

# La concordancia de *haber* existencial en la Ciudad de México

*Yolanda Lastra\**

*Universidad Nacional Autónoma de México y El Colegio de México*

*Pedro Martín Butragueño*

*El Colegio de México*

## Resumen

El artículo estudia la concordancia singular o plural entre el verbo *haber* con valor existencial y su argumento plural en 108 entrevistas recogidas en la Ciudad de México como parte de un corpus sociolingüístico estructurado por edad, género y nivel de estudios. Se consideran todos los datos relevantes, que fueron 633. Para el análisis, se sigue la metodología propuesta en el marco del “Proyecto para el estudio sociolingüístico del español de España y América (PRESEEA)”, empleando una amplia serie de variables lingüísticas y sociales. Los resultados son sometidos a un análisis estadístico riguroso y se comparan con los obtenidos en otras investigaciones; se trata de un proceso que ha recibido bastante atención en el campo de estudio de la variación lingüística del español. Las hablas de la Ciudad de México se agrupan con las de otras zonas conservadoras.

Palabras clave: variación sintáctica, haber, concordancia, sociolingüística, Ciudad de México.

\* Para correspondencia, dirigirse a: Yolanda Lastra (ylastra@unam.mx) o Pedro Martín Butragueño (pmartin@colmex.mx), Universidad Nacional Autónoma de México, Instituto de Investigaciones Antropológicas, Ciudad Universitaria, Ciudad de México 04510, México y El Colegio de México, Camino al Ajusco 20, Col. Pedregal de Santa Teresa, Ciudad de México 10740, México.

THE AGREEMENT OF EXISTENTIAL *HABER* IN MEXICO CITY

## Abstract

This paper studies the singular or plural agreement between Spanish verb *haber* with existential meaning and its plural argument, in 108 interviews collected in Mexico City as a part of a sociolinguistic corpus structured by age, gender and educational level. All relevant data (633) are considered. For the analysis, the proposed methodology in the “Project for the sociolinguistic study of Spanish from Spain and from America (PRESEEA)” is followed, using a wide range of linguistic and social variables. The results are subjected to rigorous statistical analysis and compared with those obtained in other studies. The singular/plural alternative has received considerable attention in the field of study of linguistic variation in Spanish. Mexico City Spanish is grouped with other conservative areas.

Keywords: Syntactic variation, *haber*, agreement, sociolinguistics, Mexico City.

Recibido: 20/02/2016

Aceptado: 01/07/2016

## INTRODUCCIÓN

El estatus del argumento con sentido plural que acompaña al verbo *haber* con valor existencial, y la consecuente resolución de la concordancia es uno de los problemas más clásicos en los estudios sobre variación sintáctica del español<sup>1</sup>. La selección del verbo *haber* en singular o plural no establece un cambio en las condiciones de verdad (Gómez Molina 2013: 254) y el análisis más común es la interpretación del argumento único como sujeto en vez de como objeto directo, es decir, como un caso de regularización, al interpretarse la construcción de manera semejante a los verbos intransitivos (Díaz-Campos 2014: 170), o simplificación (Bentivoglio y Sedano 2011:

<sup>1</sup> Agradecemos el apoyo prestado por Jorge Kasep, becario del Laboratorio de Estudios Fónicos, para extraer los datos del Corpus sociolingüístico de la Ciudad de México (CSCM). Una versión previa de este trabajo se leyó en el CILIR2015. XV Coloquio Internacional de Lingüística Iberorromance. Rouen, 3-5 de junio de 2015; no se entregó para considerarse en la publicación posterior derivada del coloquio. Agradecemos también los comentarios recibidos por parte del prof. Jeroen Claes en un correo del 31 de mayo de 2015.

174). Otra opción analítica es considerar que se trata de concordancia con el OD; como señala Soler (2012), “tiene un peso muy fuerte el hecho de que el elemento nominal se sustituye, no por un pronombre personal sujeto sino por el clítico de objeto directo (*lo, la, los, las*), incluso cuando hay pluralización, lo que indicaría que lo que tenemos es concordancia con el objeto directo”, como en (1):

- (1) Ahí no pasaba un coche, tampoco **los habían** ¿me entiendes? (Alcalá de Henares, CREA, grabación, poca formalidad, en Soler 2012).

La pluralización de *haber* se ha producido históricamente en el contexto de su baja transitividad, característica dada desde el propio latín. Se documenta desde antiguo, como se ve en (2a) y (2b). Existen dos tipos de concordancia de *haber* con su argumento; en la primera el verbo muestra flexión de número (2c), en la segunda se introduce el número y persona (2d), con lo que el hablante se incluye en la predicación del evento (Hernández Díaz 2006: 1146):

- (2) a. y porque en la conclusión **no hayan tonguerías ni dilaciones** [*Diálogo de Mercurio y Carón*].  
 b. No vinistes a votar cuando **hubieron elecciones** [*Martín Fierro*].  
 c. **Hubieron muchos problemas** en la presentación [radio].  
 d. **Habemos muchos** aquí que **somos inocentes** [televisión] (Hernández Díaz 2006, pp. 1146 y 1148).

Hernández Díaz sugiere las siguientes etapas de cambio: 1) innovación agazapada en el verbo auxiliar; 2) innovación con alteración menor (solo se añade un fonema); 3) innovación con alteración mayor (cambios mayores en la desinencia); 4) innovación extrema, en el presente de indicativo, de modo que *hay > han*, como en *Cada rato han accidentes* (*ibíd.*: 1151-1152).

La concordancia se manifiesta como cambio en curso en numerosas variedades del español, aunque a diferentes ritmos de avance (De Mello 1991, 1994; Bentivoglio y Sedano 2011). La cuestión ha sido bastante estudiada: además del trabajo de De Mello (1991, 1994) en once ciudades, puede verse Bentivoglio y Sedano (1989), D’Aquino (2004, 2008), Freitas (2004, 2008), Díaz-Campos (2003, 2014) para Venezuela; Montes Giraldo (1982) para Colombia; Blas Arroyo (1996) y Gómez Molina (2013) para España (con especial frecuencia en las zonas de contacto con el catalán, Claes (2014a) para las capitales antillanas, etc. En el caso de México, el problema ha sido abordado al menos por Moss (1990), Montes de Oca (1994), Hernández Díaz (2006), Soler (2012) y Serrano (2015) para la Ciudad de México, y Castillo-Trelles (2007) para Mérida, Yucatán.

La variable está siendo considerada también dentro del “Proyecto para el estudio sociolingüístico del español de España y América” (“Propuesta de codificación” de 2013, que será especialmente tenida en cuenta en este artículo), para poder contribuir a una perspectiva comparada cuya misma metodología permita establecer con exactitud los grados de avance y difusión del fenómeno.

Así, el objetivo de este trabajo es ofrecer una visión actualizada de la variación presente en la lengua hablada, al tiempo que mostrar el estatus de la variedad de español del centro de México frente a otras variedades de habla.

### HIPÓTESIS DE TRABAJO

Nuestras hipótesis de trabajo para organizar la investigación han sido dos. Por una parte, que la variedad mexicana es relativamente conservadora en contraste con los índices mostrados por otras variedades, tal como ocurre con diversos procesos lingüísticos, como por ejemplo en la expresión del sujeto pronominal (cf. Lastra y Martín Butragueño 2015). Por otro lado, que en la Ciudad de México se estaría produciendo un avance en la extensión de la concordancia con un argumento plural, lo que se percibiría tanto en las frecuencias globales, como en la difusión lingüística a través de diferentes contextos y en la difusión social del fenómeno en diversos grupos sociales, incluidos los más instruidos; nuestra expectativa es que habrían sido los grupos intermedios los encargados de empujar esta expansión.

Como se irá viendo a lo largo de la exposición, la primera hipótesis resulta ser completamente cierta, mientras que la segunda no funciona con demasiada claridad, aunque la variable sí presenta una rica dimensión social. De hecho, aunque los resultados en tiempo aparente, al comparar las generaciones, abonan la posibilidad de una cierta expansión de la concordancia plural, la comparación en tiempo real con otros estudios sugiere incluso lo contrario, una leve retracción. Es posible que la expectativa de un gran avance del proceso se haya producido por la elevada prominencia de algunos casos y por su presencia en medios de comunicación, discursos políticos, etc. El análisis de los materiales lingüísticos estructurados ha mostrado que los hechos son más moderados de lo que pudiera pensarse.

## METODOLOGÍA

En la investigación sobre *haber* existencial se han venido empleando tradicionalmente dos tipos de datos principales: los basados en diferentes cuestionarios y los obtenidos a partir de entrevistas sociolingüísticas. Como en otros casos, los cuestionarios permiten obtener datos más controlados y comparables, mientras que las entrevistas proporcionan materiales más realistas que permiten entender mejor cómo se conforma la gramática en uso (Bybee 2001, 2008). En muchas ocasiones, los datos de cuestionario tienden a producir proporciones mayores de pluralización, al proponer contextos cruciales. En ese sentido, debe tenerse cuidado al comparar unos trabajos con otros. En nuestro caso, hemos preferido, por lo pronto, el empleo de entrevistas, para poder tener una idea más clara de la magnitud de las pluralizaciones presente en la producción lingüística.

Nuestro trabajo pretende mostrar el grado de difusión de la concordancia de *haber* en la Ciudad de México a partir del estudio de 108 entrevistas sociolingüísticas que forman parte del *Corpus sociolingüístico de la Ciudad de México* (recogido entre 1997 y 2007 y constituido por entrevistas de más de 300 informantes; cf. Martín Butragueño y Lastra 2011, 2012 y 2015). Ya se ha dicho que se sigue muy de cerca las propuestas de PRESEEA (2013); puede verse asimismo Gómez Molina (2013) para una discusión detallada de la metodología sugerida en PRESEEA. Desde luego, la comparabilidad de los resultados entre diferentes trabajos es uno de los mejores argumentos al respecto.

En consonancia con esta decisión metodológica, para definir la VARIABLE DEPENDIENTE (la realización singular o plural del verbo), se consideran los casos de *haber* acompañados de una FN plural o bien de una FN singular colectiva; se incluyen las construcciones perifrásticas, con la excepción de *haber que* + V. No se consideran los ejemplos con *hay* (pues no hay ningún ejemplo de pluralización). Estas condiciones establecen el marco de variación (*envelope of variation*).

En cuanto a las VARIABLES INDEPENDIENTES LINGÜÍSTICAS, se tiene en cuenta la persona del verbo: 3a. singular, 3a plural, 1a plural (como en el ejemplo 3a); la estructura verbal, sea verbo o perífrasis (3b); el carácter humano o no de la FN (3c); la definitud de la FN (3ch); la forma léxica de la FN, sea sustantiva, pronominal o implícita (3d); la posición de la FN con respecto a *haber* (3e); las marcas de pluralidad en la FN (3f); el tiempo y modo verbales (3g); la intercalación de elementos entre el V y la FN (3h); el tipo de oración (3i); y la polaridad (3j). Se anotan dos VARIABLES INDEPENDIENTES ESTILÍSTICAS: la especialización o campo (cotidiano o técnico) (3k); y el tipo

de discurso (dialógico, argumentativo, expositivo, narrativo o descriptivo) (31). Y tres VARIABLES INDEPENDIENTES SOCIOLINGÜÍSTICAS (género, edad y nivel de estudios). Los datos se analizan cualitativa y cuantitativamente, como se irá desarrollando en las páginas siguientes.

- (3) a. ahí está la Volkswagen <~vosvagueu> adelantito/ ahí trabajaba yo [...] y **habíamos** tres o cuatro personas ahí <~ai>/ unos que vivían por allá de aquel lado y nos veníamos en el taxi [ent. 88, turno 1207, 12H] (1a. plural)<sup>2</sup>.
- b. en la superior la/ cuestión esta de/ los porros/ ya es más política// ya no se meten tanto contigo/ sí ha-/ **llegaron a haber** casos// afortunadamente/ yo en la superior/ conocí gente/ que andaba en eso/ pero ya más político [ent. 5, turno 32, 31H] (perífrasis).
- c. mi generación más o menos se salvó/ yo creo que este/ que todos/ hay un doctor/ un un este// un licenciado/ más o menos// aunque algunos sí/ ¿no?/ pues se perdieron en la borrachera y todo eso/ porque era permanente en el edificio/ ya sabe/ **había unos borrachos**/ ¿no?/ que estaban ¡toda! la vida/ tomando en en (sic)/ la puerta del edificio [ent. 13, turno 42, 32H] (FN argumento +humano).
- ch. [ah bueno] mira este <~este:> la cuestión es de que pues <~pus> yo ya no quise seguir este estudiando ¿no? (ajá) no porque no **hubieran los** <~los:>/ **los medios** ¿no? simplemente que pues <~pus>/ me empezó a gustar el dinero y quise salir a trabajar a este/ pues <~pus> a esto [ent. 77, turno 10, 11H] (FN definida).
- d. entonces <~entóns> pues <~pus> yo decía “sí/ estoy con unas personas este/ muy/ este <~este:>” [...] “importantes”/ ¿no?/ entonces <~entóns> [sí **había**] [ent. 59, turno 252, 22M] (argumento implícito).
- e. había/ que mire así haga de cuenta un-/ una hilera/ atrás de la/ de su mamá/ y del más grandecito hasta el más chiquitito/ eran **doce muchachos** los que **habían** [ent. 102, turno 368, 13H] (anteposición).
- f. cuando **habían otras personas que no tenían**/ ese señor le/ les pasaba mercancía y se la pagaban [ent. 78, turno 322, 11H] (FN con modificadores).
- g. ¿cómo me gustaría?/ pues <~pus> igual y que no **hubiera** hombres mujeriegos ¿no? bueno/ que lógico que no va a haber un hombre fiel nunca/ ¿no? [ent. 84, turno 210, 11M] (pretérito de subjuntivo).
- h. ahorita hay más eso/ se [ve más <...>] [...] pero <~pero:>/ pero siempre **habían** [...] siempre **habían este** <~este:> operativos llegaban a querer

<sup>2</sup> En los ejemplos del CSCM, se menciona entre corchetes el número de entrevista y el número de turno. En la clave final, el primer número hace referencia al nivel educativo bajo (“1”), medio (“2”) o alto (“3”); al grupo de edad, sea que se trate de jóvenes (20-34 años, “1”), adultos (35-54, “2”) o mayores (55 en adelante, “3”); el género es hombres, “H” y mujeres, “M”.

levantar [ent. 78, turno 178, 11H] (presencia de elementos interpuestos entre V y FN).

i. mira después de los diecisiete años/ que yo ya trabajé y todo no/ o sea/ pues <~pus> **yo veía que no había** permisos de nada [ent. 57, turno 96, 22M] (subordinada no relativa).

j. la manera en que me enseñaron me/ me gustó/ porque no **había**/ no había reglas/ yo llegaba un día/ y si quería entrar a ese departamento entraba [ent. 63, turno 560, 23H] (polaridad negativa).

k. en la superior la/ cuestión esta de/ los porros/ ya es más política// ya no se meten tanto contigo/ sí ha-/ **llegaron a haber** casos// afortunadamente/ yo en la superior/ conocí gente/ que andaba en eso/ pero ya más político [ent. 5, turno 32, 31H] (tema técnico).

l. era bien bonito porque/ bueno/ como que vivías en un pueblo// no había/ pavimento/ no había/ a lo mejor ni luz/ yo no me acuerdo que hubiera luz// digo en la calle/ ¿no?// **había** unos arbolotes gigantescos/ que nada más queda como uno o dos ahí en/ en el barrio [...] todos se fueron pudriendo y/ ya no hay// no había calles/ [ent. 50, turno 209, 22H] (discurso descriptivo).

## EXPOSICIÓN DE RESULTADOS

Los principales resultados obtenidos se resumen en los cuadros 1, 2 y 3. En las 108 entrevistas hubo 633 casos que satisfacen las condiciones establecidas por el marco de variación, como muestra el cuadro 1. De ellos, 561 pertenecen a los informantes (es decir, el 88.6%), y 72 a los entrevistadores y a otros participantes (11.4%). Si se considera el total de los casos, en el 8.1% de ellos el verbo *haber* se manifiesta en plural; si solo se consideran los ejemplos producidos por los 108 informantes, el porcentaje asciende muy levemente, a 8.4%. Los entrevistadores muestran, por su parte, un porcentaje muy bajo de casos, apenas del 4.4%. El 25% documentado por otros participantes es en realidad poco significativo, considerando que se trata de una pluralización entre cuatro datos posibles. De aquí en adelante nos referiremos solo a los 561 casos pertinentes producidos por los informantes, 47 de los cuales exhiben pluralización.

CUADRO 1. Resultados totales según participantes en 108 entrevistas de CSCM

| <i>Participante</i> | <i>F/N</i> | <i>%</i> |
|---------------------|------------|----------|
| Informante          | 47/561     | 8.4      |
| Entrevistador       | 3/68       | 4.4      |
| Otro participante   | 1/4        | 25.0     |
| Total=              | 51/633     | 8.1      |

El cuadro 2 presenta los resultados absolutos y porcentuales correspondientes a cada factor de cada grupo considerado, empezando por los de carácter social y siguiendo por los de naturaleza lingüística.

CUADRO 2. Resultados generales (solo informantes)

| <i>Grupo</i>            | <i>Factor</i> | <i>F/N</i> | <i>%</i> |
|-------------------------|---------------|------------|----------|
| Nivel de estudios       | Bajo          | 26/177     | 14.7     |
|                         | Medio         | 10/174     | 5.7      |
|                         | Alto          | 11/210     | 5.2      |
| Edad                    | 35-54         | 25/199     | 12.6     |
|                         | 20-34         | 9/116      | 7.8      |
|                         | 55-           | 13/246     | 5.3      |
| Género                  | Mujeres       | 21/230     | 9.1      |
|                         | Hombres       | 26/331     | 7.9      |
| Persona verbal          | 3a. plural    | 43/43      | 100.0    |
|                         | 1a. plural    | 4/4        | 100.0    |
|                         | 3a. singular  | 0/514      | 0.0      |
| Estructura verbal       | Perífrasis    | 6/51       | 11.8     |
|                         | Verbo         | 41/510     | 8.0      |
| Humanidad del argumento | Humano        | 14/132     | 10.6     |
|                         | No humano     | 33/429     | 7.7      |
| Definitud del argumento | Definido      | 14/95      | 14.7     |
|                         | Indefinido    | 33/433     | 7.1      |
| Forma del argumento     | Sustantivo    | 44/491     | 9.0      |
|                         | Implicito     | 2/42       | 4.8      |
|                         | Pronombre     | 1/28       | 3.6      |



|   |                           |        |      |
|---|---------------------------|--------|------|
| Posición del argumento                        | Pospuesto                 | 44/484 | 9.1  |
|   | Elidido                   | 2/42   | 4.8  |
|   | Antepuesto                | 1/35   | 2.9  |
| Marcas de pluralidad en el argumento          | N escueto                 | 18/198 | 9.1  |
|   | N con adjuntos            | 28/348 | 8.0  |
|   | N colectivo               | 1/15   | 6.7  |
| Tiempo (y modo) verbal                        | Préterito (S)             | 2/10   | 20.0 |
|   | Presente (I)              | 4/25   | 16.0 |
|   | Presente (S)              | 2/13   | 15.4 |
|   | Copretérito (I)           | 34/379 | 9.0  |
|   | Futuro perif. (I)         | 1/15   | 6.7  |
|   | Antepresente (I)          | 2/35   | 5.7  |
|   | Préterito (I)             | 2/73   | 2.7  |
|   | Antepretérito (S)         | 0/1    | 0.0  |
|   | Antecopretérito (I)       | 0/2    | 0.0  |
|   | Futuro morf. (I)          | 0/2    | 0.0  |
|   | Pospréterito (I)          | 0/6    | 0.0  |
| Presencia de elementos entre V y el argumento | Ausencia                  | 40/462 | 8.7  |
|   | Presencia                 | 7/99   | 7.1  |
| Tipo de oración                               | Coordinada                | 14/127 | 11.0 |
|   | Subordinada (no rel.)     | 12/130 | 9.2  |
|   | Principal o independiente | 19/273 | 7.0  |
|   | Subordinada relativa      | 2/31   | 6.5  |
| Polaridad                                     | Afirmativa                | 43/461 | 9.3  |
|   | Negativa                  | 4/100  | 4.0  |
| Tema  | Cotidiano                 | 43/461 | 9.3  |
|   | Técnico                   | 4/100  | 4.0  |
| Tipo de discurso                              | Descriptivo               | 17/124 | 13.7 |
|   | Argumentativo             | 3/31   | 9.7  |
|   | Expositivo                | 8/111  | 7.2  |
|   | Narrativo                 | 17/239 | 7.1  |
|   | Dialógico                 | 2/56   | 3.6  |
| Total=  |                           | 47/561 | 8.4  |

Por nivel de estudios, las personas con menor instrucción presentan casi tres veces más casos que los hablantes con instrucción media y alta, lo que sugiere que se trata de variación desde abajo. Obsérvese que, frente a la expectativa

generada por una de nuestras hipótesis de trabajo (*supra*), la pluralización no es empujada por las personas de nivel medio de estudios, sino por las personas de nivel bajo. Por edades las diferencias no son tan pronunciadas, siendo los individuos de edad intermedia quienes muestran una proporción mayor de pluralizaciones, aspecto que dificultaría una interpretación en tiempo aparente, pues los más jóvenes (de 20 a 34 años) no son quienes más impulsan la variante marcada. De manera leve, las mujeres van por delante de los hombres.

La persona verbal no es en realidad una variable, pues lógicamente todos los casos de 3a. singular aparecen sin pluralización. Lo interesante, en cualquier caso, es contabilizar los ejemplos de 3a. de plural frente a los de 1a. de plural, que solo son 4 y requerirán de un comentario específico *infra*. En lo que toca a la oposición entre formas verbales perifrásticas y no perifrásticas predominan con mucho las segundas, pero las primeras asignan una pequeña proporción mayor de casos plurales. Lo mismo puede decirse del rasgo humano: hay bastantes más casos de argumentos no humanos, pero los humanos muestran un porcentaje mayor. Muy clara es la división entre argumentos definidos e indefinidos, pues los primeros exhiben el doble de pluralizaciones. En cuanto a la forma del argumento, es claro que los sustantivos concentran mayor proporción de plurales. Por su parte, la posición argumental pospuesta no solo es la más preferida, sino también la que otorga un porcentaje mayor a la variante en juego. Las diferencias entre nombres escuetos, nombres con modificadores y nombres colectivos no son muy marcadas, pero se establecen precisamente en ese orden. Los pretéritos de subjuntivo y los presentes de indicativo y subjuntivo son los que presentan proporciones más abultadas de plural, pero sobre un total de casos relativamente pequeño. El copretérito de indicativo, con mucho el tiempo más común en los datos, muestra un 9% de pluralización, cercano al promedio de 8.4% de la muestra. El futuro perifrástico, el antepresente y el pretérito muestran porcentajes bajos y, finalmente, los tiempos menos comunes no presentan casos de *haber* plural. A su vez, la ausencia y la presencia de elementos entre el verbo y su argumento muestran proporciones no muy distantes. Y los tipos oracionales más propensos a la pluralización son las oraciones coordinadas y las subordinadas no relativas. La polaridad afirmativa dobla y más el porcentaje de pluralizaciones, y otro tanto ocurre con los temas de carácter cotidiano frente a los técnicos. Para terminar, son las descripciones y las argumentaciones los tipos de discurso en que más se alberga el *haber* en plural.

Es decir, si se hiciera un retrato descriptivo de la pluralización de *haber* en los casos de variación posible, éste se vería como en (4):

- (4) Estudios bajos, edad intermedia, mujeres; 3a. persona de plural, perífrasis verbales, argumento humano y definido, de carácter sustantivo, pospuesto y escueto, verbos en pretérito de subjuntivo, presentes de ambos modos y copretérito de indicativo, ausencia de elementos intermedios, coordinación oracional y subordinación no relativa, aseveraciones afirmativas, temas cotidianos desarrollados en descripciones y argumentaciones.

Por supuesto, las diferencias entre factores son pequeñas en muchos casos y es necesario establecer si son significativas o no, como se hace a continuación en el cuadro 3, en el que aparecen los grupos que fueron seleccionados en un cálculo de regresión logística escalonada<sup>3</sup>. Considérese, con todo, que un problema tradicional con esta variable sintáctica es que su baja ocurrencia dificulta establecer con claridad la importancia de algunos grupos de factores.

CUADRO 3. Grupos de factores significativos en la pluralización de *haber* (solo informantes)<sup>4</sup>

| <i>Factores</i>             | <i>F/N</i> | <i>%</i> | <i>P<sup>4</sup></i> |
|-----------------------------|------------|----------|----------------------|
| <i>a) Nivel de estudios</i> |            |          |                      |
| Bajo                        | 26/177     | 14.7     | 0.763                |
| Medio                       | 10/174     | 5.7      | 0.386                |
| Alto                        | 11/210     | 5.2      | 0.355                |
| <i>Rango= 41</i>            |            |          |                      |

<sup>3</sup> El cálculo se llevó a cabo con Goldvarb Lion (Sankoff, Tagliamonte y Smith 2012). Como es claro para el lector familiarizado con el uso de los programas VARBRUL, el grupo de factores dependiente toma como factor de aplicación la pluralización de *haber*, y como valor de no aplicación todos los demás casos. Los grupos de factores independientes, referidos al nivel de estudios, la edad, la definitud y el género, fueron seleccionados como significativos en la regresión logística escalonada llevada a cabo; la razón de verosimilitud y la significación del modelo se anotan al pie del cuadro, así como el *input*. Los grupos de factores independientes se jerarquizan por orden de importancia de arriba hacia abajo, según el rango alcanzado (el rango es la diferencia entre la probabilidad o peso logístico mayor y el menor, según la serie de factores de cada grupo de factores, también jerarquizada según si se favorece o no la aplicación del factor dependiente en juego), que coincide con el orden de selección en el modelo escalonado. “F” es frecuencia absoluta y “P” la probabilidad logística escalonada. Una probabilidad superior a 0.500 favorece el factor de aplicación, y una probabilidad inferior a 0.500 no lo favorece.

<sup>4</sup> Ofrecemos las probabilidades con tres decimales, aunque estamos conscientes de que para el tipo de discusión que se emprende bastaría con dos. Por otro lado, anotamos el “0” a la izquierda del punto decimal, como consideramos que es conveniente hacer en los textos redactados en español, es decir, “0.763” y no “.763”.

| <i>b) Edad</i>                    |        |      |       |
|-----------------------------------|--------|------|-------|
| 35-54                             | 25/199 | 12.6 | 0.680 |
| 20-34                             | 9/116  | 7.8  | 0.564 |
| 55-                               | 13/246 | 5.3  | 0.325 |
| <i>Rango= 36</i>                  |        |      |       |
| <i>c) Definitud del argumento</i> |        |      |       |
| Definido                          | 14/95  | 14.7 | 0.686 |
| Indefinido                        | 33/466 | 7.1  | 0.460 |
| <i>Rango= 23</i>                  |        |      |       |
| <i>d) Género</i>                  |        |      |       |
| Mujeres                           | 21/230 | 9.1  | 0.614 |
| Hombres                           | 26/331 | 7.9  | 0.420 |
| <i>Rango= 19</i>                  |        |      |       |

*Input*= 0.059, razón de verosimilitud= -143.715, sign.= 0.028

El cuadro 3 establece qué grupos de factores son significativos en el contexto del estudio realizado<sup>5</sup>. Es interesante que entre los cuatro grupos seleccionados, tres de ellos sean de orden social, dos de ellos jerarquizados en primer y segundo lugar de importancia: el nivel de estudios y la edad. Obsérvese que son las personas de nivel bajo de estudios y los menores de 55 años quienes favorecen en algún grado la pluralización, al mostrar probabilidades logísticas superiores a 0.500. Como se ha dicho, que sean las personas de bajos estudios, y no las de estudios medios, quienes favorecen el cumplimiento del proceso, contraviene un aspecto de las hipótesis que nos habíamos trazado, pues los grupos intermedios no patrocinan el avance de la nueva forma, como ocurre en muchos casos de cambio lingüístico. Tampoco la edad apoya muy claramente la idea de un cambio en curso, pues aunque los menores de 55 favorecen en distintos grados la pluralización, frente a las personas de más edad ( $p= 0.325$ ), lo cierto es que los más jóvenes de todos obtienen una  $p= 0.564$ , frente al mayor empuje de las personas de edad intermedia, con  $p= 0.680$ , lo que en conjunto dificulta una interpretación simple del tiempo aparente como cambio lingüístico. En otras palabras, aunque los datos dejan resquicios para seguir trabajando al respecto, el hecho es que en principio expresan más estabilidad que cambio.

<sup>5</sup> La persona verbal no se consideró en este cálculo, y el tiempo y modo verbal se simplificó en tres factores, debido a la dispersión de los datos: copretérito, pretérito de indicativo y el resto de los casos.

El único grupo lingüístico significativo es la definitud, que aparece en el tercer lugar jerárquico, con los argumentos definidos como favorecedores de la pluralización. Esto es muy importante, pues al no avalarse estadísticamente la solvencia de muchos otros grupos de factores lingüísticos, se sugiere un papel no muy activo de la difusión contextual, frente a lo previsto *supra* en un escenario de cambio, volviendo a avalar la perspectiva de estabilidad. De todos modos, como se vio en el cuadro 2, ello no significa, en términos descriptivos, que todos los factores propuestos en las variables produzcan exactamente los mismos resultados; considérese también la pequeña documentación de pluralizaciones, pese al tamaño del corpus examinado, lo que dificulta el surgimiento estadístico de efectos llamativos.

Por fin, el género también resultó incluido en el modelo, pese a las pequeñas diferencias porcentuales, con las mujeres a la cabeza del *haber* plural.

## DISCUSIÓN

Para entender el estatus de los resultados obtenidos, es necesario considerarlos en la perspectiva de otros trabajos, tal como se presenta en el cuadro 4. La primera columna consigna la referencia del estudio, la segunda la variedad considerada, la tercera el método empleado y la cuarta el índice porcentual de pluralización obtenido.

CUADRO 4. Resultados totales sobre la pluralización de *haber* en diversos trabajos

| <i>Trabajo</i>          | <i>Lugar</i>    | <i>Método</i>                                       | <i>Índice</i>   |
|-------------------------|-----------------|---|-----------------|
| Blas Arroyo (1995-1996) | Valencia (Esp.) | Prueba de interfer. (cat.-habs.)                    | 83.4%           |
| Freites (2008)          | Táchira (Ven.)  | Entrevistas   | 82.0%           |
| D'Aquino (2004)         | Caracas (Ven.)  | Entrevistas (160 ents., 706 datos)                  | 66.0%           |
| De Mello (1991/1994)    | La Paz (Bol.)   | Habla culta (83 datos)                              | 60.2%           |
| Díaz-Campos (2003)      | Caracas (Ven.)  | Entrevistas (96 ents.)                              | 54.0%           |
| Castillo-Trelles (2007) | Mérida (Méx.)   | Entrevistas-tareas y cuestionario (160 y 936 datos) | 53.0%,<br>55.0% |

|                             |                     |                                       |       |
|-----------------------------|---------------------|---------------------------------------|-------|
| Bentivoglio y Sedano (1989) | Caracas (Ven.)      | Entrevistas/habla culta (169 datos)   | 47.9% |
| Claes (2014a)               | Sto. Domingo (R.D.) | Entrevistas/cuestionario (1841 datos) | 46.7% |
| Gómez Molina (2013)         | Valencia (Esp.)     | Entrevistas (72 ents.) (275 datos)    | 46.2% |
| Claes (2014a)               | La Habana (Cub.)    | Entrevistas/cuestionario (2093 datos) | 44.6% |
| Soler (2012)                | Santiago (Chi.)     | Habla culta (≈30 ents.) (57 datos)    | 42.1% |
| Claes (2014a)               | San Juan (P.R.)     | Entrevistas/cuestionario (1655 datos) | 41.3% |
| De Mello (1991/1994)        | Lima (Per.)         | Habla culta (104 datos)               | 40.4% |
| Soler (2012)                | Caracas (Ven.)      | Habla culta (≈30 ents.) (91 datos)    | 39.6% |
| Gómez Molina (2013)         | Valencia (Esp.)     | Prueba de aceptabilidad               | 39.0% |
| De Mello (1991/1994)        | Santiago (Chi.)     | Habla culta (132 datos)               | 38.6% |
| Montes de Oca (1994)        | México (Méx.)       | Cuestionario (nivel alto)             | 36.0% |
| De Mello (1991/1994)        | Caracas (Ven.)      | Habla culta (153 datos)               | 35.9% |
| De Mello (1991/1994)        | San Juan (P.R.)     | Habla culta (95 datos)                | 30.5% |
| De Mello (1991/1994)        | La Habana (Cub.)    | Habla culta (45 datos)                | 26.7% |
| Montes de Oca (1994)        | México (Méx.)       | Cuestionario (nivel medio)            | 25.0% |
| Martín Butragueño (2015)    | Ecatepec (Méx.)     | Entrevistas (65 datos)                | 20%   |
| Soler (2012)                | Bogotá (Col.)       | Habla culta (≈30 ents.) (49 datos)    | 16.3% |
| Serrano (2015)              | México (Méx.)       | Habla popular (48 ents.) (177 datos)  | 15.8% |
| De Mello (1991/1994)        | Bogotá (Col.)       | Habla culta (127 datos)               | 15.7% |

|                                   |                     |  |      |
|-----------------------------------|---------------------|--|------|
| Soler (2012)                      | México (Méx.)       | Habla culta ( $\approx$ 30 ents.) (74 datos) | 9.5% |
| Moss (1990)                       | México (Méx.)       | Habla culta, popular y novelas (195 datos)   | 8.7% |
| Lastra y Martín B. (este trabajo) | México (Méx.)       | Entrevistas CSCM (108 ents.) (561 datos)     | 8.4% |
| Serrano (2015)                    | México (Méx.)       | Habla culta (56 ents.) (153 datos)           | 7.8% |
| De Mello (1991/1994)              | México (Méx.)       | Habla culta (92 datos)                       | 7.6% |
| Soler (2012)                      | Buenos Aires (Arg.) | Habla culta ( $\approx$ 30 ents.) (41 datos) | 7.3% |
| De Mello (1991/1994)              | Buenos Aires (Arg.) | Habla culta (82 datos)                       | 3.7% |
| De Mello (1991/1994)              | Madrid (Esp.)       | Habla culta (97 datos)                       | 0.0% |
| Soler (2012)                      | Madrid (Esp.)       | Habla culta (12 ents.) (38 datos)            | 0.0% |
| De Mello (1991/1994)              | Sevilla (Esp.)      | Habla culta (28 datos)                       | 0.0% |

Aunque hay observaciones básicamente cualitativas anteriores, el trabajo con el que arranca el estudio moderno sobre la pluralización de haber es el de Bentivoglio y Sedano de 1989 sobre Caracas. En los últimos veinticinco años se ha realizado una gran cantidad de análisis, muchos de ellos siguiendo o cuando menos considerando la metodología de este primer trabajo. Sin ningún ánimo de exhaustividad, el cuadro 4 muestra diferentes datos y resultados generales de un número amplio de países hispanohablantes y de zonas dialectales. Entre los que han empleado entrevistas, debe distinguirse entre los que analizan materiales del habla culta (muchas veces de una sincronía bastante anterior a la del estudio) y los que emplean entrevistas más claramente sociolingüísticas. Tampoco las pruebas y cuestionarios, cuando se han aplicado, son iguales, por lo que el cuadro encierra una cierta simplificación sobre la información disponible<sup>6</sup>.

<sup>6</sup> Un análisis completamente detallado de los métodos empleados y de todos los matices y variables presentes en los estudios disponibles va más allá de los límites de este artículo; nos

Otro aspecto que debe considerarse con precaución a la hora de interpretar los materiales es la cantidad de datos (desde 28 hasta 2093), así como las proporciones, pues cuando se emplean de manera combinada entrevistas y cuestionarios, los segundos suelen aportar mucha mayor cantidad de pluralizaciones, quizá por focalizar los casos más relevantes, quizá por el diferente tipo de tarea que se pide al hablante. Ciertamente, cabe también esperar mayor estabilidad en general en los resultados de los estudios apoyados en mayor número de datos relevantes.

Dicho todo esto, las diferencias documentadas en el índice de pluralización abarcan casi todas las posibilidades, desde 0.0% a 83.4%. Los casos más elevados se localizan en Valencia (España), Táchira y Caracas (Venezuela), La Paz (Bolivia) y Mérida (México). Intermedios son, por ejemplo, Santiago de Chile, San Juan de Puerto Rico o La Habana. Entre los más conservadores se cuentan, en América, Bogotá, la Ciudad de México y Buenos Aires, y en Europa, Madrid y Sevilla.

Debe subrayarse la tendencia, ya comentada, a que los trabajos realizados con pruebas muestren índices más elevados. Así, en la Ciudad de México, Montes de Oca (1994) menciona un 36% para las personas de nivel socioeducativo alto, y un 25% para las personas de nivel medio que, aunque están entre los índices más bajos para pruebas en el cuadro 4, se encuentran muy por encima de los estudios realizados total o mayormente con entrevistas: 9.5% en Soler (2012), 8.7% en Moss (1990), 8.4% en nuestro propio trabajo y 7.6% en De Mello (1991/1994). Serrano (2015) trabaja con un total de 104 entrevistas de habla culta y popular y encuentra un promedio de 12.0% (41/343 casos), con 7.8% en habla culta y 15.8% en habla popular.

Los materiales de habla culta (c. 1970) empleados por Soler, De Mello y Serrano producen índices muy semejantes a los encontrados de modo global en nuestro trabajo con los datos del CSCM para el promedio de los datos de todos los niveles educativos. Esto no avala la idea de difusión en tiempo real; por el contrario, sugieren mayormente estabilidad o incluso retracción en la presencia del fenómeno. La diferente naturaleza de los materiales de hace unos 35 años y la significación de la edad en nuestro estudio, con los promedios más moderados en las personas de más edad, podrían sin embargo permitir ciertas dudas sobre la estabilidad del fenómeno a través de las últimas décadas. Por otra parte, si se comparan los datos de Soler, basados en los materiales de la norma culta de c. 1970, con 9.5%, con nuestros

---

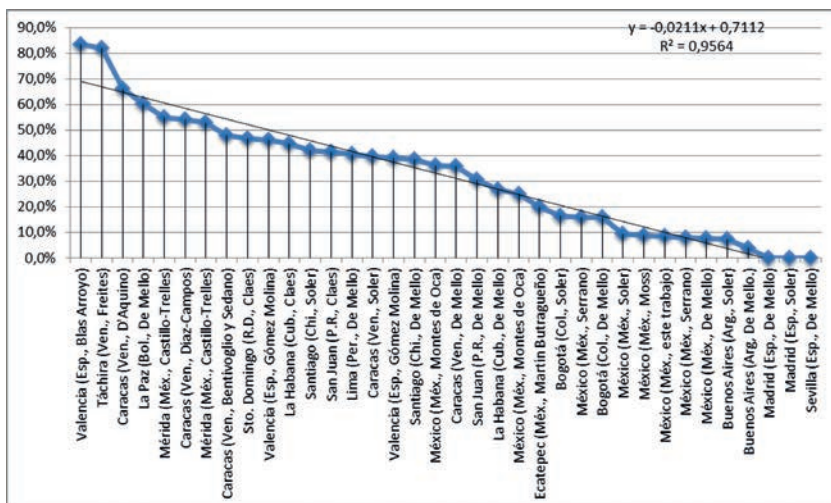
limitamos a resaltar lo que ahora nos ha parecido más relevante para establecer la comparación con nuestros propios datos.



materiales altos de c. 2005, éstos más actuales apenas llegan al 5.2%, lo que sugeriría en todo caso una retracción entre las personas de mayor nivel de estudios. Los datos de Serrano van en el mismo sentido<sup>7</sup>: su 12.0% global es ciertamente mayor que nuestro 8.4% global; el 15.8% del habla popular de los años setenta queda ahora en 14.7% en el nivel educativo bajo, y el habla culta de la generación anterior, que ofrecía un 7.8%, es apenas de 5.2% en nuestros propios datos de las personas de estudios altos. Sería tentador decir que quizá hay un patrón combinado: retracción en grupos educativos altos (y medios, con 5.7%) y mantenimiento en grupos educativos bajos, pero las pequeñas diferencias en los datos aconsejan interpretaciones cautas.

La figura 1 realiza una proyección somera de los índices expuestos en el cuadro 4.

FIGURA 1. *Esbozo geolingüístico del fenómeno*



La figura 1 proyecta de más a menos los trabajos mencionados. Llama la atención la fuerte linealidad de la curva, deducible del elevado ajuste mostrado por la  $R^2$  de la ecuación lineal, de 0.96, lo que implica una gradualidad en la distribución de los datos. Aunque la curva podría afinarse más para convertirse en una verdadera representación en geografía aparente

<sup>7</sup> Serrano (2015) sugiere de hecho que entre sus datos y los nuestros media un proceso de retracción.

(en el sentido de los diferentes testigos geográficos expuestos al considerar diversas variedades lingüísticas de un mismo proceso, sintáctico en este caso), para empezar separando los estudios según la naturaleza de los datos, llama la inversión de la concavidad hacia su parte media, produciendo un efecto general en forma de “S”, tal como suele ocurrir en diferentes situaciones de cambio lingüístico. Es claro, por otra parte, que la inferencia geográfica está lejos de ser inmediata, dado el papel desempeñado por ciertos factores, como el bilingüismo con el catalán en el caso de ciudades como Valencia, España. Fuera de ello, parece poder apuntarse una distribución general que separa el área caribeña, insular y continental, del resto del dominio hispanohablante, tal como sucede con otras variables fónicas y sintácticas (valga el caso los índices de expresión de sujetos pronominales, cf. Orozco, Shin y Carvalho 2015). Pero aun este esbozo necesitaría introducir más elementos al modelo, dado el avance del proceso en ciudades como La Paz, en Bolivia.

Tan importante como señalar las diferencias entre variedades (y entre diferentes métodos para la obtención de datos) es establecer los factores que influyen en la distribución del fenómeno. El cuadro 5 anota en la columna de la derecha los factores resaltados por algunos de los varios análisis. La columna de la izquierda anota el trabajo y la fuente de datos, así como si solo se emplea estadística descriptiva (frecuencias, f) o si se utiliza un análisis de probabilidades (p), caso este último en que se puede confiar con más claridad en la significación de los factores señalados.

CUADRO 5. Variables señaladas como significativas en la pluralización de *haber*

| <i>Investigación</i>                         | <i>Variables significativas</i>   |
|--|---|
| Lastra y Martín B. (este trabajo) (p)        | Nivel de estudios (bajo), edad (adultos y jóvenes), definitud del argumento (+definidos), género (mujeres)  |
| Gómez Molina (2013), entrevistas (p)         | Edad (jóvenes, adultos), nivel sociocultural (medio y bajo)   |
| Gómez Molina (2013), prueba de aceptabilidad | Carácter de la FN, Forma léxica de la FN, Posición de la FN, Tiempo Verbal  |
| De Mello (1991/1994), habla culta (f)        | Copretérito, argumento humano (Caracas, San Juan, Santiago, Bogotá), presencia de un verbo auxiliar, habla popular (México, especialmente de <i>habemos</i> , cf. 1991, p. 467) |

|  |  |
|--|--|
| Soler (2012), habla culta (f)                                  | Copretérito (todas las muestras), tiempos compuestos y perífrasis (Caracas, Buenos Aires, Santiago), anteposición (Caracas, Bogotá), inanimados (pero humanos en Buenos Aires y Santiago), hombres (Buenos Aires, Caracas), jóvenes (Bogotá, Caracas), carácter coloquial (Caracas) [se esperaba aumento en datos más recientes y expansión a la lengua escrita] |
| Freites (2008), entrevistas (p)                                | Argumento humano (p= .70), reforzamiento de la pluralidad  |
| Castillo-Trelles (2007), entrevistas-tareas y cuestionario (p) | Para la entrevista, sexo femenino y ausencia de cuantificadores<br>Para el cuestionario, sexo femenino, bilingüismo, jóvenes y jubilados, FN preverbal, FN +humano   |
| Díaz-Campos (2003), entrevistas (p)                            | Copretérito (p= .55) vs. pretérito (p= .16), nivel bajo  |
| D'Aquino (2004), entrevistas (p)                               | Nivel bajo (y medio), jóvenes (más que los mayores de 60), hombres (cambio en curso en su etapa final)   |
| Blas Arroyo (1995-1996)  | Catalán como lengua materna, catalán como lengua habitual, niveles socioeconómicos bajos, hablantes mayores  |
| Rivas y Brown (2012), entrevistas (p)                          | Referente humano, copretérito, entidades no delimitadas temporalmente  |
| Claes (2014a), entrevistas/ cuestionario (p)                   | Tiempo verbal, <i>priming</i> de la producción a la producción, <i>priming</i> de la comprensión a la producción, posición típica del referente de la FN en la cadena, clase social (baja)   |

Los modelos propuestos para diferentes comunidades de habla suelen ser ricos en factores sociales. Además de en nuestro estudio, el nivel socioeducativo bajo es mencionado para Valencia por Blas Arroyo (1995-1996) y Gómez Molina (2013), para Caracas por Díaz-Campos (2003) y D'Aquino (2004), para el Caribe por Claes (2014a), entre otros, aunque a veces también el nivel medio favorece y en otros casos no se ha demostrado incidencia según este grupo de factores. Los datos de Ecatepec, una zona en general de clase baja en el área metropolitana de la Ciudad de México, llevan el índice de pluralización de *haber* a un 20% (Martín Butragueño 2015).

En lo que toca a la edad, los grupos adultos y jóvenes señalados ahora para la Ciudad de México son también los favorecedores en Valencia según

Gómez Molina (2013), cuyas entrevistas son estudiadas con la misma metodología que nuestro estudio, así como según D'Aquino (2004) en Caracas. La edad tiene, en cualquier caso, una proyección menor que el nivel de estudios y pone en duda la difusión cuando menos general del proceso en tiempo aparente, aunque puede ser argumento en algunas ciudades muy concretas. Las mujeres como favorecedoras en la Ciudad de México están en coincidencia con el trabajo de Castillo-Trelles sobre Mérida (México), pero a veces se menciona a los hombres, como en Caracas (por D'Aquino 2004 y Soler 2012), aunque el género no parezca relevante en buena parte de los lugares.

Varios trabajos (aunque no el nuestro) señalan la relevancia del carácter humano de la FN<sup>8</sup>. Según Díaz-Campos (2003, p. 10 y 2014, pp. 170-172), una forma de probar que los hablantes analizan como sujeto el argumento presente es estudiar el tipo de referente,  $\pm$ humano, de la FN argumental, dado que los referentes +humanos tienden a ser tópicos, agentes y sujetos. Díaz-Campos menciona a Silva-Corvalán (1983, p. 119) para recordar que la definitud y la humanidad tienden a correlacionar con la topicidad, de modo que esos dos estarían entre los rasgos de los tópicos, que tienden a codificarse como FN sujetos. Así, Bentivoglio y Sedano (1989, pp. 73-74) muestran un 67% de FN humanas con *haber* pluralizado en Caracas, y De Mello (1991, 1994) considera el efecto del rasgo  $\pm$ humano para la forma *habían* en las 11 ciudades de su estudio, de modo que la pluralización con argumento humano es mayor en Caracas, Santiago y San Juan, aunque no se puede argumentar lo mismo en todas las demás situaciones. Para Díaz-Campos, la muestra de De Mello es pequeña y limitada al nivel culto, por lo que es difícil hacer generalizaciones y la naturaleza del referente podría no ser crítica. También nosotros estamos de acuerdo con su crítica<sup>9</sup>. Es más, en nuestros datos, el patrón se invierte incluso: en el caso particular de *habían*, solo el 38.2% (13/34) de los argumentos son humanos, mientras que 61.8% (21/34) son no humanos.

Sin embargo, lo que sí es significativo en nuestros materiales es el papel de la definitud: los argumentos definidos favorecen la pluralización con una probabilidad de 0.686 (frente al 0.460 de los indefinidos). Dado que

<sup>8</sup> Aunque, como mostró el cuadro 2, en términos puramente descriptivos, nuestro corpus de datos sí expuso una proporción mayor con argumentos humanos (10.6%) que con argumentos no humanos (7.7%).

<sup>9</sup> Los datos apuntados por De Mello (1991, p. 467, cuadro 5) acerca del nivel sociocultural en la Ciudad de México son muy pocos: apunta 3 *habían* en "culto" y 3 en "inculto", 1 *hubieron* culto por ninguno inculto, 2 y 2 *habíamos*, y ningún *habemos* culto por 5 incultos.

la definitud sería también una característica frecuentemente asociada a la topicidad, una forma de conciliar nuestros resultados con las observaciones de Díaz-Campos es establecer que la pluralización no es favorecida *per se* por el rasgo +humano, sino en general por las características habituales de la topicidad de la FN, sea que se subraye la humanidad en unos casos, o la definitud en otros<sup>10</sup>.

Los ejemplos de *habemos* merecen un comentario aparte. Como señala Hernández Díaz (2006), aquí la concordancia involucra al número y la persona, y algunos autores (como Claes 2014a, p. 347) prefieren considerarlo un caso diferente. En nuestros datos, las cuatro ocurrencias pertinentes tienen distribución social variada: uno es de nivel bajo (*Habíamos tres o cuatro personas ahí*), dos son de nivel medio (*Habíamos diferentes/ de diferentes países; Habíamos puras mujeres*) y uno de nivel alto (*Habíamos cuatro personas de Zaragoza*). Debe considerarse que pueden oírse casos, sobre todo entre las personas de ciertos estudios, donde se emplea de modo irónico.

Una consideración novedosa es el efecto de *priming* señalado por Claes (2014a), sea de producción a producción, o de comprensión a producción. En nuestros propios datos, la baja densidad de pluralizaciones en los materiales de la Ciudad de México hace dudar de que este efecto mecánico pueda ser en realidad de mucha importancia, considerando la presencia de solo 51 pluralizaciones en 3500 páginas de transcripciones. Existen además algunos ejemplos contraproducentes en ese sentido, como (5), donde la secuencia no impide la alternancia (y donde en realidad parece haber un sentido de tamaño que favorece la presencia del plural, frente al singular, donde los objetos mayores se pluralizan, y los de menor tamaño se mantienen en singular):

- (5) **había unos huesitos/** que son de chabacano así [...] pero esos/ los pintaban de colores <~colores:> [...] **habían unos huesotes grandes** que le (sic) llamaban mascotas/ tú parabas esa mascota/ en la pared/ así [ent. 90, turno 1017, 12H] (alternancia).

<sup>10</sup> El prof. Claes nos señala que “se ha observado que los SSNN definidos que aparecen con verbos existenciales deben comunicar información nueva con respecto al conocimiento/ las creencias del oyente [...]. Tal también es el caso para *haber* existencial (Claes 2014b: cap. 6.2.2). En otras palabras, si bien se observan SSNN definidos con *haber* existencial, estos no pueden funcionar [como] tópicos” (c.p.). La observación es muy interesante y supera en realidad el ámbito de discusión de este artículo. Una observación preliminar es que topicidad y carácter dado de la información son dos categorías distintas, pues de hecho existen tópicos nuevos (Krifka 2007), por lo que no es imposible seguir postulando la idea de topicidad. Para una discusión mucho más extensa del funcionamiento de los tópicos y en general de las funciones informativas, véase Martín Butragueño (en proceso, cap. 4; pueden verse borradores en <http://www.academia.edu>).

El reforzamiento de los elementos plurales (y en algún caso el no reforzamiento) ha resultado factor de peso en algunos otros estudios, pero en la Ciudad de México las marcas de pluralidad no han resultado significativas desde un punto de vista probabilístico. Es más, en términos de frecuencia, como se ha dicho, el nombre escueto (frente a los casos con diferentes adjuntos y frente a los colectivos) es el que documenta un número levemente mayor de casos.

Para Bentivoglio y Sedano, el proceso de pluralización no se completa por tres razones: *a*) la presión de la educación formal; *b*) la rareza de la pluralización de *hay*, que es la forma más frecuente; *c*) la cantidad de material fónico en *habemos* o *hubieron* facilita su identificación como ajenas a la norma por parte de las personas con más estudios, al menos en algunas áreas (2011, p. 174). Aunque los tres aspectos son válidos en los datos de la Ciudad de México, la perspectiva no parece suficiente para explicar la estabilidad del fenómeno, en especial considerando la aparente baja conciencia por parte de los hablantes (a reserva de un estudio detenido) y su presencia en ciertos discursos en principio formales, como discursos políticos (la pluralización se escucha incluso en presidentes y expresidentes mexicanos).

En cuanto a los pasos del proceso propuestos por Hernández Díaz (2006), antes mencionados, nuestros datos no apoyan plenamente la perspectiva. Aunque la Ciudad de México se encuadraría en las etapas relativamente tempranas de variación, la presencia de verbos auxiliares –la posible primera etapa– no es significativa en la pluralización (aunque sí es un poco más frecuente, 11.8% en las perífrasis, frente a 8.0% cuando no hay verbo auxiliar). En lo que toca a la segunda etapa (una alteración fónica menor) y la tercera (alteración fónica mayor), el tiempo/modo verbal no es significativo en el modelo cuando se reduce a copretérito vs. pretérito vs. otras formas; tampoco en la estadística descriptiva del cuadro 2 la idea tendría un apoyo claro, dado que los pretéritos son los que muestran una proporción mayor y las formas más frecuentes, los copretéritos, alcanzan un 9%, muy cercano al promedio general. Además, las formas de 1a. plural (*habemos*) tienen un alcance no muy amplio, pero a fin de cuentas presente. Por fin, en relación con el posible paso *hay* > *han*, hasta ahora no hemos encontrado ejemplos en el CSCM.

## CONCLUSIONES

En suma, con respecto a las hipótesis planteadas al principio del trabajo, el habla de la Ciudad de México se agrupa con las variedades conservadoras del español en lo que se refiere a la pluralización del verbo *haber* existencial. No puede confirmarse ni mucho menos que exista un intenso proceso de expansión del fenómeno, pues los testigos proporcionados por estudios anteriores muestran índices de pluralización no demasiado alejados. El tiempo aparente documentado a través de la estratificación por edad muestra a las personas de edad intermedia a la cabeza, lo que deja abierta la puerta a seguir explorando el sentido general de los datos. Una posible vía de trabajo futuro implica cruzar los datos en tiempo real y en tiempo aparente disponibles con la estratificación sociolingüística, pues los resultados actuales sugieren o son compatibles con la perspectiva de estabilidad del proceso<sup>11</sup> entre los hablantes de nivel socioeducativo bajo y de ligera retracción en los niveles medio y alto.

Una cuestión metodológica de peso es que es necesario disponer de estudios realizados estrictamente bajo los mismos supuestos, para poder llevar las comparaciones hasta sus últimas consecuencias. En ese sentido, este trabajo sigue los lineamientos de PRESEEA, ya aplicados en el análisis de las entrevistas que discute Gómez Molina (2013) para Valencia (España).

Los datos mexicanos, en cualquier caso, muestran gran riqueza social, de modo que no solo la edad, sino también el nivel socioeducativo y el género resultan significativos.

Por fin, la significación de la definitud del argumento puede sumarse a otras observaciones sobre la significación del carácter humano de tal FN, proponiendo que es quizá la administración de la topicidad en general lo que promueve el reanálisis como sujeto en vez de como objeto del argumento que acompaña al verbo existencial<sup>12</sup>.

<sup>11</sup> O incluso expansión, si se consideran los datos de Ecatepec (Martín Butragueño 2015).

<sup>12</sup> Como nos señala el prof. Claes (c.p.), existe una explicación alternativa: “la agentividad, la referencia humana, y la definitud de por sí son rasgos de los sujetos prototípicos [...]. Por lo tanto, sus resultados son muy favorables a mi hipótesis de que el factor cognitivo general conocido como la «marcadez de la codificación» hace que los hablantes seleccionen la variante pluralizada más a menudo con SSNN que se parecen más a un sujeto prototípico”. La observación en sí nos parece muy interesante, pues es compatible con los datos de que disponemos; de hecho, conduce al mismo lugar, pues el carácter topical es un correlato esperable en los sujetos (véase también la nota 10 *supra* sobre el carácter y las propiedades de los tópicos, que también pueden comunicar información nueva). No vamos a abundar más

## REFERENCIAS BIBLIOGRÁFICAS

- BENTIVOGGIO, PAOLA, y MERCEDES SEDANO. 1989. *Haber: ¿un verbo impersonal. Estudios sobre el español de América y lingüística afro-americana. Ponencias presentadas en el 45 congreso internacional de americanistas*, pp. 59-81. Bogotá: Instituto Caro y Cuervo.
- \_\_\_\_\_. 2011. Morphosyntactic variation in Spanish-speaking Latin America. *The Handbook of Spanish Sociolinguistics*. Ed. Manuel Díaz Campos, pp. 168-186. Oxford: Blackwell.
- BLAS ARROYO, JOSÉ LUIS. 1995-1996. La interferencia lingüística a debate: a propósito de un caso de convergencia gramatical por causación múltiple. *Cuadernos de Investigación Filológica* 21-22: 175-200.
- BYBEE, JOAN L. 2001. *Phonology and Language Use*. Cambridge: Cambridge University Press.
- \_\_\_\_\_. 2008. Formal universals as emergent phenomena: the origins of structure preservation. *Linguistic Universals and Language Change*. Ed. Jeff Good, pp. 108-121. Oxford: Oxford University Press.
- CASTILLO-TRELLES, CAROLINA. 2007. La pluralización del verbo *haber* impersonal en el español yucateco. *Selected Proceedings of the Third Workshop on Spanish Sociolinguistics*. Ed. Jonathan Holmquist, Augusto Lorenzino y Lotfi Sayahi, pp. 74-84. Somerville, MA: Cascadilla Proceedings Project.
- CLAES, JEROEN. 2014a. Sociolingüística comparada y gramática de construcciones: un acercamiento a la pluralización de *haber* presentacional en las capitales antillanas. *Revista Española de Lingüística Aplicada / Spanish Journal of Applied Linguistics* 27, 2: 338-364.
- CLAES, JEROEN. 2014b. *The pluralization of presentational 'haber' in Caribbean Spanish. A study in Cognitive Construction Grammar and Comparative Sociolinguistics*. Tesis doctoral. Amberes: Universidad de Amberes.
- CLAES, JEROEN. c.p. 2015. [Correo electrónico con comentarios a un borrador de este trabajo, recibido el 31 de mayo de 2015].
- D' AQUINO, GIOVANNA. 2004. *Haber* impersonal en el habla de Caracas: análisis sociolingüístico. *Boletín de Lingüística* 21: 3-26.
- D' AQUINO RUIZ, GIOVANNA. 2008. El Cambio lingüístico de *haber* impersonal, *Núcleo* 25: 103-123.
- DE MELLO, GEORGE. 1991. Pluralización del verbo *haber* impersonal en español hablado culto de once ciudades. *Thesaurus* 46: 445-471.
- \_\_\_\_\_. 1994. Pluralización del verbo *haber* impersonal en el español hablado culto. *Studia Neophilologica* 66: 77-91.
- DÍAZ-CAMPOS, MANUEL. 2003. The pluralization of *haber* in Venezuelan Spanish: A sociolinguistic change in real time. *IU Working Papers in Linguistics* 3, 5: 1-13.
- \_\_\_\_\_. 2014. *Introducción a la sociolingüística hispánica*. Oxford: Wiley Blackwell.
- FREITES, FRANCISCO. 2004. Pluralización de *haber* impersonal en el Táchira: actitudes lingüísticas. *Boletín de Lingüística* 22: 32-51.
- \_\_\_\_\_. 2008. Más sobre la pluralización de *haber* impersonal en Venezuela. El Estado Táchira. *Lingua Americana* 12, 22: 36-57.

---

en el apasionante debate sobre el correlato gramatical del proceso porque no es el objetivo de este artículo (focalizado en la variación y el cambio y en la diferenciación dialectal), ni nuestros datos permiten la discusión de si el reanálisis gramatical toma como motor al sujeto sintáctico o al tópic informativo o a las características asociadas a cada uno o en realidad a ambos, siendo que la única realidad que tenemos en las manos con los datos del CSCM es la significación estadística de la definitud.



- GÓMEZ MOLINA, JOSÉ RAMÓN. 2013. Pluralización de *haber* impersonal en el español de Valencia (España). *Verba* 40: 253-284.
- HERNÁNDEZ DÍAZ, AXEL. 2006. Posesión y existencia. La competencia de haber y tener en la posesión y haber existencial. *Sintaxis histórica del español. Primera parte: La frase verbal. Volumen 2*. Dir. Concepción Company, pp. 1053-1160. México: Universidad Nacional Autónoma de México/Fondo de Cultura Económica.
- KRIFKA, MANFRED. 2007. Basic notions of information structure. *Interdisciplinary Studies on Information Structure* 6: 13-55.
- LASTRA, YOLANDA, y PEDRO MARTÍN BUTRAGUEÑO. 2015. Subject pronoun expression in oral Mexican Spanish. *Subject pronoun expression in Spanish. A Cross-Dialectal Perspective*. Ed. Ana M. Carvalho, Rafael Orozco y Naomi Lapidus Shin, pp. 39-57. Washington, DC: Georgetown University Press.
- MARTÍN BUTRAGUEÑO, PEDRO. 2015. *Inmigración y reconstrucción lingüística: el caso de Ecatepec*, ms.
- \_\_\_\_\_ en proceso. *Fonología variable del español de México*. Vol. II: *Prosodia enunciativa*. [El cap. 4 es La construcción prosódica de la información; avances en <<http://www.academia.edu>>].
- MARTÍN BUTRAGUEÑO, PEDRO y YOLANDA LASTRA (coords.). 2011. *Corpus sociolingüístico de la Ciudad de México*. Vol. I: *Materiales de PRESEEA, nivel alto*. México: El Colegio de México.
- \_\_\_\_\_ (coords.). 2012. *Corpus sociolingüístico de la Ciudad de México*. Vol. II: *Materiales de PRESEEA, nivel medio*. México: El Colegio de México.
- \_\_\_\_\_ (coords.). 2015. *Corpus sociolingüístico de la Ciudad de México*. Vol. III: *Materiales de PRESEEA, nivel bajo*. México: El Colegio de México.
- MONTES DE OCA, PILAR. 1994. La concordancia con *haber* impersonal. *Anuario de Letras* 32: 7-35.
- MONTES GIRALDO, JOSÉ J. 1982. Sobre el sintagma *haber* + sustantivo. *Thesaurus* 37: 383-385.
- MOSS, KENDALL. 1990. Personalization of the existential *haber* in Mexican Spanish. *Deseret Language and Linguistic Society Symposium* 16: 105-112.
- OROZCO, RAFAEL, NAOMI SHIN, y ANA MARIA CARVALHO (eds.). 2015. *Subject Pronoun Expression in Spanish: A Cross-dialectal Perspective*. Washington: Georgetown University Press.
- PRESEEA. 2013. Proyecto para el estudio sociolingüístico del español de España y América. Codificación de la variable “usos del verbo *haber*” (singular/plural), en <<http://preseea.linguas.net/Methodolog%C3%ADa.aspx>>.
- RIVAS, JAVIER, y ESTHER BROWN. 2012. Concordancia variable con *haber* en español puertorriqueño. *Boletín de Lingüística* 24: 102-118.
- SANKOFF, DAVID, SALI A. TAGLIAMONTE, y E. SMITH. 2012. *Goldvarb Lion. A Multivariate Analysis Application*. Toronto: University of Toronto; University of Toronto – Ottawa: University of Ottawa, en <<http://individual.utoronto.ca/tagliamonte/goldvarb.htm>>.
- SERRANO, JULIO 2015. Evolución reciente del proceso de pluralización de *haber* presentacional en español mexicano. Acercamiento sociolingüístico, comunicación presentada al *XIII Congreso Nacional de Lingüística*. Tuxtla Gutiérrez, Chiapas: Asociación Mexicana de Lingüística Aplicada - Universidad Autónoma de Chiapas, 7-10 de octubre.
- SILVA-CORVALÁN, CARMEN. 1983. On the interaction of word order and intonation: some OV constructions in Spanish. *Discourse Perspectives on Syntax*. Ed. Flora Klein, pp. 117-140. Nueva York: Academic Press.
- SOLER ARECHALDE, MARÍA ÁNGELES. 2012. Pluralización de *haber* impersonal en el habla culta de varias ciudades hispanoamericanas. Ponencia presentada en el *III Coloquio de Cambio y Variación: Geolingüística*. México: INAH - ENAH, octubre.