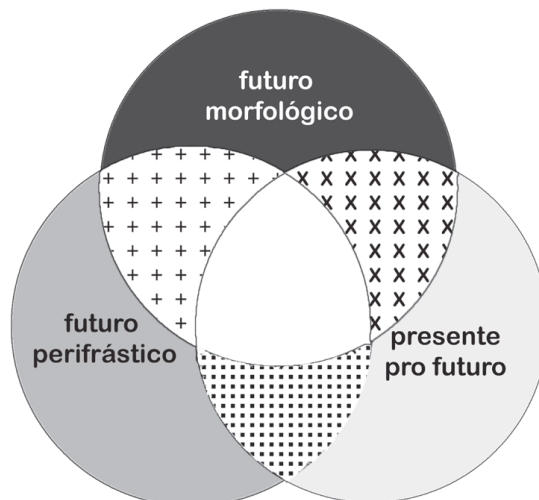
Se han cuantificado exclusivamente los usos del futuro con valor temporal y quedan excluidos los usos del futuro epistémico (“Imagino que será el miedo»), las fórmulas rutinarias (“Después, ya veremos») y las paremias o refranes. Estos son los resultados obtenidos en el estudio descriptivo:

**TABLA 3. FRECUENCIAS DE USO PARA LA ‘EXPRESIÓN DE FUTURIDAD’**

	Frecuencia	Porcentaje
Futuro morfológico	411	38,2%
Futuro perifrástico	510	47,4%
Presente pro futuro	154	14,3%
Total	1075	100,0%

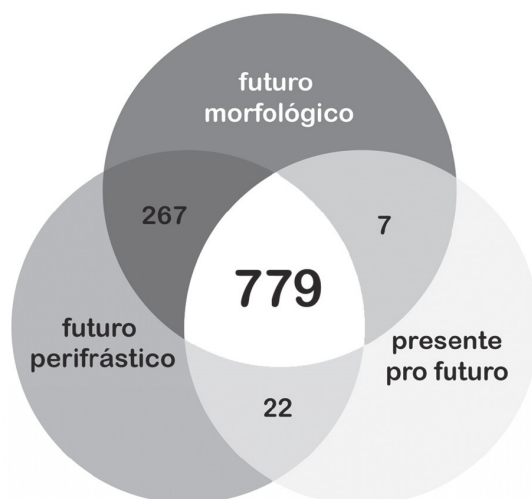
Dado que esta investigación analiza una variable lingüística con tres variantes se ha estudiado una doble confrontación: en primer lugar, las ocurrencias de las dos alternativas que han demostrado su capacidad para neutralizarse mutuamente mediante un análisis de regresión logística binomial.

**FIGURA 1. ALTERNATIVAS Y NEUTRALIZACIONES**



Y posteriormente se ha aplicado un modelo de regresión logística multinomial que solo considera como aceptables aquellas ocurrencias de las tres variantes que puedan neutralizarse con las otras dos alternativas simultáneamente. En nuestra opinión, este análisis de la expresión de futuro como fenómeno en su conjunto parece más interesante que el análisis de la alternancia entre dos formas determinadas, por más que ostenten una gran frecuencia de uso o representen un supuesto conjunto de opciones gramaticalizadas frente a una opción expresiva eventual. Para cumplir las condiciones de este modelo estadístico es necesario revisar las ocurrencias de nuestro corpus y verificar aquellas que cumplen dicha condición.

**FIGURA 2. FRECUENCIA DE OCURRENCIAS Y CAPACIDAD DE NEUTRALIZACIÓN**



Tal como puede observarse, el conjunto de 1075 ocurrencias iniciales quedaría reducido a 779 casos aceptables, que corresponde a los usos de las tres fórmulas expresivas de futuro susceptibles de neutralizarse en el discurso con cualquiera de las dos alternativas.

Según los resultados obtenidos en el doble análisis de regresión logística, tres variables independientes lingüísticas constituyen el conjunto de factores más clara e intensamente condicionantes de la preferencia de los hablantes de la comunidad de habla de Valencia por una de las opciones expresivas de futuridad: ‘especificación temporal’ (rango 0,635), ‘tipo semántico de verbo’ (rango 0.558) y ‘modalidad oracional’ (rango 0.545). En un segundo nivel de incidencia se encuentran las variables pragmáticas ‘evidencialidad’ y ‘acto de habla’, así como ‘presencia de perífrasis en el cotexto’ y ‘persona gramatical’. Finalmente hay un nutrido grupo de variables in-

dependientes de incidencia mínima: ‘estructura oracional’, ‘sexo’, ‘lengua habitual’, ‘modalidad epistémica’, ‘persona gramatical’, ‘nivel sociocultural’, ‘modo de vida’, ‘tipo de texto’, ‘campo’, ‘tipo de predicado’, ‘edad’ y ‘tenor-edad’, influyentes sobre la preferencia en diferentes niveles y de manera diversa en función de las variantes afectadas.

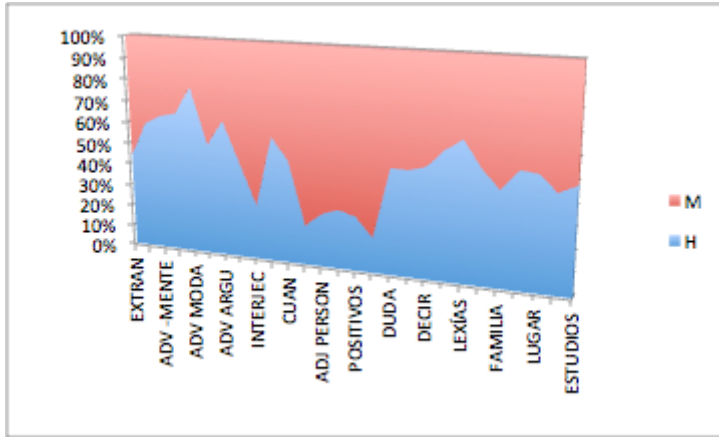
Atendiendo al patrón sociolingüístico estamos ante una variable inmersa en un proceso de cambio lingüístico que promueve la extensión de la variante perifrástica a un número de contextos cada vez más amplio y diversificado en detrimento de una alternativa morfológica. Por su parte, la acción del presente pro futuro en su competencia por áreas temporales intermedias entre ambas alternativas gramaticales podría justificar el rápido avance del futuro perifrástico, en tanto esta opción eventual de futuro podría estar disputando al futuro morfológico sus franjas temporales naturales (las áreas relativamente alejadas del momento del habla, especialmente con indicaciones temporales explícitas) para cederlas posteriormente a la variante de futuro analítico.

### **3.5. Nivel léxico-semántico**

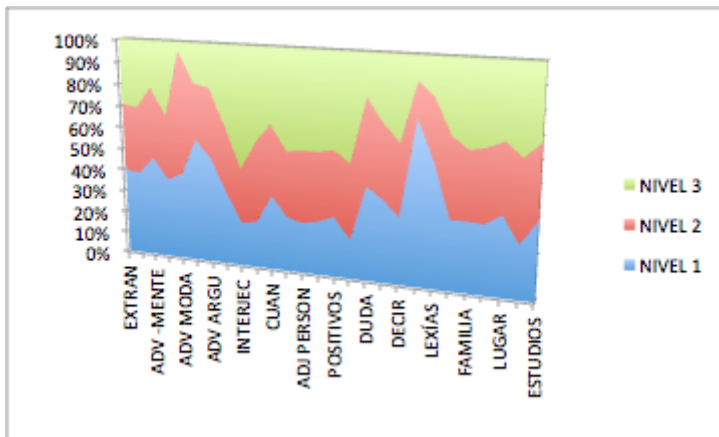
En el nivel léxico-semántico, comparando los corpus de PRESEVAL y Val. Es.Co., se ha realizado un análisis cuantitativo de la estratificación léxica según las tres variables sociales utilizadas mayoritariamente en el proyecto PRESEEA: edad, sexo y nivel de estudios. El estudio abarca: coloquialismos, deícticos, proformas, léxico específico, así como colocaciones y locuciones; además, se ha investigado la estratificación socioléxica relacionada con extranjerismos, adverbios de distinta índole: –mente, temporales, modales, conversacionales y argumentativos; proformas, cuantificadores, adjetivos, verbos, diminutivos, etc.

En líneas generales, aparece una clara estratificación en todos los casos, tanto en lo que respecta a la frecuencia de algunos vocablos, como en lo que se refiere a la presencia/ausencia de ciertos recursos léxicos.

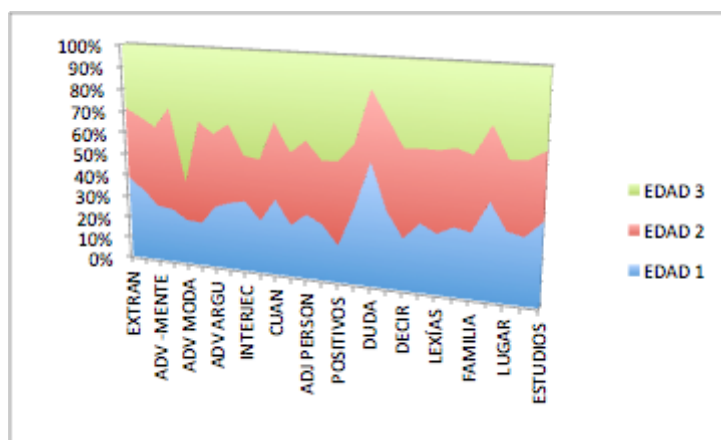
**GRÁFICO 1. FRECUENCIAS DE VOCABLOS SEGÚN SEXO: M (MUJERES), H (HOMBRES)**



**GRÁFICO 2. FRECUENCIAS DE VOCABLOS SEGÚN NIVEL DE ESTUDIOS: 1 (ALTO), 2 (MEDIO), 3 (BAJO)**



**GRÁFICO 3. FRECUENCIAS DE VOCABLOS SEGÚN EDAD: 1 (JOVEN), 2 (ADULTO), 3 (SENIOR)**



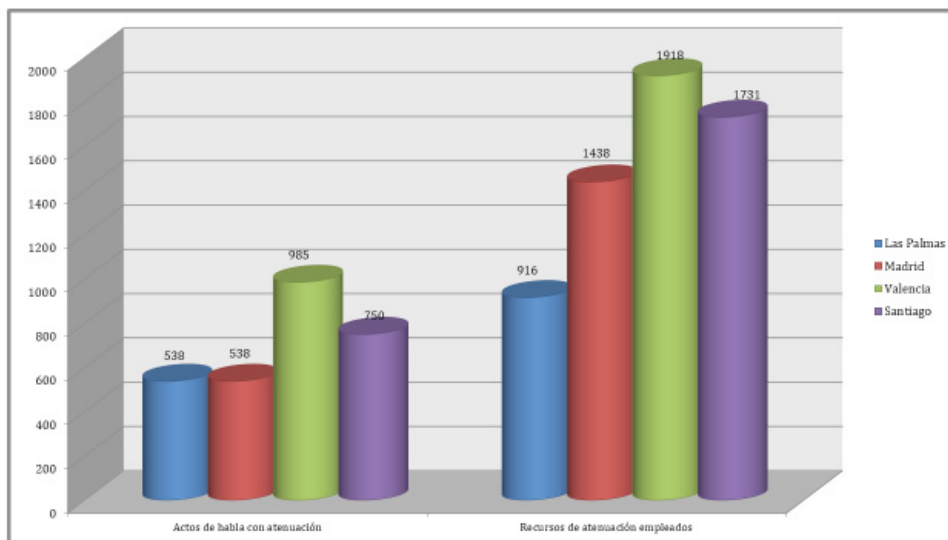
El análisis conjunto de los resultados anteriores nos permite extraer las siguientes conclusiones: la variable sociológica de mayor significatividad es la de sexo (H, hombre; M, mujer), seguida por la de nivel de instrucción (1, alto; 2, medio; 3, bajo) y, en menor medida, la edad (1, 20-34 años; 2, 35-55 años; 3, mayores de 56 años). Y desde la perspectiva estadística inferencial, el primer grupo de edad correlaciona con el nivel sociocultural alto mientras que el tercer grupo de edad correlaciona con el sociolecto bajo.

### 3.6. Nivel pragmático

En el nivel pragmático se han llevado a cabo análisis sociolingüísticos sobre dos fenómenos: ‘la atenuación’ y ‘las fórmulas de interacción (formas de tratamiento)’.

La atenuación ha sido uno de los fenómenos discursivos priorizados en el PRE-SEEA. Se ha realizado un estudio coordinado y comparado entre los equipos de Valencia, Alcalá-Madrid, Las Palmas y Santiago de Chile, donde quedan patentes diversos patrones geolectales.

**GRÁFICO 4. FRECUENCIA DE USO COMPARADAS DE ACTOS DE HABLA Y DE ESTRATEGIAS DE ATENUACIÓN**



Para la caracterización lingüística y sociolingüística de la atenuación se ha tomado como unidad de análisis el acto de habla. La variable lingüística corresponde al número de estrategias de atenuación por acto de habla (se han establecido 25 tipos de recursos lingüísticos). La hipótesis es que la utilización de mayor número de recursos supone un grado mayor de atenuación.

En cuanto al patrón sociolingüístico de Valencia se observa que la mayor atenuación aparece en las mujeres (52.3%), los adultos (41.6%) y hablantes del sociolecto medio (52.4%). Por su parte, la estrategia más usada para atenuar es la impersonalización.

### 3.7. Transferencias y cambio de código

Por último, dada la situación de dos lenguas en contacto, tan importante como la identificación de las correlaciones sociolingüísticas es la investigación de los casos de transferencias y de cambio de código. Los resultados del estudio descriptivo señalan la presencia de transferencias lingüísticas (interferencias, 12.6%, y préstamos, 11,9%) y de cambios de código (75.5%), bien sean de etiqueta, intraoracional o interoracional.

Los factores sociales con mayor relevancia en cada uno de dichos fenómenos lingüísticos pueden observarse en las tablas siguientes según el sexo, la edad, el nivel de estudios y la lengua habitual de los hablantes.

**TABLA 4. FENÓMENOS DE CONTACTO LINGÜÍSTICO SEGÚN SEXO**

	Interferencias	Préstamos	Cambios de código
Hombres	25 (38.5%)	37 (60.6%)	231 (58.9%)
Mujeres	40 (61.5%)	24 (39.4%)	157 (41.1%)
Total	65	61	388

**TABLA 5. FENÓMENOS DE CONTACTO LINGÜÍSTICO SEGÚN EDAD**

	Interferencias	Préstamos	Cambios de código
20-34 años	9 (13.8%)	6 (9.8%)	84 (21.6%)
35-55 años	25 (38.5%)	21 (34.4%)	143 (36.8%)
>55 años	31 (47.7%)	34 (55.8%)	161 (41.6%)
Total	65	61	388

**TABLA 6. FENÓMENOS DE CONTACTO LINGÜÍSTICO SEGÚN NIVELES SOCIOCULTURALES**

	Interferencias	Préstamos	Cambios de código
Alto	9 (13.8%)	26 (42.6%)	89 (22.9%)
Medio	16 (24.6%)	10 (16.4%)	112 (28.9%)
Bajo	40 (61.6%)	25 (41.0%)	187 (48.2%)
Total	65	61	388

**TABLA 7. FENÓMENOS DE CONTACTO LINGÜÍSTICO SEGÚN TIPO DE HABLANTE**

	Interferencias	Préstamos	Cambios de código
Castellanohablante	15 (23%)	31 (51%)	99 (25.5%)
Bilingüe	50 (77%)	30 (49%)	289 (74.5%)
Total	65	61	388

**REFERENCIAS**

- Albelda Marco, Marta (2004). «Cortesía en diferentes situaciones comunicativas. La conversación coloquial y la entrevista semiformal». En D. Bravo y A. Briz (eds.), *Pragmática sociocultural. Estudios sobre cortesía en español*. Barcelona: Ariel, pp. 109-134.
- Albelda Marco, Marta (2012). «Variación sociolingüística en las estrategias de atenuación del corpus PRESEEA-VALENCIA del sociolecto alto». *Actas del XVI Congreso ALFAL*. Alcalá de Henares, pp. 1857-1866
- Albelda Marco, Marta (2013). «La atenuación: tipos y estrategias». En J. R. Gómez Molina (coord.), *El español de Valencia. Estudio sociolingüístico*. Berna, Peter Lang, pp. 315-343.
- Briz Gómez, Antonio (coord.) (1995). *La conversación coloquial. Materiales para su estudio*. Anejo XVI de *Cuadernos de Filología*. Valencia: Universidad de Valencia.

- Buzón García, José María (2013). «La expresión de la futuridad: variación y cambio lingüísticos». En J. R. Gómez Molina (coord.), *El español de Valencia. Estudio sociolingüístico*. Berna, Peter Lang, pp. 145-182.
- Estellés, María y Adrián Cabedo (2017). «La atenuación fónica en entrevistas (proyecto PRESEEA) y en conversaciones (corpus Val.Es.Co): Un estudio de campo». *Linred, Lingüística en la Red*, 15: 1-16.
- Gómez Devís, María Begoña (2013). «Fenómenos frecuentes en el contacto lingüístico castellano-catalán (en su variedad valenciano)». En J. R. Gómez Molina (coord.), *El español de Valencia. Estudio sociolingüístico*. Berna: Peter Lang, pp. 347-370.
- Gómez Molina José R. (1998). *Actitudes lingüísticas en una comunidad bilingüe y multilectal. El área metropolitana de Valencia*. Anejo XXVIII de *Cuadernos de Filología*. Universidad de Valencia.
- Gómez Molina José R. (2001). *El español hablado de Valencia. Materiales para su estudio. I. Nivel sociocultural alto*. Anejo XLVI de *Cuadernos de Filología*. Universidad de Valencia.
- Gómez Molina José R. (2004). «Transferencia y cambio de código en una comunidad bilingüe: Área metropolitana de Valencia». *Contextos*, XVII-XVIII 33-36: 309-360.
- Gómez Molina José R. (2005). *El español hablado de Valencia. Materiales para su estudio. II. Nivel sociocultural medio*. Anejo LVIII de *Cuadernos de Filología*. Universidad de Valencia
- Gómez Molina José R. (2006). «Gramática y uso de las perífrasis verbales de modalidad obligativa». *Actas del XVI Congreso de la Sociedad Chilena de Lingüística*. Universidad Austral de Chile, pp. 237-250.
- Gómez Molina José R. (2007). *El español hablado de Valencia. Materiales para su estudio. III. Nivel sociocultural bajo*. Anejo LXI de *Cuadernos de Filología*. Universidad de Valencia.
- Gómez Molina José R. (2008a). «Castellano y catalán en contacto». En M. T. Echenique y E. Casanova (eds.), *El deber de les paraules. Les aportacions de Germà Colon a la Romanística*. Valencia: Universidad de Valencia, pp. 217-228.
- Gómez Molina, José R. (2008b). «Estudio sociolingüístico de las perífrasis ‘deber + infinitivo’ y ‘deber de + infinitivo’ en el español hablado». *Homenaje al Dr. César Hernández Alonso*. Universidad de Valladolid: 867-880.
- Gómez Molina, José R. (2011a). «Las construcciones con ‘de/Ø + que + verbo en forma personal’ en las cláusulas integradas». *Actas de las IX Congreso Internacional de Lingüística General*. Valladolid: Universidad de Valladolid, pp. 1009-1025.
- Gómez Molina José R. (2011b). «La preposición ‘de’ como mecanismo comunicativo en las construcciones ‘Ø/de + que + verbo en forma personal’». *Oralia*, 14: 345-376.
- Gómez Molina José R. (2012.) «Variación y cambio fónicos de la /d/ intervocálica en el español de Valencia. Proyecto PRESEEA-PRESEVAL». *Lingüística Española Actual*, 34(2): 167-196.

- Gómez Molina José R. (2013a). «Pluralización de *haber* impersonal en el español de Valencia». *Verba*, 40: 253-284.
- Gómez Molina José R. (2013b). «Interacción de factores internos y externos en la pronunciación de la /d/ intervocálica». En J. R. Gómez Molina (coord.), *El español de Valencia. Estudio sociolingüístico*. Berna: Peter Lang, pp. 31-68.
- Gómez Molina José R. (2013c). «Las construcciones ‘de/Ø’ + que + verbo en forma personal». En J. R. Gómez Molina (coord.), *El español de Valencia. Estudio sociolingüístico*. Berna: Peter Lang, pp. 183-222.
- Gómez Molina José R. (2013d). «Usos del verbo ‘haber’ impersonal». En J. R. Gómez Molina (coord.), *El español de Valencia. Estudio sociolingüístico*. Berna: Peter Lang, pp. 109-143.
- Gómez Molina José R. (2013e). «Norma y usos de las perífrasis ‘deber + infinitivo’ y ‘deber de + infinitivo’». En J. R. Gómez Molina (coord.), *El español de Valencia. Estudio sociolingüístico*. Berna: Peter Lang, pp. 71-108.
- Gómez Molina José R. y María Begoña Gómez Devís (2010). «Mantenimiento y elisión de la /d/ intervocálica en el español de Valencia». *Verba*, 37: 89-122
- Gómez Molina José R. y María Begoña Gómez Devís (2016). «Vaya valla. El yeísmo en el español de Valencia». *Boletín de Filología de la Universidad de Chile*, 51(2): 49-88.
- Gómez Molina, José R. y José M. Buzón García (2015). *Variabilidad en el paradigma verbal de futuro. El español de Valencia y de otras sintopías*. Berna, Peter Lang.
- Gómez Molina, José R. y María Begoña Gómez Devís (2020). «The agreement of existential *haber* in three varieties of spoken Spanish». En I. Molina, F. Paredes y A. M. Cestero (eds.), *Sociolinguistic patterns and processes of convergence and divergence in Spanish*. Spanish in Context. Monographic Issue, 17(2), pp. 273-293.
- Gómez Molina, José R., Isabel Molina y Florentino Paredes (2012). «La variación de la /d/ intervocálica en comunidades de habla conservadoras de España: los casos de Madrid y Valencia». *Actas del XVI Congreso Internacional de la ALFAL*. Alcalá de Henares: Universidad de Alcalá, pp. 1975-1986.
- Gómez Molina, José R., José M. Buzón García y Begoña Gómez Devís (2017). *Actitudes Lingüísticas en Valencia y su Área Metropolitana. Estudio longitudinal y análisis de tendencias*. Valencia, Tirant Lo Blanch.
- Lope Blanch Juan M. (1968). «El proyecto de estudio coordinado de la norma lingüística culta de las principales ciudades de Iberoamérica y de la Península Ibérica». *El Simposio de México del PILEI: Actas, informes y comunicaciones*. México, U.N.A.M, pp. 222-233.
- Lope Blanch Juan M. (ed.) (1977). *Estudios sobre el español hablado en las principales ciudades de América*. México: U.N.A.M.
- Lope Blanch Juan M. (1986). *El estudio del español hablado culto. Historia de un proyecto*. México: U.N.A.M.
- Real Academia Española y Asociación de Academias de la Lengua Española (2005). *Diccionario panhispánico de dudas*. Santillana Ediciones Generales.

- Roselló Verdeguer, Jorge (2013). «Las formas de tratamiento». En J. R. Gómez Molina (coord.), *El español de Valencia. Estudio sociolingüístico*. Berna: Peter Lang, pp. 277-314.
- Roselló Verdeguer, Jorge (2017a). «El uso de los pronombres átonos en el corpus PRESEEA-VALENCIA. Los casos de leísmo». *LinRed: Lingüística en la Red*, 15(35).
- Roselló Verdeguer, Jorge (2017b). «El uso de tú y usted en el área metropolitana de Valencia. Un enfoque variacionista». *ELUA. Estudios de Lingüística. Universidad de Alicante*, 31: 285-309.
- Roselló Verdeguer, Jorge (2018). «Los pronombres clíticos en el español hablado de Valencia. Un estudio sociolingüístico». *Diálogo de la Lengua*, X: 18-55.
- Roselló Verdeguer, Jorge (2021). «Factores determinantes en la expresión del sujeto pronominal en el español hablado de Valencia». *LinRed. Lingüística en la Red*, XVII: 1-24.
- Silva Corvalán, Carmen (1994). «Direcciones en los estudios sociolingüísticos de la lengua española». *Actas del Congreso de la Lengua Española* (Sevilla, 7-10 de octubre de 1992). Madrid: Instituto Cervantes, pp. 399-415.
- Terrádez Gurrea, Marcial (2013). «El corpus PRESEVAL. Estratificación y frecuencias léxicas». En J. R. Gómez Molina (coord.), *El español de Valencia. Estudio sociolingüístico*. Berna: Peter Lang, pp. 225-245.



**PROCESOS DE VARIACIÓN Y CAMBIO EN MADRID:  
EL PAPEL DE LAS IDENTIDADES EN LA FORMACIÓN DE REGIOLECTOS**

Isabel Molina Martos  
*Universidad de Alcalá*

## 1. INTRODUCCIÓN<sup>1</sup>

En estas páginas nos ocuparemos de algunas dinámicas de variación y cambio relativas a la realidad sociolingüística urbana, entendiendo por «ciudad» un concepto que implica dos dimensiones interrelacionadas: una física, espacial, y otra referida a los hablantes, a la comunidad de habla. Es esta segunda la que tradicionalmente ha interesado a la sociolingüística, en la medida en que los hablantes son quienes se comunican dentro de las coordenadas del espacio físico y, por tanto, quienes generan y vehiculan la variación y el cambio lingüístico. La comunidad sociolingüística es, en este sentido, la *comunidad de habla*, un concepto que ha evolucionado a lo largo de las décadas, con implicaciones teóricas y metodológicas para el desarrollo de la teoría sociolingüística. En sus investigaciones pioneras, Labov (1972) concebía la comunidad de habla a partir de las «actitudes lingüísticas» compartidas por los hablantes ante los hechos variables. La comunidad ya no quedaba delimitada por un conjunto de hablantes cuyos usos lingüísticos podían ser más o menos equivalentes, sino porque estos valoraban de manera semejante los hechos de habla, esto es, las variables lingüísticas activas en su comunidad.

---

<sup>1</sup> Este trabajo ha sido posible gracias al proyecto de I+D+i, convocatoria 2019, *Agenda 2050. El español del centro-norte de España: procesos de variación y cambio espaciales y sociales*, Ref. PID2019-104982GB-C51, financiado por el Ministerio de Ciencia, Innovación y Universidades de España y por la Agencia Estatal de Investigación (MCIU/AEI/10.13039/501100011033). Agradezco a mi amigo Pedro Martín Butragueño la atenta lectura de este texto.

Estas primeras investigaciones, que Eckert (2012) clasificó como «primera ola de la sociolingüística», manejaban las categorías sociales clásicas de la sociología: grupos etarios, sexo, niveles de instrucción, clase social, profesión o ingresos, entre otras. Desde entonces, la sociolingüística ha experimentado un desarrollo significativo, tanto en su concepción de la comunidad de habla como al postular nuevas unidades sociales de análisis que contribuyan a explicar la variación sociolingüística. Las maneras de entender la ciudad y la comunidad de habla se han refinado sustancialmente, no solo porque la sociolingüística como disciplina ha evolucionado, sino porque las realidades en sí mismas también han cambiado. Las ciudades, como espacios en los que conviven y se comunican personas, han ganado en complejidad. Los movimientos migratorios han desencadenado el aumento demográfico de las ciudades en detrimento de los núcleos rurales hasta desbordar el espacio tradicional y dejar obsoleto el viejo concepto de ciudad; estas han pasado a convertirse en macro-urbes con características nuevas y una mayor complejidad social que, a su vez, proyecta una mayor diversidad lingüística: la convivencia de hablantes de distintas procedencias, dialectales y lingüísticas, ha modificado sustancialmente los patrones sociolingüísticos que estuvieron vigentes al menos hasta mediados del siglo pasado.

La sociolingüística en el siglo XXI centra una parte importante de su reflexión en estas macro-ciudades transformadas por efecto de los flujos migratorios en entidades distintas en todos los sentidos. La macrourbe se compone de muchas partes diferenciadas por su composición social, por los distintos orígenes de los ciudadanos que conviven, por la distancia física y comunicativa entre sectores, etc. En consecuencia, el concepto laboviano de comunidad de habla ya no es suficiente para soportar una metodología de análisis que conduzca a una imagen realista de las nuevas dinámicas de variación y cambio. Las macrourbes modernas incluyen una multiplicidad de comunidades de habla, al tiempo que son caldo de cultivo de otras nuevas que parten de las estructuras preexistentes y las transforman en algo completamente distinto. Los residentes en estos conglomerados urbanos están expuestos a una diversidad de experiencias vitales que coexiste en paralelo, pero no son comunes a todos ellos; de hecho, lo más frecuente hoy día es que muchos de los ciudadanos ni hayan conocido ni vayan a conocer nunca amplias zonas de la ciudad y vivan ajenos a una multiplicidad de sectores o distritos que son parte de esta (cf. Martín Butragueño, en este volumen). La falta de comunicación distingue las nuevas realidades urbanas de los espacios tradicionales, donde las experiencias compartidas daban a la comunidad de habla una cierta unidad que hoy ha desaparecido. La ausencia de contacto con una buena parte de las realidades que conforman el conglomerado urbano, la incomunicación entre amplios sectores de la población urbana, determina dinámicas diferenciadas de variación lingüística y de cambio que necesariamente obligan a modificar las tradicionales unidades de análisis.

La existencia separada que viven los distintos colectivos sociales es particularmente pronunciada entre la población migrante que, por otra parte, es un compo-

nente esencial de la comunidad urbana. Asimismo, también las viejas dinámicas de integración social se han visto modificadas en muy poco tiempo. Así, por ejemplo, las investigaciones sociolingüísticas que se hicieron en Madrid hacia 1990 reflejando el asentamiento de las migraciones de los años 1960-70, dejaban ver una cierta progresión en la integración sociolingüística de estos migrantes con los madrileños autóctonos. Aquellos eran españoles originarios de las dos Castillas, de Andalucía, de Extremadura, de Aragón, etc., cuyos hijos, ya nacidos en Madrid, apenas se diferenciaban en su forma de hablar de los otros madrileños: su habla era también madrileña, pero con rasgos característicos de la periferia urbana (Martín Butragueño, 2002; Molina, 2015). Esas investigaciones todavía reflejaban las dinámicas sociolingüísticas de la vieja ciudad que había estado vigente a lo largo del siglo XX. Con el cambio de siglo y la llegada de nuevas oleadas migratorias de distintos lugares del mundo, la estructura urbana tradicional se ha modificado hasta transformarse en un conglomerado de realidades sociales y lingüísticas diferenciadas, cada vez menos interconectadas entre sí, donde «la marginación y la minorización pueden ser tales que ni siquiera haya espacios de convivencia física» (Martín Butragueño, en este volumen).

Para estudiar las nuevas realidades urbanas, la sociolingüística del siglo XXI emplea metodologías y herramientas analíticas que dan cuenta de aspectos parciales de la diversidad urbana, cuya suma aporta una imagen comprensiva de la complejidad y el dinamismo social. En este contexto son muy significativos los estudios de paisaje lingüístico que ponen de relieve ese carácter de conglomerado característico del entorno urbano (Landry y Bourhis, 1997; Scollon y Scollon, 2003; Bloommaert, 2013). A grandes rasgos, los nuevos enfoques estudian la ciudad desde planos diferenciados, en los que cabe distinguir los estudios del nivel macro, centrados en las grandes estructuras urbanas, de los que ponen el foco en la variación individual. Unos y otros deben entenderse como complementarios, pues de la suma de las distintas perspectivas debería aflorar una visión más realista de la realidad urbana de este siglo.

## **2. EL PESO DE LA VARIACIÓN INDIVIDUAL EN LA VARIACIÓN Y EL CAMBIO LINGÜÍSTICO**

La dimensión de análisis macro, centrada en las grandes estructuras urbanas, es solidaria de la variación individual, pues la una se proyecta en la otra, de manera que la variación individual repercute y retroalimenta las dinámicas macroestructurales. En concreto, el análisis de las identidades se revela como una dimensión productiva, directamente relacionada con el cambio lingüístico, pues el posicionamiento ante las nuevas identidades es uno de sus factores desencadenantes, el que vincula lo social con lo lingüístico. Se trata de una concepción compleja de la identidad, en la que esta no se ve como suma de categorías, sino como resultado de su intersección, que lleva a una construcción conjunta a partir de todas ellas.

También la sociolingüística ha evolucionado desde sus concepciones tradicionales de identidad hasta otras más recientes. Las investigaciones de la primera ola definían la identidad a partir de categorías sociodemográficas como la edad, el género o la clase social (Trudgill, 1972); en una segunda etapa, la identidad se entiende con base en las prácticas sociales de los hablantes (Milroy, 1980; Eckert, 1989); por último, los estudios recientes consideran la variación individual y sus significaciones sociales como un mecanismo de construcción de identidades (Podesva, 2007).

### **2.1. El individuo en la construcción de identidades. Agentividad, complejidad y dinamismo**

En los estudios de la tercera ola, enfocados en la variación individual, la identidad pasa a ser la unidad de análisis, concebida como una categoría pluridimensional, más acorde con la complejidad propia de la psique humana, que proyecta identidades múltiples, incluyendo aspectos como la identidad de género, la clase social, la raza, la ideología política, la condición de inmigrante y un largo etcétera. A su vez, la identidad no es estática, sino esencialmente dinámica (Bucholtz y Hall, 2004), pues la misma complejidad de identidades que define la diversidad de espacios sociales urbanos induce a adoptar nuevas actitudes sociales y gestos que identifican a cada individuo con ciertos grupos o colectivos. En este sentido, la identidad se vincula con el cambio lingüístico, al ser una característica construida y representada socialmente a partir de prácticas semióticas entre las que se incluyen las prácticas lingüísticas; el hablante se concibe como agente, con capacidad de influir en sus comunidades de prácticas y, eventualmente, en la comunidad de habla y en el conglomerado urbano.

El vínculo entre los significados sociales que puede proyectar una variable y la construcción de la identidad está en la indexicalidad. Así, en el marco dialectal de las variedades del castellano europeo, las realizaciones más o menos tensas o relajadas de la consonante /-s/ en posición de coda vehiculan significados sociales variados y cambiantes en el tiempo y en el espacio. La pronunciación aspirada [h] de la /s/ en el Madrid de los años 60 podía indexar la procedencia migrante de quien la utilizaba, mientras que en los años 80, para los andaluces occidentales de nivel social medio y medio-alto, pasó a vehicular significados ligados a valores posiblemente próximos a la «elegancia», por contraste con otras realizaciones relajadas de la sibilante como la asimilación o la elisión. En suma, el distinto posicionamiento identitario del hablante en la interacción indexicaliza y codifica con significados diferentes las mismas realizaciones consonánticas, que según para quién y en qué situación, pueden denotar «cosmopolitismo», «movilidad social», etc. y vincularse a personajes o tipos sociales distintos como «hablante meridional», «joven de barrio de la periferia madrileña», «jerga juvenil», etc.

En conclusión, la variable sociolingüística no es portadora de un significado social fijo y estable, sino que evoca valores indexicales ya existentes, al tiempo que tiene la capacidad de forjar otros nuevos. Esta idea de indexicalidad resulta especial-

mente útil para abordar la cuestión de la construcción de identidades en interacción y superar la visión tradicional de la comunidad de habla estable y homogénea.

### 3. PROCESOS DE VARIACIÓN Y CAMBIO: EL PAPEL DE LAS IDENTIDADES EN LA FORMACIÓN DE REGIOLECTOS

Una de las variables más estudiadas del español, por su carácter panhispánico, es la variable *-s/* en coda, un hiperdialectalismo que se toma como referencia para dividir las variedades geolectales del español en variedades conservadoras –las que mantienen una alta probabilidad de realización plena de */-s/*– y variedades innovadoras –las que presentan altos índices de debilitamiento y/o pérdida de este segmento–. Esta división, que se reproduce casi sin excepción en cada país o área regional (Moreno Fernández, 2004; Martín Butragueño, 2014; Fernández de Molina y Hernández-Campoy, 2018), ha desarrollado en toda su geografía y extensión social, un amplio abanico de indexicalidades.

En el español de España, la pronunciación aspirada [h] connota significaciones que pueden jerarquizarse desde las más generales a las más específicas. El significado de ‘hablante meridional’ tal vez sería el que, de forma general, primero se viene a la mente al escuchar a un español producir realizaciones aspiradas de la sibilante. Sin embargo, su campo indexical es mucho más amplio en cada comunidad de habla. En Andalucía o Extremadura, la aspiración no significa ‘de origen migrante’, como sucede en algunos distritos de Madrid, donde también puede indicar que ‘el hablante vive en algún barrio de la periferia sur de la capital’, o que quien la utiliza tiene ‘bajo nivel de instrucción’ o un ‘nivel socioeconómico bajo’. Asimismo, la realización plena [s] indiza otros valores, además de un nivel educativo alto, como pedantería o aspiración al ascenso social entre otros posibles.

A diferencia de Madrid, en Andalucía el campo indexical de la aspiración es radicalmente distinto: en la región meridional, donde la solución más frecuente es la elisión de la sibilante, el sonido aspirado puede evocar al ‘pijo andaluz’, al ‘sevillano’ o a un ‘joven estudiante universitario’, entre otros significados. Estos campos indexicales tan diferenciados se explican, en parte, por los diferentes modelos referenciales que operan en los espacios de variación geolectal en España. El concepto de «estándar» resulta muy útil para adentrarse en las dinámicas del cambio lingüístico por su proyección en la estratificación sociolingüística. En la dimensión diastrática, la influencia del modelo lingüístico prestigioso ejerce potencialmente de fuerza «centrípeta» que explica por qué se dan procesos de convergencia y divergencia dialectal, relacionados a su vez con identidades referenciales, regionales y locales.

Por otra parte, como se apuntaba más arriba, las comunidades urbanas tradicionales han dejado de ser la referencia en el siglo XXI, un hecho que se pone de manifiesto en Madrid, donde se activan identidades referenciales heterogéneas. En ello ha incidido el diverso origen de los hablantes, que se integran en comunidades de habla

múltiples e intercomunicadas entre sí a través de la emigración y de la movilidad espacial y social, muy acentuada entre las generaciones jóvenes.

Al analizar el continuo dialectal que se desarrolla entre un extremo más conservador y otro más innovador, los estudios sociolingüísticos de variedades dialectales conservadoras como Madrid, frente a otras innovadoras como las andaluzas, ponen de relieve la importancia del prestigio y de la asunción de nuevas identidades en la evolución de estas variedades. Un ejemplo de esto puede verse en los movimientos demográficos internos de la población madrileña y andaluza de origen español y su incidencia en los procesos de nivelación y koineización.

### 3.1. Variación dialectal y sociolingüística de la *-s/*

En la tabla 1, se muestran algunos datos de la tendencia general al debilitamiento y a la elisión de la *-s/* en comunidades de habla españolas del centro y del sur de la península ibérica. El continuo de variación se refleja espacialmente en el eje nortesur: las variedades centro septentrionales se acercan más a la norma del estándar nacional, favoreciendo la proximidad de la realización plena de la sibilante, en tanto que las variedades meridionales se alejan progresivamente de dicho estándar, con predominio de las variantes debilitadas o elididas<sup>2</sup>.

**TABLA 1. DISTRIBUCIÓN DE LAS PRINCIPALES VARIANTES DE */s/* POSTNUCLEAR EN ALGUNAS ÁREAS DIALECTALES DEL ESPAÑOL DE ESPAÑA (VIDA, VILLENA Y MOLINA 2023)**

			<b>s</b>	<b>h</b>	<b>Ø</b>
Áreas Conservadoras	<b>Madrid-Bº Sal.</b>	<b>(Molina Martos 2006)</b>	82.0	16.0	1.0
	<b>Madrid-Vallecas</b>	<b>(Molina Martos 2015)</b>	67.6	29.5	2.8
	Alcalá de Henares	(Blanco Canales 2005)	64.0	34.0	2.0
	Getafe	(Martín Butragueño 2002)	59.0	35.0	6.0
	Toledo	(Molina Martos 1998)	52.0	36.0	12.0
	Mérida	(Fernández de Molina 2016)	8.8	69.8	21.4
Áreas Innovadoras	Las Palmas	(Samper Padilla 1990)	3.0	64.3	32.7
	Linares (Jaén)	(Gómez Serrano 1994)	2.0	48.1	49.9
	<b>Málaga</b>	<b>(Vida Castro 2004)</b>	1.5	31.1	67.5
	<b>Granada</b>	<b>(Tejada Giráldez 2015)</b>	1.4	28.1	70.5

La variación de la */-s/* también se proyecta en la estructura socioeconómica. Así, en las variedades conservadoras españolas, como Madrid, Alcalá de Henares y Getafe (Molina Martos, 2006, 2015), se observa que las variantes debilitadas (principalmente aspiradas) son más probables en los hablantes de nivel sociocultural bajo,

<sup>2</sup> Se marcan en negrita las comunidades estudiadas a partir de los corpus del proyecto PRESEEA.

mientras que la sibilancia es promocionada por los niveles socioculturales altos. Sin embargo, las pronunciaciones del barrio de Salamanca en Madrid son significativamente más conservadoras que las de las comunidades de la periferia metropolitana. En parte, el motivo es que los estudios de Getafe y Vallecas tienen en cuenta a la población inmigrante procedente de áreas innovadoras meridionales.

En las variedades intermedias –aquellas en las que la variante predominante es la aspiración– se observan patrones sociales diferenciados. En Las Palmas, la variante elidida es favorecida por los hablantes de nivel sociocultural bajo, mientras que la aspiración se erige como variante prestigiosa. En Mérida, en cambio, la realización sibilante es más frecuente entre los hablantes de mayor nivel social, mientras que la elisión y la aspiración son características de los hablantes de menor nivel sociocultural. Por último, en las variedades más innovadoras de Granada y Málaga, donde la solución predominante es la elisión, las variantes aspirada y plena son más frecuentes en los niveles socioculturales altos y en los estilos de habla formales.

A la vista de este comportamiento general se puede afirmar que existe una jerarquía del prestigio relativo de las variantes: [s] > [h] > Ø, de tal modo que la aspiración pasa a cubrir el espacio de [s] en las variedades intermedias e innovadoras.

### 3.2. La comparación de los análisis de Madrid y Andalucía en el tratamiento de la /s/ en coda

En Madrid, el centro de la capital es claramente conservador, con un índice de mantenimiento del 82% en el barrio de Salamanca (Molina Martos, 2006), pero este conservadurismo se reduce progresivamente, dejando ver una gradación geográfica y sociolingüística de acuerdo a la posición más o menos periférica respecto al centro de Madrid, a la mayor presencia de migrantes llegados en la segunda mitad del s. XX y al nivel socioeconómico de la comunidad en cuestión (véanse los mapas 1 y 2).

TABLA 2. /-s/ EN LAS COMUNIDADES DE HABLA MADRILEÑAS

%	[s]	[h]	[σ]	[ø]
MADRID-Bº SALAMANCA	82	14	2	1
MADRID-VALLECAS	67,6	26,4	3,1	2,8
ALCALÁ DE HENARES	64	30	4	2
GETAFE	59	28	7	6

FIGURA 1. DISTRITOS DE LA CIUDAD MADRID

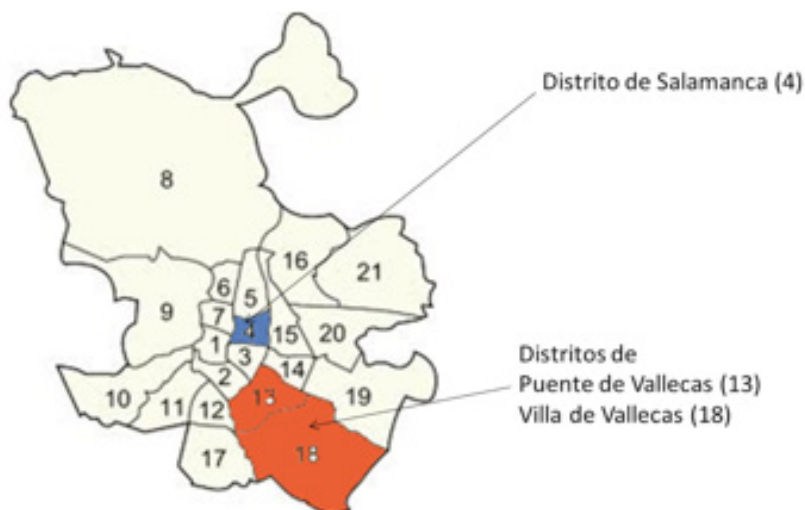
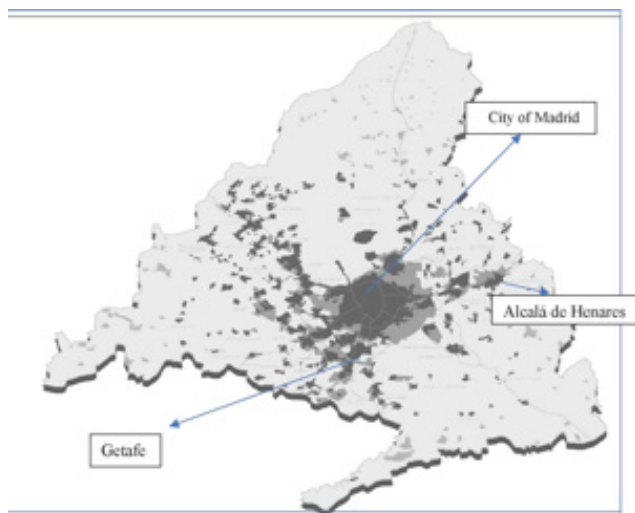


FIGURA 2. COMUNIDAD DE MADRID.



Estas diferencias están directamente relacionadas con un campo de indexicalidad que a su vez es el reflejo de identidades diversas. Así, en el siglo XX, para integrarse en el estándar madrileño, las poblaciones de Vallecas, Alcalá y Getafe debían realizar un movimiento de retracción respecto a sus pronunciaciones de origen, más innovadoras, en convergencia con la única variante con prestigio abierto en Madrid, representada en la conservación de la [s]. Este cambio dialectal reflejaba un cambio en su *identidad referencial*. El contraste entre las comunidades de habla madrileñas

del entorno metropolitano dibujaba un movimiento de convergencia lingüística del medio rural con el urbano, en el que los barrios de la periferia funcionaban como puente. Cuando el hablante rural emigraba a Madrid, se instalaba en barriadas de la periferia como Vallecas o Carabanchel, donde modificaba su habla a imitación de la madrileña. A su vez, los madrileños de la periferia tomaban como referencia los patrones de prestigio transmitidos por los medios de comunicación y la escuela, más próximos a la población autóctona de nivel socioeconómico más alto, radicada en el centro de la ciudad (representada por el barrio de Salamanca en el corpus de PRESEEA de Madrid). El movimiento de convergencia se ordenaba de acuerdo a la siguiente secuencia:

Medio rural > periferia urbana > centro de la ciudad
--

En términos generales, los inmigrantes incorporaban los rasgos de la comunidad de acogida para integrarse lingüísticamente y librarse del estigma de hablantes rurales. La integración lingüística, de la que participaron inmigrantes de diferentes orígenes, se resolvía mediante la pérdida progresiva de las marcas dialectales.

Pero el proceso de integración sociolingüística ha ganado en complejidad con el paso del tiempo y han surgido nuevas identidades que reafirman el sentido de pertenencia al barrio. Los análisis de la ciudad de Getafe (Martín Butragueño, 2002) y del distrito de Vallecas (Molina, 2015), que se hicieron con veinte años de diferencia, marcan dos etapas diferenciadas en este proceso. Los datos de Getafe, de 1990, mostraban que, aunque todos los getafeños favorecían la conservación de la [s] plena, los más jóvenes lo hacían de forma más pronunciada: lejos de conformarse con sumarse al cambio local, superaban a sus homólogos madrileños en la realización de la variante prestigiosa. Martín Butragueño clasificaba esta pauta como un *cambio en la estructura comunitaria*, que reproduce el deseo de una rápida integración lingüística en la comunidad de acogida por parte de la población con más recursos para reconocer el estándar. En cambio, en el estudio de la población inmigrante de Vallecas, ya en el siglo XXI, no se manifiesta de manera tan clara el deseo de convergencia con la población madrileña del centro de la ciudad. Entre los vallecanos, el desarrollo conservador de la /-s/ se había frenado hasta estabilizarse en unos índices de mantenimiento del 65%-70%, con soluciones más relajadas que en el centro de la ciudad. Este freno en la convergencia con el estándar se explica como un cambio identitario de los vallecanos, quienes habrían dejado de verse a sí mismos como inmigrantes en la capital para reconocerse como madrileños de pleno derecho. Sus índices de conservación, a medio camino entre la relajación de sus comunidades de origen y la alta conservación del centro de la capital, les permite indexicalizar diferencias identitarias entre la periferia y el centro de Madrid. Los vallecanos recurren a una mayor frecuencia de la aspiración [h] de la que es propia del barrio de Salamanca para indicar que esta

pronunciación les define e indexicaliza que proceden de la periferia de la ciudad y que son descendientes de españoles originarios de las regiones meridionales. Estos resultados deben interpretarse a la luz de los cambios sociales que ha experimentado la ciudad en el nuevo milenio, relativos a la llegada de inmigración de origen no español y al desarrollo de una identidad colectiva de barrio en Vallecas. La inmigración española deja paso a la de otros países: una nueva generación de inmigrantes extranjeros ha ocupado su puesto en la escala social y ha contribuido a consolidarlos como población autóctona. Los vallecanos ya no son andaluces o extremeños recién llegados a la capital, sino propiamente madrileños para quienes el origen autóctono o inmigrante de sus padres o de sus abuelos no determina diferencias significativas en cuanto a la lengua, ni marca curvas de hipercorrección que traduzcan sus deseos de integración lingüística. Los vallecanos del siglo XXI han interiorizado que, aunque desciendan de manchegos o andaluces, son madrileños y parte de esta comunidad de habla. Todo ello explica que, al mismo tiempo que mantienen un patrón estable conservador, convergente con el del barrio de Salamanca, utilicen la aspiración [h] de la /-s/ para marcar estilísticamente su discurso, tiñéndolo con el sello distintivo, innovador, de la periferia del sureste de Madrid.

Pero Madrid no es una isla lingüística: precisamente su papel como ciudad de acogida de migrantes españoles de distinto origen ha determinado una dinámica dialectal y sociolingüística generadora de otros cambios lingüísticos. La mayor movilidad geográfica y social de las poblaciones en el siglo XXI está determinando una nueva evolución en las variedades dialectales, que hoy en día se muestran interconectadas a través de sectores de la población que mantienen su residencia y su identidad a caballo entre distintas comunidades de habla.

Si nos limitamos a la variación geolectal que se muestra en la tabla 3, se observa cómo el conservadurismo fónico disminuye en su avance hacia el sur:

**TABLA 3. ESTRATIFICACIÓN DE -(S) EN MADRID Y ANDALUCÍA ORIENTAL**

%	[s]	[h]	[σ]	[ø]
MADRID-Bº SALAMANCA	82	<b>14</b>	2	1
MADRID-VALLECAS	67,6	<b>26,4</b>	3,1	2,8
GRANADA (ANDALUCÍA)	0,6	<b>20,4</b>	0,8	76,9
MÁLAGA (ANDALUCÍA)	1,5	<b>30,9</b>	0,2	67,5

Pero, igual que en Madrid, es preciso saber cómo se estratifica socialmente la variable en las comunidades andaluzas para determinar la dirección del proceso de nivelación. Según se desprende de los análisis de los equipos PRESEEA de Granada y Málaga, las variedades urbanas de estas dos ciudades oscilan entre dos pulsiones: una geográfica, que separa Andalucía oriental y occidental (las variedades innovadoras del andaluz occidental se alejan progresivamente de las más conservadoras del andaluz oriental) y la otra vertical, que refleja la urbanización y las corrientes centripetas en

las sociedades contemporáneas y aleja los sociolectos de clase media en las ciudades de las variedades vernaculares rurales y de los sociolectos de clase trabajadora urbana (Villena, 2013). Esta doble tensión ha favorecido el desarrollo en Andalucía oriental de una *nueva variedad intermedia* entre el estándar nacional y las variedades vernaculares (Villena, 2008; Villena y Vida, 2015, 2017). Los protagonistas en el desarrollo de esa variedad emergente son los *hablantes urbanos jóvenes de clase media* de Andalucía oriental (Granada y Málaga), quienes han adoptado rasgos lingüísticos del estándar castellano, al tiempo que mantienen otras marcas meridionales identitarias. Así, un hablante prototípico de esa variedad intermedia usaría rasgos próximos al estándar nacional: /x/ tensa, escisión de /θs/ y variantes aspiradas de la /-s/ en coda silábica. Dicha variedad se expande especialmente entre los hablantes de estatus social medio y alto, los hablantes instruidos y las mujeres (Villena y Vida, 2017).

Este perfil sociolingüístico del hablante de la nueva variedad intermedia andaluza debe ponerse en relación con la población vallecana de la periferia sur de Madrid, puesto que en el continuum dialectal del castellano del centro y del sur peninsular, los dos extremos se juntan: por una parte, la población inmigrante de la periferia sureste de Madrid ha consolidado en su sistema fónico la variante [h], que emplea con una frecuencia superior (26,4%) a la que es característica del centro de la ciudad (14%) y que, entre los más jóvenes, aumenta hasta el 27,6%. La aspiración reivindica una identidad que aun siendo genuinamente madrileña, se redefine con marcas propias. Estos hablantes de la periferia de Madrid son próximos a los hablantes urbanos jóvenes de clase media de Andalucía oriental, quienes en contra de la tendencia general andaluza a la elisión de la (-s) en coda, optan por soluciones más conservadoras –aspiración– para no ser identificados como procedentes del medio rural o como pertenecientes a niveles sociales bajos. Se trata de hablantes con mayor movilidad geográfica y social, una generación que ha salido de su lugar de origen para estudiar fuera de Andalucía o que se ha desplazado para trabajar a Madrid.

La adopción de este rasgo denota la emergencia de una nueva identidad mixta de individuo meridional orientado al estándar y conforma una variedad nueva que representa la reunión de los dos modos más importantes de hablar español en España: por un lado, el septentrional del castellano conservador y, por otro, el meridional, identificado con el andaluz innovador. Los dos extremos del continuo dialectal centro-sur, se juntan en estos hablantes jóvenes, unos descendientes de inmigrantes y residentes en la periferia sur de Madrid, y otros, jóvenes andaluces orientales de clase media. Se ha producido, por tanto, una convergencia (o tal vez una coincidencia) entre la clase media de Andalucía oriental y los niveles socioeconómicos medio-bajos y de origen migrante de Madrid, configurando una especie de dialecto terciario centromeridional (Villena y Vida, 2017). Este proceso conduce a la formación de una variedad koinética intermedia caracterizada por el mantenimiento de los rasgos no marcados propios de las variedades innovadoras meridionales y la adopción de rasgos centrales y estándar especialmente prominentes.

## REFERENCIAS

- Blanco Canales, Ana (2005). *Estudio sociolingüístico de la ciudad de Alcalá de Henares*. Alcalá de Henares: Universidad de Alcalá.
- Bloommaert, Jan (2013). *Ethnography, Superdiversity and Linguistic Landscapes*. Bristol: Multilingual Matters.
- Bucholtz, Mary y Kira Hall (2004). «Theorizing identity in language and sexuality research». *Language in Society*, 33(4): 469-515. <https://doi.org/10.1017/S0047404504334020>
- Cestero Mancera, Ana M. y Francisco Moreno Fernández (eds.) (2017). *LinRed, monográfico del PRESEEA. Procesos de variación y cambio en el español de España. Estudios sobre el corpus PRESEEA, XV*. [www.linred.es/monográficos]
- Eckert, Penelope (1989). *Jocks and Burnouts: Social Categories and Identity in the High School* (Vol. 19, Número 6). Teachers College Press.  
<http://www.jstor.org/stable/2073258?origin=crossref>
- Eckert, Penelope (2012). «Three Waves of Variation Study: The Emergence of Meaning in the Study of Sociolinguistic Variation», *Annual Review of Anthropology*, 41(1): 87-100. <https://doi.org/10.1146/annurev-anthro-092611-145828>
- Fernández de Molina, Elena (2016). «El nivel social como indicador de la variación de /-s/ en el habla de Mérida (Badajoz)». *Dialectología*, 16: 93-116.
- Fernández de Molina, Elena y Juan M. Hernández-Campoy, (2018). «Geographic Varieties of Spanish». En K. Geslin (ed.), *The Cambridge Handbook of Hispanic Linguistics*. Cambridge: Cambridge University Press, pp. 496-528.
- Gómez Serrano, Antonio (1994). *Aspectos del habla de Linares (Jaén)*. Málaga: Universidad de Málaga.
- Labov, William (1972). *Sociolinguistic Patterns*. Pennsylvania: University of Pennsylvania Press.
- Landry, Rodrigue y Richard Bourhis (1997). «Linguistic landscape and ethnolinguistic vitality: An empirical study». *Journal of Language and Social Psychology* 16(1): 23-49.
- Martín Butragueño, Pedro (2002). «La dimensión de las hipótesis sustantivas: datos sociolingüísticos del sur de Madrid». *Variación sociolingüística y teoría fonológica*. México, D.F.: El Colegio de México.
- Martín Butragueño, Pedro (2014). «Obstruyentes oclusivas: tensión y relajamiento como motor de la variación fónica. Notas sobre la variación en las oclusivas orales». En P. Martín Butragueño (ed.), *Fonología variable del español de México*. Vol. I: *procesos segmentales*, México, D.F.: El Colegio de México, pp. 245-272.
- Milroy, Lesley (1980). *Language and Social Networks*, Oxford: Basil Blackwell.
- Molina Martos, Isabel (1998). *La fonética de Toledo. Contexto geográfico y social*. Alcalá de Henares: Universidad de Alcalá.

- Molina Martos, Isabel (2006). «Innovación y difusión del cambio lingüístico en Madrid». *Revista de Filología Española* 86(1): 127-149.
- Molina Martos, Isabel (2015). «La variable sociolingüística /s/ en el distrito de Vallecas (Madrid)». En A. M. Cestero Mancera, I. Molina y F. Paredes (eds.), *Patrones sociolingüísticos de Madrid*, Bern: Peter Lang, pp. 177-250.
- Moreno Fernández, Francisco (2004). «Cambios vivos en el plano fónico del español: variación dialectal y sociolingüística», R. Cano Aguilar (ed.), *Historia de la lengua española*. Barcelona: Ariel, pp. 973-1009.
- Podesva, Robert J. (2007). «Phonation type as a stylistic variable: The use of falsetto in constructing a persona». *Journal of Sociolinguistics*, 11(4): 478-504. <https://doi.org/10.1111/j.1467-9841.2007.00334.x>
- Samper Padilla, José Antonio (1990). *Estudio sociolingüístico del español de Las Palmas de Gran Canaria*. Las Palmas de Gran Canaria: La Caja de Canarias.
- Scollon, Ron y Scollon, Suzie Wong (2003). *Discourses in place: Language in the material world*. New York: Routledge.
- Tejada Giráldez, M<sup>a</sup> de la Sierra (2015). *Convergencia y divergencia entre comunidades de habla: a propósito de la /s/ implosiva. Contribución al estudio de los patrones sociolingüísticos del español de Granada*. [Tesis doctoral]. Granada: Universidad de Granada.
- Trudgill, Peter (1972). «Sex, covert prestige and linguistic change in the urban British English of Norwich». *Language in Society*, 1(2): 179-195.
- Vida Castro, Matilde. (2004). *Estudio sociofonológico del español hablado en la ciudad de Málaga*. Alicante: Universidad de Alicante.
- Vida Castro, Matilde; Villena Ponsoda, Juan A. y Molina Martos, Isabel (2023). «Variación de la /-s/ postnuclear en español: patrones sociolingüísticos y geolectales». *Círculo de Lingüística Aplicada a la Comunicación* 94: 23-35. <https://dx.doi.org/10.5209/clac.84701>
- Villena Ponsoda, Juan Andrés. (2008). «Sociolinguistic patterns of Andalusian Spanish». *International Journal of the Sociology of Language*. 193/194: 139-160.
- Villena Ponsoda, Juan Andrés (2013). «Actos de identidad: ¿por qué persiste el uso de los rasgos lingüísticos de bajo prestigio social? Divergencia geográfica y social en el español urbano de Andalucía». En R. Guillén y R. Millán (coords.), *Sociolingüística Andaluza 16. Estudios descriptivos y aplicados sobre el andaluz*, Sevilla: Universidad de Sevilla, pp. 173-208.
- Villena Ponsoda, Juan Andrés y Matilde Vida Castro (2015). «Between local and standard varieties: Horizontal and vertical convergence and divergence of dialects in Southern Spain». En I. Buchstaller y B. Siebenhaar, *Language Variation – European Perspectives VI*, Amsterdam: John Benjamins, pp. 125-140.
- Villena, Juan Andrés y Matilde Vida (2017). «Variación, identidad y coherencia en el español meridional, sobre la indexicalidad de las variables convergentes del español de Málaga». En *LinRed, monográfico del PRESEEA. Procesos de variación y cambio en el español de España. Estudios sobre el corpus PRESEEA*, ed. Ana M<sup>a</sup> Cestero & Francisco Moreno. XV. [[www.linred.es/monograficos](http://www.linred.es/monograficos)]



**EL GRUPO PRESEEA-GRANADA:**  
**ESTUDIOS, RESULTADOS Y SUGERENCIAS HACIA EL FUTURO**

Juan Antonio Moya Corral  
*Universidad de Granada*

## **1. PROYECTOS**

Los trabajos sobre el español de Granada se inician con anterioridad al año 1995. Ese mismo año la Junta de Andalucía reconoció el Grupo de investigación del Departamento de Lengua Española de la Universidad de Granada, *Estudios de español actual* (HUM430), y empiezan las primeras publicaciones.

En tal contexto se pone en marcha el proyecto HAGA (*Habla de Granada*) que merece una subvención de la Dirección General de Investigación Científica y Técnica del Ministerio de Educación y Ciencia de España para el trienio 1995-98 (PB94-1467-C02-02). El objeto perseguido por el proyecto era el de conocer de primera mano la situación del español en la comunidad de habla de la ciudad de Granada. Muy pronto aparecieron los primeros trabajos: Moya y García Wiedemann (1995), Moya (1996, 1997 y 2000), Martínez y Moya (1997).

Una vez conocidos los cambios que se estaban operando en el habla urbana, surgió la necesidad de indagar en la capacidad de irradiación de tales cambios en el habla de su área de influencia. Es así como nació el Proyecto FORDIAL (2000-2003), proyecto sobre la *FORMación DIAlectal*. Se pusieron en marcha varias tesis doctorales, una en un punto con ceceo: Pinos Puente; otra en una localidad seseante: Alhama de Granada. Se analizaron tres grupos de informantes: uno de residentes en la localidad de origen, que nos darían la pauta de los usos lingüísticos rurales; otro de instalados en Granada, que nos informaron del proceso de adecuación a la norma urbana; y un tercero de *commuters*, es decir, hablantes que residen habitualmente en el pueblo de origen, pero que mantienen un

contacto frecuente con Granada (diario, semanal, varias veces al mes, etc.). Este último grupo nos permitía conocer la importancia de otros factores en el proceso de convergencia con el habla granadina (Melguizo, 2005; Molina Serrato, 2003).

Poco después llegó la incorporación del grupo granadino al proyecto PRE-SEEA. Verdaderamente, Granada estuvo atenta, desde un principio, a los trabajos del proyecto panhispánico. Ya en el Congreso de ALFAL de Las Palmas de Gran Canaria (1996) manifestó su deseo de adhesión, pero no fue hasta 2004, con la concesión del proyecto ESEGRA, cuando se produjo la incorporación definitiva y el inicio de una activa cooperación.

A partir de ese momento, los trabajos coordinados se sucedieron sin solución de continuidad: tras el proyecto ESEGRA, *Estudio sociolingüístico del español de Granada* (2004-2007) llegaron el ESCEGRA, *Estudio sociolingüístico del corpus del español de Granada* (2007-2010) y el PASOS-Granada, *Patrones sociolingüísticos del español de Granada* (2011-2015), en todos ellos el IP era quien suscribe. En estos años tuvo lugar mi jubilación y el cambio generacional, pero los trabajos no cesaron ni por un momento. El proyecto ECOPASOS-Granada, *Estudio complementario de los patrones sociolingüísticos del español de Granada* (2016-2019), cuyo IP fue Antonio Manjón-Cabeza Cruz, supuso la continuación del proyecto anterior. Siguieron casi de manera simultánea el proyecto AGENDA 2050 (2020-2024), cuyos IPs son Antonio Manjón-Cabeza Cruz y Francisco José Sánchez García, y el PALER-Granada, *Proyecto para el estudio de la adecuación de los emigrantes residentes en Granada* (2020-2022) coordinado por Antonio Manjón-Cabeza Cruz y Esteban Tomás Montoro del Arco. Existen, no obstante, otros proyectos coordinados con otros grupos y universidades, ajenos a PRESEEA.

## 2. CORPUS

Fruto del proyecto ESEGRA surgió el *Corpus oral para su estudio sociolingüístico del español hablado de Granada* (Moya Corral, coord., 2007, 2008, 2009)<sup>1</sup>.

El corpus consta de una muestra representativa con afijación uniforme de hablantes estratificada según las variables de sexo, edad y grado de instrucción. El número total de informantes es de 54, lo que supone la investigación de tres sujetos por casilla, tal como se observa en la tabla 1:

---

<sup>1</sup> [https://estudiosdeespanolactual.ugr.es/seccion\\_libre/corpus-preseea-granada/](https://estudiosdeespanolactual.ugr.es/seccion_libre/corpus-preseea-granada/).

**TABLA 1: MUESTRA-TIPO POR CUOTAS DEL PROYECTO PRESEEA-GRANADA**

	Generación 1 (20-34 años)		Generación 2 (35-54 años)		Generación 3 (más de 55 años)		Total
	Hombres	Mujeres	Hombres	Mujeres	Hombres	Mujeres	
Grado de instrucción 1	3	3	3	3	3	3	18
Grado de instrucción 2	3	3	3	3	3	3	18
Grado de instrucción 3	3	3	3	3	3	3	18
Total	9	9	9	9	9	9	54

El proyecto AGENDA 2050 ha generado un segundo *Corpus oral para su estudio sociolingüístico del español hablado en Granada*, que ya está disponible en: [https://estudiosdeespanolactual.ugr.es/seccion\\_libre/corpusproyecto-agenda-2050/](https://estudiosdeespanolactual.ugr.es/seccion_libre/corpusproyecto-agenda-2050/).

La muestra obedece al mismo diseño que la anterior representada en el gráfico 1. Ambos corpus permiten el estudio en tiempo real de los cambios del español en Granada.

Por su parte, el Proyecto PALER ha permitido reunir varios corpus que atienden a la procedencia de los inmigrados residentes en Granada. Un buen grupo de ellos ya está acabado:

1. ECUGRA. *Corpus de ecuatorianos residentes en Granada* (2019)
2. GECUGRA. *Corpus de ecuatoguineanos residentes en Granada* (2020)
3. RUMGRA. *Corpus de rumanos residentes en Granada* (2021)
4. POLGRA. *Corpus de polacos residentes en Granada* (2021)
5. MARGRA. *Corpus de marroquíes residentes en Granada* (2021)

Cada corpus está constituido por 30 entrevistas, de las cuales la mitad está hecha por vernáculos granadinos y la otra mitad por miembros de la propia comunidad migrante (en español). El corpus ECUGRA está disponible en:

[https://estudiosdeespanolactual.ugr.es/seccion\\_libre/corpus-ecugra/](https://estudiosdeespanolactual.ugr.es/seccion_libre/corpus-ecugra/)

Están en preparación:

6. *Corpus de venezolanos residentes en Granada*
7. *Corpus de chinos residentes en Granada*

### 3. ESTUDIOS

Los trabajos del grupo de investigación granadino atienden a dos frentes diferentes: uno coincide con los estudios priorizados por el Proyecto PRESEEA, que, naturalmente, son comunes a todos los subproyectos inscritos en el marco del proyecto panhispánico; el otro lo constituyen trabajos que, si bien no son prioritarios para

PRESEEA, son fundamentales para el conocimiento del habla de la comunidad en la que estamos incluidos: Granada y su área de influencia.

### **3.1. Estudios priorizados en el marco PRESEEA**

En este apartado hay que incluir tres grupos de estudios cuya clasificación responde al estado de desarrollo en que se encuentran:

#### **3.1.1. Finalizados. *Todos ellos con varias publicaciones de los miembros del grupo***

- a) -d- intervocálica
- b) *Haber* concordado
- c) Sujeto pronominal
- d) Tabú
- e) Reformuladores
- f) Atenuación

#### **3.1.2. Iniciados. *Son estudios que se están desarrollando en trabajos de doctorado***

- g) Presente narrativo
- h) Prosodia

#### **3.1.3. No iniciados**

- i) Interrogativos de control de contacto
- j) Evidencialidad en acervo común

Estos trabajos permiten conocer el estado en que se encuentran los fenómenos que afectan al español en su conjunto y, además favorecen la comparación del estado de la cuestión en las numerosas zonas en las que el proyecto PRESEEA cuenta con grupos asociados: España y América

### **3.2. Estudios NO priorizados en PRESEEA, pero prioritarios en la comunidad de habla en la que está inscrito el subproyecto**

#### **3.2.1. *Fonética***

- a) Sibilantes explosivas
- b) Realización de /x/

- c) Realización de /tʃ/
- d) Realización de /-s/ implosiva
- e) Realización de /-r/ y /-l/ implosivas
- f) La /-n/ anorgánica

### 3.2.2. Gramática

- g) Alternancia morfemática y valores de diminutivos
- h) Subida de clíticos en complejos verbales
- i) Perífrasis de obligación
- j) Usos de *sino* + verbo en forma personal

## 4. PRIMERAS CONCLUSIONES

Los trabajos incluidos en el apartado 3.2.1. permiten recoger la especificidad del español del área en que se encuentra inscrito el grupo granadino: Granada y, en general, el andaluz oriental.

Así, por ejemplo, tales trabajos nos han permitido asistir a un rápido y brusco proceso de transformación del español hablado en Granada. Los análisis en tiempo aparente nos han acercado el momento descrito por los maestros del *ALEA* y han permitido aproximarnos a los agentes que estuvieron en la base de dicha transformación. Según el magno informe de Alvar, Llorente y Salvador (1961-73), el español de Granada en 1953, año en que se recogen las encuestas del *ALEA*, se caracterizaba, esencialmente, por seseo, aspiración de /x/ y realización fricativa de /tʃ/; sin embargo, los resultados de los análisis llevados a cabo por los miembros de grupo de investigación granadino muestran una situación radicalmente distinta, en la que los rasgos dominantes son la distinción fonológica de s/θ, la fricación no tensa de /x/ y la africada palatal /tʃ/. Hemos comprobado que este volumen de cambios está asociado a causas comunes vinculadas, en gran medida, a un conjunto de factores sociohistóricos, complejos, que tienen que ver con la existencia de una guerra (la guerra civil española de 1936-39) y una intensa postguerra en la que tuvieron lugar condicionantes que afectaron a una generación que se vio obligada a vivir en un contexto muy cambiante y adoptar, en consecuencia, hábitos sociales y lingüísticos muy diferentes a los que presentaba la generación precedente (Moya y Sosinski, 2015; Moya, 2018, 2022). El grupo social que inicia todo este proceso no se somete, estrictamente, a los límites de lo que conocemos por generación, pues suele extenderse por un espacio temporal más amplio; puede, incluso, incluir la generación siguiente. Se trata, en realidad, de una cohorte social fuertemente motivada y dotada de la capacidad necesaria para transmitir y consolidar sus propias pautas de conducta.

Hemos comprobado, igualmente, que estos cambios detectados en Granada tienen un reflejo similar en más puntos de Andalucía: Málaga y Huelva (Villena, 1996; Regan, 2017). Sin embargo, es de destacar que otras comunidades de habla también andaluzas, Córdoba y Sevilla en particular, se muestren ajenas a tales cambios y, por el contrario, conserven sus formas tradicionales (Uruburu, 1996; Santana, 2017). Entendemos que la divergencia mostrada por estos puntos es debida, muy probablemente, a que las formas tradicionales del habla de dichas comunidades se han constituido en identitarias de la zona (vid. Moya, 2022). Ello, sin embargo, no ha impedido que los procesos sociales asociados con las clases altas y con la convergencia, a que tienden estos estratos prestigiosos, muestren un repunte moderado de las formas del estándar. Pero es evidente que los cambios profundos, que la mencionada, y activa, cohorte social introdujo en Granada, Málaga y Huelva, encontraron una fuerte resistencia en comunidades donde las formas tradicionales estaban muy asentadas (Moya, 2022). Córdoba y Sevilla (Menéndez Pidal, 1962) sesean y aspiran desde el momento mismo de su conquista (1246 y 1248 respectivamente).

Los resultados de /-s/ implosiva y las consecuencias acarreadas en el área granadina y andaluza en general han sido otro de los objetos de especial atención para Grupo de investigación granadino. La pérdida de este fonema ha ocasionado trastornos de carácter no solo semántico, *pisto/pito*, de número *casa/casas*, o persona *canta/cantas*, sino también sintácticos relacionados con la construcción de la frase nominal: «las hermanas bailarinas» vs. «la hermana bailarina». La situación creada por la pérdida de /-s/ ha dado ya lugar a dos fenómenos de especial relevancia en el entorno andaluz –sus efectos, no obstante, se extienden incluso por áreas limítrofes–: las vocales proyectadas del oriente y el nacimiento de una africada prepalatal sorda [tʰ] en el centro y occidente. Pero no serían extrañas otras modificaciones más profundas, que pueden afectar a la estructura sintáctica de la lengua, vinculadas con problemas de concordancia en el sistema de la frase nominal: [lah- ermána bailarína] ‘las hermanas bailarinas’; [loh- ermáno meyíθo] ‘los hermanos mellizos’ (Moya, 2018).

#### 4. FINAL

El grupo de investigación *Estudios de español actual* ha trabajado de la mano de PRESEEA durante los 25 años que se han conmemorado en el congreso de Alcalá de Henares de 2023 y ha cumplido fielmente no solo con los cometidos propuestos en el interior del proyecto sino también con las obligaciones que le venían impuestas por su inscripción en una zona de profundo cambio lingüístico, como es Granada, el oriente andaluz y, por extensión el andaluz en general y el español en su conjunto.

Desde nuestro conocimiento actual del habla urbana, consideramos que se hace necesario un acercamiento paralelo al habla rural desde una perspectiva sociolingüística. Son varios los proyectos que se interesan por el habla rural, pero, en general, desde propuestas dialectales, históricas; sin embargo, la situación del mundo rural, el abandono del cam-

po o el aislamiento en el que se encuentra inmerso están dando lugar a la aparición de formas muy avanzadas que es necesario valorar, conocer mejor y determinar las razones históricas, sociales, psicológicas, económicas, etc. que están actuando en su desarrollo. No cabe duda de que la Sociolingüística cuenta con herramientas idóneas –unas ya activadas y otras que habrá de encontrar– para el estudio de este espacio abandonado en el que la lengua circula por territorios más cercanos al proceder juglaresco y popular que al de la tensión social que se observa en los espacios urbanos. La trayectoria de las hablas rurales conduce más a la fragmentación que a la conservación del idioma. Los estudios lingüísticos pueden colaborar con otros movimientos sociales cuya finalidad es la conservación y la integración social de un mundo que se desintegra.

La Sociolingüística puede y debe acercarse al mundo rural para dar a conocer sus luces y sus sombras.

## REFERENCIAS

- Alvar, Manuel, Antonio Llorente Maldonado de Guevara y Gregorio Salvador (1961-73). *Atlas lingüístico y etnográfico de Andalucía*. Granada: CSIC.
- Martínez Moya, María Dolores y Juan Antonio Moya Corral (1997). «Reacciones actitudinales hacia la variación dialectal en hablantes granadinos». *LEA*, 22(2): 137-160.
- Melguizo Moreno, Elisabeth (2005). «Formación dialectal: una propuesta teórica. Análisis sociolingüístico de s/θ en un grupo de inmigrantes procedentes de Pinos Puente e instalados en Granada». *Interlingüística*, 15(2): 981-990.
- Menéndez Pidal, Ramón (1962). «Sevilla frente a Madrid. Algunas precisiones sobre el español de América». *Miscelánea homenaje a André Martinet, III*. La Laguna, pp. 99-165.
- Molina Serrato, Francisca (2003). «La acomodación a la norma granadina. A propósito de un grupo de alhameños instalados en Granada». En J. A. Moya Corral y M. I. Montoya Ramírez (eds.), *Variación lingüística y enseñanza de la lengua española. Actas de las VIII Jornadas sobre la enseñanza de la lengua española*. Granada: Editorial Universidad de Granada, pp. 293-305.
- Moya Corral, Juan Antonio (1996). «La confusión s/θ en la ciudad de Granada». En Gutiérrez Díaz, Francisco, (ed.), *El Español, Lengua Internacional (1492-1992). Acta del I Congreso Internacional de AESLA, Granada, 23-26 de septiembre de 1992*. Murcia: Universidad, pp. 403-407.
- Moya Corral, Juan Antonio (1997). «Desarraigo social y cambio lingüístico: el ejemplo de Granada». En A. Narbona Jiménez y M. Roperó Muñoz, *Actas del congreso del Habla Andaluza [Sevilla del 4 al 7 de marzo, 1997]*. Sevilla: Seminario Permanente del Habla Andaluza, pp. 623-634.
- Moya Corral, Juan Antonio (2000). «La irradiación de la norma lingüística de Granada hacia su área metropolitana». En R. Morales Raya (ed.), *Homenaje a*

- la profesora María Dolores Tortosa Linde. Granada: Editorial Universidad de Granada, pp. 393-402.
- Moya Corral, Juan Antonio (2007, 2008, 2009). *Corpus oral para su estudio sociolingüístico del español hablado en Granada*. Granada: Editorial Universidad de Granada.
- Moya Corral, Juan Antonio (2018). «Sobre el equilibrado reajuste de las hablas andaluzas». *Itinerarios*, 28: 35-66.
- Moya Corral, Juan Antonio (2022). «Destrucción y (re)construcción. Convergencia e identidad». En C. Díaz Alayón, (ed.), *Studia philologica in honorem José Antonio Samper*. Academia Canaria de la Lengua y Arco/Libros-La Muralla S.L., pp. 771-781.
- Moya Corral, Juan Antonio y Emilio J. García Wiedemann (1995). *El habla de Granada y sus barrios*. Granada: Universidad.
- Moya Corral, Juan Antonio y Marcin Sosinski (2015). «La inserción social del cambio. La distinción s/θ en Granada. Análisis en tiempo aparente y en tiempo real». *LEA*, 37(1): 33-72.
- Regan, Brendan (2017). «A study of ceceo variation in Western Andalusia (Huelva)». *Studies on Hispanic and Lusophone Linguistics*, 10(1): 119-160.
- Santana Marrero, Juana (2016-7). «Factores externos e internos influyentes en la variación de /θs/ en la ciudad de Sevilla». *Analecta malacitana*, 39: 143-177.
- Villena Ponsoda, Juan A. (1996). «Convergence and divergence in a standard-dialect continuum: Networks and individuals in Malaga». *Sociolinguística*, 10: 112-137.
- Uruburu, Agustín (1996). «La lengua hablada en Córdoba». *Revista Española de Lingüística Aplicada*, 11: 225-250.

# UNA MIRADA AL PASADO, PRESENTE Y FUTURO DEL PROYECTO PRESEEA EN LAS PALMAS DE GRAN CANARIA

Marta Samper Hernández  
*Universidad de Las Palmas de Gran Canaria*

## 1. INTRODUCCIÓN

La finalidad de la presente aportación es trazar una panorámica, necesariamente sucinta por razones de espacio, sobre la trayectoria del equipo de la Universidad de Las Palmas de Gran Canaria dentro del *Proyecto para el Estudio Sociolingüístico del Español del España y de América* (PRESEEA). La historia del proyecto está estrechamente ligada, como es sabido, a la institución grancanaria, ya que su nacimiento, a partir de una idea expuesta tres años antes en Veracruz, tuvo lugar en el congreso de la ALFAL celebrado en la Facultad de Filología de la ULPGC en 1996. Sin embargo, el equipo palmense no se incorporó de inmediato a la nueva empresa debido a que sus miembros se encontraban ligados a otros proyectos que requerían exclusiva dedicación: además del *Proyecto Panhispánico de Disponibilidad Léxica*, coordinado por López Morales, y del *Proyecto de Difusión del Español en los Medios*, dirigido por Raúl Ávila, destaca entre ellos el *Proyecto de la norma Culta Hispánica «Juan M. Lope Blanch»*, que es considerado, a pesar de las evidentes diferencias que existen entre ambos, el «antecedente más claro y cercano» de PRESEEA (Moreno Fernández, 2021: 10). El hecho de que Las Palmas de Gran Canaria haya sido –junto con Madrid y Sevilla– una de las comunidades de habla españolas que han formado parte de estos dos grandes proyectos ha supuesto una enorme ventaja, ya que ha permitido contar con dos corpus de muestras orales de actuación que han posibilitado provechosos trabajos comparativos en tiempo real. Así, se han publicado investigaciones sobre la evolución de variables fonológicas, como la /d/ intervocálica (Samper Padilla y Hernández Cabrera, 2010) o la /s/ implosiva (Samper Padilla y Hernández Cabrera, 2016a), o sociopragmáticas,

como la atenuación (Samper Hernández, 2018) o el uso de los marcadores del discurso (Hernández Cabrera y Samper Hernández, 2014; Hernández Cabrera, 2015). En el año 2004 la comunidad de habla palmense se integró en el marco del PRESEEA mediante el proyecto *Estudio sociolingüístico del español de Las Palmas de Gran Canaria*, financiado por el Ministerio de Educación y Ciencia, en coordinación con los equipos de Universidad de Valencia, la Universidad de Málaga, la Universidad de Granada, la Universidad de Alcalá y la Universidad de Lérida. A partir de ese momento, el grupo de investigación canario ha estado integrado en todos los proyectos nacionales coordinados que se han ido enlazando, de manera ininterrumpida, hasta el actual *Procesos de variación y cambio espaciales y sociales. Español de España: Agenda 2050 (VARES-AGENDA50)*, que finalizará en mayo de 2024. Todo ello ha permitido que el español de Canarias, al menos en su variante palmense, haya pasado de ser considerado «el peor conocido de todos los españoles» (Alvar López, 1962: 60) a contar con amplias y profundas descripciones en todos los planos de análisis lingüístico. Dedicaremos las próximas páginas al comentario de algunos de los patrones sociolingüísticos hallados a partir de las entrevistas del corpus PRESEEA-Las Palmas.

## 2. EL CORPUS PRESEEA-LP: PASADO, PRESENTE Y FUTURO

Las grabaciones que componen el corpus PRESEEA de Las Palmas de Gran Canaria comenzaron a realizarse en el año 2005 con las pertenecientes al nivel de instrucción alto, y se completaron en 2010, con las de los hablantes sin estudios o con estudios básicos. En total, se entrevistó a 72 hablantes, todos nacidos y residentes en la capital grancanaria. Como se puede observar, se trata de un número superior a los 54 informantes que corresponden a una población de 376 953 habitantes como era la de Las Palmas de Gran Canaria en el año 2004; sin embargo, el equipo palmense prefirió adaptar la muestra a las características de la diseñada para las ciudades de más de 500 000 habitantes con la finalidad de asegurar la fiabilidad de los datos que se obtuvieran en el análisis de los diferentes fenómenos, sobre todo teniendo en cuenta que la ciudad contaba con un área metropolitana de más de 600 000 habitantes. Para la selección de los informantes se siguió un criterio por cuotas con afijación uniforme, con la preestratificación de las variables sexo (36 hombres y 36 mujeres), edad (24 hablantes por grupo etario: de 20 a 34; de 35 a 54; y más de 55 años) y grado de instrucción (24 sujetos sin estudios o con estudios básicos; 24 con estudios medios; y 24 con formación universitaria).

De acuerdo con las pautas de PRESEEA (Moreno Fernández, 2021), se obtuvieron 45 minutos de grabación por informante –aunque en muchos casos las entrevistas se extendieron por encima de la hora– y se favoreció la aparición de diferentes secuencias discursivas: narración, descripción, exposición, argumentación y diálogo.

Aunque en un principio se subieron a la página de PRESEEA (<https://preseea.uah.es/>) 18 muestras orales de 10 minutos cada una, desde 2023 se encuentran alojadas 18 grabaciones completas, todas ellas con sus correspondientes transcripciones.

Con respecto al futuro cercano, se prevé la recogida de un nuevo corpus, de características similares al ya existente, a partir del año 2025. Esto posibilitará la realización de provechosos estudios en tiempo real que comprueben la marcha de los fenómenos lingüísticos que se han ido describiendo a partir de las muestras de habla anteriores.

### 3. ANÁLISIS DE VARIABLES SOCIOLINGÜÍSTICAS. ALGUNOS PATRONES DE COMPORTAMIENTO EN EL ESPAÑOL DE LAS PALMAS DE GRAN CANARIA A PARTIR DE LOS RESULTADOS DE PRESEEA-LP

#### 3.1. Variables sociolingüísticas contempladas

A lo largo de los casi 20 años de PRESEEA en Las Palmas de Gran Canaria se han analizado variables de todos los niveles lingüísticos: algunas ya han sido ampliamente descritas, otras se encuentran aún en proceso de estudio. Se exponen en la tabla 1, de manera esquemática, cuáles son algunas de estas variables:

**TABLA 1. VARIABLES ANALIZADAS Y EN PROCESO DE ANÁLISIS EN EL CORPUS PRESEEA-LP.**

Variación fónica	-/d/-, -/s/, -/r/
Variación gramatical	Pluralización de <i>haber</i> impersonal, uso del diminutivo, distribución de clíticos con perífrasis verbales, leísmo, presencia/ausencia de sujeto, futuro, pronombres de retoma
Variación léxica	disponibilidad léxica en diversas muestras de hablantes y desde diferentes perspectivas
Variación sociopragmática	atenuación, marcadores discursivos

Del mismo modo, a partir del año 2018 el equipo palmense comenzó a analizar las creencias y actitudes lingüísticas de diversos grupos de hablantes grancanarios hacia las diferentes variedades normativas del español, dentro del proyecto internacional PRECAVES XXI (*Proyecto para el estudio de las creencias y actitudes hacia las variedades del español en el siglo XXI*) coordinado por Cestero Mancera y Paredes García (Universidad de Alcalá). Tras una primera fase en que se recogieron y analizaron las respuestas de estudiantes universitarios de Filología (Hernández Cabrera y Samper Hernández, 2018, 2021), en una segunda etapa se está trabajando con una muestra de 108 sujetos distribuidos por cuotas con afijación uniforme, en consonancia con las pautas metodológicas de PRESEEA (Hernández Cabrera y Samper Hernández, 2022; Samper Hernández, en prensa a y b). Tanto la realización de estas últimas encuestas como los análisis correspondientes constituyen un objetivo específico del proyecto coordinado en marcha desde junio de 2020.

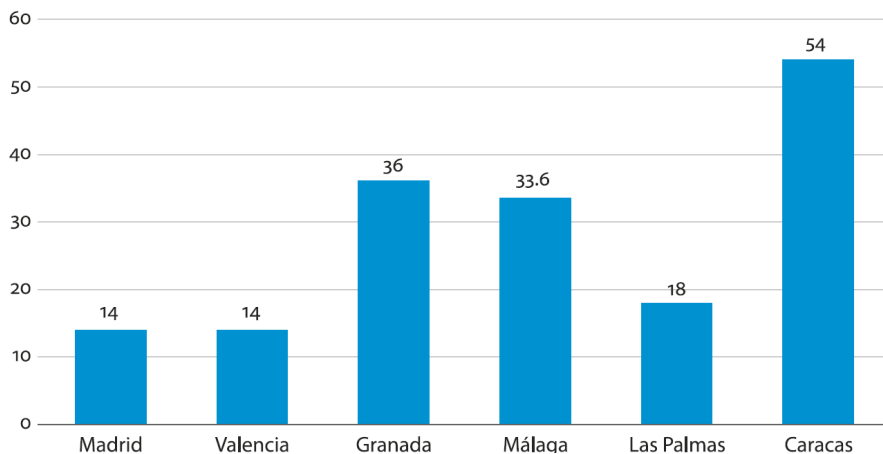
## 3.2. Algunos patrones sociolingüísticos en el español de Las Palmas de Gran Canaria

### 3.2.1. Realización de la /d/ intervocálica

En coordinación con muchos de los equipos de PRESEEA, la variable /d/ intervocálica ha sido una de las priorizadas en el estudio del español palmense, de manera que ya se cuenta con un amplio número de publicaciones que describen en profundidad el comportamiento de este fonema según diferentes condicionantes lingüísticos, sociológicos y situacionales (Samper Padilla y Hernández Cabrera, 2010, 2011; Samper Padilla *et al.*, 2010, 2011; Malaver Arguinzones y Samper Padilla, 2016; Samper Padilla y Samper Hernández, 2020). Además, dos de los miembros del equipo de Las Palmas han sido los responsables, junto con Irania Malaver, de elaborar la *Guía PRESEEA de estudio de la /d/ intervocálica* (Samper Padilla *et al.*, 2021).

Los resultados obtenidos en los diferentes trabajos permiten afirmar que en la comunidad de habla palmense la elisión de la -/d/- no constituye un fenómeno en progreso. Sucede precisamente lo contrario: en la capital grancanaria parece estarse produciendo, desde hace años, una regresión de la pérdida del fonema, en beneficio de una pronunciación relajada. Se enumeran a continuación algunas de las conclusiones más destacadas obtenidas por el equipo palmense en los trabajos citados:

**FIGURA 1. ÍNDICES DE ELISIÓN DE -/D/- EN DIFERENTES COMUNIDADES DE HABLA ANALIZADAS EN EL PROYECTO PRESEEA SAMPER PADILLA Y SAMPER HERNÁNDEZ (2020: 225)<sup>1</sup>**



<sup>1</sup> Los trabajos de los que se han extraído los datos de estas comunidades de habla pueden consultarse en Samper Padilla y Samper Hernández (2020: 225, nota 8).

- a) Se encuentra un índice moderado de elisión de la /d/ intervocálica, sobre todo si se comparan los porcentajes hallados con los de otras zonas dialectales. En la figura 1 se pueden encontrar los índices de pérdida del fonema en el entorno general de diferentes comunidades de habla estudiadas en el marco del proyecto:
- b) Este comportamiento moderado se encuentra en todos los grupos sociales contemplados: incluso entre los hablantes con estudios básicos el porcentaje hallado en el entorno amplio (22,6%) resulta menor que los registrados por los universitarios de las comunidades de habla malagueña y granadina.
- c) Cuando se contrastan los datos de las grabaciones de PRESEEA –realizadas entre los años 2005 y 2010– con los correspondientes a las entrevistas recogidas en los años 80 del siglo XX por Samper Padilla (1990), se observa que la diferencia entre los dos niveles de instrucción extremos en cuanto a la elisión de la -d/-, si bien sigue siendo importante, se ha reducido considerablemente. Este acercamiento entre los grupos extremos se produce porque en las muestras más recientes los hablantes del nivel de estudios básico eliden menos que en las más antiguas, donde se hallaba un porcentaje del 46,9%. Por tanto, se pone de manifiesto un avance progresivo de la restitución de la /d/ intervocálica, algo que ya señalaba Samper Padilla en 1990 a partir de los datos procedentes de la variable edad: «Estos resultados indican que la pérdida de la dental es un fenómeno en regresión en el español de Las Palmas, de acuerdo con las investigaciones previas [las de Catalán (1960) y Trujillo Carreño (1980)] que señalaban una restitución moderna de -d/-» (Samper Padilla, 1990: 272).
- d) En el mismo sentido, el análisis según la variable generacional en los materiales de PRESEEA muestra una pérdida más frecuente del fonema por parte de los hablantes mayores, al contrario de lo que ocurre en comunidades como Caracas, Granada y Málaga (Samper Padilla y Malaver Arguinzones, 2016: 331; Samper Padilla y Samper Hernández, 2020: 229).
- e) La elisión de /d/ en Las Palmas de Gran Canaria se produce mayoritariamente en el indefinido *todo* (y sus variantes) cuando actúa como determinante, algo que supone una evidente diferencia con lo que sucede en las comunidades peninsulares, donde el proceso de debilitamiento de la dental tiene su manifestación más relevante en los participios en *-ado*. Este hecho confirma lo señalado por Alvar López en 1972, cuando se refería a «la creencia, muy extendida [entre los hablantes de Las Palmas] de que los peninsulares dicen siempre *-ao* y los canarios, también siempre, *-ado*» (Alvar López, 1972: 86) y guarda una evidente relación con los índices de frecuencia de aparición del participio, mucho mayores en las comunidades peninsulares que en Canarias. No es extraño, de acuerdo con esta última explicación, que la mayor elisión de la -d/- en el determinante *todo* concuerde con lo que ocurre también en Caracas.

### 3.2.2. La pluralización de haber impersonal

Entre los trabajos dedicados por el equipo palmense a las variables gramaticales, destacan los que abordan el fenómeno de la pluralización del verbo impersonal *haber* (Samper Padilla y Hernández Cabrera, 2012, 2016b; Hernández Cabrera, 2016). Corresponde también a Hernández Cabrera la elaboración de la guía de estudio de esta variable, en coordinación con Gómez Molina (Gómez Molina y Hernández Cabrera, 2021).

Los resultados de los análisis del corpus PRESEEA-LP señalan una escasa pluralización del verbo *haber* impersonal en la comunidad de habla palmense, en contra de lo que se había afirmado en numerosísimos estudios previos (Almeida y Díaz Alayón, 1989, Ortega Ojeda, 1996, entre otros). Hernández Cabrera (2016: 144) afirma, en este sentido: «los investigadores del español de Canarias han considerado, de forma casi unánime, que el uso personal de *haber* en las islas es un fenómeno generalizado, incluso entre hablantes de los niveles sociales más altos». En la investigación de Hernández Cabrera (2016) a partir de las 72 entrevistas de PRESEEA las únicas variables independientes que ofrecieron resultados estadísticamente significativos son las que se recogen en la tabla 2. Como puede observarse, el rasgo + humano de la frase nominal favorece la pluralización del verbo impersonal; en lo que concierne a los condicionantes sociales, los grupos que la favorecen son los informantes mayores de 55 años y los hablantes de nivel de instrucción bajo y, sobre todo, medio.

**TABLA 2. VARIABLES SELECCIONADAS POR GOLDVARB PARA EXPLICAR LA PLURALIZACIÓN DE *HABER*. HERNÁNDEZ CABRERA (2016: 159)**

Variables independientes		Peso probabilístico
2. Rasgo semántico de la FN	- hum	.406
	+ hum	.594
14. Edad	1ª generación	.439
	2ª generación	.442
	3ª generación	.618
15. Nivel de estudios	Universitario	.303
	Medio	.644
	Primario	.560
Input 0.418 Log likelihood = -214.008 Significance = 0.014 (subida)/0.083 (bajada)		

Los resultados obtenidos podrían estar evidenciando un cambio en marcha hacia la regresión del fenómeno en Las Palmas de Gran Canaria, dado que la pluralización no es favorecida ni por el nivel educativo alto ni por las dos primeras generaciones. A estos datos habría que sumar el correspondiente a la variable sexo, no significativo estadísticamente pero sí relevante, ya que las mujeres son las que utilizan las formas plurales en menor medida.

La conclusión a la que se llega tras el análisis de las muestras de actuación orales en el formato de entrevista semidirigida debe compararse con los resultados que se extraen de la aplicación de diferentes pruebas de aceptabilidad. Samper Padilla y Hernández Cabrera (2012) partieron de las respuestas de dos grupos de informantes: por un lado, 39 hablantes de la tercera generación de nivel educativo medio y, por otro, 39 estudiantes de primer curso universitario. Los índices de aceptación de las formas plurales recogidos en la tabla 3 muestran unos resultados que contrastan con lo hallado en las entrevistas semiformales y que concuerdan en mayor medida con lo que se había señalado en estudios previos sobre el español de Canarias: dejando a un lado las formas de primera persona, en ciertos tiempos verbales los porcentajes son realmente elevados, sobre todo entre los hablantes mayores. Se observa, eso sí, una disminución evidente del grado de aceptación entre los informantes de menor edad, algo que resulta especialmente notable en el caso del indefinido *hubieron*.

**TABLA 3. PORCENTAJES DE ACEPTACIÓN DE DISTINTAS FORMAS PERSONALES DE *HABER* SEGÚN LA VARIABLE GENERACIONAL EN GRAN CANARIA. SAMPER PADILLA Y HERNÁNDEZ CABRERA (2012: 751)**

Habemos	15,4	7,7
Habíamos	38,5	30,8
Hubieron	71,8	48,7
Han habido	74,3	64,1
Habían	76,9	46,1
Suelen haber	79,5	61,5
Habrán	79,5	71,8
Habrían	82	71,8
T	64,7	50,3

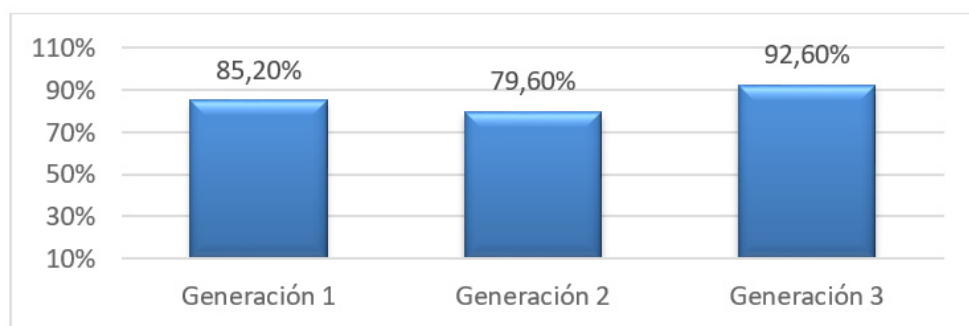
Muy recientemente, González Obra (2023) ha aplicado, de manera experimental, otra prueba de aceptabilidad a una muestra de 18 hablantes grancanarios distribuidos por cuotas con afijación uniforme. Las oraciones propuestas se limitan a las formas simples y compuestas de tercera persona; no aparecen, por tanto, casos de primera persona del plural ni de perífrasis verbales. Los resultados generales (tabla 4) muestran unos porcentajes de aceptación muy llamativos, ya que siempre se sitúan por encima del 80%:

**TABLA 4. PORCENTAJES DE ACEPTACIÓN DE DISTINTAS FORMAS PERSONALES DE *HABER* EN GRAN CANARIA: DATOS GENERALES. ADAPTADO DE GONZÁLEZ OBRA (2023)**

Hubieron	83,3%
Habían	83,3%
Han habido	86,1%
Habrían	88,9%
Hayan	88,9%
T	86,1%

Cuando se atiende a la variable generacional en estos mismos 18 sujetos (figura 2), se observa que estos índices se encuentran en todos los grupos de edad, si bien se aprecia un ligero descenso entre los hablantes de 35 a 54 años, que quizás pueda deberse a su mayor implicación «en el mundo de la competencia profesional, económica y de ascenso en la escala social» (Silva-Corvalán y Enrique-Arias, 2017: 124). En cualquier caso, los datos empíricos no corroboran la existencia de un cambio en marcha hacia la regresión de la pluralización del verbo *haber*.

**FIGURA 2. PORCENTAJES DE ACEPTACIÓN DE LAS FORMAS PLURALES DE *HABER* SEGÚN LA VARIABLE GENERACIONAL EN GRAN CANARIA. ADAPTADO DE GONZÁLEZ OBRA (2023: 18)**

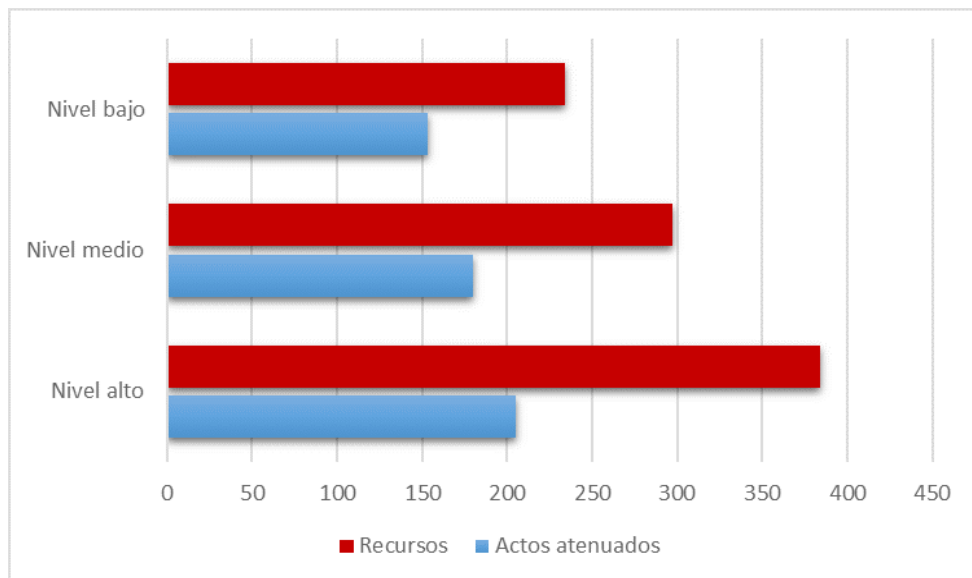


Los contundentes índices de aceptación obtenidos por González Obra en 2023 con una muestra de 18 hablantes grancanarios subrayan la necesidad de realizar nuevas investigaciones que comprueben la evolución del fenómeno en las muestras orales de actuación que previsiblemente serán recogidas, como se apuntó más arriba, a partir de 2025.

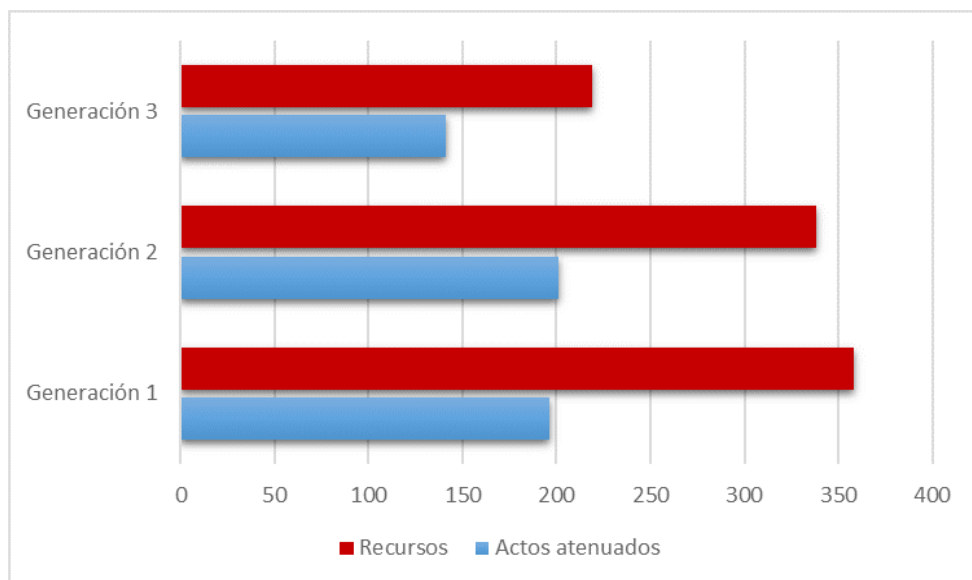
### 3.2.3. *Atenuación*

Como ejemplo de las aportaciones del equipo palmense que han abordado variables discursivas, se exponen en las siguientes líneas, de forma muy resumida, algunas conclusiones sobre la categoría sociopragmática de la atenuación (Samper Hernández, 2013, 2017, 2018, 20120; Albelda *et al.*, 2020). En el análisis de 9 horas de entrevista semidirigida a 18 informantes distribuidos de manera simétrica según los condicionantes extralingüísticos sexo, edad y nivel de instrucción, se ha hallado un uso de la estrategia mitigadora inferior al encontrado en otras comunidades como Madrid (Vallecas), Santiago de Chile o Valencia. No obstante, como se aprecia en las figuras 3 y 4, podríamos estar hablando de una situación de cambio en marcha en la capital grancanaria, ya que los entrevistados que recurren más a esta estrategia son los pertenecientes a la primera y segunda generación y los universitarios.

**FIGURA 3. ATENUACIÓN SEGÚN NIVEL DE INSTRUCCIÓN. ADAPTADO DE SAMPER HERNÁNDEZ (2017)**



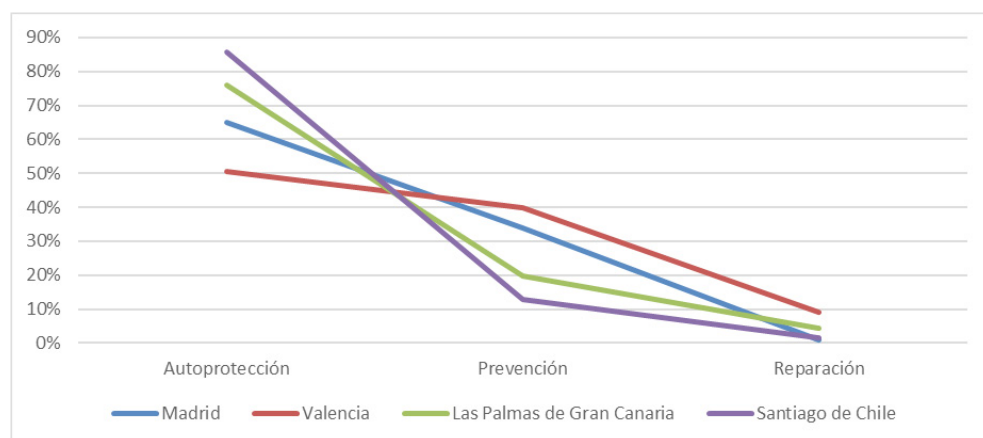
**FIGURA 4. ATENUACIÓN SEGÚN EDAD. ADAPTADO DE SAMPER HERNÁNDEZ (2017)**



En lo que respecta a las funciones y tipos de atenuación, se encuentra una evidente proximidad entre Las Palmas y Santiago de Chile, frente a otras comunidades

de habla como Valencia y Madrid (Vallecas). En la figura 5 se puede comprobar que en las dos comunidades del español atlántico se atenúa sobre todo para proteger la imagen propia y mucho menos para prevenir la amenaza a la imagen ajena. Tanto en Madrid como, sobre todo, en Valencia, los porcentajes de una y otra función resultan mucho más cercanos. Por otro lado, en Las Palmas y Santiago se utilizan más recursos relacionados con el contenido proposicional del acto de habla (atenuación del *dictum*), registrados en proporciones claramente menores en las comunidades madrileña y valenciana.

**FIGURA 5. FUNCIONES DE LA ATENUACIÓN EN 4 COMUNIDADES DE HABLA. SAMPER HERNÁNDEZ (2020: 918)**



### 3.2.4. La variación léxica en Gran Canaria: algunas pinceladas

Se trata de un aspecto ampliamente estudiado por el equipo palmense a través de la metodología de la disponibilidad léxica, importada al mundo hispánico en los años 70 del siglo XX por López Morales. Los trabajos han atendido tanto al léxico disponible de los estudiantes preuniversitarios (Samper Padilla y Hernández Cabrera, 1997), en consonancia con el ambicioso proyecto panhispánico coordinado por el mismo López Morales, como al vocabulario que tienen disponible -listo para ser usado si la conversación lo requiere- otros grupos de hablantes: escolares (Samper Hernández, 2009), mayores de 50 años (Cruz Ventura, 2016) o individuos distribuidos a partir de los criterios metodológicos de PRESEEA (Medina Peñate, 2018). También se han hecho incursiones en el léxico disponible de niños de 6 años, con la novedad metodológica fundamental de que las encuestas fueron orales y estuvieron enfocadas desde una perspectiva lúdica (Samper Padilla *et al.*, 2019; Samper Hernández *et al.*, 2020). Dada la diversidad de estos estudios, no es posible extraer unas conclusiones

globales sin entrar en consideraciones más profundas, por lo que animamos a las personas interesadas a que consulten las diferentes aportaciones canarias a esta fructífera línea de investigación.

#### 4. A MODO DE CONCLUSIÓN

Estos casi 20 años de participación del equipo de Las Palmas en PRESEEA han supuesto un verdadero avance en la descripción sociolingüística del español de la capital grancanaria. Sin duda, el hecho de trabajar dentro de un proyecto coordinado, con unas directrices metodológicas rigurosas y que favorecen, además, la comparación entre las diferentes comunidades de habla ha tenido mucho que ver en ello. En este artículo se han repasado algunos de los patrones sociolingüísticos descritos en las diferentes investigaciones que se han desarrollado desde el año 2004, pero es mucho lo que queda aún por descubrir acerca de una variedad que, a pesar de contar con pocos hablantes, tiene una evidente importancia por su papel de puente entre España y América. Especial interés reviste la recolección de nuevas muestras de actuación orales, que permitirá provechosos estudios en tiempo real que ayudarán a confirmar o a refutar las conclusiones que se han ido aportando hasta el momento a partir de los materiales anteriores. Con esta apasionante perspectiva de futuro, PRESEEA se dibuja, efectivamente, como una verdadera joya para el equipo palmense.

#### REFERENCIAS

- Albelda Marco, Marta, Ana M.<sup>a</sup> Cestero Mancera, Silvana Guerrero González y Marta Samper Hernández (2020). «Variación sociopragmática y geolectal en el uso de la atenuación», *Lengua y habla*, 24: 1-53.
- Almeida, Manuel y Carmen Díaz Alayón (1989). *El español de Canarias*, Santa Cruz de Tenerife: editado por los autores.
- Alvar López, Manuel (1962). *Dialectología española*, Madrid: C.S.I.C. (Cuadernos Bibliográficos, VII).
- Alvar López, Manuel (1972). *Niveles socio-culturales en el habla de Las Palmas de Gran Canaria*, Las Palmas de Gran Canaria: Cabildo Insular de Gran Canaria.
- Catalán, Diego (1960). «El español canario. Entre Europa y América», *Boletín de Filología*, 19: 317-337.
- Cruz Ventura, Sarai (2016). *La disponibilidad léxica de hablantes canarios adultos y comparación con los corpus juveniles*, [Tesis doctoral], Gran Canaria: Universidad de Las Palmas de Gran Canaria.
- Gómez Molina, José Ramón y Clara E. Hernández Cabrera (2021). *Guía PRESEEA de Estudio del verbo impersonal haber. Guías PRESEEA de Estudios*. Disponible en: <https://preseea.uah.es/documentos-preseea-de-investigacion>

- González Obra, Patricia (2023). *Pluralización del verbo haber impersonal: usos y actitudes de hablantes canarios*, [Trabajo de fin de título], Gran Canaria: Universidad de Las Palmas de Gran Canaria.
- Hernández Cabrera, Clara E. (2015). «Marcadores discursivos de Las Palmas de Gran Canaria», En A. Valencia y A. Viguera, (coords.), *Más sobre marcadores hispánicos: usos de España y América en el corpus de estudio de la norma culta*, Ciudad de México: UNAM, pp. 257-301.
- Hernández Cabrera, Clara E. (2016). «Variación de haber impersonal en el español de Las Palmas de Gran Canaria». *ELUA: Estudios de Lingüística*. Universidad de Alicante, 30: 141-162.
- Hernández Cabrera, Clara E. y Marta Samper Hernández (2014). «Marcadores del discurso de Las Palmas de Gran Canaria», En A. Valencia (coord.), *Marcadores discursivos en la Norma Culta Hispánica: 1964-2014*. Cuadernos de la ALFAL, 5, pp. 163-190.
- Hernández Cabrera, Clara E. y Marta Samper Hernández. 2018. «Creencias y actitudes de los jóvenes universitarios canarios hacia las variedades cultas del español», En A. M.<sup>a</sup> Cestero y F. Paredes (eds.), *Percepción de las variedades cultas del español: creencias y actitudes de jóvenes universitarios hispanohablantes*. Boletín de Filología, 53(2), pp. 179-208.
- Hernández Cabrera, Clara E. y Marta Samper Hernández (2021). «Creencias y actitudes de los jóvenes universitarios canarios hacia la variedad andaluza», En J. Santana y A. Manjón-Cabeza (eds.), *Percepción del andaluz culto: creencias y actitudes de jóvenes hispanohablantes y estudiantes de ELE*. Philologia Hispalensis, 35 (1), pp. 95-116.
- Malaver Arguinzones, Irania y José A. Samper Padilla (2016). «Estudio de la /d/ intervocálica en los corpus PRESEEA». *Boletín de Filología*, 51(2): 325-345.
- Medina Peñate, Inmaculada (2018). *El léxico disponible en el marco del PRESEEA: investigación en Las Palmas de Gran Canaria y en Madrid*, [Tesis doctoral], Alcalá de Henares: Universidad de Alcalá.
- Moreno Fernández, Francisco (2021). «Metodología del «Proyecto para el estudio sociolingüístico del español de España y de América» (PRESEEA)». *Documentos PRESEEA de investigación. Documentos de trabajo I*. Disponible en: <https://preseea.uah.es/documentos-preseea-de-investigacion>
- Ortega Ojeda, Gonzalo (1996). *La enseñanza de la lengua española en Canarias. Algunas perspectivas*, La Laguna: Instituto de Estudios Canarios.
- Samper Hernández, Marta (2009). *Evolución de la disponibilidad léxica en estudiantes grancanarios de enseñanza Primaria y Secundaria*, Las Palmas de Gran Canaria: Cabildo Insular de Gran Canaria.
- Samper Hernández, Marta (2013). «La atenuación lingüística en el español de Las Palmas de Gran Canaria», *Lingüística Española Actual*, 35(2): 325-348.
- Samper Hernández, Marta (2017). «Análisis sociolingüístico de la atenuación en el español de Las Palmas de Gran Canaria», En M. Albelda y W. Mihatsch (eds.),

- Atenuación e intensificación en géneros discursivos*. Madrid: Iberoamericana Vervuert, pp. 153-168.
- Samper Hernández, Marta (2018). «Un cambio en tiempo real: la atenuación entre hablantes universitarios de Las Palmas de Gran Canaria», *Rilce: Revista de Filología Hispánica*, 34(3): 1259-1279.
- Samper Hernández, Marta (2020). «Estudio sociolingüístico de los mecanismos atenuantes utilizados en entrevistas semidirigidas de Las Palmas de Gran Canaria». *Revista Signos*, 53(104): 910-934.
- Samper Hernández, Marta (En prensa a). «Creencias y actitudes de los canarios hacia la variedad lingüística chilena». *Lenguas Modernas*, 61.
- Samper Hernández, Marta (En prensa b). «Los canarios ante la variedad andaluza: creencias y actitudes lingüísticas», En R. Cruz, J. Santana e I. C. Santos, *El andaluz y su prestigio. Creencias y actitudes lingüísticas de los hispanohablantes hacia la variedad andaluza*, Granada: Comares.
- Samper Hernández, Marta, Clara E. Hernández Cabrera y José A. Samper Padilla (2020). «Incidencia del factor sexo/género en la disponibilidad léxica de niños grancanarios de 6 años». En F. Gimeno (coord.), *Lengua, sociedad y cultura: Estudios dedicados a Alberto Carcedo González*. Alicante: Universidad de Alicante, pp. 287-312.
- Samper Padilla, José A. (1990). *Estudio sociolingüístico del español de Las Palmas de Gran Canaria*, Las Palmas de Gran Canaria: La Caja de Canarias.
- Samper Padilla, José A. y Clara E. Hernández Cabrera (1997). «El estudio de la disponibilidad léxica en Gran Canaria: datos iniciales y variación sociolingüística». En J. Dorta y M. Almeida (coords.), *Contribuciones al estudio de la lingüística hispánica: homenaje al profesor Ramón Trujillo* (vol. 2). Santa Cruz de Tenerife: Montesinos, pp. 229-240.
- Samper Padilla, José A. y Clara E. Hernández Cabrera (2010). «Un estudio de variación fónica en dos corpus grancanarios». En R. M.<sup>a</sup> Castañer y V. Lagüéns (coords.), *De moneda nunca usada. Estudios dedicados a José M.<sup>a</sup> Enguita Utrilla*, Zaragoza: Institución Fernando El Católico, pp. 527-538.
- Samper Padilla, José A. y Clara E. Hernández Cabrera (2011). «La elisión de -d- intervocálica en los informativos de la televisión canaria: comparación con los datos de PRESEEA», En R. Ávila (ed.), *Variación del español en los medios*, Ciudad de México: El Colegio de México, pp. 289-312.
- Samper Padilla, José A. y Clara E. Hernández Cabrera (2012). «En torno a los usos personales de haber en el español de Las Palmas de Gran Canaria». En T. Jiménez, B. López, V. Vázquez y A. Veiga (coords.), *Cum corde et in nova grammatica. Estudios ofrecidos a Guillermo Rojo*, Santiago de Compostela: Universidad de Santiago de Compostela, pp. 743-754.
- Samper Padilla, José A. y Clara E. Hernández Cabrera (2016a). «Sobre un cambio en el español de Las Palmas de Gran Canaria: de [s] a [h] en el contexto prevocálico tónico». En A. López, A. Narbona y S. del Rey (eds.), *El español a través del*

- tiempo: estudios ofrecidos a Rafael Cano Aguilar*, Sevilla: Universidad de Sevilla, pp. 119-124.
- Samper Padilla, José A. y Clara E. Hernández Cabrera (2016b). «La pluralización de haber en dos niveles educativos de las Palmas de Gran Canaria», En A. Bañón, M.<sup>a</sup> M. Espejo, B. Herrero y J. L. López (eds.), *Oralidad y análisis del discurso: homenaje a Luis Cortés Rodríguez*, Almería: Universidad de Almería, pp. 577-594.
- Samper Padilla, José A., Clara E. Hernández Cabrera y Marta Samper Hernández (2019). «Disponibilidad léxica en niños grancanarios de 6 años. Comparación con México y Costa Rica». *Ogigia: Revista electrónica de estudios hispánicos*, 25, 139-163.
- Samper Padilla, José A., Clara E. Hernández Cabrera, Magnolia Troya Déniz, Ana M.<sup>a</sup> Pérez Martín y Marta Samper Hernández (2010). «Un estudio de variación fónica en el marco del PRESEEA: realizaciones de /d/- entre los hablantes de nivel universitario de Las Palmas de Gran Canaria». *Lingüística Española Actual*, 32(1): 5-34.
- Samper Padilla, José A., Clara E. Hernández Cabrera, Magnolia Troya Déniz, Ana M.<sup>a</sup> Pérez Martín y Marta Samper Hernández (2011). «Elisión de /d/ en los niveles de estudios primario y universitario de Las Palmas de Gran Canaria. Comparación con otras comunidades de habla», En J. J. de Bustos Tovar, R. Cano Aguilar, E. Méndez García de Paredes y A. López Serena (eds.), *Sintaxis y análisis del discurso hablado en español. Homenaje a Antonio Narbona*, Vol.1, Sevilla: Universidad de Sevilla, pp. 409-424.
- Samper Padilla, José A., Irania Malaver Arguinzones y Marta Samper Hernández (2021). «Guía PRESEEA de estudio de la /d/ intervocálica». *Guías PRESEEA de Estudios*. Disponible en: <https://preseea.uah.es/documentos-preseea-de-investigacion>
- Samper Padilla, José A. y Marta Samper Hernández (2020). «The weakening of intervocalic /d/ in the Spanish of Las Palmas de Gran Canaria. Comparison with other speech communities». *Spanish in Context*, 17(2): 221-246.
- Silva-Corvalán, Carmen y Andrés Enrique-Arias (2017). *Sociolingüística y pragmática del español* (segunda edición), Washington D.C.: Georgetown University Press.
- Trujillo Carreño, Ramón (1980). *Lenguaje y cultura en Masca. Dos estudios*. Santa Cruz de Tenerife: Interinsular Canaria.

**PRESEEA-SEVILLA:**  
**10 AÑOS DE SOCIOLINGÜÍSTICA EN LA CAPITAL HISPALENSE**

Juana Santana Marrero  
*Universidad de Sevilla*

**1. INTRODUCCIÓN<sup>1</sup>**

En octubre de 2023 se cumplieron 10 años de la incorporación plena de la ciudad de Sevilla al *Proyecto de estudio sociolingüístico del español de España y América* (PRE-SEEA), bajo la coordinación de quien escribe estas líneas. Parece que fue ayer cuando el profesor Manuel Ariza me ofreció la posibilidad de liderar esta iniciativa, que hasta el momento apenas había dado frutos en la capital hispalense. Se trataba de una apuesta que encajaba a la perfección dentro de las líneas de investigación que, durante décadas, habíamos venido desarrollando en el grupo de investigación *Sociolingüística andaluza: estudio sociolingüístico del habla de Sevilla*<sup>2</sup> (HUM 141), por lo que nos sentíamos afortunados de poder aprovechar la oportunidad que se nos brindaba. Desde entonces hasta ahora hemos pasado por diferentes momentos que incluyen, en síntesis, la enorme ilusión por formar parte del mayor proyecto de sociolingüística hispánica que existe en nuestros días, la gran responsabilidad de estar a la altura del nivel de exigencia de una iniciativa de esta naturaleza y el esfuerzo que supone sacar adelante las diferentes tareas que comprende el análisis de la lengua oral desde la perspectiva de la variación social: empezando por las labores de campo de recopilación del corpus, pasando por la trans-

---

<sup>1</sup> Este trabajo es parte del 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).

<sup>2</sup> <http://sociolingüisticaandaluza.snlc.es/>

cripción y el etiquetado de las muestras y terminando con el análisis de los fenómenos lingüísticos.

En este texto haremos un breve repaso de las principales etapas y facetas del equipo PRESEEA de la ciudad de Sevilla. Las limitaciones de espacio con las que contamos nos han obligado a ser necesariamente breves y, por tanto, a no poder exponer en detalle los resultados de las investigaciones desarrolladas a lo largo de esta última década. Por ello, invitamos a los lectores que así lo deseen a consultar los títulos que aparecen al final de este capítulo.

En Sevilla, como hemos adelantado, nos incorporamos a PRESEEA en 2013, algo más tarde que los otros equipos españoles (Granada, Las Palmas de Gran Canaria, Madrid, Málaga y Valencia). Esta circunstancia, lejos de ser una desventaja, nos permitió empezar nuestra participación con grandes dosis de confianza, pues, siempre que se la pedíamos, contábamos con la generosa ayuda que nos brindaban los colegas de esas otras ciudades. Ellos marcaban la senda que nosotros teníamos que seguir y podíamos aprender de los maestros que nos precedían.

Nuestro primer gran reto, que constituye también la primera etapa del equipo PRESEEA-Sevilla, fue la incorporación al proyecto coordinado *Estudios complementarios de los patrones sociolingüísticos de España*, junto a Granada, Las Palmas de Gran Canaria, Madrid y Málaga. Concretamente, nosotros constituimos el subproyecto *Patrones sociolingüísticos del español de Sevilla* (Ref. Ref. FFI2015-68171-C5-3-P)<sup>3</sup>, que se prolongó desde 2016 hasta 2019<sup>4</sup>. Una vez finalizado ese periodo, formamos, junto al equipo de Las Palmas de Gran Canaria, el subproyecto *Agenda 2050. El español de Sevilla y Las Palmas de Gran Canaria: procesos de variación y cambio espaciales y sociales* (Ref. PID2019-104982GB-C54), en vigencia hasta mayo de 2024, que pertenece al proyecto coordinado *Procesos de variación y cambio espaciales y sociales. Español en España: agenda 2050*. Ambas etapas comprenden dos procesos fundamentales: la recopilación y edición del corpus PRESEEA-Sevilla, que llevamos a cabo en el primer periodo, y el análisis de las unidades lingüísticas, aspecto que ha estado presente en los dos estadios.

## 2. EL CORPUS PRESEEA-SEVILLA

La recolección de materiales del habla de la ciudad hispalense, siguiendo los criterios metodológicos de PRESEEA, era una necesidad prioritaria para poder realizar los análisis. En el grupo de investigación *Sociolingüística andaluza: estudio sociolingüístico del habla de Sevilla* contábamos con unas muestras orales, grabadas y transcritas, que se habían recogido siguiendo parámetros sociolingüísticos de sexo, edad y nivel

---

<sup>3</sup> Proyecto de Excelencia financiado por el Ministerio de Economía y Competitividad del Gobierno de España y por el Fondo Europeo de Desarrollo Regional.

<sup>4</sup> Para las aportaciones de esta primera etapa puede consultarse Santana Marrero (2020a).

de estudios (Lamíquiz Ibáñez *et al.*, 1983; Lamíquiz Ibáñez *et al.*, 1987; Pineda Pérez *et al.*, 1992). No obstante, dado que eran materiales de los años setenta y ochenta de la centuria precedente, y dado que, además, se habían recopilado siguiendo unas normas específicas que se alejaban de las habituales en PRESEEA<sup>5</sup>, resultaba evidente que teníamos que emprender la tarea de recolectar el nuevo corpus del habla de la capital andaluza. Para ello no empezamos de cero. En primer lugar, el equipo de trabajo que coordinaba el profesor Ariza había conseguido hacer varias grabaciones, de las cuales pudimos aprovechar algunas. En todos los casos, eran audios que necesitaban ser transcritos y etiquetados. En segundo lugar, en el año 2011 nos habíamos comprometido con el *Proyecto de la norma culta hispánica Juan M. Lope Blanch*, con el que veníamos trabajando desde la segunda mitad del siglo XX, a actualizar algunas entrevistas del sociolecto alto. Por tanto, pudimos aprovechar esos materiales, acomodándolos a los requisitos de transcripción de PRESEEA<sup>6</sup>. A pesar de lo dicho, la tarea que quedaba por hacer era ingente, porque partíamos de unas pocas entrevistas que, en el mejor de los casos, estaban transcritas, pero de ningún modo reunían los requisitos formales de este macroproyecto, por lo que había que adaptarlas a los nuevos criterios.

Atendiendo a lo dicho, entre los años 2016 y 2019 tuvimos que compaginar las labores de recolección y edición del corpus con la del análisis. Aunque, a priori, resulta contradictorio realizar los estudios sin contar con las muestras de habla, pudimos salvar la situación con distintas estrategias. A este respecto, parte de las investigaciones que vieron la luz durante esos primeros años se basaban, por necesidad, en uno o dos sociolectos, de forma paralela al ritmo que íbamos recogiendo el corpus. Además, durante este primer periodo trabajamos a partir de documentos internos, todavía no preparados para la publicación, pero sí suficientemente válidos para desarrollar los estudios que teníamos comprometidos. Finalmente, tras la edición y publicación del corpus PRESEEA-Sevilla (Repede, 2019a, 2019b, 2019c), que está compuesto por 72 entrevistas semidirigidas, durante la segunda fase del proyecto hemos podido dedicarnos plenamente a la tarea de la descripción sociolingüística del habla de los sevillanos.

### 3. EL ANÁLISIS DEL HABLA DE SEVILLA

#### 3.1. Nivel fonético-fonológico

Como no podía ser de otra manera, tratándose de la variedad andaluza, les hemos prestado atención a diversos aspectos del nivel fonético-fonológico. Concretamente,

---

<sup>5</sup> Por ejemplo, la transcripción seguía las pautas ortográficas, sin etiquetado, y los parámetros para establecer las casillas sociolingüísticas eran distintos a los que se utilizan en PRESEEA.

<sup>6</sup> Contar con los materiales del habla culta de Sevilla de mediados del siglo pasado nos ha permitido realizar análisis en tiempo real, al contrastarlos con las entrevistas recientes que conforman el corpus PRESEEA-Sevilla (Santana Marrero, 2019a, 2021a, 2023b). En esta misma línea, hemos realizado también algún trabajo en el sociolecto bajo (León-Castro Gómez *et al.*, 2022).

nos hemos ocupado tanto de segmentos fónicos en el ataque silábico, como es el caso de la *-d/-* intervocálica (Jiménez Fernández, 2015, 2019, 2020a, 2020b; Jiménez Fernández *et al.*, 2020), el fonema dental /θ<sup>s</sup>/- (Santana Marrero, 2016a, 2016-2017, 2017, 2020b, 2022a) y el fonema prepalatal /tʃ/- (Santana Marrero, en prensa a), como de segmentos situados en la coda silábica, donde ubicamos la *-s/* (Jiménez Fernández *et al.*, 2023) y las líquidas *-r/* y *-l/* (Jiménez Fernández, en prensa). En esta ocasión nos centraremos solo en los dos primeros fenómenos mencionados, pues se trata de análisis ya finalizados en todos los sociolectos.

El estudio de la *-d/-*, en el que se consideraron las variantes elisión, aspiración y mantenimiento, puso de manifiesto altas cotas de omisión. Por sectores sociales, se trata de una solución registrada fundamentalmente en los hablantes jóvenes y en los del sociolecto bajo. Estamos, por tanto, ante un patrón no prestigioso que tiene visos de continuidad en esta comunidad de habla, ya que se mantiene con fuerza entre los informantes de menos edad. El entorno fónico *-ado* favoreció especialmente la variante innovadora Ø.

Por su parte, la materialización del fonema /θ<sup>s</sup>/ reveló una pauta similar a las otras ciudades andaluzas analizadas hasta el momento (Granada, Huelva y Málaga). A diferencia de lo que se pensaba inicialmente, Sevilla se suma a la mimetización de la solución convergente con la variedad centro-norteña, escisión /θ/-/s/, que en un principio era más característica de las zonas orientales. De hecho, la realización dento-interdental [θ] de los segmentos que en la escritura corresponden con <c+e, i> y <z+vocal> aventajó ligeramente, de forma global, a la variante local seseante. Nuestros datos arrojaron, además, resultados de interés en lo que respecta a la distribución sociolingüística. Concretamente, el alófono conservador [θ] fue liderado por las mujeres, por los entrevistados más jóvenes y, principalmente, por los del sociolecto alto. Estamos, por tanto, ante un patrón de prestigio que se acerca a las realizaciones que históricamente se han asociado a las formas estandarizadas, propias del español septentrional. Por otro lado, las soluciones vernáculas seseante y ceceante, que también tuvieron cabida en el habla de los sevillanos, presentaron un panorama sociolingüístico diferente. Específicamente, la variante seseante tuvo mayor presencia entre los hombres y los hablantes de los sociolectos bajo y medio, mientras que la ceceante registró un número menor de casos, anclados principalmente en los entrevistados de más edad y en el sociolecto bajo. No deja de ser llamativo, además, que la mayor parte de los informantes alternaron las soluciones fonéticas dento-interdental y seseante en sus intervenciones, por lo que fue la pauta más extendida en toda la muestra. Estamos ante un importante proceso de cambio que está experimentándose en la pronunciación de los hablantes de la ciudad de Sevilla.

### 3.2. Niveles morfosintáctico y léxico-semántico

En el plano morfosintáctico destacan las investigaciones realizadas sobre las formas pronominales átonas de tercera persona (Repede, 2017a, 2020a) y sobre el

pronombre personal sujeto (Repede, 2015, 2019d)<sup>7</sup>. El análisis de los pronombres personales átonos de tercera persona indicó que en Sevilla prevalece el sistema etimológico, aunque con un tímido avance del leísmo. No obstante, este último solo se documentó con referente personal. Estamos ante un cambio que se está iniciando en la ciudad hispalense y que tiene connotaciones prestigiosas, en tanto que fue impulsado por las mujeres, por los jóvenes y por el nivel educativo alto. Por otro lado, el análisis de la presencia / ausencia del pronombre personal sujeto solo se ha llevado a cabo en informantes con estudios universitarios, donde se registraron unos índices de mantenimiento no demasiado elevados. Esta variante parece estar condicionada por criterios como la ambigüedad de la forma verbal o la persona gramatical, entre otros (especificidad o la correferencialidad). En cuanto a la influencia de los factores sociales, fueron las mujeres y la segunda generación quienes más utilizaron los sujetos explícitos.

En lo que respecta a la parcela léxico-semántica, la naturaleza de los materiales de PRESEEA dificulta bastante su estudio, por lo que apenas ha habido aportaciones en este plano. A este respecto, el análisis de las combinaciones léxicas (Castillo Carballo, 2020) permitió corroborar la premisa de que, en las actualizaciones de las diferentes unidades fraseológicas, fundamentalmente colocaciones, locuciones y, en menor medida, fórmulas rutinarias, la polaridad informalidad-formalidad se suele resolver a favor de esta última en el contexto de las entrevistas semidirigidas. A su vez, se localizaron, especialmente en el sociolecto alto, expresiones propias de registros formales, incluidas las tecnolectales.

### 3.3. Nivel pragmático-discursivo

En el equipo de PRESEEA-Sevilla se han hecho abundantes contribuciones a esta parcela del análisis lingüístico, que en los últimos años se ha instalado plenamente en los estudios de variación espacial y social en español. Aquí se engloban nuestras investigaciones sobre la atenuación, sobre los marcadores del discurso y sobre las formas de tratamiento.

Los trabajos sobre la atenuación que hemos llevado a cabo en la capital hispalense se han centrado, por un lado, en mecanismos específicos, como es el caso de los recursos de impersonalización (León-Castro Gómez, 2015, 2020a; León-Castro Gómez *et al.*, 2018, 2020; Repede, 2017b; Repede *et al.*, 2019), las construcciones *como que* y *un poco* (Repede, 2020b, 2023a, en prensa a), el discurso referido

---

<sup>7</sup> Además, hay una investigación en marcha sobre las perífrasis verbales de infinitivo que verá la luz en forma de tesis doctoral. En ella se incluyen aspectos de especial interés, como la representatividad de *ir a* + infinitivo para expresar futuridad o la distribución entre *deber* y *deber de* + infinitivo. En Sobral Olivera (2023a, 2023b, en prensa) se exponen algunos resultados parciales.

(Repede, 2018, 2019e, 2019f, 2021) y el diminutivo (Castillo Carballo *et al.*, 2018; León-Castro Gómez, 2020b, en prensa; León Castro-Gómez *et al.*, 2022); y, por otro lado, en el análisis de forma conjunta de todas las estrategias localizadas en el habla de Sevilla (Repede, 2022b, 2023b, 2023e, 2024).

En cuanto a los mecanismos de impersonalización, el estudio de *uno*, *tú* y *se* ha puesto de manifiesto su uso para diferentes funciones, que van desde la generalización (inclusiva o exclusiva) hasta el encubrimiento pragmático del hablante, cuya finalidad consiste en proteger su imagen, buscar acuerdo social o criticar. Si *uno* fue más frecuente en la tercera generación y los hombres, *tú* y *se* aparecieron en hombres y mujeres de la segunda generación. También dentro de las estrategias particulares de mitigación lingüística se sitúan los estudios sobre *como que* y *un poco*. Se ha podido observar que ambas construcciones se utilizaron ampliamente en aquellos contextos interactivos donde los hablantes buscan principalmente salvaguardar y proteger la propia imagen. Por factores sociales, *como que* fue muy común en el discurso de las mujeres, de los hablantes de la primera generación y del nivel medio, mientras que *un poco* prevaleció en los informantes más jóvenes y en los que tienen más formación académica. Por su parte, el discurso referido como mecanismo atenuador se convierte en una estrategia de desfocalización del centro deíctico mediante tres mecanismos: enajenación, impersonalización y despersonalización. Por factores sociales, las mujeres sevillanas fueron las que más emplearon la enajenación. En cambio, los hombres mitigaron a través del discurso reproducido con valor impersonalizador y despersonalizador. Asimismo, los informantes jóvenes coinciden con los adultos en el empleo de este recurso con valor impersonalizador y con los mayores en su uso con función de enajenación. Finalmente, el discurso referido impersonalizador es más frecuente en el sociolecto bajo, mientras que la despersonalización y la enajenación se dan más en los hablantes de instrucción media. En cuanto al diminutivo, los materiales de Sevilla nos permitieron comprobar el predominio de la forma *-ito* y su utilización mayoritaria con función atenuativa. Este tipo de sufijación fue más recurrente entre las mujeres, los jóvenes y los hablantes de un nivel de instrucción bajo.

Para finalizar la atenuación, el análisis conjunto de todos los mecanismos que se dieron en PRESEEA-Sevilla indican que la estrategia más empleada fue la impersonalización del sujeto, seguida muy de cerca por las fórmulas fático-apelativas. Por criterios sociales, las principales diferencias localizadas en los usos según el sexo se dieron en el empleo de marcadores del discurso correctores y el empleo de fórmulas de control de contacto, ya que el primero fue un recurso más utilizado por las mujeres y el segundo por los hombres. Además, la primera y tercera generación priorizaron la impersonalización del sujeto, mientras que la generación adulta atenúa, sobre todo, para implicar al interlocutor en lo dicho por el hablante, esto es, mediante las fórmulas fáticas.

En lo que respecta a los marcadores del discurso, por un lado, les hemos prestado atención a dos variables priorizadas en PRESEEA, los apéndices de control de contacto (Santana Marrero, 2019a, 2019b) y los reformuladores (Santana Marrero,

2022b, 2023a, 2023b, 2023c, 2023d). Por otro lado, hemos llevado a cabo el estudio de los ordenadores del discurso (Santana Marrero, 2024) y de los introductores de cita (Repede, 2022a, 2023f), dado el interés que revisten estos últimos fenómenos en muestras de oralidad no planificada como las que manejamos.

En cuanto a los marcadores de reformulación, los recapitulativos constituyeron el grupo más numeroso entre los hablantes sevillanos, seguidos por los explicativos. A su vez, los reformuladores *más empleados para estas dos funciones fueron: bueno y en fin* recapitulativos y *o sea y es decir* explicativos. Para la rectificación, el tercer tipo de reformulación más empleado, destacó *bueno* rectificativo. De forma global, los hombres y los informantes del sociolecto alto fueron los colectivos que más incluyeron la reformulación en sus intervenciones. Asimismo, se apreciaron interesantes procesos de cambio que afectan a *bueno* y *en fin* recapitulativos, el primero más recurrente en los jóvenes y el segundo en los informantes de más edad. En cuanto a los ordenadores del discurso, hemos podido comprobar la importante presencia que tienen estas fórmulas en la oralidad no planificada, adquiriendo una singularidad distinta al texto escrito. En este sentido, la mayoría de las series enunciativas no llevaron marca explícita para indicar las diferentes partes, sino que fueron más recurrentes las fórmulas de continuidad (lideradas por *después* y *luego*), seguidas de las de inicio (*primero* fue el ordenador más común) y, a mucha distancia, las formas de cierre, con pocas variantes (*por lo demás* y *al final*). Al igual que sucedía en los reformuladores, los varones y los hablantes con más formación académica mostraron más tendencia a incluir estas unidades en sus parlamentos. En tercer lugar, los sevillanos entrevistados emplearon nueve marcadores del discurso diferentes para introducir cita directa: *en plan (de)*, *como (que)*, *porque*, *pues*, *entonces*, *ya*, *bueno*, *pero* y *eh*. Se advirtió cómo unos estaban más especializados en referir palabras (*en plan*), otros en la reproducción de pensamientos (*como que*) y otros en el discurso estereotipado (*pues*). Los grupos sociales más activos en introducir estas estrategias en sus intervenciones fueron las mujeres, la primera generación y los hablantes del sociolecto alto.

Además de lo expuesto, hemos analizado marcadores del discurso específicos, que presentaron especial relevancia en nuestro corpus: *en plan* (Repede, 2020c), *luego / después* (Santana Marrero, 2016b, 2021a, 2021b) y *vamos a ver* (Sobral Olivera, 2023c).

Finalmente, los resultados del análisis de las formas de tratamiento (Zerva, 2016, 2020) indican que predominaron las fórmulas de familiaridad, independientemente del contexto situacional, y sin que se apreciaran diferencias significativas según el nivel educativo de los informantes.

#### 4. OTRAS INICIATIVAS PARALELAS A PRESEEA

Junto a lo dicho, no sería justo terminar esta exposición sin aclarar que actualmente tenemos en marcha dos proyectos que están en estrecha convivencia con PRESEEA-

Sevilla. Uno de ellos aborda el estudio de los procesos de interferencia lingüística que se dan en las comunidades de inmigrantes que residen en la capital hispalense. Hasta el momento, en Sevilla hemos ido avanzando en dos sentidos. En primer lugar, hemos iniciado la recolección de muestras de habla de inmigrantes hispanohablantes (en estrecha conexión con el *Corpus dinámico del español de la inmigración*, CORDIESIN): por ahora están finalizadas las de peruanos (Repede, 2022c), colombianos y nicaragüenses (<https://corpasmigrasev.snlt.es/>) y están iniciadas las de ecuatorianos. En segundo lugar, hemos comenzado las tareas de análisis de algunos fenómenos lingüísticos de interés como los marcadores introductores de discurso indirecto, en ocasiones en contraste con los resultados de PRESEEA-Sevilla (Repede, 2023c, 2023d, en prensa b).

La otra iniciativa en la que participamos es el *Proyecto para el estudio de las creencias y actitudes hacia las variedades del español en el siglo XXI* (PRECAVES XXI), coordinado por Ana María Cestero Mancera y Florentino Paredes García, catedráticos de la Universidad de Alcalá. Siguiendo sus directrices, y tomando como punto de partida una metodología común, en Sevilla hemos llevado a cabo varios análisis en las dos etapas del proyecto, una primera basada en el sondeo de jóvenes universitarios (Santana Marrero, 2018a, 2018b, 2020c) y una segunda que parte de una muestra de hablantes distribuidos según los criterios de estratificación de PRESEEA (Santana Marrero, 2022c, 2022d, 2023e, 2023f, en prensa b). La percepción que tienen los sevillanos de su variedad propia o de la centro-norteña nos ayuda a comprender algunos de los procesos de convergencia fonética hacia esta última que, según hemos constatado a partir del análisis de los materiales de PRESEEA-Sevilla, se están produciendo en la ciudad hispalense.

## 5. CONCLUSIÓN

Tras esta breve exposición no podemos deshacernos de la sensación positiva de haber avanzado en el análisis del habla de la ciudad de Sevilla desde la óptica de la variación social. No obstante, siguen quedando múltiples fenómenos por estudiar, lo que nos anima a continuar trabajando en nuevas metas investigadoras. En definitiva, en PRESEEA-Sevilla afrontamos con entusiasmo los retos que están por venir, siempre con el objetivo final de conocer el mosaico sociolingüístico que define el habla de la capital hispalense.

## REFERENCIAS

- Castillo Carballo, María Auxiliadora (2020). «Construcciones sintagmáticas con restricción en entrevistas semidirigidas». En D. Repede y M. León-Castro Gómez (eds.), *Patrones sociolingüísticos del español hablado en la ciudad de Sevilla*. Berna: Peter Lang, pp. 33-50.

- Castillo Carballo, María Auxiliadora y Juan Manuel García Platero (2018). «La apreciación afijal en contextos de sociolecto bajo en el habla de Sevilla». *Philologica Canariensis*, 24: 77-88. Disponible en: <https://ojsspdic.ulpgc.es/ojs/index.php/PhilCan/article/view/885>
- Jiménez Fernández, Rafael (2015). «Sobre el mantenimiento y la pérdida de la -d- intervocálica en el habla de Sevilla (sociolecto bajo)». En J. Santana Marrero (dir.), M. León-Castro Gómez y A. Zerva (eds.), *La variación en el español actual. Estudios dedicados al profesor Pedro Carbonero. Sociolingüística andaluza*, 17. Sevilla: Editorial Universidad de Sevilla, pp. 193-210.
- Jiménez Fernández, Rafael (2019). «Elisión de la /d/ intervocálica en hablantes de nivel sociocultural bajo de Sevilla». *Lengua y Habla*, 23: 258-285. Disponible en: <http://erevistas.saber.ula.ve/index.php/lenguayhabla/article/view/15672>
- Jiménez Fernández, Rafael (2020a). «Estudio comparativo de la variación de la dental -d/- en la ciudad de Sevilla». En D. Repede y M. León-Castro Gómez (eds.), *Patrones sociolingüísticos del español hablado en la ciudad de Sevilla*. Berna: Peter Lang, pp. 65-92.
- Jiménez Fernández, Rafael (2020b). «Variación fonológica de la /d/ intervocálica en el sociolecto alto de Sevilla, RILCE». *Revista de Filología Hispánica*, 36(2): 674-707. DOI: 10.15581/008.36.2.674-707
- Jiménez Fernández, Rafael (En prensa). «Variación sociofonética de las líquidas /r/ y /l/ en posición postnuclear en el español hablado de Sevilla (sociolecto alto)». en M. León-Castro Gómez y R. Jiménez Fernández (coords.), *Procesos de variación y cambio en el español de Sevilla y Las Palmas de Gran Canaria*. Madrid/Frankfurt: Iberoamericana/Vervuert.
- Jiménez Fernández, Rafael y Marta León-Castro Gómez (2020). «Elisión de la /d/ intervocálica en el sociolecto medio de Sevilla». *Lingüística Española Actual*, 42(2): 251-276.
- Jiménez Fernández, Rafael y Marta León-Castro Gómez (2023). «Variación en la pronunciación de la /s/ implosiva: datos del nivel sociocultural alto en el corpus PRESEEA-Sevilla». *Boletín de Filología de la Universidad de Chile*, 58(2): 371-398. Disponible en: <https://revistateoria.uchile.cl/index.php/BDF/article/view/73297>
- Lamíquiz Ibáñez, Vidal y Miguel Ángel Pineda Pérez (1983). *Encuestas del habla urbana de Sevilla. Nivel culto. Sociolingüística andaluza*, 2. Sevilla: Publicaciones de la Universidad de Sevilla.
- Lamíquiz Ibáñez, Vidal y Miguel Ropero Núñez (1987). *Encuestas del habla urbana de Sevilla. Nivel popular. Sociolingüística andaluza*, 4. Sevilla: Publicaciones de la Universidad de Sevilla.
- León-Castro Gómez, Marta (2015). «Variación en la segunda persona del singular como estrategia de indeterminación referencial en el habla de Sevilla: sociolecto bajo». En J. Santana Marrero (dir.), M. León-Castro Gómez y A. Zerva (eds.), *La variación en el español actual. Estudios dedicados al profesor Pedro Carbonero. Sociolingüística andaluza*, 17. Sevilla: Editorial Universidad de Sevilla, pp. 211-232.

- León-Castro Gómez, Marta (2020a). «Uno como mecanismo de impersonalización en PRESEEA: un estudio contrastivo entre Granada y Sevilla». *Tonos Digital*, 38(1). Disponible en: <https://digitum.um.es/digitum/handle/10201/86348>
- León-Castro Gómez, Marta (2020b). «Variación en el empleo del diminutivo en el habla urbana de Sevilla: estudio sociolingüístico sobre el corpus PRESEEA». En D. Repede y M. León-Castro Gómez (eds.), *Patrones sociolingüísticos del español hablado en la ciudad de Sevilla*. Berna: Peter Lang, pp. 93-116.
- León-Castro Gómez, Marta (En prensa). «Los sufijos *-ito/-illo* en el habla de Sevilla. Análisis en tiempo real». En M. León-Castro Gómez y R. Jiménez Fernández (coords.), *Procesos de variación y cambio en el español de Sevilla y Las Palmas de Gran Canaria*. Madrid/Frankfurt: Iberoamericana/Vervuert.
- León-Castro Gómez, Marta y Doina Repede (2018). «El pronombre indefinido *uno* como estrategia de impersonalización: un estudio en el corpus oral PRESEEA-Sevilla». *Nueva Revista del Pacífico*, 69: 67-89. Disponible en: <http://www.nuevarevistadelpacifico.cl/index.php/NRP/article/view/130>
- León-Castro Gómez, Marta y Doina Repede (2020). «Aproximación sociolingüística de los impersonalizadores *tú* y *se* en el corpus oral PRESEEA-Sevilla». En D. Repede y M. León-Castro Gómez (eds.), *Patrones sociolingüísticos del español hablado en la ciudad de Sevilla*. Berna: Peter Lang, pp. 117-140.
- León-Castro Gómez, Marta y Rafael Jiménez Fernández (2022). «La alternancia *-ito/-illo* en hablantes sevillanos de nivel educacional bajo: un estudio en tiempo real». *Literatura y Lingüística*, 45: 543-569. Disponible en: [https://www.scielo.cl/scielo.php?pid=S0716-58112022000100543&script=sci\\_arttext](https://www.scielo.cl/scielo.php?pid=S0716-58112022000100543&script=sci_arttext)
- Pineda Pérez, Miguel Ángel y Manuel Ollero Toribio (1992). *Encuestas del habla urbana de Sevilla. Nivel medio. Sociolingüística andaluza*, 6. Sevilla: Publicaciones de la Universidad de Sevilla.
- Repede, Doina (2015). «El sujeto nominal pospuesto: ¿variante o construcción independiente?». En J. Santana Marrero (dir.), M. León-Castro Gómez y A. Zerva (eds.), *La variación en el español actual. Estudios dedicados al profesor Pedro Carbonero. Sociolingüística andaluza*, 17. Sevilla: Editorial de la Universidad de Sevilla, pp. 233-250.
- Repede, Doina (2017a). «El uso de las formas pronominales átonas en el corpus PRESEEA-Sevilla». *Borealis: An International Journal of Spanish Linguistics*, 6(1): 200-221. Disponible en: <https://septentrio.uit.no/index.php/borealis/article/viewFile/4088/3850>
- Repede, Doina (2017b). «Análisis del discurso reproducido en el corpus PRESEEA-Sevilla: el sociolecto bajo». En J. M. Santos Rovira (ed.), *Variación lingüística e identidad en el mundo hispanohablante*. Lugo: Axac, pp. 55-74.
- Repede, Doina (2018). «El uso del discurso referido en el corpus PRESEEA-Sevilla: el sociolecto alto». *Anuario de Letras*, 6(1): 187-218. Disponible en: <https://revistas-filologicas.unam.mx/anuario-letras/index.php/al/article/view/1481/0>

- Repede, Doina (2019a). *El español hablado en Sevilla. Corpus PRESEEA-Sevilla. Vol. I. Hablantes de instrucción alta*. Sevilla: Editorial Universidad de Sevilla.
- Repede, Doina (2019b). *El español hablado en Sevilla. Corpus PRESEEA-Sevilla. Vol. II. Hablantes de instrucción baja*. Sevilla: Editorial Universidad de Sevilla.
- Repede, Doina (2019c). *El español hablado en Sevilla. Corpus PRESEEA-Sevilla. Vol. III. Hablantes de instrucción media*. Sevilla: Editorial Universidad de Sevilla.
- Repede, Doina (2019d). «Condicionantes sociolingüísticos de la expresión del sujeto pronominal en el sociolecto alto del corpus PRESEEA-Sevilla». *Revista de Investigación Lingüística*, 22: 397-423. DOI: <https://doi.org/10.6018/rii.390031>
- Repede, Doina (2019e). «Estudio sociolingüístico del discurso reproducido en el corpus oral PRESEEA-Sevilla». *Signo y Seña*, 35: 194-214. DOI: <https://doi.org/10.34096/sys.n35.6944>
- Repede, Doina (2019f). «Discurso directo y discurso indirecto: estudio sociolingüístico en el corpus *Encuestas del habla urbana de Sevilla*». *Lingüística y Literatura*, 40: 42-64. DOI: <https://doi.org/10.17533/udea.lyl.n76a02>
- Repede, Doina (2020a.) «Uso y funciones del pronombre átono *le(s)*: estudio en el corpus oral PRESEEA de Sevilla». En D. Repede y M. León-Castro Gómez (eds.), *Patrones sociolingüísticos del español hablado en la ciudad de Sevilla*. Berna: Peter Lang, pp. 141-168.
- Repede, Doina (2020b). «El aproximador *como que* con valor atenuador en el corpus oral PRESEEA-Sevilla». *Itinerarios. Revista de Estudios Lingüísticos, Literarios, Históricos y Antropológicos*, 32: 245-261. Disponible en: <https://itinerarios.uw.edu.pl/resources/html/article/details?id=222531>
- Repede, Doina (2020c). «La construcción *en plan* en el español hablado de Sevilla: uso, función y distribución social». *Tonos Digital*, 38(1). Disponible en: <https://digitum.um.es/digitum/handle/10201/86340>
- Repede, Doina (2021). «El discurso referido como recurso atenuador en la ciudad de Sevilla». *Cultura, Lenguaje y Representación*, 25: 285-303. DOI: <https://doi.org/10.6035/CLR.2021.25.16>
- Repede, Doina (2022a). «Marcadores del discurso como introductores de cita directa: aproximación sociopragmática en el corpus PRESEEA-Sevilla». *Anuario de Estudios Filológicos*, XLV: 285-310. DOI: <https://doi.org/10.17398/2660-7301.45.285>
- Repede, Doina (2022b). «Estudio sociopragmático de los mecanismos atenuadores en el corpus oral PRESEEA-Sevilla». *Círculo de Lingüística Aplicada a la Comunicación*, 92: 153-167. DOI: <https://doi.org/10.5209/clac.71776>
- Repede, Doina (ed.) (2022c). *Corpus oral para el estudio sociolingüístico de hablantes peruanos en la ciudad de Sevilla (PERUSEV)*. Sevilla: Editorial Universidad de Sevilla.
- Repede, Doina (2023a). «La locución *un poco* en el corpus oral PRESEEA-Sevilla: funciones discursivas y distribución social». *Forma y Función*, 36(1): 77-100. DOI: <https://doi.org/10.15446/fyf.v36n1.97379>

- Repede, Doina (2023b). «La atenuación entre hablantes de instrucción alta en el corpus PRESEEA-Sevilla: usos contrastivos en sexo y edad». *Estudios de Lingüística. Universidad de Alicante*, 39: 5-19. DOI: <https://doi.org/10.14198/ELUA.21199>
- Repede, Doina (2023c). «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. DOI: <https://doi.org/10.25115/oralia.v26i1.8269>.
- Repede, Doina (2023d). «Estudio piloto sobre los marcadores del discurso como introductores de cita directa en el corpus oral PERUSEV». *Lingüística*, 39(2): 11-32. DOI: <https://doi.org/10.5935/2079-312x.20230015>
- Repede, Doina (2023e). «La atenuación lingüística en la ciudad de Sevilla: el caso del corpus oral PRESEEA». En Z. Camus (ed.), *Lenguaje y sociedad. Acercamientos, teorías, problemas y perspectivas*. Sevilla: Editorial Universidad de Sevilla, pp. 129-147.
- Repede, Doina (2023f). «Los marcadores citativos en español: distribución sociolingüística y geolectal». *Estudios y aplicaciones de PRESEEA. Sociopragmática y enseñanza de español. Normas. Revista de Estudios Lingüísticos Hispánicos*, 13: 281-295. DOI: <https://doi.org/10.7203/Normas.v13i1.27730>
- Repede, Doina (2024). «Análisis sociolingüístico de los recursos atenuadores en hablantes de Sevilla con nivel de instrucción alta». En G. F. Sánchez, D. Castilleja, J. Jiménez-Salcedo, C. Huerdo-Moreno y M. Bonet Bofill (eds.), *Variedades del español: aproximaciones desde la sociolingüística, la pragmática, la traducción y la interpretación*. Leiden: Editorial Brill, pp. 62-81.
- Repede, Doina (En prensa a). «Análisis sociopragmático de «un poco» en las entrevistas semidirigidas». *Onomázein*, 63.
- Repede, Doina (En prensa b). «Estrategias introductorias y tipo de reproducción de las citas directas en el corpus oral PERUSEV». *Revista Española de Lingüística Aplicada*.
- Repede, Doina y Marta León-Castro Gómez (2019). «Procedimientos de impersonalización en la ciudad de Sevilla: el caso del pronombre uno». *Cuadernos de Investigación Filológica*, 46: 77-102. DOI: <https://doi.org/10.18172/cif.3890>
- Santana Marrero, Juana (2016a). «Seseo, ceceo y distinción en el sociolecto alto de la ciudad de Sevilla: nuevos datos a partir de los materiales de PRESEEA». *Estudios sobre la lengua española hablada en el mundo hispánico en su variedad geográfica y social con materiales del PRESEEA. Boletín de Filología de la Universidad de Chile*, 51(2): 255-280. DOI: <http://dx.doi.org/10.4067/S0718-93032016000200010>
- Santana Marrero, Juana (2016b). «Luego y después: uso discursivo y variación». *RILCE. Revista de Filología Hispánica*, 32(2): 513-535. DOI: [10.15581/008.32.2.513-35](https://doi.org/10.15581/008.32.2.513-35)

- Santana Marrero, Juana (2016-2017). «Factores externos e internos influyentes en la variación /θ<sup>s</sup>/ en la ciudad de Sevilla». *Analecta Malacitana*, XXXIX: 143-177. Disponible en: <http://www.revistas.uma.es/index.php/analecta/article/view/5613/5262>
- Santana Marrero, Juana (2017). «Variación de las realizaciones de /θ<sup>s</sup>/ en el sociolecto bajo de la ciudad de Sevilla: datos de PRESEEA-SE». *Procesos de variación y cambio en el español de España. Estudios sobre corpus PRESEEA. Lingüística en la Red*, XV: 1-17. Disponible en: [http://www.linred.es/monograficos\\_pdf/LR-monografico15-articulo2.pdf](http://www.linred.es/monograficos_pdf/LR-monografico15-articulo2.pdf)
- Santana Marrero, Juana (2018a). «Creencias y actitudes de los jóvenes universitarios sevillanos hacia las variedades cultas del español». *Percepción de las variedades cultas del español: creencias y actitudes de jóvenes universitarios hispanohablantes. Boletín de Filología de la Universidad de Chile*, 53(2): 115-144. DOI: <http://dx.doi.org/10.4067/S0718-93032018000200115>
- Santana Marrero, Juana (2018b). «Creencias y actitudes de jóvenes universitarios sevillanos hacia las variedades normativas del español de España: andaluza, canaria y castellana». *Pragmática Sociocultural / Sociocultural Pragmatics*, 6(1): 71-97. DOI: <http://doi.org/10.1515/soprag-2018-0003>
- Santana Marrero, Juana (2019a). «Apéndices interrogativos en el corpus PRESEEA-Sevilla: sociolectos alto y bajo». *Oralia. Análisis del Discurso Oral*, 22(1): 111-127. Disponible en: <https://ojs.ual.es/ojs/index.php/ORALIA/article/view/6704/5431>
- Santana Marrero, Juana (2019b). «Apéndices interrogativos en el habla de Sevilla: un estudio en tiempo real». *Itinerarios. Revista de Estudios Lingüísticos, Literarios, Históricos y Antropológicos*, 29: 297-318. DOI: 10.7311/ITINERARIOS.29.2019.13
- Santana Marrero, Juana (2020a). «Patrones sociolingüísticos en la ciudad de Sevilla: estudios iniciales y situación actual». En D. Repede y M. León-Castro Gómez (eds.), *Patrones sociolingüísticos del español hablado en la ciudad de Sevilla*. Berna: Peter Lang, pp. 11-32.
- Santana Marrero, Juana (2020b). «Patrones sociolingüísticos de pronunciación en la ciudad de Sevilla: variación sociofonética de *seseo* y *ceceo*». En D. Repede y M. León-Castro Gómez (eds.), *Patrones sociolingüísticos del español hablado en la ciudad de Sevilla*. Berna: Peter Lang, pp. 169-208.
- Santana Marrero, Juana (2020c). «Percepción de las variedades andaluza y castellana de los jóvenes sevillanos: un análisis contrastivo». *Onomázein*, 50: 71-89. DOI: 10.7764/onomazein.50.05
- Santana Marrero, Juana (2021a). «Luego y después como marcadores del discurso en la norma culta de Sevilla: un estudio en tiempo real. *Marcadores discursivos luego, después y donde en la norma culta hispánica*». *Cuadernos de la ALFAL*, 13(1): 166-192. ISSN: 2218-0761. Disponible en: [https://www.mundoalfal.org/sites/default/files/revista/13\\_1\\_cuaderno\\_009.pdf](https://www.mundoalfal.org/sites/default/files/revista/13_1_cuaderno_009.pdf)

- Santana Marrero, Juana (2021b). «*Luego y después* marcadores del discurso: estudio en el corpus PRESEEA-Sevilla». Comunicación presentada en el *XIV Congreso Internacional de Lingüística General*, Universidad de Sevilla, 23 al 25 de junio 2021. Sevilla, España.
- Santana Marrero, Juana (2022a). «Convergence patterns in the city of Seville: sociolinguistic study of *seseo*». *Spanish in Context*, 19(1): 122-145. DOI: <https://doi.org/10.1075/sic.20003.san>
- Santana Marrero, Juana (2022b). «Los marcadores de reformulación en el corpus PRESEEA de Sevilla». *Estudios de variación lingüística: homenaje a Juan Andrés Villena Ponsoda. Lingüística en la Red*, XIX: 1-22. Disponible en: <https://revistas.publicaciones.uah.es/ojs/index.php/linred/article/view/1876/1067>
- Santana Marrero, Juana (2022c). «How do Andalusian journalism students perceive Andalusian and Castilian linguistic varieties?». *Revista Española de Lingüística Aplicada*, 35(2): 565-595. DOI: <https://doi.org/10.1075/resla>
- Santana Marrero, Juana (2022d). «Creencias y actitudes de los sevillanos hacia la variedad castellana». En A. M.<sup>a</sup> Cestero Mancera y F. Paredes García (eds.), *Percepción de la variedad castellana: creencias y actitudes en el siglo XXI*. Alcalá de Henares: Editorial Universidad de Alcalá, pp. 237-282. DOI: <https://doi.org/10.37536/AHJD6439>
- Santana Marrero, Juana (2023a). «Análisis sociopragmático de los reformuladores explicativos en el habla de Sevilla». *Oralia. Análisis del Discurso Oral*, 26(1): 147-177. Disponible en: <https://ojs.ual.es/ojs/index.php/ORALIA/article/view/7900/7786>
- Santana Marrero, Juana (2023b). «Los marcadores de reformulación en la norma culta de Sevilla: análisis desde una perspectiva longitudinal». *Oralidad Culta Hispanoamericana. Estudios de Filología, Marcadores Discursivos y Dialectalismos. Cuadernos de la ALFAL, volumen especial*, 2 (agosto de 2023): 89-112. DOI: <https://doi.org/10.5935/2218-0761.2023vspe05>
- Santana Marrero, Juana (2023c). «Análisis funcional y sociolingüístico de los principales marcadores de reformulación en el corpus PRESEEA de Sevilla». *Boletín de Filología de la Universidad de Chile*, 58(2): 479-512. Disponible en: <https://boletinfilologia.uchile.cl/index.php/BDF/article/view/73307/75187>
- Santana Marrero, Juana (2023d). «Distribución sociolingüística de los marcadores de reformulación recapitulativa en el habla de Sevilla». *Estudios y aplicaciones de PRESEEA. Sociopragmática y enseñanza de español. Normas. Revista de Estudios Lingüísticos e Hispánicos*, 13: 269-280. DOI: <https://doi.org/10.7203/Normas.v13i1.27729>
- Santana Marrero, Juana (2023e). «Creencias y actitudes de los sevillanos hacia la variedad lingüística chilena». *Lenguas Modernas*, 62: 155-182. Disponible en: <https://lenguasmodernas.uchile.cl/index.php/LM/article/view/73189/75056>
- Santana Marrero, Juana (2023f). «La variedad andaluza según los sevillanos: datos de PRECAVES XXI». En R. Cruz Ortiz, J. Santana Marrero e I. C. Santos

- Díaz (eds.), *El andaluz y su prestigio. Creencias y actitudes lingüísticas de los españoles hacia la variedad andaluza en el siglo XXI*. Granada: Comares, pp. 121-158.
- Santana Marrero, Juana (2024). «Los ordenadores del discurso: estudio en el corpus PRESEEA-Sevilla». *Revista Signos. Estudios de Lingüística*, 57(114): 129-150.
- Santana Marrero, Juana (En prensa a). «Estudio sociolingüístico de la pronunciación de /tʃ/ en la ciudad de Sevilla». En M. León-Castro Gómez y R. Jiménez Fernández (coords.), *Procesos de variación y cambio en el español de Sevilla y Las Palmas de Gran Canaria*. Madrid/Frankfurt: Iberoamericana/Vervuert.
- Santana Marrero, Juana (En prensa b). «Reconocimiento y valoración de las variedades cultas hispanoamericanas por parte de los sevillanos: datos del proyecto PRECAVES XXI». *Círculo de Lingüística Aplicada a la Comunicación*.
- Sobral Olivera, Daniel (2023a). «La alternancia *cantaré/voy a cantar* en el corpus PRESEEA Sevilla». Comunicación presentada en el *XV Congreso Internacional de Lingüística General*. Universidad Complutense de Madrid, 14 al 16 de junio 2023. Madrid, España.
- Sobral Olivera, Daniel (2023b). «*Debopensar lo que tengo que hacer*: comportamiento y extensión de las perífrasis verbales de infinitivo obligatorias por medio de *deber (de), tener que y haber que + infinitivo*». Comunicación presentada en el *Simposio Internacional PRESEEA. 25 Años de Sociolingüística Hispánica*. Universidad de Alcalá, 12 al 13 junio de 2023. Alcalá de Henares, España.
- Sobral Olivera, Daniel (2023c). «La transformación de *vamos a ver* en PRESEEA Sevilla: de perífrasis verbal a marcador del discurso». Comunicación presentada en las *XI Jornadas de Iniciación a la Investigación en Lengua y Literaturas Hispánicas*. Universidad Complutense de Madrid, 27 al 30 marzo de 2023. Madrid, España.
- Sobral Olivera, Daniel (En prensa). «La planificación del porvenir: alternancia entre el futuro morfológico y el futuro perifrástico en PRESEEA Sevilla». En M. León-Castro Gómez y R. Jiménez Fernández (coords.), *Procesos de variación y cambio en el español de Sevilla y Las Palmas de Gran Canaria*. Madrid/Frankfurt: Iberoamericana/Vervuert.
- Zerva, Adamantía (2016). «La cortesía y las formas de tratamiento en el español actual». En P. Pavlakis *et al.* (eds.), *Estudios y homenajes hispanoamericanos, vol. IV*. Madrid: Ediciones del Orto, pp. 421-430.
- Zerva, Adamantía (2020). «Conciencia sociolingüística de los hablantes sevillanos sobre el uso de *tú/usted*». En D. Repede y M. León-Castro Gómez (eds.), *Patrones sociolingüísticos del español hablado en la ciudad de Sevilla*. Berna: Peter Lang, pp. 209-226.



**CONTRIBUCIÓN DEL GRUPO PRESEEA-MÁLAGA A LA CONSTRUCCIÓN  
DEL MAPA SOCIOLINGÜÍSTICO DE VARIACIÓN Y CAMBIO EN ESPAÑA  
DESDE UN ENFOQUE COMPARATIVO Y MULTIDIMENSIONAL**

Juan Andrés Villena  
*Universidad de Málaga*

Antonio Manuel Ávila  
*Universidad de Málaga*

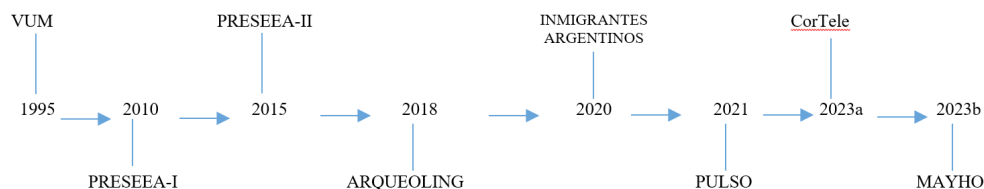
## **1. INTRODUCCIÓN**

Las aportaciones del grupo PRESEEA-Málaga al desarrollo del Proyecto PRE-SEEA en España son relevantes y pueden encontrarse en los diferentes trabajos publicados desde el inicio de la coordinación de PRESEEA en este país, aunque por motivos de espacio se señalan a continuación solo los referidos a los últimos tres años (Ávila, 2021a, b y c; 2022a, b y c; en prensa; Ávila *et al.*, 2020, 2021, 2023; Ávila y Rodríguez Cruces, 2020; Ávila y Segura, 2022; Sánchez y Ávila, 2020; Díaz, 2022; Díaz y Villena, en prensa; Molina, 2020, 2021, 2022, en prensa; Santos y Ávila, 2021, 2023; Vida, 2020, 2021, 2022; Vida *et al.*, 2023; Villena y Vida, 2021; Villena *et al.*, 2021, 2022, en prensa; von Essen, 2020a, b; 2021, 2022, en prensa).

Procuraremos, a continuación, señalar las principales aportaciones metodológicas realizadas sobre esta ciudad andaluza destinadas a facilitar los análisis de procesos de cambio lingüístico en tiempo real y aparente y a presentar las líneas de investigación actualmente en desarrollo.

Tras más de 30 años de estudio y diseño del mapa lingüístico de la ciudad de Málaga, poseemos en la actualidad diferentes corpus de datos que quedan esquematizados en la Figura 1:

**FIGURA 1. DESARROLLO CRONOLÓGICO DE LOS CORPUS RECOGIDOS EN EL ENTORNO PRESEEA MÁLAGA**



El trípode que sostiene la estructura está formado por los corpus *Vernáculo Urbano Malagueño* (VUM) (1995), PRESEEA-I (2010) y PRESEEA-II (2015), complementados con información geolingüística y datos del Corpus Telemático (CorTele) en marcha, el corpus de Arqueología lingüística (ARQUEOLING), el de inmigración (Inmigrantes argentinos) y el *Corpus para el estudio de los cambios sociolingüísticos de la Mujer de Ayer y Hoy* (MAYHO). Estas recopilaciones de datos, realizadas en diferentes momentos históricos, y cada una con sus propias características, ha permitido realizar análisis que han facilitado la identificación y el estudio de los procesos de cambio lingüístico en curso y los cambios aparentes en diferentes variedades del español hablado en la ciudad de Málaga.

Con todos los datos recogidos en los últimos 30 años hemos podido realizar una auténtica sociolingüística comparada del español, abordando aspectos como la fonología, la morfosintaxis, la atenuación o la variación léxica. Se han realizado estudios comparativos de la variación de /s/ en la coda silábica y /d/ intervocálica, así como la variación del pronombre personal sujeto y los pronombres átonos en España y América. No obstante, la sucesión de las bases de datos señaladas ha supuesto desafíos constantes y ha provocado que se ejecuten tareas a diferentes velocidades.

En cualquier caso, la organización de las bases de datos señaladas nos ha permitido establecer los objetivos principales actuales de nuestra investigación:

1. Conocer y analizar las variedades intermedias en España a partir del análisis de estas variedades en la ciudad de Málaga.
2. Construir el mapa sociolingüístico de variación y cambio en la ciudad de Málaga elaborado mediante análisis comparados de datos procedentes de los diferentes corpus.
3. Estudiar otras variedades de la comunidad de habla que permitan establecer el panorama sociolingüístico general (lenguas de herencia/contacto de lenguas).
4. Diseñar nuevos corpus adaptados a las nuevas realidades tecnológicas (1) telemático de datos polarizados (CorTele, acro- y basilectal) y sociales (2) *Mujer Ayer y Hoy* (MAYHO).
5. Estudiar en tiempo real determinados fenómenos lingüísticos, lo que nos permite complementar los análisis en tiempo aparente.

6. Comprender los procesos lingüísticos por medio de procesos cognitivos subyacentes a través de estudios comparativos de actitudes/creencias y en el diseño de análisis acústico-perceptivos tanto de las variedades autóctonas, como de las de inmigración.

7. Por último, la difusión académica y social nos preocupa especialmente.

## 2. DESARROLLO DE LOS TRABAJOS

No obstante, como hemos señalado, dada la variedad de tareas que se derivan de los objetivos anteriores, no es fácil conseguir un desarrollo idéntico para todas ellas. Así, podemos diferenciar entre trabajos que se encuentran en fase de ejecución intermedia y avanzada.

En fase intermedia:

A. El corpus telemático basado en chats de *WhatsApp* (*CorTele*). El propósito es el (re)diseño de la investigación a partir de un corpus telemático de datos polarizados. Se trata de un corpus basado en chats de *WhatsApp* (Molina, en prensa), que supone un desafío metodológico importante. Recoge datos de las variedades acro- y basilectales (Molina y Ávila, en prensa).

B. Los estudios fonéticos sobre el vocalismo en la ciudad de Málaga. En el marco del estudio del vocalismo se ha diseñado un estudio piloto para explorar las diferencias entre los datos de Granada y Málaga (trabajo de tesis doctoral de Belén Reyes). Se llevan a cabo, asimismo, estudios perceptivos en ambas comunidades. No solamente se trata de completar un análisis sociofonético de datos conversacionales aún inédito, sino de contribuir también al estudio de la variación y el cambio del consonantismo en la coda silábica.

C. Los estudios de variación sobre la omisión del pronombre personal sujeto. Constituye el Proyecto de Tesis Doctoral de Alba Segura, miembro del Equipo de Trabajo (Ávila y Segura, 2022). El trabajo consiste en el análisis de la variación sintáctica desde la sociolingüística cognitiva y persigue la búsqueda de los factores lingüísticos, sociales, pragmáticos, discursivos, fonéticos y otros propios del hablante, que provoquen la ausencia o presencia del pronombre personal sujeto en el español hablado de la ciudad de Málaga a partir del estudio de los corpus VUM (1995) y PRESEEA-II (2015). Además, para cubrir otro de los objetivos principales del trabajo –comprobar si la ambigüedad fonética que provoca la elisión de /-s/ final en las formas verbales pudiera tener algún efecto en la presencia/ausencia del pronombre sujeto– se realiza un experimento de percepción que permite determinar si realmente se puede considerar la ambigüedad fonética como factor predictor o si existen otras pistas acústicas que ayuden al hablante oyente de la comunidad de habla a reconocer la persona verbal incluso en caso de omisión. Está previsto también un estudio experimental a partir de una recogida complementaria de datos mediante *eye-tracker*, a partir de la cual se relacionará la velocidad de localización

de la información con variables lingüísticas, principalmente discursivas, como la correlación y la distancia del referente, y variables relacionadas con la ambigüedad, tanto morfológica como fonética.

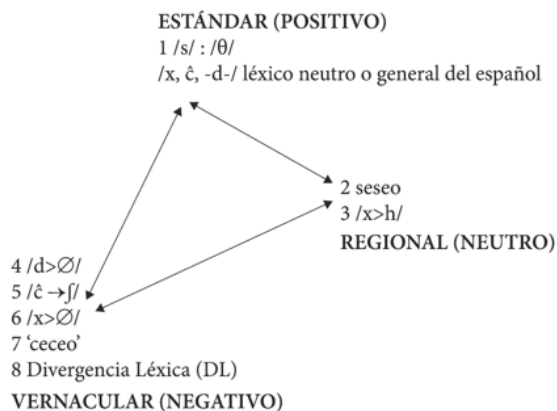
D. El estudio de las creencias/actitudes que subyacen en los cambios lingüísticos se lleva a cabo tanto en la comunidad de habla (PRECAVES, Santos y Ávila), como en la comunidad de inmigrantes (Actitudes en contexto de inmigración). Clara von Essen (2020b) ha diseñado un estudio de MGT sobre la comunidad de inmigrantes argentinos en Málaga, complemento de su estudio general sobre los patrones fónicos de la comunidad migrante argentina en la ciudad de Málaga. La investigación de los procesos cognitivos en la variación léxica abre una perspectiva nueva (Proyecto PULSO), al igual que la percepción de los cambios fonológicos complementa los análisis fonéticos acústicos desarrollados previamente (Vida, 2020, 2021; Molina, 2022).

E. Los estudios de variación sobre la atenuación. El trabajo sobre mitigación ha producido resultados comparados en el nivel sociocultural alto (Ávila, 2021; Ávila y Rodríguez Cruces, 2020). La investigación se extiende ahora al resto de los datos del corpus gracias a la tesis doctoral de María Rodríguez Cruces, que trabaja con informantes de los corpus VUM (1995), PRESEEA-II (2015) y MAYHO (2023). De esta manera, además de comparar el estudio con otros previos, se puede realizar una investigación diacrónica y comprobar si ha cambiado el uso de la atenuación en la ciudad. Como novedad, se trabaja en la reformulación de algunas hipótesis, fruto de los avances previos realizados por diversos grupos PRESEEA. Por ejemplo, se incluye el cambio de tema como posible estrategia atenuadora, se atiende a la teoría de los tipos de actos ilocucionarios de Searle y se presta atención al empleo de ciertos fenómenos fonéticos y prosódicos que pudieran estar intrínsecamente relacionados con la mitigación lingüística.

En fase avanzada:

A. El estudio de la variedad intermedia. Es el núcleo de la investigación. Tal y como se muestra en la Figura 2, el propósito es descubrir y describir la coherencia en el uso de los rasgos que indexicalizan la síntesis de la convergencia hacia el estándar, por un lado, y la reafirmación de la variedad andaluza, tanto en el plano variacionista y sociofonético acústico (Molina, 2020; Vida, 2020, 2021; Villena y Vida, 2020; Villena *et al.*, 2022), como en el plano léxico (Ávila, 2022c, 2022b), por otro lado.

**FIGURA 2. PROCESOS DINÁMICOS CONVERGENCIA-DIVERGENCIA EN EL ESPAÑOL DE MÁLAGA**



Fuente: Ávila 2023.

B. Los estudios sobre el corpus de Arqueología Lingüística (Arqueoling, estudios en tiempo real). Es la fuente de datos en tiempo real de la segunda mitad del siglo XX que nos permite observar cuándo se inician los fenómenos lingüísticos que se observan en la actualidad. El diseño de los archivos se organizó y concretó durante el año 2021 (Villena *et al.*, 2021). Posteriormente, Rocío Cruz (UGR) trabajó con algunos archivos en el marco de la ayuda postdoctoral Margarita Salas. Disponemos de datos en archivos de audio (radio, TV, registros personales) y documentos audiovisuales que nos permiten realizar estudios transversales sobre los datos «arqueológicos» y de los corpus VUM (1995), PRESEEA-I (2010) y PRESEEA-II (2015), correspondientes a sincronías sucesivas (Vida, en prensa). Se complementarán con datos geolingüísticos (ALEA) y los procedentes del corpus telemático CorTele (Molina y Ávila, en prensa), donde se da cuenta de la situación del corpus, organización de registros, datos cuantitativos, método de transcripción y alineación, etc. La investigación en curso de microcorpus (administrativos; coloquiales privados y semipúblicos) basados en registros sonoros, donados al proyecto de distintas fuentes, permitirán el análisis de diversas variables lingüísticas para confirmar hipótesis sobre la formación y el desarrollo de la variedad intermedia.

C. Los estudios sobre la variación léxica (riqueza, centralidad, disponibilidad, amplitud): variación léxica, estudios sobre el léxico disponible y Proyecto PULSO (*Observación del Pulso Social en Andalucía a través del análisis léxico, PULSO Andalucía*). *Primera fase*. Este Proyecto, que sintetiza todas las líneas de análisis de la variación léxica desarrolladas en la Universidad de Málaga, está financiado por la Junta de Andalucía: UMA20-FEDERJA-013. Se pretende captar la situación del vocabulario en el proceso de modernización y globalización, paralelo a los procesos

fonológicos y morfosintácticos reflejados en otros objetivos. Se recogen datos en el marco metodológico de la disponibilidad léxica, con el propósito de llevar a cabo una aproximación al condicionamiento social de los repertorios léxico-conceptuales individuales en áreas seleccionadas de actualidad social, política e ideológica (Ávila, 2021b, 2022b, 2022 en prensa, 2023 en prensa; Ávila *et al.*, 2020, 2022a, 2022b). Se insiste en los fundamentos teóricos y metodológicos para el estudio de la competencia léxica (Villena *et al.*, 2022, en prensa), con insistencia en la construcción de herramientas y software accesible online (Ávila *et al.*, 2021).

D. Los estudios de la variación de los pronombres átonos en España y América. Díaz y Villena (2022, en prensa) analizan datos del CORPES para el estudio de la variación de los pronombres átonos en España y América. Díaz (2022) ha iniciado el análisis con los verbos psicológicos tipo-*molestar* y su contraste con los verbos tipo-*gustar* y los transitivos tipo-*matar*, mediante modelos logísticos bayesianos.

E. El estudio de la sociolingüística de la migración. María Clara von Essen compara los resultados del análisis sobre la variedad de los inmigrantes argentinos en Málaga (estudio acústico-perceptivo de *ye* y de las fricativas coronales; estudio complementario del léxico y de la morfología verbal) con los datos de las variedades de Málaga (von Essen, 2020a, 2020b, 2021, 2022, 2022 en prensa). La investigadora ha iniciado también el diseño del estudio comparado de la entonación en ambas comunidades.

### 3. FINAL

Aunque algunos aspectos del proyecto se encuentran en diferentes etapas de desarrollo y se han afrontado diversos desafíos, los resultados obtenidos hasta ahora son prometedores y contribuyen al conocimiento sociolingüístico del español en España. Se espera que el proyecto continúe avanzando y brinde nuevas perspectivas sobre la variación y el cambio lingüístico en el contexto español. Es fundamental difundir los resultados en el ámbito académico y en la sociedad para promover un mayor conocimiento y comprensión de la diversidad lingüística en España.

En el momento actual nos preocupa especialmente la transferencia de los resultados de nuestros trabajos a la sociedad. En este sentido, estamos trabajando en los siguientes ámbitos:

*Arqueoling.* Los resultados del estudio arqueológico lingüístico pondrán a disposición de la comunidad muestras de su pasado, junto a interpretaciones de los hechos. El trabajo es colaborativo: sociedad/equipo de investigación (web específica con datos y resultados publicados).

*DispoCen.* Una herramienta para el cálculo de la Centralidad léxica. Propuesta de estudios futuros: crear una corriente de investigación sobre variación léxica en PRESEEA que podría iniciarse con un estudio común de riqueza léxica basado en experiencias anteriores.

*Videos de difusión educativa y social.* Existe una base de datos de archivos audiovisuales de las variedades y fenómenos analizados. La comunidad educativa, en particular, podría estar interesada en disponer de registros de este tipo que ayudasen a comprender la diversidad sociolingüística.

*Cursos de formación* de estudiantes egresados y de estudiantes de máster y doctorado en metodología de análisis cuantitativo. Se aprovechan las iniciativas de la Junta de Andalucía (Iniciativa Juvenil) y de la UMA (becas de formación en investigación; becas de colaboración).

Hemos encauzado todo lo anterior a través de un Proyecto de Innovación Educativa avalado por la Universidad de Málaga construido a partir de los siguientes objetivos:

A. Establecer la comparación del *antes* y del *ahora* del rol social de la mujer presente en el corpus PRESEEA a través del estudio de sus voces. El acceso al trabajo profesional y la formación ha impulsado el uso preferente de determinadas variantes. Cada grupo de trabajo estudia un rasgo característico (fonético-fonológico, morfológico, léxico, sintáctico o pragmático-discursivo) de las mujeres de *ayer* y de *hoy* acudiendo a los materiales recogidos en los corpus. El estudio de esta evolución de carácter sociolingüístico se presenta apasionante cuando se correlacionan las transformaciones lingüísticas con la descripción que aportan los archivos sonoros sobre el papel de la mujer en la sociedad.

B. Permitir al alumnado acceder a las diferentes fases de elaboración del corpus PRESEEA para realizar análisis contrastivos con perspectiva histórica. Los participantes acceden a la información a partir de las escuchas de las entrevistas que contienen los corpus. Gracias a ellas, pueden entender cómo la historia ha moldeado diferentes rasgos de la pronunciación de las mujeres locales y realizan análisis etnográficos y antropológicos para construir la historia de la mujer a lo largo de estos 25 años.

C. Motivar al alumnado a aprender técnicas y manejo de herramientas para la construcción, preparación y análisis de corpus de habla: programas de edición de audio (*WavePad, Audacity*), de anotación y transcripción (ELAN) o de análisis acústicos (*Praat*). La incorporación de los métodos de investigación a los procedimientos de enseñanza-aprendizaje estimula la curiosidad que lleva a la construcción del conocimiento y permite, finalmente, el desarrollo del espíritu crítico.

D. Establecer en PRESEEA una línea de trabajo que genere futuros proyectos de innovación educativa.

E. Incentivar el desarrollo de trabajos de investigación de estudiantes universitarios en el contexto PRESEEA, en especial, de tesis doctorales.

## REFERENCIAS

Ávila, Antonio (2021a). La atenuación discursiva en el contexto de la acomodación comunicativa. Análisis de casos. *Pragmalingüística*, 29: 27-43.

- Ávila, Antonio (2021b). «Distribución social del léxico dialectal en la ciudad de Málaga. Reflexiones en torno a la utilidad del empleo de dialectalismos en el aula de lenguas extranjeras». *Entrelinguas* 6(1): 11-25.
- Ávila, Antonio (2021c) «¿Es legítimo (y necesario) escribir las variedades no estándares de las lenguas? El caso del español de Andalucía». En I. Piechnik y M. Wicherek (eds.), *Langues romanes non standard*, Kraków: Uniwersytet Jagiellonski w Krakowie: Biblioteka Jagiellonska, pp. 43-57. ISBN: 978-83-67127-110.
- Ávila, Antonio (2022a). «En los orígenes del estudio de la variación lingüística Interpretación optimista de los resultados de un estudio cuantitativo de variación léxica en la ciudad de Málaga». *Spanish in Context*, 19(2): 189-216. DOI: <https://doi.org/10.1075/sic.20016.avi>
- Ávila, Antonio (2022b). «Algunas percepciones categoriales compartidas por preuniversitarios andaluces sobre la crisis del coronavirus y sus consecuencias. ¿Deberíamos preocuparnos? Un acercamiento desde la disponibilidad y la centralidad léxica». *Tejuelo: Didáctica de la Lengua y la Literatura* 35(3): 17-42.
- Ávila, Antonio (2022c). «Explicación multidimensional de la dinámica sociolingüística en Andalucía. El caso del español de Málaga». *Revista Española de Lingüística Aplicada (RESLA)*, 36(1): 60-86
- Ávila, Antonio (2023). «Avance de una propuesta para el desarrollo de la tradición lexicoestadística hispánica: el índice de centralidad léxica». *Nueva Revista de Filología Hispánica*, 71(1): 3-29.
- Ávila, Antonio (en prensa). «A lexical-cognitive approach to the consequences of the coronavirus crisis for the emotions and perceptions of the Spanish population aged 65 and over». *European Journal of Applied Linguistics*.
- Ávila, Antonio, Inmaculada Santos y Ester Trigo (2020). «Análisis léxico-cognitivo de la influencia de los medios de comunicación en las percepciones de universitarios españoles ante la COVID-19». *Círculo de Lingüística Aplicada a la Comunicación*, 84: 85-95. DOI: <https://doi.org/10.5209/clac.70701>
- Ávila, Antonio y María Rodríguez (2020). «Estudio de las funciones de la atenuación en hablantes de Málaga con nivel de instrucción alto. Aproximación sociolingüística». *Pragmática sociocultural/Sociocultural Pragmatics*. <https://www.degruyter.com/document/doi/10.1515/soprag-2020-0013/html>
- Ávila, Antonio, José María Sánchez Sáez y Nana Odishelidze (2021) «DispoCen. Mucho más que un programa para el cálculo de la disponibilidad léxica», *ELUA, Estudios de Lingüística*, 35: 9-36.
- Ávila, Antonio y Alba Segura (2022). «Estudio de las variables predictoras de la expresión del sujeto pronominal en el corpus PRESEEA Málaga. Nivel de instrucción bajo». *Anuario de Letras*, UNAM.
- Ávila, Antonio, Inmaculada Santos y Ester Trigo (2023). «Perceived Impacts of COVID-19 by Spanish University Students: Changes in the Physical and Social Environments». En S. D. Brunn y D. Gilbreath (eds.), *COVID-19 and a World of Ad Hoc Geographies*, eBook: Springer, pp. 2059-2076.

- Díaz, Francisco (2022) «¿El ruido la molesta o El ruido le molesta? Sobre la alternancia dativo/acusativo con los verbos psicológicos en el español de Málaga (nivel culto)». *LinRed*, 19. <https://revistas.publicaciones.uah.es/ojs/index.php/linred/article/view/1870>
- Díaz, Francisco y Juan Andrés Villena (2024). «¿Los atlas le aburren o los atlas la aburren? Distribución del uso de los clíticos pronominales no reflexivos de tercera persona en España e Hispanoamérica (CORPES XXI)». En I. Molina, E. Hernández, P. Martín Butragueño y E. Mendieta (eds.), *Caminos y palabras. Estudios de variación lingüística dedicados a Pilar García Mouton*, Valencia: Tirant Humanidades, pp. 549-582.
- Molina, Álvaro (2020). «Percepción comunitaria de la distinción fonemática de la /s/ y la /θ/ en hablantes andaluces. Estudio en la ciudad de Málaga». *Spanish in Context* 17(3): 415-437. <https://benjamins.com/catalog/sic.19007.mol>
- Molina, Álvaro (2021) *Fundamentos acústico-perceptivos de la escisión prestigiosa de /θ/*. *Estudio sociofonético en Málaga*. Berna: Peter Lang Series: Studien zur romanischen Sprachwissenschaft und interkulturellen Kommunikation. <https://www.peterlang.com/document/1140460>
- Molina, Álvaro (2022). «Variación fonético-perceptiva en el español peninsular. Test intercomunitario en las ciudades de Madrid y Málaga. Estudio preliminar». En A. Ávila, M. Vida y J. L. Jiménez Ruiz (eds.), *Estudios de variación lingüística: Homenaje a Juan Andrés Villena. Anexo Monográfico, LinRed, XIX*. Open Access. <https://revistas.publicaciones.uah.es/ojs/index.php/linred/article/view/1865>
- Molina, Álvaro (en prensa). «Voice messages on WhatsApp as telematic corpora in Spanish dialectology. The PRESEEA-Málaga Project». En E. Pustka, C. Quijada Van den Berghe y V. Weiland (eds.), *Corpus Dialectology: from methods to theory (French, Italian, Spanish)*. John Benjamins Series: Studies in Corpus Linguistics.
- Sánchez, José María, Antonio Ávila (2020). *DispoCen. Licencia SafeCreative*. Número normalizado: 2012046136184. Málaga, Universidad de Málaga, 12/2020. <https://www.youtube.com/watch?v=IU5VfUvG4Ag>
- Santos, Inmaculada y Antonio Ávila (2021). «Creencias y actitudes lingüísticas de los universitarios malagueños hacia la variedad andaluza», *Philologia Hispalensis*, 35(1), 127-147. ISSN 1132-0265 / eISSN 2253-8321, de la Facultad de Filología de la Universidad de Sevilla. <https://revistascientificas.us.es/index.php/PH/article/view/12010>
- Santos, Inmaculada y Antonio Ávila (2023). «Creencias y actitudes de los malagueños hacia la variedad castellana: análisis según variables sociodemográficas». En A. Cestero y F. Paredes (coords.), *La percepción de la variedad castellana. Creencias y actitudes lingüísticas en el siglo XXI*, Alcalá de Henares: Universidad de Alcalá, pp. 203-236.
- Vida, Matilde (2020). «Procesos de estandarización y divergencia dialectal en las variedades meridionales del español de España». En F. Dubert-García, V. Míguez y X. Sousa (eds.), *Varietades lingüísticas en contacto na Península Ibérica*. Consello da Cultura Galega, pp. 127-154.

- Vida, Matilde (2021). «Desgaste dialectal y patrimonio lingüístico regional. Prestigio e identidad cultural en la formación de la variedad coínética andaluza». En A. Pérez Jiménez (Ed.) *Patrimonio filológico: contribuciones y nuevas perspectivas*. Berna: Peter Lang, pp. 37-54.
- Vida-Castro, Matilde (2022). «On competing indexicalities in southern Peninsular Spanish. A sociophonetic and perceptual analysis of affricate [ts] through time». *Language Variation and Change*, 34(2): 137-163. DOI: doi:10.1017/S0954394522000084
- Vida, Matilde, Juan Andrés Villena e Isabel Molina (2023). «Variación de la /-s/ postnuclear en español: Patrones sociolingüísticos y geolectales». *Círculo de Lingüística Aplicada a la Comunicación*, 94: 23-35.
- Villena, Juan Andrés y Vida, Matilde (2021). «Variation, identity and indexicality in Southern Spanish. On the emergence of a new variety in Urban Andalusia». En M. Cerruti y S. Tsiplakou (eds.), *Intermediate language varieties. Konai and regional standards in Europe*. John Benjamins, pp. 149-182. DOI: <https://doi.org/10.1075/silv.24>
- <https://benjamins.com/catalog/silv.24.07vil>
- Villena, Juan Andrés, Antonio Ávila, y Yorgos Sinoaidis (2021). *Las voces de Málaga. Reconstrucción histórica y archivo sonoro de la Ciudad Lingüística*. Málaga: Universidad de Málaga. Área de Lingüística General. Documento inédito.
- Villena, Juan Andrés, Matilde Vida y Álvaro Molina (2022). «Coherence in a levelled variety. The Case of Andalusian». En K.V. Beaman y G. Guy (eds.). *The Coherence of Linguistic Communities: Orderly Heterogeneity and Social Meaning*. New York: Routledge, pp. 239-257. DOI: 10.4324/9781003134558-18. <https://www.taylorfrancis.com/chapters/edit/10.4324/9781003134558-18/coherence-levelled-variety-juan-andrés-villena-ponsoda-matilde-vida-castro-álvaro-molina-garc%C3%ADa>
- Villena, Juan Andrés, Antonio Ávila y José María Sánchez (en prensa) «Individual Lexical Breadth and its associated measures. A contribution to the calculation of individual lexical richness». *New Perspectives on Spanish Lexical Development*, De Gruyter.
- von Essen, María Clara (2020a): «On the different ways of being a bi-dialectal immigrant. The case of Argentineans in Spain». *Lengua y Migración/Language & Migration* 12(2): 7-43. ISBN: 1889-5425
- <http://lym.linguas.net/Download.axd?type=ArticleItem&id=262>
- von Essen, María Clara (2020b): «Significado social de las actitudes lingüísticas, la red social y las variables de pequeña escala en los estudios de inmigración: combinación de métodos cualitativos y cuantitativos». *Iberoromania*, 91: 93-132. ISSN: 1865-9039. DOI: <https://doi.org/10.1515/iber-2020-0008>
- von Essen, María Clara (2021). *Identidad y contacto de variedades. La acomodación lingüística de los inmigrantes rioplatenses en Málaga*. Colección: Studien zur romanischen Sprachwissenschaft und interkulturellen Kommunikation. Berlin:

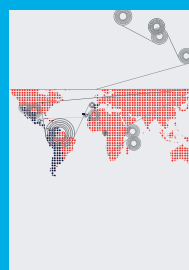
- Peter Lang. DOI: <https://doi.org/10.3726/b18490>. ISBN: 978-3-631-85341-2.
- von Essen, María Clara (2022). «La influencia del entorno fonológico en la caracterización fonética de la consonante ye en contextos de inmigración». *Lingüística en la red (LinRed)*, *Estudios de variación lingüística: Homenaje a Juan Andrés Villena*, Núm. XIX. ISSN: 1697-0780. DOI: <https://doi.org/10.37536/linred.2022.XIX.1867>
- von Essen, María Clara (en prensa): «Acomodación léxica en el contacto de dialectos del español. El caso de los rioplantenses en Málaga». En R. Kailuweit, S. Schlumpf-Thurnherr (eds.), *Migración, pluricentrismo y acomodación. Nuevas perspectivas desde la lengua española*, Baden-Baden: Nomos Verlag.



# La sociolingüística del español en España y América: 25 años de estudios coordinados

Isabel Molina Martos, Pedro Martín Butragueño,  
Gema Cuesta Chorro, Nicolás Moreno Martín de Nicolás  
(eds.)

En 2024 se cumplió un cuarto de siglo desde que se pusiera en marcha la iniciativa de una investigación sociolingüística coordinada del español. Este objetivo se ha logrado en el marco del “Proyecto para el Estudio Sociolingüístico del Español de España y América” (PRESEEA), desde cuyas bases metodológicas se ha recopilado el mayor corpus de datos de la lengua española en las principales capitales del mundo hispanohablante. A su vez, este gran corpus ha sido la base para una enorme cantidad de trabajos de investigación. Partiendo de los fundamentos establecidos en el PRESEEA, la monografía aporta una visión de conjunto de la investigación sociolingüística a través de dieciséis capítulos relativos a una selección de comunidades hispanohablantes (Buenos Aires, Ciudad de México, Granada, La Habana, Las Palmas, Madrid, Málaga, Medellín, Pamplona, Santiago de Chile, Sevilla, Valencia). En algunos de estos, se repasa el estado de la investigación sociolingüística en distintas comunidades urbanas hispanohablantes, mientras que en otros se ilustra la evolución teórica y metodológica de la disciplina mediante el análisis de los fenómenos lingüísticos variables del español.



Universidad  
de Alcalá

EDITORIAL  
UNIVERSIDAD DE ALCALÁ