

UN HITO CLAVE EN LA EVOLUCIÓN DE LAS PERÍFRASIS
MODALES DE INFINITIVO: ANÁLISIS SOCIOLINGÜÍSTICO DE LA
ALTERNANCIA *HABER DE/TENER QUE* + INFINITIVO EN TEXTOS
DE INMEDIATEZ COMUNICATIVA DEL SIGLO XIX
(A Key Milestone in the Evolution of Modal Infinitival Periphrases: A
Sociolinguistic Analysis of the Alternation *haber de / tener que* in
Nineteenth-Century Texts Reflecting Spoken Language
("Communicative Immediacy"))

JOSÉ LUIS BLAS ARROYO
MARGARITA PORCAR MIRALLES
JAVIER VELLÓN LAHOZ
Universitat Jaume I

RESUMEN

En el marco de un proyecto de sociolingüística histórica para el estudio de las perífrasis modales de infinitivo en textos de inmediatez comunicativa, el presente trabajo ofrece un análisis variacionista acerca del proceso de variación y cambio que afecta a las construcciones *tener que* y *haber de* + infinitivo en el siglo XIX. La elección de esta centuria obedece al interés por analizar una etapa clave en este proceso, ya que parece tratarse de la antesala del cambio radical que, a lo largo del siglo XX, llevará a la marginación de las perífrasis con *haber* en contextos escritos y dialectales ajenos a la lengua de uso común. Contrariamente a lo que sugieren otros estudios basados en tradiciones discursivas formales, los datos de esta investigación concluyen que dicho cambio se ha activado ya con relativa intensidad en el siglo XIX, como se desprende del hecho de que *tener que* llegue a superar incluso a *haber de* en la expresión de contenidos modales, si bien esta última mantiene todavía una proporción nada desdeñable de los usos temporales de futuro (no modales), característicos de dicha perífrasis en etapas previas del español. Entre los entornos modales en que tiene lugar preferentemente dicho cambio, observamos cómo la expansión de *tener que* ocurre preferentemente en algunos contextos lingüísticos –oraciones negativas, primera persona, verbos de lengua y movimiento, así como en algunos subparadigmas de la conjugación– y extralingüísticos –cartas familiares e íntimas, jóvenes y hombres de clases subalternas–. Este último dato sugiere el posible origen de esta evolución favorable a las perífrasis con *tener* como un cambio desde abajo. Como contrapartida, un análisis dialectal muestra a los hablantes catalanes como los principales artífices del mantenimiento de *haber de*, como consecuencia de probable proceso de convergencia lingüística entre las dos lenguas de su repertorio verbal.

PALABRAS CLAVE: Variacionismo, sociolingüística histórica, cambio lingüístico, perífrasis verbales, modalidad.

ABSTRACT

As part of a historical sociolinguistic study of modal infinitival periphrases in texts reflecting spoken language, this paper presents a variationist analysis about the process of variation and change affecting *haber de* and *tener que* + infinitive in the nineteenth century. The choice of this time period reflects the fact that this century seems to be a key moment in the abrupt change that, in modern times, has reduced *haber de* to some restricted contexts such as written and formal registers as well as to some dialectal areas. Contrary to what is suggested by other studies based on more formal discursive traditions, the data from this research allow us to conclude that the acceleration of this change has already become relatively strong in the nineteenth century. Among the structural contexts that favour this change, we note some specific loci such as negative sentences, first person, verbs of speech and movement, as well as some paradigms of the verbal conjugation. Moreover, some extralinguistic factors seem to be highly relevant in this process: i.e. the more informal styles –family and intimate letters–, as well as use by young men from the lower classes. This latter finding suggests the social diffusion of *tener que* as a change from below. On the contrary, a dialectal analysis shows that the Catalan speakers played a major role in the maintenance of *haber de*, probably as a result of linguistic convergence between Spanish and Catalan.

KEY WORDS: Variationism, Historical Sociolinguistics, Language change, Verbal periphrases, Modality.

1. INTRODUCCIÓN

Los siguientes fragmentos corresponden a las cartas de amor que una poetisa hispano-cubana, Gertrudis de Avellaneda, escribía a un amigo a mediados del siglo XIX. Al margen del tono apasionado que destilan, nada fuera de lugar, por otro lado, en el periodo en que fueron escritas las cartas, desde el punto de vista lingüístico nos interesa llamar la atención acerca de la utilización alternativa de dos expresiones perifrásticas (*tener que* y *haber de* + infinitivo) en contextos que, por lo demás, son muy similares:

- (1) ...cuando me cubro la cara o hablo con quien la tiene cubierta, *tengo que decir* tú con los labios a gentes que no son nada para mi corazón (*Avellaneda: epistolario de amor y de amistad*).
- (2) ...sacar las cartas de D^a. Amadora de Almonte a alguna criada, o al mozo, lo cual quiero evitar, porque *habría de decirles* el nombre mencionado (*Avellaneda: epistolario de amor y de amistad*).
- (3) ¡Cepeda! tú eres lo que has sido [...] el más amable de los hombres y el más querido de los amigos: esto eres todavía y esto *tienes que ser* mientras yo viva (*Avellaneda: epistolario de amor y de amistad*).

- (4) ...para dar lugar a que el ingenio y el encanto del misterio obren sobre el corazón de una mujer poeta, y puedan inspirarla un afecto que *ha de ser* para ella la caja de Pandora (*Avellaneda: epistolario de amor y de amistad*).

Obsérvese, en efecto, cómo, además de escritos por la misma mano, en todos los casos las construcciones perifrásticas coinciden en la expresión de similares contenidos semánticos, relacionados con las nociones de obligación o necesidad. Asimismo, existe una coincidencia plena entre los verbos principales de la perífrasis (*decir, ser*), auxiliados en unos casos con *tener* y en otros con *haber*. Pero las equivalencias no acaban ahí, ya que, incluso en los ejemplos (1) y (2), el sujeto de las perífrasis resulta ser el mismo: la 1ª persona del singular. En suma, lo que revelan estos fragmentos es que, sin negar la posible existencia de algunos casos de incompatibilidad, como los que analizaremos más abajo (ver apartado 4.2), lo cierto es que las potenciales diferencias entre las perífrasis parecen neutralizarse a menudo en el discurso, y los hablantes acuden a ellas como variantes representativas de una misma unidad lingüística. Como es sabido, esta unidad es categorizada por la sociolingüística bajo el concepto de variable lingüística, dada la posibilidad de sus diferentes variantes de alternar en los mismos contextos lingüísticos, estilísticos y sociales, aunque no necesariamente con la misma frecuencia (Tagliamonte 2011). Precisamente, el objetivo de la sociolingüística variacionista, cuyos principios seguimos en este trabajo, es averiguar cuál es la jerarquía explicativa de esos contextos (factores) cuando todos son considerados al mismo tiempo y de la manera más objetiva posible.

Hoy sabemos que ambas perífrasis han compartido durante siglos una amplia esfera de la modalidad, y que, hasta tiempos recientes, *haber de* ha sido mayoritaria. Con todo, y como parte de un largo proceso de sustitución entre las dos formas auxiliares, *tener que* ha terminado imponiéndose sobre su competidora, hasta relegarla en la actualidad a los registros formales de la lengua escrita, así como –en menor medida– a algunas variedades dialectales en el seno de la lengua oral (Gómez Torrego 1988, 1999; Hernández García 1998; Fernández de Castro 1999; Sinner 2003; Martínez Díaz 2003, 2008; Blas Arroyo 2004, 2011; García Fernández 2006, 2013; Hernández Díaz 2006; López Izquierdo 2008).

Mucho menos es, sin embargo, lo que conocemos acerca de la evolución experimentada por aquellos factores que han contribuido a ese cambio lingüístico a lo largo de la historia del español. Por lo demás, buena parte de lo escrito acerca de este tema se ha concentrado en la cuestión semántica, y más concretamente en los matices modales (y en algún caso temporales) que caracterizan a las dos perífrasis, pero apenas sabemos nada acerca de la potencial influencia de otros factores, sean estos lingüísticos o extralingüísticos. Y sin embargo, diversos estudios anteriores acer-

ca de la variación en otros paradigmas de las perífrasis verbales de infinitivo (Blas Arroyo y Vellón 2014, Blas Arroyo y Porcar en prensa, Blas Arroyo y González 2014 a), nos han demostrado que tal influencia es más que una hipótesis, las más de las veces, incluso, por encima de la propia modalidad.

En el marco de un proyecto de sociolingüística histórica sobre textos cercanos al polo de la inmediatez comunicativa (Oesterreicher 2004), en lo que sigue ofrecemos una investigación variacionista acerca de la alternancia entre *tener que* y *haber de* + infinitivo en un estadio particularmente destacado como el siglo XIX¹. De hecho, algunos artículos que aportan datos cuantitativos habían señalado ya esta centuria como un estadio inicial, aunque todavía muy incipiente, del cambio favorable a *tener que* que tendrá lugar de manera brusca a lo largo del siglo XX (López Izquierdo 2008, Martínez Díaz 2003). Nuestro interés aquí, sin embargo, no estriba tanto en comprobar el grado de coincidencia con estas cifras, extraídas, por lo general, de corpus compuestos mayoritariamente por tradiciones discursivas formales, cuanto analizar exhaustivamente el contexto variable que envuelve tales usos, y comprobar si existen factores que condicionan de forma significativa dicha variación (incluida la modalidad). Además, la consideración conjunta de todos estos factores nos permitirá desvelar la posible interacción o la (in)dependencia entre unos y otros, misión harto difícil con un análisis univariante.

A continuación presentamos los principales detalles metodológicos de la investigación (apartado 3), como paso previo al análisis de los resultados que arroja un estudio estadístico multivariante (apartado 4). Antes, sin embargo, realizaremos una breve síntesis de los principales hitos en el proceso de cambio que ha afectado a las dos perífrasis en la historia del español.

2. *HABER DE Y TENER* + INFINITIVO: BREVE HISTORIA DE UNA ALTERNANCIA SINTÁCTICA

Los estudios sobre gramática histórica han reconocido que la irrupción de *tener* como verbo auxiliar de perífrasis modales de obligación, en distribuciones análogas a *haber*, se relaciona con el proceso de gramaticalización y neutralización semántica de ambos verbos en el periodo comprendido entre los siglos XII y XVI. Como ha señalado Hernández Díaz

¹ El presente trabajo se incluye dentro del Proyecto de Investigación "Variación y cambio lingüístico a través de textos de inmediatez comunicativa: un proyecto de sociolingüística histórica" (P1-1B2013-01) financiado por el Ministerio de Ciencia y Tecnología al equipo de investigación que dirige el primero de los autores (periodo 2011-2013). Para una revisión de los principales caracteres y objetivos de este proyecto de sociolingüística histórica, véase Blas Arroyo (en prensa).

(2006), múltiples factores y de índole diversa (semántica, pero también sintáctica y pragmática) son los que determinaron ese significativo cambio lingüístico. Sin duda, contribuyó al desgaste de *haber* lexical la generalización como auxiliar en la formación de tiempos compuestos y su extensión como verbo existencial e impersonal. No hay que olvidar, por otro lado, su tránsito también hacia la transmisión de valores modales, además de los temporales, junto con el infinitivo. A este respecto, ya Seifert (1930: 237) subrayaba las dificultades de un debilitado *haber* para mantener sus usos cuando otro verbo, *tener*, con una semántica afín apuntaba hacia una posesión más activa e intensa. Por su parte, Garachana y Rosemeyer (2011: 39) han subrayado que este proceso de sustitución representa un claro ejemplo de que el cambio en la gramática puede originarse precisamente a partir de una mera sustitución en el ámbito del léxico: una vez establecidas las bases de la equivalencia entre verbos, los hablantes no parecen diferenciar el nivel léxico del gramatical.

En lo formal, las perífrasis con *habere* no precisaban de nexo de unión entre sus elementos en latín clásico, pero en el latín tardío y en protorro-mance aparecen ya las preposiciones *a* y *de* con tal función, con el fin de matizar su sentido modal y diferenciarlo de la temporalidad futura que se estaba gestando a partir de la gramaticalización de la antigua perífrasis. Así explica Gili Gaya (1970: 112) dicho proceso:

Sabido es que se perdió en romance el futuro latino (*amabo*) y para expresar el tiempo futuro se formó la perífrasis *amar he, has, ha, hemos, heis, han*, es decir, el infinitivo seguido del presente del verbo *haber*; después se aglutinó éste a aquél (*amaré*, etc.) y así quedó formado el nuevo futuro. El tiempo venidero de una acción se expresaba por la obligación presente de realizarla: *amar he* significaba primitivamente *he de amar*. A medida que en *amaré* se fue sintiendo únicamente el valor de tiempo futuro, hubo que echar mano de un medio que significase obligación en cualquier tiempo, y para ello se inició y extendió el uso de *haber de cantar* (*he de cantar*) frente a *cantaré* (*cantar he*).

Lapesa, por su parte (2000: 882), se hace eco de los variados matices significativos de *haber de*, así como de sus usos con valor prospectivo o, simplemente, como equivalente de una forma simple: las llamadas perífrasis *pleonásticas*, en las que la perífrasis no indica obligación, solo acción (v. gr. *Cuemo lo mandó mio Çid, assi lo an todos ha far* (v. 322) = lo hacen). En síntesis, la perífrasis *haber* (*a, de*) + infinitivo tuvo tres valores principales en la época medieval: modal de obligación/necesidad; temporal de futuro y perífrasis pleonástica.

Al igual que en otras lenguas romances (v.gr. en portugués y en algunos dialectos suditálicos), durante la Baja Edad Media, como ya apuntamos más arriba, se consolidaron en castellano las perífrasis obligativas con el auxiliar *tener*, en el contexto general de la sustitución del verbo *haber* como verbo pleno. Para Yllera (1980: 111), la forma con la que aparecieron estas perífrasis refleja un proceso inverso al de las construcciones con

haber. “en el caso de *aver* con infinitivo, la fórmula más antigua fue *aver* + inf., seguido por *aver a* + inf. y, finalmente, *aver de* + inf [...] En las perífrasis formadas por *tener* e infinitivo, el orden de aparición es inverso”. *Tener de*, la primera variante perifrástica que apunta tímidamente en el siglo XIII, selecciona esta preposición por influjo de *aver de*, preferente en la segunda mitad de este siglo respecto a *aver a*². A partir del siglo XV se registran ejemplos de *tener* en perífrasis sin nexo o bien con la conjunción *que*, fórmula esta última que se irá consolidando durante las dos centurias siguientes, hasta imponerse en el español moderno.

El caso de *tener que* + infinitivo es singular entre las perífrasis personales con infinitivo, en cuanto que es la única construcción en la que el nexo no es una preposición (Pountain 2001). Olbertz (1998: 250) recuerda además que, en torno a esta combinación, se configuran dos tipos de estructuras sintácticas, como las de (5) y (6):

- (5) Juan Gabriel, acuérdate que tienes muchos problemas que repasar
 (6) Juan Gabriel, acuérdate que tienes que repasar muchos problemas

Algunos autores han demostrado que solo las cláusulas del tipo representado en el ejemplo (6), sin ningún antecedente entre *tener* y *que*, pueden considerarse verdaderas perífrasis, a diferencia de las cláusulas de relativo como la que aparece en (5), un criterio que compartimos (Gutiérrez 1980, Olbertz 1998, Sinner 2003). Sin embargo, no es descartable que el origen de (6) pudiera estar en una elipsis a partir de (5): tener (algo) que (está por) hacer (Gómez Torrego 1988, Pountain 2001).

En el siguiente cuadro, López Izquierdo (2008: 793) realiza un recuento acerca de la frecuencia de las diferentes perífrasis modales de infinitivo extraídas del *Corpus del español*³. Las cifras son reveladoras de la existencia de un cambio notable en el ámbito de estas perífrasis entre el período de orígenes y la actualidad.

TABLA 1. *Haber / tener de / que* + infinitivo: frecuencias totales
 (Fuente: López Izquierdo 2008: 793)

	XIII	XIV	XV	XVI	XVII	XVIII	XIX	XX
<i>haber de</i>	2804	1179	2339	17643	17329	4963	8458	1492
<i>tener de</i>	858	205	307	1249	1082	284	274	91
<i>tener que</i>	30	13	59	666	644	536	2406	7793

² La selección y frecuencia de uso mayoritaria de la preposición *de* también con *tener* refuerza la idea de la analogía gramatical surgida a partir de la equiparación léxica de los verbos posesivos: se produce la suplantación total de una construcción perifrástica por la otra (Garachana y Rosemeyer 2011: 41).

³ No se incluyen aquí las ocurrencias de la perífrasis *haber que*, de valor impersonal, y, por lo tanto, sin posibilidad de alternar con otras en otros muchos contextos.

En opinión de la autora, las tendencias que se observan en 1600, y que culminan los procesos evolutivos previos, se mantienen hasta el XIX. A partir de este momento y hasta la actualidad, si bien se conservan todas estas perífrasis, las frecuencias de uso serán muy dispares entre sí. Y en esta misma línea apuntan los datos aportados por Martínez Díaz (2003) a partir de diferentes corpus escritos. Así, en el XVIII y XIX, *haber de* domina sobre *tener que* en todo tipo de textos. En el XX, sin embargo, se invierten las proporciones.

Al margen de la distribución cuantitativa, otro aspecto de interés al considerar la evolución diacrónica de las perífrasis se centra en la equivalencia o distinción funcional entre ambas. Si partimos de la base de una sustitución léxica de *aver* por *tener* también en estas construcciones verbales, es lógico esperar que no estén muy alejados sus contextos de uso, al menos en sus primeras apariciones. No obstante, a tenor de la opinión de muchos gramáticos, se reconoce, al menos hasta llegado el siglo XX, una determinada distribución complementaria a la que, curiosamente, no es ajena la propia semántica de los verbos auxiliares. En opinión de Stengaard (2003: 1150), por ejemplo, en el español antiguo las construcciones con el verbo *aver* parecen tener un denominador semántico común, el factor deagentivizante. En sus palabras: "... por medio de las perífrasis con *aver*, el sujeto de la acción expresada por el infinitivo o pierde su posible papel de sujeto-agente, o se refuerza el papel de sujeto-receptor o paciente implicado por la acción verbal en cuestión". Este efecto de sentido se pone en relación con la propia semántica del verbo *haber*, que denota una posesión no agentiva o receptiva, donde el sujeto no ejerce control sobre lo poseído, a diferencia del verbo con el que contiene: *tener* (Seifert 1930)⁴. Dicha agentividad, o participación del sujeto, es interpretada por Martínez Díaz (2008: 916) en términos de necesidad, o lo que es lo mismo, con *tener* es el sujeto quien juzga la imposición de la obligación; con *haber*, dicha obligación obedece a factores externos al mismo sujeto.

De este modo, las perífrasis con *aver*, según Yllera (1980: 100-101) tenían un valor modal amplio en los albores del romance, e indicaban tanto la necesidad como la obligación, así como matices más perfilados (obligación moral, necesidad atenuada, ley o costumbre), en alguno de los cuales alternó con *deber*. La construcción con *tener*, sin embargo, expresaba una acción cuyo cumplimiento era considerado necesario por el propio agente. A partir del XV, sin embargo, parece ir produciéndose una "confusión" semántica plena que lleva a una simple alternancia estilística de

⁴ Desde un acercamiento cognitivo, Garachana (1997) precisa esta oposición en términos de prototipicidad de la posesión, según la cual *haber* experimentó un vaciado semántico (control figurado de lo poseído) que no alcanzó a *tener*. Y en otro trabajo (Garachana y Rosemeyer 2011), se señala que la agentividad del sujeto es mayor en las construcciones con *tener que* que con *haber de*.

variantes: mientras que las construcciones con *aver* se emplean en contextos en los que la obligación externa es menos coercitiva o, en ocasiones, motivada por una fuerza mayor ineludible, *tener* amplía sus valores hacia la expresión tanto de la necesidad individual como de la obligación externa. Según Stengaard (2003: 1151), la incorporación del factor deagentivizante a la semántica de *tener* sería determinante para motivar esta alternancia. Y aún más: hacia finales del XVIII, *tener* incorporará también, de forma más aislada, empleos epistémicos, como los que se observan en las perífrasis con *deber* y *haber de* + infinitivo (López Izquierdo 2008).

En el siglo XIX, periodo de crucial interés en este trabajo, los gramáticos enfatizan la equivalencia entre *haber de* y *tener que*, incluso en lo que atañe a la expresión de la temporalidad futura. El valenciano Vicente Salvá (1835) sostenía en varios apartados de su *Gramática de la lengua castellana según ahora se habla* que: “La frase *haber de amar* equivale a *tener que amar*; *tener necesidad de amar*, es decir que comunica el valor de futuro a todo tiempo con que se junta” (p. 185), “*haber de* es *tener que* en todos en los tiempos” (p. 186)”, “*Haber de* significa también la necesidad de hacer una cosa, pues *Habías de acudir a la cita*, en nada se diferencia de *tenías que acudir a la cita*”. Y similar equivalencia establecía por las mismas fechas la *Gramática* de la Real Academia en su novena edición (1870: 62), tanto en la expresión de la futuridad como de la necesidad-obligación. Del mismo modo, Bello (1859: 316-318) destaca que: “*Haber de* significa necesidad, deber (...) *Pero solemos emplear esta frase con el solo objeto de significar tiempo futuro*”. Y más adelante puntualiza: “Úsase la misma sustitución de *tener* a *haber* en formas compuestas del auxiliar, la preposición *de* y un infinitivo: *tengo de salir*, frase en que se indica una determinación decidida de la voluntad, una resolución” (316-318). Por otra parte, la expresión “tengo que escribir” la hace equivalente a “debo”, “tengo la precisión de escribir” (378).

En el siglo XX, cuando el dominio de *tener que* ya se ha consumado, autores como R. Seco ([1930] 1967) no llegan tampoco a diferenciarlas y en su gramática llega a homologarlas, incluso, con los significados y usos de otra perífrasis deóntica: *deber* + infinitivo.

Más recientemente, Gómez Torrego (1988, 1999) retoma la distinción basada en la necesidad subjetiva, más acentuada con *tener*, y la obligación externa. Aun así, este autor admite que, salvo en algunos contextos restringidos (con infinitivos compuestos y en algunos contextos exclamativos), los usos en la práctica de ambas construcciones son idénticos, con la diferencia de que las perífrasis con *tener* “resultan más enérgicas y, por tanto, más coloquiales que las de *haber*. Salvo estas diferencias de registro, estos dos tipos de perífrasis suelen ser sinónimos en lo que a la obligación se refiere” (1988: 78).

Esta diferencia de registro es confirmada por el análisis que realiza

Martínez Díaz (2003, 2008). Los datos obtenidos demuestran que *haber de* presenta todavía un empleo regular en los registros formales de la lengua escrita, así como en algunas zonas dialectales del español, por la influencia convergente del contacto lingüístico con el catalán⁵. Para la autora, pues, el proceso histórico que ha llevado a la sustitución de *haber por tener* todavía hoy no ha concluido.

En conclusión, el recorrido evolutivo de ambas perífrasis parece haber conducido a la neutralización de las diferencias que un día pudieron existir en el ámbito de la obligación-necesidad y, sobre esta equivalencia, ha terminado imponiéndose en nuestros días *tener que* sobre *haber de*. Ahora bien, como ha demostrado sobradamente la sociolingüística variacionista, la equivalencia en el discurso de dos variantes de una misma variable lingüística, no implica que la variación responda necesariamente al azar. De hecho, las más de las veces la distribución de aquellas suele ser el resultado de una conjunción compleja de factores lingüísticos, pero también estilísticos y sociales, de cuya incidencia es posible dar cuenta en términos probabilísticos con las herramientas metodológicas adecuadas. Por lo demás, la influencia de tales factores puede variar con el tiempo, un aspecto que suele pasarse por alto en algunas aproximaciones al cambio lingüístico. Por eso, en este trabajo nos interesa especialmente saber qué ocurrió en el siglo XIX con las dos perífrasis en liza. Y ello porque el periodo decimonónico se sitúa en un momento clave en el largo proceso de sustitución de *haber de*, perífrasis modal por excelencia hasta ese momento, por la hasta entonces mucho más restrictiva *tener que*. En este marco, ¿qué aportan nuestros datos, extraídos esta vez de textos escritos más cercanos al polo de la inmediatez comunicativa, y por lo tanto, a la lengua oral? ¿Cuál es la correlación de fuerzas entre ambas perífrasis? ¿Es el siglo XIX el momento en que se desencadena el cambio, que se acelerará extraordinariamente en la centuria siguiente, hasta arrinconar a *haber de* a los usos formales de la lengua escrita o a ciertos empleos dialectales? Y, sea cual sea la respuesta a estos interrogantes: ¿qué factores lingüísticos y extralingüísticos propician esa evolución y cuáles, por el contrario, suponen todavía un freno?

Antes de responder todos estos interrogantes, resumimos los principales objetivos teóricos del presente estudio, así como los detalles metodológicos más relevantes.

3. OBJETIVOS Y METODOLOGÍA

El presente estudio forma parte de un proyecto de sociolingüística histórica (vid. nota 1), para el que se ha compilado un amplio corpus com-

⁵ En el mismo sentido, Wesch (1997), Hernández García (1998), Sinner (2003), Porcar (2012).

puesto por textos próximos al polo de la inmediatez comunicativa, escritos por individuos de diferente extracción social y dialectal. Con todo, en este último caso, el análisis se limita a textos redactados por españoles o por individuos nacidos fuera de España, pero que pasaron la mayor parte de su vida en este país. Asimismo, se distinguen textos de diferentes registros, desde los asuntos más íntimos o familiares, a otros de naturaleza menos privada, con diversos grados intermedios. Para el siglo XIX, dicho corpus consta de 28 obras, de las cuales 23 corresponden a epistolarios y los cinco restantes a textos autobiográficos (libros de cuentas, memorias, diarios, etc.; ver la relación completa en el anexo 1). El conjunto, que da voz a más de ciento cincuenta locutores diferentes⁶, se cifra en 490.014 palabras.

A partir de este corpus, nuestros objetivos se concretan en los siguientes puntos:

- a) comprobar la distribución de uso de las dos perífrasis, *haber de y tener que* + infinitivo, en textos de inmediatez comunicativa, más cercanos, por consiguiente, a la oralidad de épocas pasadas que otras tradiciones discursivas utilizadas tradicionalmente en el estudio diacrónico de la lengua.
- b) averiguar si, en el español decimonónico, se advierten factores lingüísticos, estilísticos o sociales que justifiquen la selección preferente de cada una de esas variantes;
- c) determinar cuál es la jerarquía explicativa de esos factores.

Por medio de un programa de concordancias (*Wordsmith* v. 4), se seleccionaron todas las ocurrencias de las dos variantes en el corpus. A continuación, se procedió a codificarlas de acuerdo con más de una veintena de factores de diferente naturaleza, ensayados previamente en otros estudios variacionistas acerca de estas mismas unidades lingüísticas (cf. Balasch 2008, 2012, Blas Arroyo y Porcar en prensa; Blas Arroyo y González 2014, Blas Arroyo y Vellón 2014)⁷:

- 1) *Lingüísticos*, entre los que distinguimos tres tipos:
 - Fonológicos: contexto fónico siguiente (primer fonema del verbo principal o de la palabra que sigue a *tener* ('tengo ya que dejarte'), con excepción del propio enlace, ya sea *de* o *que*).

⁶ Ello es posible dada la multiplicidad de remitentes de cartas en algunos de los epistolarios incluidos en el corpus, la mayoría procedentes de la emigración a América (*Historias de América, Noticias desde Cuba, Cartas de emigrantes vallisoletanos, Desde la otra orilla, El hilo que une...*). Otros, sin embargo, son la obra de un solo autor.

⁷ Para la delimitación exacta de todos estos factores, tanto de naturaleza lingüística como extralingüística, se remite al lector a los apartados correspondientes de la sección 4.

- Morfosintácticos: 1) persona y número del grupo verbal, 2) tiempo y modo, 3) sintaxis del sujeto (omitido, explícito), 4) forma del verbo auxiliar (simple, compuesto); 5) sintaxis del verbo principal (simple, compuesto); 6) tipo de cláusula (subordinadas, otras); y 7) modalidad oracional (afirmativas, negativas, otras).
 - Semántico-pragmáticos: 1) valores semánticos (modales, temporales (futuro)), 2) matices modales, 3) grado de animacidad del sujeto (humanos, no humanos), 4) modo de acción del verbo principal (estativos, movimiento, lengua, resto (dinámicos), 5) modalización contextual (entornos formalmente intensificados, resto); 6) grado de (im)personalidad semántica (activas, pasivas, impersonales).
- 2) Factores *estilísticos*. A este respecto se distinguen dos parámetros: a) temática principal de los textos (familiar, semifamiliar, otros), y b) grado de relación entre los interlocutores (estrecha/distante/otras).
- 3) Factores *extralingüísticos*: a) periodo (1^a mitad/2^a mitad del siglo XIX); b) región; c) sexo; d) edad; y e) estrato social.

Para el análisis estadístico nos servimos del programa de regresión logística Goldvarb 3.0, habitual en las investigaciones variacionistas. Mediante su uso se pueden calcular las diferencias frecuenciales entre las diferentes variantes y sus contextos, y, lo que todavía es más relevante, el grado de significación y la jerarquía explicativa de los factores expuestos, cuando todos son considerados al mismo tiempo. Ello permite descubrir, además, relaciones de (inter)dependencia e interacción entre unos factores y otros, lo que sería imposible con una estadística meramente descriptiva.

4. RESULTADOS Y ANÁLISIS

4.1. Generales

El número total de ocurrencias de las dos perífrasis en el corpus asciende a 282, una cifra reducida, sin duda, aunque no tan inesperada en los estudios de variación sintáctica, cuyas variables tienen, por lo general, una baja recurrencia en el discurso, especialmente si se las compara con las de naturaleza fonológica (Blas Arroyo 2005, Tagliamonte 2011). De ellas, 162 (56.7%) corresponden a *haber de*, frente a 120 (42.3%) a *tener que*⁸. En términos absolutos, observamos, pues, cómo en el siglo XIX, las

⁸ Agrupamos también bajo esta misma forma las tres únicas ocurrencias de la perífrasis *tener de* que hemos localizado en el corpus. Parece evidente que, pese a su notable vitalidad en épocas pasa-

ocurrencias de la perífrasis con el verbo *haber* con sujeto personal tienen todavía un notable protagonismo, aunque a distancia ya del que tuvieron en épocas anteriores de la lengua. Una comparación con recuentos de otros periodos no ofrece lugar a dudas. Así, un estudio sobre textos de naturaleza similar del siglo XVI (Blas Arroyo y Porcar b en prensa), muestra que, en el español clásico, la proporción de las perífrasis con *tener* (*tener que* y *tener de*) es apenas de 1 a 10 (N = 168) con respecto a la muy mayoritaria *haber de* (N = 1584), y queda incluso por debajo de las que utilizan *deber* como verbo auxiliar (N = 326).

Nuestros datos del siglo XIX no coinciden, sin embargo, con los que arrojan otros cálculos realizados en los últimos tiempos a partir de la consulta de corpus de grandes dimensiones, basados en tradiciones discursivas más formales. Este es el caso, por ejemplo, de los recuentos efectuados por López Izquierdo (2008: 793) tras la búsqueda de estas perífrasis en el *Corpus del español* (Davies 2006). El resultado de esa búsqueda arroja un total de 8458 ocurrencias para *haber de* (76%), frente a 2680 (24%) para las combinaciones perifrásticas con *tener*. Estas proporciones son tan diferentes a las obtenidas en nuestro estudio que postulamos como hipótesis la posibilidad de que obedezcan a la diferente naturaleza discursiva de los textos sobre los que se asientan ambos corpus. Mientras que en el monumental *Corpus del español*, existe una prelación indudable de tradiciones discursivas literarias y formales (al menos hasta el siglo XX, en el que se habilita una mayor variación de registros), los textos que conforman el corpus de esta investigación se sitúan en el otro extremo del *continuum* distancia-inmediatez comunicativa.

La hipótesis acerca del mayor acercamiento de estos textos al polo de la inmediatez, y en definitiva, a los usos coloquiales de épocas pretéritas, tiene un corolario relevante para lo que aquí nos interesa. Y es que, contrariamente a lo que pudiera desprenderse del estudio de las fuentes tradicionales en el estudio diacrónico del español, en la lengua común del siglo XIX la presencia de *tener que* está ya fuertemente asentada, en especial en algunos contextos como los que analizaremos en los siguientes apartados,

das (Yllera 1980, Blas Arroyo y González 2014), a estas alturas de la historia, dicha perífrasis había perdido ya gran parte de sus empleos tradicionales, quedando relegada a usos dialectales muy ocasionales.

TABLA 2: Frecuencias de uso de las perífrasis *haber de* y *tener que* + infinitivo en el corpus (datos generales y distribución por valores semánticos)

Valores semánticos	<i>Tener que</i>		<i>Haber de</i>	
	N	%	N	%
Modales (Deónticos)	115	56.1	90	43.9
Modales (Epistémicos)	4	40	6	60
Modales (Otros)	1	11.1	8	88.9
Temporales	–	–	58	100
Σ	120	42.3	162	56.7

4.2. Entre la modalidad y la futuridad

Como puede apreciarse en la tabla 2, junto a los contenidos modales, mayoritarios en el corpus (80%), encontramos también otros en los que las perífrasis no van mucho más allá de la mera expresión de significados que apuntan a un futuro más o menos definido. Ciertamente es que en la propia significación modal se halla implícito muchas veces un valor prospectivo, pues estas perífrasis poseen un valor aspectual ingresivo (o incoativo) que, por naturaleza, se dirige hacia la futuridad⁹. Sin embargo, en ocasiones, como las ejemplificadas a continuación, el hablante prescinde de lo modal, y en su enunciado no parece pretender otra cosa que la mera formulación de hechos que sitúa en un futuro, sea con respecto al momento de la enunciación, como en (7), sea en relación con un momento pretérito, como en (8). Y aunque minoritarios, estos valores meramente prospectivos suponen una proporción nada desdeñable (20%) de todos los usos perifrásticos localizados en el corpus;

- (7) y esperando ocasión oportuna para pagarle esta deuda de gratitud. Vamos que se *ha de alegrar* con la nueva que va a oír (*Cartas de San Enrique de Ossó*)
- (8) en esto viene el que *había de ser* mi confesor [...] Esta noche tampoco salgo de la habitación... (*Diario de mi prisión*)

Conviene recordar que, en latín, las perífrasis con *habeo*, que están en el origen de la castellana, comenzaron teniendo un significado obligatorio para terminar con un valor exclusivamente temporal. Este proceso no se completó en la misma medida con *haber de*, pero en el castellano antiguo y clásico

⁹ Por otro lado, se ha observado también que el significado prospectivo se desprende de algunos rasgos comúnmente asociados a la obligación, como la intencionalidad y la determinación que presiden muchos de esos contenidos deónticos (Rojo 1974: 86). Como señala Sinner (2003: 200): “El desarrollo del significado obligatorio o volitivo hacia el de futuridad es una evolución fácilmente comprensible, ya que en principio no es otra cosa que la continuación lógica de la determinación expresada por el hablante”.

co dicho valor temporal tuvo, ciertamente, una notable vitalidad. Según Sáez Godoy (1968), la perífrasis *haber de* + infinitivo pasó de ser empleada con valor de futuro en un 17% en los *Entremeses* cervantinos y un 18% en las comedias de Lope de Vega, a tan solo un 1% en el teatro de la segunda mitad del siglo XX. Por lo demás, algunas lenguas romances próximas conservan aún este valor futuro en las perífrasis con *haber*. Es el caso del portugués europeo (no así del brasileño, donde se considera anticuado) o, más aún para lo que aquí nos interesa, del gallego y el astur-leonés. Junto a estas variedades de contacto (Sinner 2003), los valores de futuro de la construcción se han detectado también en diversas hablas andaluzas (Zamora 1960), así como en el español de América, donde pervive todavía en algunas áreas dialectales (Kany 1969; Steel 1982; De Bruyne 1993; Westmoreland 1997).

En el español clásico, los significados prospectivos no modales aparecen también, aunque de forma más ocasional, en las perífrasis con *tener*, y más concretamente, en aquellas que se sirven de la preposición (*tener de*) como enlace entre el auxiliar y el verbo principal (Yllera 1980, Blas Arroyo y González 2014 b). Sin embargo, nuestros datos del siglo XIX muestran, sin ningún género de dudas, una asociación categórica del futuro con *haber de*, y quedan fuera de esta esfera semántica las perífrasis con *tener*.

La ausencia de variación en esta parcela de la futuridad hace que, en lo que sigue, centremos exclusivamente nuestro interés en la esfera de los contenidos modales, que suponen, como vimos, el 80% de todas las ocurrencias de las perífrasis en el corpus. Y en este nuevo cómputo, la tabla 3 muestra cómo *tener que* (53.5%) supera ya incluso a *haber de* (46.5%). Dicha tabla contiene también los datos cuantitativos correspondientes al cruce con los factores lingüísticos y extralingüísticos reseñados más arriba (ver apartado 4). Con todo, y por razones de espacio, en los siguientes apartados nos ocuparemos solo de aquellos parámetros que han resultado significativos a la luz del análisis estadístico realizado.

TABLA 3: Contribución de los factores lingüísticos, estilísticos y sociales para la selección de la perífrasis *tener que* y *haber de* + infinitivo en contextos modales (Goldvarb 3.0)

	P. (<i>tener/haber</i>)	%	N
Total N: 120/104			
Media Corregida: .554			
Matices (modales)			
Obligación externa		[57.9/42.1	77/56]
Obligación interna		[42.9/57.1	12/16]
Necesidad/conveniencia		[60.5/39.5	26/17]
Conjetura		[40/60	4/6]
Otros		[11.1/88.9	1/8]
<i>Rango</i>			

Modalidad oracional			
Afirmativas	.49/.51	52.4/47.6	99/90
Negativas	.87/.13	83.3/16.7	20/4
Otras	.03/.97	9.1/90.9	1/11
<i>Rango</i>	84		
Tiempo/modo			
Presente indicativo	.34/.66	41.7/58.3	55/77
Imperfecto indicativo	.37/.63	48/52	12/13
Resto	.82/.18	79.1/20.9	53/14
<i>Rango</i>	48		
Modo de acción			
Estativos	.39/.61	43.3/56.7	26/34
Lengua	.58/.42	61/39	25/16
Movimiento	.76/.24	68.9/31.1	31/14
Resto	.38/.62	49.4/50.6	38/39
<i>Rango</i>	38		
Persona y número			
1ª persona.	.71/.29	69.1//30.9	47/21
2ª persona	.42/.58	51.2/48.8	22/21
3ª persona	.38/.62	39.6/60.4	40/61
Formas no personales	.57/.43	91.7/8.3	11/1
<i>Rango</i>	33		
Grado de (im)personalidad semántica			
Oraciones activas	.53/.47	57.6/42.4	118/87
Oraciones pasivas/impersonales	.20/.80	10.5/89.5	2/17
<i>Rango</i>	33		
Grado de animacidad (3ª persona)			
Humano	.52/.48	45/55	32/39
No humano	.44/.56	27/73	8/22
<i>Rango</i>	8		
Tipo de cláusula			
Subordinadas		[50.3/49.7	72/71]
Otras		[58.8/41.2	47/33]
<i>Rango</i>			
Modalización contextual			
Contextos intensificados		[60.3/39.7	38/25]
Contextos neutros		[51.6/48.4	82/77]
<i>Rango</i>			
V Principal (simple/compuesto)			
Simple		[43.2/56.8	114/150]
Compuesto		[33.3/66.7	6/12]
<i>Rango</i>			
Tenor (temática / relación interlocutores)			
Cartas familiares/íntimas	.64/.36	66.7/33.3	66/33
Cartas no familiares	.52/.48	44.2/55.8	19/24
Textos no epistolares	.32/.68	42.7/57.3	35/47
<i>Rango</i>	32		

Edad			
Jóvenes	.68/.32	64.1/35.9	50/28
Adultos	.38/.62	45.5/54.5	56/67
<i>Rango</i>	30		
Sexo			
Hombres		[52.3/47.7	101/92]
Mujeres		[61.3/38.7	19/12]
<i>Rango</i>			
Estratificación social			
NSC alto		[50.4/49.6	70/69]
NSC medio		[61.8/38.2	21/13]
NSC bajo		[64.4/35.6	29/16]
<i>Rango</i>			
Regiones			
Monolingües		[57.3/46.3	65/56]
Bilingües (catalán)		[47.6/52.4	10/13]
Bilingües (resto)		[43.5/56.5	20/22]
<i>Rango</i>			

Convergence at iteration: 10. Cells: 145
 Log likelihood= - 105.195 Significación: .007

4.3. Valores modales

Como señalábamos en el apartado anterior, pese al hallazgo de valores genuinamente temporales entre las combinaciones perifrásticas con *haber*, son los contenidos modales los verdaderamente prioritarios en el español del siglo XIX. De entre estos valores es, sin duda, la asociación con la esfera de lo deóntico la más destacada, ya que representa nada menos que el 91% de todas las ocurrencias de la variable. Con todo, junto a estos advertimos también algunos casos aislados (N = 8; 4%) en los que las expresiones verbales se tiñen de un cierto énfasis expresivo para comunicar estados de sorpresa, indignación, recriminación, etc. (Gómez Torrego 1999: 3356). Así ocurre, por ejemplo, con el siguiente fragmento, donde quien escribe se vale de una pregunta retórica para mostrar la contradicción entre lo expresado por los dos periodos de la oración:

- (9) Si yo me acusé de dicha falta ¿cómo *he de presentar* descargos? (*Carta familiar de D. José Butrón*).

Poco representativos son también los enunciados con valor epistémico (4,5%), más frecuentes en español con las perífrasis construidas con *deber* (*de*) + infinitivo (Gómez Torrego 1999: 3353), y en las que el hablante da cuenta de contenidos que interpreta en el ámbito de lo probable o aproximado. Esos usos aparecen ya desde antiguo para la perífrasis con *haber*

(Yllera 1980), pero, al decir de López Izquierdo (2008: 802), en el caso de *tener que* “comenzarán a extenderse a partir de finales del siglo XVIII y sobre todo durante el XIX”¹⁰. Los de (10) y (11) son dos de los pocos ejemplos representativos:

- (10) ... conquese sale por consecuencia que mi ermana *havia de tener* oculto lo que menos ese caudal (*Documentos norteños*).
- (11) dejándose ver al mismísimo don José Butron tal como *tuvo que ser* en vida; esto es, caballero antiguo devoto á su Dios y á su Rey, comedido en sus actos, generoso hasta el despilfarro en aras de su patria... (*Extracto de las cartas recibidas por E. J. Butrón*).

En relación a la modalidad deóntica, donde se concentra el grueso de las perífrasis, diversos autores han intentado precisar los matices que se advierten en su seno (Keniston 1937; Yllera 1980; Olbertz 1998; Gómez Torrego 1988, 1999; Fernández 1999; García Fernández 2006; López Izquierdo 2008; Martínez Díaz 2008), una tarea no exenta de dificultades y colmada de subjetivismo, como ya tuvimos ocasión de comprobar (ver apartado 3). Así las cosas, y en un intento de huir en la medida de lo posible de este último, hemos dividido el eje deóntico en diversas categorías que atienden a dos parámetros que encontramos de manera recurrente en la bibliografía, aunque, hasta donde llega nuestro conocimiento, nunca se hayan combinado entre sí para su comprobación empírica. Se trata de:

- a) el grado de obligación/necesidad impuesta; y
- b) el agente que impone dicha obligación/necesidad.

Como resultado de la combinación entre ambos ejes, nos encontramos con los siguientes valores principales¹¹:

¹⁰ Utilizamos aquí el término epistémico en su acepción más tradicional, y no en otros usos más aislados, como los que utiliza Martínez Díaz (2008) al hablar de “obligación epistémica”. Para esta autora, dicha obligación se deriva de la presencia de un agente que orienta u obliga, pero desde una perspectiva subjetiva y no externa.

¹¹ La distinción entre los matices obligativos tampoco ha resultado sencilla en este trabajo. Sin embargo, queremos dejar constancia de que, en la codificación de todos y cada uno de los ejemplos, han intervenido dos de los autores del trabajo de manera independiente. Superada esta fase, los casos en que hubo discrepancia (apenas un 10%) fueron sometidos con posterioridad a la evaluación del tercer autor. En los casos más dudosos, en que fue imposible llegar a un acuerdo, ya fuera por las insuficiencias del contexto situacional o por cualquier otro motivo, los ejemplos quedaron sin codificar para este factor concreto. Por último, quisiéramos recordar que, en línea con lo argumentado por Martínez Díaz (2008: 1285) en lo que sigue entendemos la modalidad obligativa en un sentido amplio, es decir, “como expresión de la subjetividad del enunciado, de modo que el sujeto de la enunciación puede, o no, coincidir con el sujeto del enunciado”. Lo hacemos así, ya que, en caso contrario, nos veríamos impelidos a considerar tan solo los casos de obligación expresados por la primera persona del singular.

1. *Necesidad u obligación subjetiva autoimpuesta (interna)*. Se incluyen en este matiz los deberes basados en el convencimiento, en la determinación o intención del sujeto por motivos religiosos, éticos, filosóficos, de gratitud, respeto, o de cualquier otro criterio ideológico o personal. Se basa, pues, en las convicciones subjetivas íntimas o en el deseo del sujeto, lo que acerca estas perífrasis a las de carácter volitivo (Roca Pons 1980: 73, Yllera 1980: 114). Los de (12) y (13) son algunos ejemplos representativos:

- (12) Basta por ahora lo otro, siendo sólo tú el que *ha de decidir* sobre este particular (*José Madrazo a sus hijos*)
- (13) ...pero es necesario consolarse porque es una carrera que todos *tenemos que pasar* ... (*Historias de América*)

2. *Obligación de agente orientado o externa*. Los contenidos de obligación, necesidad ineludible o conveniencia imperativa y coaccionante, son de carácter *externo* al agente de la acción verbal. Se trata, pues, de enunciados directivos, entre los que distinguimos cuatro matices principales:

a) Obligación impuesta por norma, acuerdo, convención social, código legal, etc., con independencia de que estén o no recogidos por escrito:

- (14) ... para el efecto *tendrás que entenderte* con el comandante general de marina (*Historias de América*).

b) Mandato u orden externa para obligar a realizar una acción:

- (15) Si llegamos a tener dichos escritos, al instante se *han de imprimir* (*Epistolario José de Azara*).

c) Obligación impuesta por las circunstancias externas, esto es, la idea de una necesidad ajena a la voluntad del sujeto.

- (16) ...pues si te hallases aquí no te faltaría nada a la vera de tu hermano que *tiene que servirse* de gente extraña, pudiendo estar juntos (*Cartas de emigrantes escritas desde Cuba*).

d) Inevitabilidad. Incluimos aquí aquellos usos en los que la idea de futuro posee tal seguridad por parte del hablante que su realización se considera (modalmente, pues) necesaria e inevitable, como en (17):

- (17) ... y espero no te olvidarás de obra tan necesitada y que tanta gloria *ha de dar* a Dios y que se va multiplicando de un modo admirable (*Cartas de San Enrique de Ossó*).

3. *Necesidad o conveniencia*. Lo que prima ahora es la *necesidad o conveniencia* consideradas subjetivamente por el sujeto, de tal manera que este último controla o decide imponer(se) una determinada obligación no tanto por cuestiones morales o convicciones íntimas como en los casos de (12) y (13), cuanto por razones de mera conveniencia o por el beneficio que de ella pudiera derivarse. De este modo, el poder coercitivo de tales deberes es mucho menor que el expresado en los casos anteriores de 2 c. Así, frente a (16), donde el hablante reconoce que se ve obligado a tratar a personas ajenas a su familia dadas las circunstancias (vive a miles de kilómetros de aquella), en (18) y (19) los remitentes de sendas cartas expresan la necesidad o la conveniencia que atañe, desde su punto de vista respectivo, a cuestiones relacionadas con la amistad (18) o la riqueza de un país (19):

- (18) Su grata me ha animado mucho: los amigos *se han de conocer* en las verdades que dicen... (*Cartas de San Enrique de Ossó*).
- (19) y si no se puede trabajar en el campo que es donde *tiene que salir* la riqueza del país, estamos de más, apaga y vámonos (*Cartas desde América*). 11

Asimismo, incluimos en este apartado algunos usos “fáticos” que se han advertido en el uso de estas perífrasis, y en los que el hablante parece apoyarse (Gómez Manzano 1992: 160, habla de ellos como de “una especie de muletilla”) en su intento por “entrar en comunicación” (Gómez Torrego 1999: 3354), un hecho especialmente pertinente en el género epistolar que sirve como base principal del corpus. Estos son algunos casos representativos, en los que puede advertirse cómo el remitente marca, a través de estas construcciones verbales, algunas fases en el proceso comunicativo que mantiene con su destinatario:

- (20) ...*has de saber* que no hay por aquí pardillo que no haga esta y otras muchas reflexiones (*Epistolario José de Azara*)
- (21) Nada más *tenemos que decirte* sobre el particular, sino que manos a la obra (*Historias de América*)

Como puede apreciarse en la tabla 3, en el corpus se aprecia una notable asociación entre los valores modales expresivos –etiquetados como “Otros” y ejemplificados anteriormente en (9)– y *haber de*, si bien las ocurrencias de estos son muy escasas en el corpus (apenas 8 ejemplos, de los cuales 7 corresponden a esta última perífrasis), por lo que resulta difícil obtener conclusiones más fundadas. Sin embargo, los demás matices modales se reparten de manera mucho más equilibrada entre las dos variantes perifrásticas. Aun así, la comparación de frecuencias muestra una leve inclinación favorable al uso de *haber de* para la expresión de obligaciones derivadas de las convicciones interiores del sujeto a partir de sus valores morales o éticos (57.1%). Por el contrario, *tener que* aparece más

unida a las obligaciones de agente orientado o externas (58%), pero también, paradójicamente, a la expresión –subjativa– de la necesidad o conveniencia de hacer las cosas (60%). Sea como sea, estas diferencias no resultan significativas en la muestra analizada, por lo que no cabe descartar la influencia del azar en su configuración.

En definitiva, aunque nada garantiza que las cosas no pudieran ser de otra manera con un corpus más amplio, lo cierto es que en este no se cumplen las asignaciones que, desde una perspectiva funcional más o menos categórica, han intentado vincular los empleos modales –paradigmáticos– de cada perífrasis. De este modo, nuestros datos no avalan una supuesta invasión de los ámbitos expresivos de una perífrasis por los de su competidora, sino más bien la influencia diversa del contexto variable en el empleo de unas construcciones verbales que, para los hablantes, forman parte de un mismo dominio semántico.

4.4. Modalidad oracional

En diversos estudios anteriores, hemos tenido ocasión de comprobar la relevancia de este factor, que se postula como determinante para explicar la variación entre diversas variantes en el seno de las perífrasis modales de infinitivo. Así, en el análisis acerca de la alternancia entre las variantes prepositiva y no prepositiva de *deber* + infinitivo, vimos cómo las primeras aparecían significativamente más en los contextos de polaridad negativa, y ello tanto en el periodo clásico (Blas Arroyo y Porcar en prensa), como en el español moderno (Blas Arroyo y Vellón 2014). Y lo mismo hemos advertido en el castellano áureo con la alternancia entre *tener que* y *tener de* (Blas Arroyo y González 2014b), cuya primera variante se asocia también de manera destacada a esos mismos entornos negativos.

Estos mismos resultados se confirman ahora en la variación que nos ocupa, y nuevamente con las oraciones negativas como las más vinculadas a *tener que*, con un grado de significación que figura entre los más elevados del estudio (.87; 83%). En el extremo opuesto se sitúan las oraciones no enunciativas, como interrogativas directas, exclamativas, etc., que, pese a su baja recurrencia en el discurso, dejan un testimonio suficientemente explícito de ser uno de los puntos de la gramática menos favorables a esta variante en el siglo XIX (.03; 9%), y uno de los favoritos, por el contrario, de *haber de* (.97; 91%). Entre ambos polos se sitúan las oraciones enunciativas afirmativas, con mucho las más frecuentes en el corpus (84%), pese a lo cual su influencia en la selección de una u otra variante es prácticamente neutra.

4.5. Tiempo y modo verbales

Tan solo nueve paradigmas de la conjugación aparecen representados en el corpus, aunque estos lo hagan de una manera muy desequilibrada muestralmente. De este modo, y como cabía esperar, observamos cómo el presente de indicativo representa por sí solo el 59% de los ejemplos de la variable, seguido a considerable distancia por el imperfecto (11%), el pasado simple (7,4%), el futuro (6%) –todos ellos de indicativo– y las formas no personales (6%). El resto no llega al 4%, con la presencia aislada, de mayor a menor, del pretérito perfecto de indicativo (4%), del presente de subjuntivo (4%), del condicional (2.2%) y del imperfecto de subjuntivo (1%).

Un primer análisis de frecuencias muestra que son, precisamente, estas formas menos habituales en el discurso las que manifiestan una preferencia más clara por *tener que*, con casos incluso categóricos (pretérito perfecto indicativo), y otros donde, sin llegar a serlo, se aprecia una inclinación clara por esta perífrasis. Así ocurre con las formas no personales (92%), el condicional (89%), el pasado simple (87%), el presente de subjuntivo (78%) y, en menor medida, el futuro (60%). Con todo, las cifras absolutas de estos tiempos son poco representativas, de ahí que, para el análisis multivariante hayamos recodificado este factor con la presencia de tres únicos contextos: 1) presente e 2) imperfecto, ambos de indicativo, y 3) resto de formas de la conjugación.

El resultado de este análisis confirma claramente la asociación de esos paradigmas minoritarios con *tener que* (.82; 81.5%). Por el contrario, el tiempo más repetido en el corpus, el presente de indicativo, supone todavía un freno importante para el avance de esta perífrasis (.34; 41.7%), ya que en el siglo XIX se alía en mayor medida con su competidora, *haber de* (.66; 58.3%). Este obstáculo se ve alentado por las cifras, también negativas, del segundo paradigma más común en el corpus, el imperfecto de indicativo, cuya asociación con *haber de* parece también evidente (.63; 48%).

En consecuencia, a pesar de los avances incuestionables de su competidora en diversas áreas de la gramática, el mantenimiento de *haber de* en el siglo XIX parece venir propiciado por su elevada presencia en aquellos paradigmas de la conjugación más frecuentes en el discurso, lo que todavía representa un importante freno para su retroceso, ya evidente, sin embargo, en otros contextos.

4.6. Modo de acción

En un trabajo anterior acerca de la variación entre las perífrasis *tener que/de* en los siglos de oro (Blas Arroyo y González 2014b), tuvimos ya la

ocasión de comprobar la estrecha asociación entre la primera de esas variantes y los verbos de lengua (.92; 87%), y en particular *decir*, una forma que en el 95% de las ocasiones seleccionaba *tener que*. Aunque con un poder explicativo menor esta vez (.58; 61%), dicha asociación positiva se aprecia todavía en el XIX. En el caso de *decir*, uno de los verbos de lengua más habituales, la preferencia por *tener que* es también ahora manifiesta, con 8 ejemplos (72%), frente a tan solo 3 (28%) para *haber de*. Pese a ello, son los verbos de movimiento los que ocupan una posición más alta en la escala favorable a *tener que*, con valores probabilísticos (.76; 69%) particularmente elevados.

En el extremo opuesto se sitúan los demás verbos dinámicos (.38; 49%), así como los de naturaleza estativa (.39; 43%), más favorables a la selección de *haber de*. Un análisis más pormenorizado confirma estos resultados en el caso de algunos verbos representativos, más frecuentes en el discurso que la media. Es el caso, por ejemplo, de *ser* entre los estativos, del que 7 de las 10 ocasiones en que aparece como verbo principal en el seno de una perífrasis lo hace junto al auxiliar *haber*. Más equilibrado, aunque igualmente favorable a esta perífrasis, resulta *hacer*, el más habitual entre los verbos dinámicos, cuyas 18 ocurrencias en el corpus lo hacen también mayoritariamente (10) con *haber*.

Así las cosas, estos datos confirman solo parcialmente la hipótesis de Garachana y Rosemeyer (2011), según la cual, la consolidación de *tener que* en el ámbito de la obligación va ligada a la selección preferente de verbos no estativos. Al menos en el siglo XIX, dicha consolidación tiene lugar preferentemente en dos paradigmas de tales verbos, los de lengua y movimiento, pero no en el resto, que, a este respecto, se comportan como los verbos estativos.

4.7. Grado de animacidad

Como se recordará, uno de los factores pragmáticos que interesaba analizar en el estudio era la potencial influencia del grado de animacidad del sujeto en la selección de las perífrasis, distinguiendo a este respecto entre sujetos humanos y no humanos¹². Un primer dato al respecto es que, con una u otra perífrasis los primeros constituyen la gran mayoría (76%), un hecho por lo demás lógico si nos enfrentamos a contenidos deónticos en los que la obligación requiere prototípicamente de un agente. Asimismo, hemos considerado como factor independiente el grado de (im)personalidad semántica de esos sujetos, diferenciando así entre los

¹² Como es lógico, en este caso el análisis se limita a la tercera persona, la única en la que es posible esta distinción semántica.

que aparecen en oraciones activas –todavía más mayoritarios en el corpus (91%)– y aquellos que surgen en frases pasivas o impersonales (9%).

En relación con el grado de animacidad, los datos muestran una leve inclinación de los sujetos humanos por las construcciones con *tener que* (.52), frente a los no humanos (.44). Hay que reconocer que estas diferencias, aunque significativas estadísticamente, son menores y en cierto modo se hallan lastradas por la baja recurrencia de ambos contextos en el corpus, ya que se restringen exclusivamente a la 3ª persona. Estas mismas restricciones muestrales condicionan la relevancia del grado de (im)personalidad del sujeto, pese a lo cual el factor es también seleccionado como significativo. Y ello, fundamentalmente, por las diferencias de mayor calado que en este caso se aprecian entre los dos contextos analizados. Así, la tabla 3 muestra claramente cómo tan solo un 10.5% (.20) de las oraciones pasivas o impersonales lo hacen con *tener que*, entornos, en suma, que parecen mucho más favorables a *haber de* (.80; 89.5) en los textos del siglo XIX.

Un factor relacionado, que también recibe el aval estadístico (rango 33), es la persona y número del sujeto, con diferencias significativas entre unos contextos y otros. De este modo, advertimos que la 1ª persona¹³ representa uno de los terrenos mejor abonados para el empleo de *tener que*. Nada menos que en el 68% (valor P. 71) de todas las ocasiones encontradas en el corpus, el hablante acude a esta perífrasis para expresar diferentes contenidos modales. Paradójicamente, a estas le siguen las formas no personales (.62), mucho menos frecuentes en el discurso, pero que de forma casi categórica (91.7) se construyen con *tener y* no con *haber* (solo un ejemplo de esta última en el corpus).

Responsable del mantenimiento de *haber de* en cifras todavía elevadas es, sin embargo, la 3ª persona, que en el corpus manifiesta una preferencia sólida por esta perífrasis (.62; 60.4%), frente a su competidora (.38; 39.6%). Con todo, un análisis más detenido muestra que dicha preferencia se localiza de manera muy destacada en la 3ª persona del singular, donde la preferencia por *haber de* alcanza cifras muy elevadas (95,5%), cercanas, pues, a la selección categórica. El correspondiente reanálisis, en el que este factor queda aislado del resto, confirma la significación altísima de esta asociación entre la perífrasis y la 3ª personal del singular (.91).

4.8. Factores extralingüísticos

Aunque, por lo general, menos influyente en la esfera de la variación sintáctica que en el nivel fónico (Blas Arroyo 2005: 85ss.), el resultado de

¹³ Agrupamos aquí tanto la 1ª persona del singular como la del plural, aunque la representación de esta última sea mucho más reducida en el corpus (apenas 4 ejemplos). El mismo criterio se sigue para las demás personas.

dicha variación puede ser también la consecuencia del influjo de otros factores no estructurales a los que nos referiremos en esta sección.

En la práctica, así ocurre en el presente caso con el eje estilístico de la variación, donde encontramos algunas diferencias significativas que serán, a nuestro juicio, premonitorias de un proceso que se acelerará durante el siglo xx. Para la delimitación de este eje hemos procedido a cruzar los factores considerados inicialmente en la investigación, y que, como se recordará, correspondían a: 1) la temática principal de los textos; y 2) el grado de relación entre los interlocutores. El resultado de tal combinación ofrece un *continuum* estilístico integrado por tres grupos suficientemente alejados entre sí como para permitir una comparación razonable. Se trata de:

- a) Cartas de contenido familiar o íntimo entre personas a las que unen estrechos lazos de convivencia, ya sea de origen familiar (en la mayoría de los casos), ya de amistad, amor, etc.
- b) Cartas en las que prima una temática no familiar, y en las que, además, existe una clara distancia entre los interlocutores en los ejes de la intimidad o la solidaridad.
- c) Textos autobiográficos (memorias, diarios, declaraciones judiciales)¹⁴.

El resultado de este análisis muestra diferencias interesantes y significativas estadísticamente. Y es que, como se aprecia en el siguiente gráfico, existe una distancia notable entre algunos de los contextos analizados. Así, mientras que las cartas familiares o entre íntimos favorecen (.64) el empleo de la perífrasis *tener que*, lo contrario sucede en los textos también de naturaleza autobiográfica (.32). Entre ambos extremos se sitúan las cartas de temática y relación no familiar, con cifras probabilísticas intermedias (.52), que hablan de una influencia prácticamente neutra.

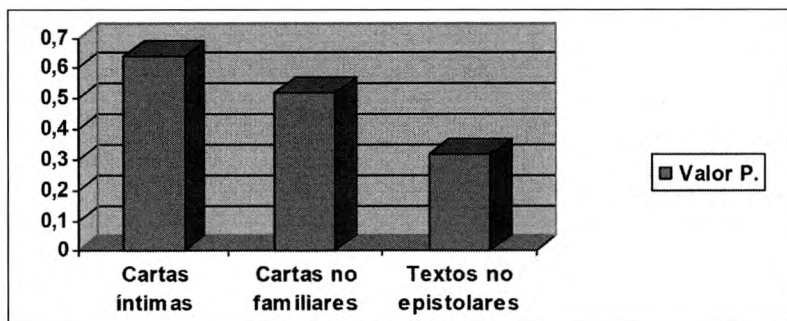


GRÁFICO 1: Probabilidades asociadas a la selección de *tener que* en el corpus.

¹⁴ Para facilitar las comparaciones, aquellas ocurrencias que no cayeran prototípicamente en esos tres grupos no se codificaron para el presente factor.

Este resultado podría considerarse un anticipo del cuadro sociolingüístico que la variable lingüística ofrecerá un siglo más tarde. Así, observamos cómo el avance más significativo de *tener que* se produce en los contextos más cercanos a la oralidad, como los que se derivan de las cartas más íntimas. En el polo opuesto, se sitúan aquellas tradiciones discursivas que, aun cercanas también al polo de la inmediatez comunicativa (memorias, diarios, etc.), poseen un grado de interactividad menor, y en cuya redacción sus autores han podido tener más presente el canon de la norma escrita. Ello justificaría además la posición intermedia ocupada por las cartas no familiares, que a este respecto se sitúan a mitad de camino entre los dos extremos mencionados: por un lado, se trata de textos epistolares, con ciertas dosis, pues, de interactividad, pero, por otro, en ellas se espera un mayor control del estándar escrito, como consecuencia de la mayor distancia entre los interlocutores y una temática menos íntima.

Una centuria más tarde, este cuadro se habrá acelerado hasta el punto de reservar prácticamente la perífrasis *haber de* a los contextos escritos más formales. Con todo, se ha destacado que este proceso encuentra todavía algún freno, desde luego en la lengua escrita, pero también, en alguna medida, en la lengua oral, en algunas comunidades de habla. Así, se ha señalado, por ejemplo, que la pervivencia de esta perífrasis en algunas regiones del español de América es mayor que en el español europeo (Seco 1986: 214, De Bruyne 1993: 552), en especial como forma alternativa a la expresión del futuro (Steel 1982: 182, Westmoreland 1997: 379). Con todo, también en el seno del español peninsular encontramos datos que sugieren una mayor retención de *haber de* en algunas zonas dialectales, como consecuencia de procesos de convergencia lingüística con lenguas peninsulares en cuyo repertorio verbal dicha combinación tiene también un papel destacado. Este es el caso, por ejemplo, del catalán, lengua, además, en la que no existe la perífrasis equivalente con *tener*, como no sea la consecuencia de un castellanismo sintáctico que, aun extendido en nuestros días (*tenir que*), es abiertamente rechazado por la normativa (Hernández García 1988; Sinner 2003; Blas Arroyo 2004, 2011; Martínez Díaz 2008). Y aunque en otro sentido, lo mismo sucede con el gallego, el astur-leonés o el aragonés, lenguas en cuyo paradigma verbal encontramos ambas perífrasis, si bien no siempre con valores y usos paralelos a los del castellano (Rojo 1974; Álvarez 1983; Sinner 2003; Sánchez Pitarch 2005-2006). De hecho, en estas lenguas la perífrasis *haber (de/a) + infinitivo* muestra numerosos empleos temporales de futuro, hasta el punto de que, en algunas de ellas —el caso del gallego—, constituyen los usos más representativos, por encima de los contenidos modales deónticos (Sinner 2003).

Inicialmente, la procedencia dialectal diversa de los textos que componen el corpus nos permite examinar estas diferencias. De ahí que hayamos codificado todas las ocurrencias de la variable por regiones dialecta-

les, distinguiendo a este respecto entre variedades monolingües y bilingües. Con todo, dentro de estas últimas hemos aislado dos regiones, con el fin de evaluar la potencial influencia del contacto de lenguas en los materiales del siglo XIX. Se trata, por un lado, de las áreas del ámbito lingüístico catalán, cuya lengua autóctona comparte usos obligativos de *haber de* con el castellano, pero carece de la otra variante; y en segundo lugar, del resto de regiones bilingües, cuyas lenguas disponen en su repertorio verbal de ambas formas perifrásticas.

Conforme a lo esperado, un primer análisis de los datos muestra una mayor preferencia por *haber de* (53.8%) en las áreas bilingües que en las monolingües (46,3%), si bien estas diferencias son reducidas y no resultan significativas a la luz de la muestra analizada. Por lo demás, en el seno de las hablas bilingües apenas hay distancia entre las cifras correspondientes al ámbito lingüístico catalán (52.4%) y al resto (56.5%). Ahora bien, un examen más detenido de estos datos permite apreciar algunas diferencias reveladoras en el seno de las comunidades de habla catalana. Y es que, descontadas las ocurrencias procedentes de áreas más castellanizadas¹⁵, y restringiendo el análisis a aquellas hablas donde el proceso de conservación de la lengua autóctona ha sido mayor (desde luego en el siglo XX, pero con más motivo aún en el XIX), nos encontramos con que los ejemplos procedentes de autores catalanes son mucho más decididamente favorables a *haber de*. Así las cosas, un reanálisis del factor dialectal, en el que oponemos las ocurrencias catalanas del corpus al resto, convierte a las primeras en altamente significativas, con el primer grupo como uno de los más favorecedores a las perífrasis con *haber* (.84; 65%).

Igualmente prometedoras parecen las diferencias relacionadas con la *edad* de los autores de los textos. Para la configuración de este factor hemos realizado una interpretación de carácter émico, antes que cronológico (Chambers 1995), dadas las dificultades que encontramos para acceder a la edad exacta de muchos hablantes en el momento de la escritura. De este modo, hemos acudido a una clasificación binaria, en la que los hablantes quedan clasificados en dos grupos en función del momento vivencial por el que atravesaban cuando redactaron sus textos. Para ello, nos hemos servido tanto de lo que se dice en el interior de esos textos, o de informaciones biográficas disponibles por otras vías (prólogos, biografías, etc.), como –en especial, en el caso de las cartas– de datos indirectos, relacionados con el propio acto comunicativo.

Con los datos disponibles, observamos que el avance de *tener que* a expensas de *haber de* tiene como protagonistas principales a los más

¹⁵ Así ocurre con los ejemplos extraídos de epistolarios escritos por figuras sobresalientes de la cultura y la política valencianas, como Sorolla, Navarro Rerter o Cirilo Amorós, representantes de una burguesía que ya en el siglo XIX había abrazado el proceso castellanizador en mayor medida que la catalana.

jóvenes (.68; 64%), mientras que los genolectos más adultos se muestran más conservadores (.37; 44%), un resultado, por lo demás, en consonancia con lo observado tantas veces en los procesos de cambio lingüístico en los que compiten variantes tradicionales y novedosas (Tagliamonte 2011). A este respecto, es interesante observar la interacción entre este factor y el nivel sociocultural, un factor que, considerado de manera aislada, no ofrece diferencias significativas (con todo la clase baja supera a la alta en 14 puntos porcentuales en el uso de *tener que*), pero cuyo cruce con la edad permite vislumbrar un posible cambio desde abajo en la distribución de estas perífrasis en el siglo XIX. Así, y como puede apreciarse claramente en el gráfico siguiente, el avance entre las generaciones adultas de la perífrasis *tener que* + infinitivo se ajusta a una distribución lineal perfecta, con las clases más bajas en cabeza de dicho proceso (64%), seguidas por los niveles socioculturales intermedios (54%), y con las élites sociales como principal freno (30%) a dicha evolución. Pese a ello, las diferencias muestrales entre estos subgrupos son elevadas, por lo que hay que tomar estos resultados –además, no significativos estadísticamente– con precaución.

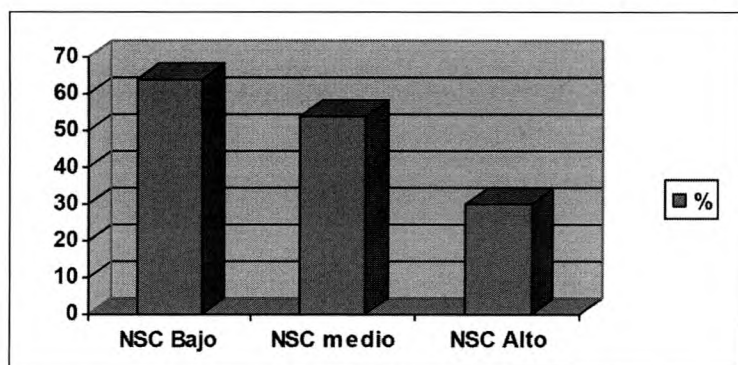


GRÁFICO 2: Distribución de los usos de *tener que* entre las generaciones adultas de tres grupos sociales (%)

Tampoco resultan significativas las diferencias relacionadas con el *sexo* de los autores, un hecho agravado aquí por el serio desequilibrio que supone la sobrerrepresentación en la muestra de los hombres (85%) sobre las mujeres (15%). Sea como sea, es digno de destacar la escasa diferencia entre los datos frecuenciales de unos (42%) y otras (46%) en el cómputo global de las variantes. Sin embargo, en el grupo de los hombres se aprecia una diferencia destacada entre los jóvenes (51%) y los más adultos (35%) en su inclinación por *tener que*, un nuevo dato que –con todas la precauciones reseñadas– contribuye a diseñar la existencia de un posible cambio desde abajo en el siglo XIX en el seno de las perífrasis modales de infinitivo.

5. CONCLUSIONES

La perspectiva variacionista en el estudio de la diacronía parte de un enfoque distinto a los tradicionales intentos de diferenciación funcional y semántica entre formas lingüísticas potencialmente alternantes en el discurso. Lejos de considerar apriorísticamente cada una de esas como portadoras de tal o cual significado “propio”, y por consiguiente, dicha alternancia como fruto de la “confusión” a la que los hablantes las habrían sometido, un estudio como el presente muestra cómo tales formas pueden conceptualizarse como variantes de una misma variable lingüística. Ciertamente, la distribución de estas variantes queda condicionada por las presiones del sistema lingüístico, estilístico y social, como hemos tenido ocasión de comprobar a través del correspondiente análisis cuantitativo; pero, al mismo tiempo, hemos constatado que las dos perífrasis aquí analizadas son susceptibles de aparecer prácticamente en los mismos contextos, de manera que los hablantes acuden a ellas como parte de una misma parcela semántica de su repertorio verbal.

En el presente caso, ese “prácticamente” es una precaución necesaria para confirmar que, en el siglo XIX, *tener que* y *haber de* compiten como dos variantes alternativas en la esfera de la modalidad, pero no así en la temporal, ámbito que queda restringido a la segunda perífrasis. En estos usos, *haber de* forma parte de otra variable sintáctica, como es la expresión de la futuridad verbal, en competencia, pues, con los futuros morfológico (*-rê*) y perifrástico (*ir a* + infinitivo).

Delimitados, pues, al dominio de la modalidad, los datos de la investigación sugieren que, ya en el siglo XIX, la variante tradicionalmente minoritaria, *tener que*, supera a su rival, tanto en términos absolutos como en muchos contextos particulares. Con todo, la representación de esta última es todavía elevada, más aún si la comparamos con las proporciones que obtiene en la actualidad, cuando aparece restringida a ciertos registros formales de la lengua escrita, y en menor medida, al habla oral de algunas regiones dialectales. En el plano estructural, el mantenimiento de *haber de* se ve favorecido por la influencia positiva que ejercen algunos contextos lingüísticos particularmente frecuentes en el discurso. Así ocurre con el presente e imperfecto de indicativo, los verbos estativos, y no pocos de los dinámicos, y los sujetos de 3ª persona del singular, entornos que figuran entre los más frecuentes en el corpus y que, junto con algunos más ocasionales –oraciones pasivas e impersonales, así como las no enunciativas–, representan los contextos estructurales más favorables a la selección de las perífrasis con *haber*. En sentido contrario, la extensión de *tener que* tiene lugar preferentemente entre las enunciativas negativas, el resto de los paradigmas de la conjugación diferentes al presente e imperfecto de indicativo, los verbos de lengua, y más aún de movimiento, así como la 1ª persona.

Por el contrario, de esos mismos datos no se deriva una influencia significativa de los matices modales, a los que se ha ceñido con tesón la bibliografía lingüística en el tratamiento de este fenómeno de variación sintáctica. Dejando al margen algunos casos aislados, en los que las perífrasis contienen valores expresivos o epistémicos, y ciñéndonos a la esfera de lo deóntico, con mucho la más vinculada a las dos perífrasis, el estudio empírico no muestra patrones de variación definidos. Ciertamente, se aprecia una mayor selección de *tener que* en los entornos semánticos de necesidad o conveniencia sentidas voluntariamente por un agente, mientras que las obligaciones de carácter interno (moral, religioso, etc.) se asocian preferentemente con *haber de*. Sin embargo, cuando esas mismas obligaciones son de agente orientado o naturaleza externa más o menos coactiva, el mayor protagonismo recae en las perífrasis con *tener*. Con todo, conviene recalcar que, además de equívocas, estas diferencias no son significativas a la luz de la muestra analizada.

Un análisis del contexto variable que rodea a las perífrasis no estaría completo sin el concurso de otros factores, de naturaleza extralingüística esta vez y que, no por menos relevantes estadísticamente, dejan de tener su importancia para explicar la variación. Este es el caso, por ejemplo, del eje estilístico, donde hemos advertido una diferencia significativa entre dos extremos, representados por las cartas de contenido más familiar e íntimo, por un lado, y los textos de naturaleza autobiográfica, por otro. La nula interactividad de estos últimos y quién sabe si la mayor conciencia lingüística de unos autores, que quizá no descartaban en su fuero interno la eventual publicación de esas memorias en el futuro, podrían explicar tales diferencias, que apuntan hacia condicionamientos diferentes de las normas características de la lengua escrita sobre estas diversas tradiciones discursivas. Por lo demás, la investigación muestra también la incidencia de algunos factores extralingüísticos, como la procedencia dialectal, con los locutores catalanes como principales arietes en el mantenimiento de *haber de*, a consecuencia, con toda seguridad, de un fenómeno de convergencia lingüística que lleva al empleo preferente de aquellos recursos que son comunes a las dos lenguas en contacto, castellano y catalán. Del mismo modo, los datos muestran una mayor difusión de la variante en expansión, *tener que*, entre los sectores más jóvenes de la población, con diferencias significativas con respecto a los más adultos. Además, la extensión preferente entre algunos sectores de estos últimos, como las clases bajas y los hombres, sugiere la posibilidad de que el siglo XIX sea el germen de un cambio desde abajo favorable a esa perífrasis que se acelerará en la siguiente centuria.

Por último, quisiéramos resaltar el valor que, para el estudio de la diacronía, puede tener el análisis de corpus cuya principal fuente son los textos cercanos al polo de la inmediatez comunicativa, y por tanto, más afines a la oralidad de otras épocas. En la práctica, datos como los que hemos

analizado aquí insinúan que el diagnóstico acerca de ciertos fenómenos de variación podría variar significativamente en función de las tradiciones discursivas empleadas como base empírica de las investigaciones. A este respecto, en otro lugar hemos sugerido ya cómo algunas diferencias en torno a la distribución de las variantes prepositiva y no prepositiva de las perífrasis con *deber* entre diversos estudios sobre idénticos cortes temporales podrían tener esa causa (cf. Balasch 2008, 2012; Blas Arroyo y Porcar en prensa). Y por lo que al presente estudio se refiere, las diferencias entre nuestros recuentos y los efectuados previamente a partir de materiales de naturaleza preferentemente formal o literaria son tan abultadas que validan aún más dicha hipótesis.

CORPUS

ÁLVAREZ BLÁZQUEZ, JOSÉ M^a (1956): “Memorias de un menestral curioso”, *Museo de Pontevedra XIII*, 61-102 (xviii).

ARBELO GARCÍA, ADOLFO I. (2011): *Correspondencia canario-americana: familia y redes sociales (siglos XVIII y XIX)*, Santa Cruz de Tenerife: Ediciones Idea.

BLASCO MARTÍNEZ, ROSA M. y CARMEN RUBALCABA PÉREZ (2003): *Para hablarte a tan larga distancia.... Correspondencia de una familia montañesa a ambos lados del Atlántico (1855-1883)*, Santander: Estudio.

BUTRÓN Y CORTÉS, J. (1897): *Carta familiar de D. José Butron y Cortés dirigida a su hijo el Guardia Marina D. Rafael Butron y Pareja*. Madrid: Librería de Fernando Fé.

-- (1897): “Extracto de las cartas recibidas por D. Emilio José Butrón acusando recibo de la carta familiar de su señor abuelo”, en E. J. Butrón y de la Serna (ed.), *Carta familiar de D. José Butron y Cortés dirigida a su hijo el Guardia Marina D. Rafael Butron y Pareja*, Madrid: Librería de Fernando Fé.

CHÁVEZ PALACIOS, JULIÁN (1998): *La pérdida de Filipinas narrada por un soldado extremeño (1896-1899). Memorias del sargento Deogracias González Hurtado*, Cáceres: Editoría Regional de Extremadura.

DE PAZ SÁNCHEZ, MANUEL (2002): “Al recibo de la presente’: Algunas cartas personales de emigrantes canarios del siglo XIX”, *Tebeto: Anuario del Archivo Histórico Insular de Fuerteventura*, 15, 493-505.

FONTANELLA DE WEINBERG, M. (1993): *Documentos para la historia lingüística de Hispanoamérica*, Vol. I, Madrid: BRAE.

GENER, EDUARDO (1976): *Diario de viaje de un comerciante gaditano (1829)*, Cádiz: Instituto de Estudios Gaditanos.

GENOVÉS AMORÓS, VICENTE (1981): *El epistolario de Navarro Reverter con Cirilo Amorós*, Valencia: Federico Domenech.

GIMENO PUJOL, MARÍA D. (2010): *El epistolario de José Nicolás de Azara (1730-1804). Estudio, edición y notas*, Madrid: Castalia (selección de cartas del periodo 1800-1804).

GÓMEZ DE AVELLANEDA, GERTRUDIS (1989): *Poesías y epistolario de amor y de amistad*, Madrid: Castalia.

GÓMEZ SEIBANE, SARA (2011): “Textos para la historia del castellano en Galicia”, *Moenia* 17, 367-420.

IBARRA Y MANZONI, AURELIANO (1995): *Diario de mi prisión. Desde el día 8 de Noviembre de 1866 hasta el día 21 de Marzo de 1867*, Elx: Ajuntament d'Elx.

ISASI, CARMEN Y JOSÉ LUIS RAMÍREZ LUENGO (2013): *Una muestra documental del castellano norteño en el siglo XIX*, Lugo: Axac.

LÓPEZ ÁLVAREZ, JUACO (2000): "Cartas desde América: La emigración de asturianos a través de la correspondencia: 1864-1925", *Revista de Dialectología y Tradiciones Populares*, LV/ 1: 81-120.

MADRAZO, JOSÉ (1998): *José de Madrazo: Epistolario*, Madrid: Fundación Marcelino Botín.

MÁRQUEZ MACÍAS, R. (1994): *Historias de América. La emigración española en tinta y papel*, Huelva: Ediciones Ertoil.

MARTÍNEZ, M^a LUISA (2007): *Noticias desde Cuba: cartas de emigrantes vallisoletanos en la segunda mitad del siglo XIX*, Valladolid: Universidad de Valladolid.

MARTÍNEZ MARTÍN, LAURA (2010): 'Asturias que perdimos, no nos pierdas'. *Cartas de emigrantes asturianos en América (1863-1936)*, Gijón: Museu del Pueblu d'Asturies.

MARTÍNEZ MARTÍNEZ, CARMEN (2007): *Desde la otra orilla: cartas de Indias en el Archivo de la Real Chancillería de Valladolid*, León: Universidad de León.

OSSÓ, ENRIQUE DE Y GLORIA VOLPE (1997): *Cartas de San Enrique de Ossó y Cervelló al Dr. Félix Sardá y Salvany*, Barcelona: STJ.

OTAZU, FAUSTO (1995): *Fausto de Otazu a Inigo Ortes de Velasco: cartas, 1834-1841*, Álava: Diputación Foral de Álava.

PÉREZ MURILLO, M^a DOLORES (1999): *Cartas de emigrantes escritas desde Cuba*, Cádiz: Universidad de Cádiz.

PINO Y ROZAS, JOAQUÍN (2001): *Un epistolario virreinal: cartas familiares de Don Joaquín del Pino Rozas, VIII Virrey del Río de la Plata*, Madrid: Fundación Rafael del Pino.

PRADO, MERCEDES (1994): *Dario de Regoyos: sus cartas inéditas*, Bilbao: Bere-kintza.

RUBALCABA PÉREZ, CARMEN (2006): "Libros de cuentas de Pedro Jado", *Entre las calles vivas de las palabras*, Santander: Ediciones Trea.

TOMÁS, FACUNDO (2007): *Epistolarios de Joaquín Sorolla*. Vol. 1. *Correspondencia con Pedro Gil Moreno de Mora*, Barcelona: Anthropos.

REFERENCIAS BIBLIOGRÁFICAS

ÁLVAREZ CACCAMO, CELSO (1983): "A influencia do galego sobre o sistema verbal e no uso de certas perífrases verbais do castelán de Galiza", *Grial*, 82, 423-442.

BALASCH, SONIA (2008): "Debe (de) ser: evolución de la variación", en M. Westmoreland y J. A. Thomas (eds.), *Selected Proceedings of the 4th Workshop on Spanish Sociolinguistics*, Somerville MA: Cascadilla Proceedings Project, 109-119.

— (2012): "Revisión amplia de la variación *deber (de) + Infinitivo* en la diacronía del español", en Emilio Montero (ed.), *Actas del VIII Congreso Internacional de Historia de la Lengua Española*, I, Meubook, 607-620.

BELLO, ANDRÉS (1859): *Gramática de la lengua castellana destinada al uso de los americanos*, Valparaíso: Universidad de Chile.

BLAS ARROYO, JOSÉ LUIS (2004): "El español actual en las comunidades del ámbito lingüístico catalán", en R. Cano (coord.), *Historia de la Lengua Española*, Barcelona: Ariel, 1065-1086.

- (2005): *Sociolingüística del español*, Madrid: Cátedra.
- (2011): “Spanish in Contact with Catalan”, en M. Díaz Campos (ed.), *Handbook of Hispanic Sociolinguistics*, Oxford: Blackwell, 374-394.
- (en prensa): “Tras las huellas de la variación y el cambio lingüístico a través de textos de inmediatez comunicativa: fundamentos de un proyecto de sociolingüística histórica”, *Actas del IX Congreso Internacional de Historia de la lengua española*, Cádiz, Septiembre 2012.
- y JUAN GONZÁLEZ (2014a): “La alternancia *deber/deber de* + infinitivo en el siglo XVI. Factores condicionantes en un fenómeno de variación sintáctica a partir de un corpus epistolar”, *Spanish in Context*, 11,1, 76-96.
- y JAVIER VELLÓN (2014): “En los entresijos de la gramática: la supervivencia de *deber de* + inf. en el español moderno (siglos XVIII-XX)”, *Revista de Filología Española*, XCIV (1), 9-38.
- y MARGARITA PORCAR (2014): “Patrones de variación y cambio en la sintaxis del Siglo de Oro. Un estudio variacionista de dos perífrasis modales en textos de inmediatez comunicativa”, *RILCE. Revista de Filología Hispánica* (en prensa).
- y JUAN GONZÁLEZ (2014b): “*Qué tengo que/de hacer?*”: variación y cambio lingüístico en el seno de las perífrasis de infinitivo a partir de textos escritos de impronta oral en el español clásico”, *Studies in Hispanic and Lusophone Linguistics*, 7.2., 241-274.
- CHAMBERS, JACK. K. (1995): *Sociolinguistic Theory*, Oxford: Blackwell.
- DAVIES, MARK (2006): *A frequency dictionary of Spanish*, New York: Routledge.
- DE BRUYNE, JACQUES (1993): *Spanische Grammatik*, Tübingen: Niemeyer.
- FERNÁNDEZ DE CASTRO, FÉLIX (1999): *Las perífrasis verbales en el español actual*, Madrid: Gredos.
- GARACHANA CAMARERO, MAR (1997): “Acerca de los condicionamientos cognitivos y lingüísticos de la sustitución de ‘aver’ por ‘tener’”, *Verba. Anuario Galego de Filoloxía*, 24, 203-235.
- y MALTE ROSEMEYER (2011): “Rutinas léxicas en el cambio gramatical. El caso de las perífrasis deónticas e iterativas”, *Revista de Historia de la Lengua Española*, 6, 35-60.
- y ESTHER ARTIGAS (2012): “Corpus digitalizados y palabras gramaticales”, *Scriptum Digital*, 1, 37-65.
- GARCÍA FERNÁNDEZ, LUIS (dir.) (2006): *Diccionario de perífrasis verbales*, Madrid: Gredos.
- (2012): *Las perífrasis verbales en español*, Madrid: Castalia.
- GILI GAYA, SAMUEL (1970¹³): *Curso superior de sintaxis española*, Barcelona: Vox.
- GÓMEZ TORREGO, LEONARDO (1988): *Las perífrasis verbales*, Madrid: Arco/Libros.
- (1999): “Los verbos auxiliares. En I. Bosque y V. Demonte (dirs.), *Gramática Descriptiva de la Lengua Española*, III, Madrid: Espasa-Calpe, 3223-3390.
- GUTIÉRREZ ORDÓÑEZ, SALVADOR (1980): “Tengo que vender unos libros / Tengo unos libros que vender”, *Verba*, 7, 389-396.
- HERNÁNDEZ DÍAZ, AXEL (2006): “Posesión y existencia. La competencia de *haber* y *tenery haber* existencial”, en C. Company (coord.), *Sintaxis histórica de la lengua española*, Vol. 1, Tomo II, México: Fondo de Cultura Económica, 1055-1164.
- HERNÁNDEZ GARCÍA, CARMEN (1998): “Una propuesta de clasificación de la interferencia lingüística a partir de dos lenguas en contacto: el catalán y el español”, *Hesperia*, 1, 61-80.

- KANY, CHARLES E. (1969): *Sintaxis hispanoamericana*, Madrid: Gredos.
- KENISTON, HAYWARD (1936): "Verbal Aspect in Spanish", *Hispania*, 19, 163-176.
- LAPESA, RAFAEL (2000): *Estudios de morfosintaxis histórica del español*, ed. de R. Cano y M^a T. Echenique, Madrid: Gredos.
- LÓPEZ IZQUIERDO, MARTA (2008): "Las perífrasis modales de necesidad: emergencia y renovación", en C. Company y J. G. Moreno (coords.), *Actas del VII Congreso Internacional de Historia de la Lengua española*, I, Madrid: Arco/Libros, 789-806.
- MARSÁ, FRANCISCO (1986): *Diccionario normativo y guía práctica de la lengua española*, Barcelona: Ariel.
- MARTÍNEZ DE NOBOA, ANDRÉS (1839): *Nueva Gramática de la lengua castellana según los principios de la filosofía gramatical*, Madrid: Eusebio Aguado.
- MARTÍNEZ DÍAZ, EVA (2003): "La frecuencia de uso de *haber y tener* en las estructuras perifrásticas de obligación. Algún fenómeno de variación en el español de Cataluña", *Interlingüística*, 14, 681-694.
- (2008): "La expresión de la modalidad de obligación en el Corpus del español conversacional de Barcelona y su área metropolitana (2001)", en A. Moreno (coord.): *El valor de la diversidad (meta)lingüística. Actas del VIII congreso de Lingüística General*. Referencia: <<http://www.lllf.uam.es/clg8/actas/index.html>>.
- OLBERTZ, HELLA (1998): *Verbal periphrases in a Functional Grammar of Spanish*, Berlin: Mouton de Gruyter.
- OESTERREICHER, WULF (2004): "Textos entre inmediatez y distancia comunicativas. El problema de lo hablado-escrito en el Siglo de Oro", en R. Cano (coord.), *Historia de la lengua española*, Barcelona: Ariel, 729-769.
- PORCAR, MARGARITA (2012): "Lenguas en contacto en la homilética del XVIII. El caso de las perífrasis de modalidad deóntica", *Actas del IX Congreso Internacional de Historia de la Lengua española*, Cádiz, Septiembre 2012 (en prensa).
- POUNTAIN, CHRISTOPHER J. (2001): *A History of the Spanish Language through Texts*, New York: Routledge.
- REAL ACADEMIA ESPAÑOLA (1870): *Gramática de la lengua castellana*, Madrid: Rivadeneyra.
- (1973): *Esbozo de una nueva gramática de la lengua española*, Madrid: Espasa-Calpe.
- ROCA PONS, JOSÉ (1980): *Introducción a la gramática*, Madrid: Vergara Ed.
- ROJO, GUILLERMO (1974): *Perífrasis verbales en el gallego actual*, *Verba. Anuario Galego de Filoloxía*, Anejo 2.
- SALVÁ, VICENTE (1835): *Gramática de la lengua castellana según ahora se habla*, París-Valencia, edición autor.
- SÁEZ-GODOY, LEOPOLDO (1968) "Algunas observaciones sobre la expresión del futuro en español", en *Actas del XI Congreso Internacional de Lingüística y Filología Románicas*, IV, Madrid: CSIC, 1875-1889.
- SÁNCHEZ PITARCH, FERNANDO (2005-2006): "Breve descripción del aragonés de Agüero", *Archivo de Filología Aragonesa*, Vol. 61-62, 185-218.
- SECO, MANUEL (1986): *Diccionario de dudas y dificultades de la lengua española*, Madrid: Espasa-Calpe.
- SECO, RAFAEL (1967⁹): *Manual de gramática española*. Madrid: Aguilar.
- SEIFERT, EVA (1930): "*Haber y tener* como expresiones de la posesión en español", *Revista de Filología Española*, XVII, 233-276 y 345-385.
- SINNER, CARSTEN (2003): "Valores y usos de *haber de + infinitivo y tener que + infi-*

nitivo en diferentes variedades del castellano”, en C. Pusch y A. Wesch. (eds.), *Perífrasis verbals en les llengües (ibero-)romàniques*, Madrid/Frankfurt: Iberoamericana-Vervuert, 193-206.

STEEL, BRIAN (1982): “Algunos apuntes para un nuevo diccionario de americanismos”, en M. Perl (ed.), *Estudios sobre el léxico del español en América*, Leipzig: Enzyklopädie, 176-197.

STENGAARD, BIRTE (2006): “Función y evolución de la construcción *aver* + preposición + infinitivo”, en J. L. Girón y J. J. de Bustos (coords.), *Actas del VI Congreso Internacional de Historia de la Lengua Española*, Vol. 1, Madrid: Arco/Libros, 1147-1152.

TAGLIAMONTE, SALI (2011): *Variationist Sociolinguistics: Change, Observation, Interpretation*, Oxford: Wiley-Blackwell Publishers.

WESCH, ANDREAS (1997): “El castellano hablado en Barcelona y el influjo del catalán. Esbozo de un programa de investigación”, *Verba. Anuario Galego de Filoloxía*, 24, 287-312.

WESTMORELAND, MAURICE (1997): “The dialectalization of Spanish future tense usage”, *Word*, 48/3, 375-395.

YLLERA, ALICIA (1980): *Sintaxis histórica del verbo español: Las perífrasis medievales*, Zaragoza: Universidad de Zaragoza.

ZAMORA VICENTE, ALONSO (1960): *Dialectología española*, Madrid: Gredos.