

REVISTA  
DE  
FILOLOGÍA ESPAÑOLA

---

Volumen XCI N.º 1 enero-junio 2011 Madrid (España) ISSN: 0210-9174

---

*DEBER (DE) + INFINITIVO: ¿UN CASO DE  
VARIACIÓN LIBRE EN ESPAÑOL? FACTORES  
CONDICIONANTES EN UN FENÓMENO  
DE ALTERNANCIA SINTÁCTICA*

JOSÉ LUIS BLAS ARROYO  
Universitat Jaume I

1. INTRODUCCIÓN

Las manifestaciones de variación gramatical que aparecen en los ejemplos siguientes representan uno de los casos más palmarios de enfrentamiento entre las prescripciones normativas y el uso lingüístico común. En efecto, en enunciados como los de (1) y (2), extraídos del *Macrocorpus sociolingüístico de Castellón y sus comarcas (MCSCS)*, que sirve como base empírica al presente estudio, las autoridades normativas llevan tiempo insistiendo en que las variantes del verbo *deber* con y sin preposición “deberían” servir para la expresión de significados modales, y más en concreto, para la distinción entre valores epistémicos y deónticos, respectivamente. Y en efecto, de acuerdo con estas reglas, en (1) el hablante intenta recordar cuántos alumnos había en su clase de párvulos cuando era niño, y aventura una cifra (*el doble de los que habían realmente en aquel colegio...*), no sin las lógicas dosis de conjetura dado el tiempo transcurrido. El empleo normativo de la perífrasis en incidencia indirecta (*deber de*) en este caso es correlato del que utiliza el informante de (2), esta vez sin preposición, al argumentar que las madres deberían —estarían “obligadas a”— dejar de trabajar durante los primeros meses de la vida de un niño.

(1) [...] y la época que: umm/ que nosotros estábamos allí *debían de haber* pues el doble de niños de los que habían realmente en aquel colegio (*MCSCS*, 18).

(2) [...] pues yo pienso que *debería dejar de trabajar* los meses que se necesita mucho a una madre, los primeros meses, a lo mejor hasta un año (*MCSCS*, 129).

Ahora bien, con la misma claridad con que la Academia y otras instituciones normativizadoras defienden como necesaria esta distinción funcional, en las últimas décadas se ha reiterado desde diferentes perspectivas teóricas y metodológicas que la pretendida diferenciación modal opera hasta cierto punto en la lengua escrita —y a menudo tampoco—, pero no así en la comunicación oral. En la práctica, y de regreso al corpus del que obteníamos los ejemplos anteriores, el mismo hablante que en (1) se servía de la perífrasis con preposición para la expresión de la probabilidad, acudía un poco más adelante —en (3)— a la otra variante con idéntico valor modal. Y como contrapartida, el informante que utilizaba la combinación *deber* + infinitivo para manifestar la obligación de las madres de permanecer con sus hijos emplea a renglón seguido la perífrasis alternativa para continuar con su argumentación, como podemos apreciar en (4).

(3) Y: entonces la— los párvulos sí que me acuerdo [...] *debíamos ser* lo menos cincuenta o sesenta en aquella clase de párvulos (MCSCS, 18).

(4) [...] si tú tienes un hijo con otra persona y trabajan las dos personas, pienso que *deben de cuidar* los dos, la pareja, de su hijo ¿no? (MCSCS, 129).

Por otro lado, algunos lingüistas han llamado la atención acerca de una tendencia constante —y en algunos casos muy avanzada— a la desaparición de este fenómeno de variabilidad sintáctica, del que, como resultas, tan solo quedaría la perífrasis sin preposición. De ser ello cierto, dicha evolución habría alcanzado su punto más avanzado en el español de América, mientras que las variedades a este lado del Atlántico mostrarían un comportamiento algo más conservador, aunque presidido por tendencias similares.

Pese a lo anterior, muy poco es lo que sabemos acerca de la posible incidencia de otros factores lingüísticos, estilísticos o sociales que puedan servir como revulsivo, o en caso contrario, como freno a las tendencias anteriormente expuestas. Descartada la oposición modal epistémico-deóntica como explicativa de la distinción entre las dos perífrasis verbales, cabe preguntarse si, como reza el título de este artículo, nos encontramos, efectivamente, ante un mero hecho de variación libre, o si, por el contrario, factores de diversa naturaleza pueden contribuir a explicar las tendencias que a este respecto muestra el español en el presente estadio de lengua. Muy probablemente a esta falta de conocimiento contribuye la parquedad con que operan en el discurso las manifestaciones de la variación sintáctica recogidas en corpus de lengua oral. Y más aún si, como ocurre en este caso, los hablantes disponen de otras perífrasis, como es el caso de *tener que*, *haber de* o *haber que*, con similares matices significativos que las que nos ocupan en estas páginas. De ahí que cuando se dispone de datos, estos sean, por lo general, escasos y poco representativos, y a menudo limitados a un determinado sociolecto, lo que impide ir más allá que el mero recuento de fre-

cuencias y el señalamiento de tendencias generales a partir de la comparación de cifras no siempre homologables.

El presente estudio pretende contribuir a llenar ese vacío en la interpretación de este fenómeno de variación sintáctica. Para ello presentamos los datos de una investigación empírica a partir de los materiales contenidos en el *Macrocorpus sociolingüístico de Castellón y sus comarcas (MCSCS)*, un extenso corpus de habla oral y representativo de todo el espectro social en las comunidades de habla castellanenses. Ello permite la aplicación de métodos estadísticos multivariantes y, por consiguiente, una considerable afinación a la hora de interpretar los datos y analizar su relevancia.

A partir de dicho estudio, confirmamos que la variación que nos ocupa no se halla relacionada con la pretendida distinción modal, pero sí con otros hechos de modalización, como los mecanismos de intensificación y atenuación, cuyos efectos se hacen sentir en direcciones opuestas: los primeros, frenando el proceso de pérdida de la preposición, que, sin embargo, es acelerado por los segundos. Estos factores lingüísticos y comunicativos, junto con otros de menor relevancia y significación, serán examinados con detalle en las páginas siguientes, que quedan estructuradas de la siguiente manera. En el epígrafe 2 realizamos un breve repaso acerca de la historia de esa variación, al que seguirá en § 3 la descripción del aparato metodológico empleado en el estudio. En el largo apartado 4 examinaremos pormenorizadamente los datos obtenidos tras el análisis empírico, en el que desempeñan un papel relevante tanto los resultados del aparato estadístico multivariante como las potenciales interacciones entre los diversos factores considerados. Por último, en el capítulo 5 sintetizaremos las principales conclusiones derivadas de la investigación.

## 2. HISTORIA Y PRESENTE DE UN HECHO DE VARIACIÓN SINTÁCTICA EN ESPAÑOL

Antes de comenzar este repaso de las tendencias evolutivas que el mencionado fenómeno de variación ha experimentado en español, debemos dejar claro que no terciaremos en la tradicional polémica acerca del carácter perifrástico o no de las combinaciones con el verbo *deber* + infinitivo (sin preposición), dado que la solución a este problema, por destacado que sea desde el punto de vista gramatical, no contradice su alternancia ubicua con la perífrasis con *de* —esta vez sí, reconocida por todos—. Valga con recordar que, en relación con este tema, se enfrentan gramáticos como Gili Gaya —quien excluía *deber* + infinitivo del catálogo de perífrasis—<sup>1</sup> y otros lingüistas que, más recientemente,

<sup>1</sup> «El verbo *deber* en su acepción propia de ‘hallarse obligado’, ‘tener obligación’, no es auxiliar; *deber* + infinitivo no es, por consiguiente, una frase verbal de las que estamos estudiando, en el uso que los gramáticos tienen por correcto. En cambio, *deber de* + infinitivo altera el sen-

han avalado su interpretación perifrástica, amparándose para ello en criterios fundamentalmente sintácticos<sup>2</sup>.

Perífrasis o no —aunque nosotros nos serviremos del término en estas páginas— lo cierto es que las dos construcciones verbales se vienen utilizando desde hace siglos para la expresión tanto de la modalidad deóntica, relacionada con la obligación y/o la necesidad, como de la epistémica, vinculada a los contenidos conjeturales<sup>3</sup>. En uno de los estudios más detallados sobre el tema desde el punto de vista histórico, Yllera calificaba la distinción académica como superficial, ya que en su opinión «no existe diferencia entre el sentido de *dever* y el de *dever de* durante la Edad Media e incluso posteriormente ambos pueden expresar obligación o probabilidad»<sup>4</sup>. De este modo, si en la Edad Media el significado principal de *dever* + infinitivo era indicar ‘obligación, necesidad o conveniencia’, ya en el siglo XV también era relativamente frecuente su empleo para expresar suposición o conjetura. Al mismo tiempo, la mencionada autora encontraba testimonios tan tempranos como el siglo XIII acerca de la modalidad deóntica expresada mediante la perífrasis *deber de*. Por su parte, un observador agudo como Cuervo —partidario, no obstante, de mantener las distinciones normativas por su valor funcional— indicaba que, por analogía con otras perífrasis obligativas, *deber* admitió desde siempre las preposiciones *a* o *de*, y a partir del siglo XIV, preferentemente la segunda de ellas. Y concluía: «arraigado el uso de

---

tido propio del verbo *deber* y lo convierte, por lo tanto, en auxiliar» (Samuel Gili Gaya, *Curso superior de sintaxis española*. Barcelona, Bibliograf, 1961, pág. 113). Por su parte, otros autores han acudido al concepto de verbo semiauxiliar para dar cuenta del verbo *deber* en sus combinaciones sin preposición, con lo que este aparece situado en un paradigma diferente al de la perífrasis con *de* (vid. José Roca Pons, *Estudios sobre las perífrasis verbales del español*, Madrid, Real Academia Española, 1958; Hertha Olbertz, *Verbal Periphrasis in a Functionnal Grammar of Spanish*, New York, Mouton de Gruyter, 1998).

<sup>2</sup> Estos indicios incluyen el comportamiento similar al de otras perífrasis de infinitivo, como la transformación pasiva total o la interrogativa con *qué*, la imposibilidad de transformar la perífrasis en una construcción ecuacional o la conmutación del infinitivo por un elemento equivalente funcionalmente, dado que no se trata de un complemento del verbo auxiliar (para el desarrollo y ejemplificación de estos y otros argumentos en la misma línea, vid. Pilar Gómez Manzano, *Perífrasis verbales con infinitivo (valores y usos en la lengua hablada)*, Madrid, UNED, 1992; Magnoña Troya, *Perífrasis verbales de infinitivo en la norma lingüística culta de Las Palmas de Gran Canaria*, Anejo LVI del BRAE, Madrid, Real Academia Española y Universidad Palmas de Gran Canaria, 1998; Leonardo Gómez Torrego, «Los verbos auxiliares. Las perífrasis verbales de infinitivo», *Gramática Descriptiva de la Lengua Española*, Ignacio Bosque y Violeta Demonte, dirs., Madrid, Espasa Calpe, 1999, págs. 3.223-3.390; Francisco Fernández de Castro, *Las perífrasis verbales en el español actual*, Madrid, Gredos, 1999; Luis García Fernández, dir., *Diccionario de perífrasis verbales*, Madrid, Gredos, 2006.

<sup>3</sup> César Hernández, *Gramática funcional del español*, Madrid, Gredos, págs. 386-387; Pilar Díez de Revenga y Dolores Igualada, «Construcciones modales en documentos notariales del siglo XIII», *Glossae*, 2, 1989-1990, págs. 119-132; Alicia Yllera, *Sintaxis histórica del verbo español: Las perífrasis medievales*, Zaragoza, Universidad de Zaragoza, 1980.

<sup>4</sup> Yllera, *op. cit.*, pág. 140.

*deber de* en el sentido de *haber de* para indicar necesidad, se extendió la analogía a los casos en que éste denota inducción, conjetura, duda»<sup>5</sup>.

Ahora bien, como ha subrayado Balasch<sup>6</sup> recientemente, sorprende la escasa atención que se ha prestado a los patrones evolutivos de esa aparente «confusión» entre las dos perífrasis, y, más concretamente, al grado en que diversos factores han podido influir históricamente en la distribución de las dos variantes con los valores modales reseñados. Mediante la aplicación de una metodología variacionista a textos antiguos, esta autora ha comprobado, por ejemplo, como aparte de algunas preferencias estilísticas significativas en algunos casos, entre los siglos XVII y XIX la variante *deber* + infinitivo no solo se extiende considerablemente a expensas de la variante con preposición, sino que —y lo que es más relevante— reemplaza a esta última, primero en los contextos deónticos y luego en los epistémicos.

De vuelta al presente, en la actualidad compiten dos actitudes hacia este fenómeno de variación sintáctica. Por un lado, se encuentra la de aquellos gramáticos que, de una forma más o menos explícita, apuestan por mantener la distinción funcional entre las dos perífrasis, cuyo origen preceptivo hay que remontar al *Diccionario de Autoridades*. En tiempos más recientes, el *Esbozo* (1973) académico reconocía la confusión entre las dos construcciones tanto en el uso corriente actual como en épocas previas, pero seguía apostando por conservar la oposición. Una recomendación que ha sido asumida posteriormente por gramáticas e instituciones normativizadoras diversas, así como por los modernos manuales de estilo de algunos medios de comunicación (v. g. *El País*, *ABC*, *TVE*, *Telemadrid*) que condenan la no diferenciación. Representante de esta actitud es Gómez Torrego quien ha señalado que esta confusión se halla en parte relacionada con el hecho de que «pocos hablantes conocen la diferencia entre estas dos construcciones perífrásticas, por lo que suelen emplearlas indistintamente [...]. Hay que hacer un esfuerzo para mantener esta diferencia, pues así se evitará incurrir en posibles ambigüedades»<sup>7</sup>. Con todo, un poco más adelante este mismo autor matiza que «en la lengua oral (¡nunca en la escrita!) se puede permitir la omisión de la preposición de en los casos de posibilidad. No se justifica, por el contrario, intercalar de en los casos de obligación» (las cursivas son nuestras)<sup>8</sup>. Más recientemente, en la *Nueva Gramática de la lengua español-*

<sup>5</sup> Rufino José Cuervo, *Diccionario de construcción y régimen de la lengua castellana*, Bogotá, Instituto Caro y Cuervo, 1893/1994, pág. 808.

<sup>6</sup> Sonia Balasch, «*Debe (de) ser*: evolución de la variación», *Selected Proceedings of the 4th Workshop on Spanish Sociolinguistics*, Maurice Westmoreland y Juan Antonio Thomas, eds., Somerville, MA, Cascadilla Press, 2008, págs. 109-119.

<sup>7</sup> Leonardo Gómez Torrego, *Manual de español correcto. Vol. II*, Madrid, Arco/Libros, 1989, págs. 188-189.

<sup>8</sup> Gómez Torrego, *op. cit.*, págs. 189. Véanse en el mismo sentido, favorable a mantener la distinción funcional entre las dos perífrasis: Rafael Seco, *Manual de gramática española*, Madrid,

la<sup>9</sup>, las Academias de la lengua española navegan entre dos aguas en relación con este tema, al complementar las indicaciones gramaticales —«suele analizarse ‘deber + infinitivo’ como modal radical y ‘deber de + infinitivo’ como epistémico» (pág. 2.143)— y normativas —«para expresar obligación se recomienda la variante sin preposición» (pág. 2.144)—, con el reconocimiento de la notable variación existente en el mundo hispánico —«existen [...] numerosas excepciones a esta regla general» (pág. 2.143)— y eso que las ejemplificaciones parten en todos los casos de textos escritos y no orales.

No en vano, incluso quienes abogan por mantener la distinción como un recurso funcional óptimo para deshacer ambigüedades, reconocen la frecuente confusión que se produce en la lengua cotidiana. A este respecto los autores del *Diccionario de perfrasis* recuerdan que «la mayor parte de los hablantes utilizan ambas construcciones indistintamente, aunque algunos de ellos tiendan a utilizar (deber + infinitivo) como epistémica, pero no (deber de + infinitivo) como deóntica»<sup>10</sup>. Estas impresiones han sido puestas a prueba empíricamente en los últimos años por diversas investigaciones, en las que se recogen las tendencias que operan en distintas comunidades de habla del mundo hispánico. Más aun, autoras como Gómez Manzano señalan a partir de esos datos —en su caso extraídos del habla culta de Madrid— que:

el análisis de estos casos *no nos ha permitido descubrir ninguna razón o norma para saber cuándo es más frecuente el uso con preposición o sin ella*. Hay un claro predominio de los casos sin preposición, *pero podemos hablar de variación libre en cuanto al empleo de una u otra*, ya sea para expresar la obligatoriedad ya sea para expresar la hipótesis o probabilidad (la cursiva es nuestra)<sup>11</sup>.

Una opinión que, implícitamente, viene a refrendar Fernández de Castro desde otra óptica, cuando de forma retórica se plantea el siguiente interrogante: «¿tiene sentido hablar de ‘confusión’ entre dos cosas que en realidad nunca han sido distintas?»<sup>12</sup>.

---

Aguilar, 1967, págs. 172-173; Manuel Seco, *Diccionario de dudas y dificultades de la lengua española*, Madrid, Espasa Calpe, 1986, pág. 136. Por otro lado, el lector interesado puede encontrar testimonios muy ilustrativos sobre la condena hacia la confusión que, desde el siglo XIX en adelante, se han propalado desde diversos ámbitos académicos y normativos en José Samper, Clara Hernández y Magnolia Troya, «El uso de deber + infinitivo y deber de + infinitivo en la norma lingüística culta de América y España», *Boletín de Filología (Homenaje a Ambrosio Rabanales)*, XXXVII, 1998-1999, pág. 1.061, y Fernández de Castro, *op. cit.*, pág. 174.

<sup>9</sup> Real Academia Española/Asociación de Academias de la Lengua Española, *Nueva gramática de la lengua española*, Madrid, Espasa Calpe, 2009.

<sup>10</sup> García Fernández, *op. cit.*, pág. 113.

<sup>11</sup> Gómez Manzano, *op. cit.*, pág. 162.

<sup>12</sup> Fernández de Castro, *op. cit.*, pág. 175.

Los estudios empíricos acerca de la variación entre nuestras dos perífrasis rara vez han ido más allá de la consideración de la eventual —y casi siempre desmentida— oposición entre las modalidades deóntica y epistémica. Con todo, en la bibliografía se han aventurado algunas hipótesis para explicar, al menos parcialmente, la “confusión” entre las dos variantes. Náñez<sup>13</sup>, por ejemplo, recuerda como en la lengua alternan a veces locuciones verbales con un significado idéntico o muy parecido con la preposición y sin ella (*saber (de) matemáticas*), y que la Academia considera correctas. Por otro lado, el empleo de la preposición *de* se vería impulsado por el hecho de que «al quedar la frase más llena con *de (debe de + infinitivo)* indirectamente resulta más expresiva y adecuada para la expresión de la obligación» (pág. 242), una solución que, no obstante, este autor censura, frente a la mayor flexibilidad mostrada hacia los usos epistémicos de *deber + infinitivo*<sup>14</sup>. Por su parte, otros autores han avalado las tesis que defendiera en su día Erica García<sup>15</sup> en su conocida interpretación sobre el (de)queísmo, y que, para lo que aquí nos interesa, supone justificar la extensión de la perífrasis sin preposición por el mayor esfuerzo cognitivo que supone insertar un elemento que no hacerlo<sup>16</sup>. Por otro lado, desde este punto de vista, la mayor perceptibilidad de la partícula conduce a su estigmatización en mayor medida que la que supondría omitir esa misma unidad en los contextos en que se recomienda su uso.

Pese al interés de estos y otros argumentos similares, su falta de concreción en estudios empíricos ha impedido confirmar su validez en el habla real. Por lo demás, muy poco es lo que sabemos acerca de la posible relación entre este hecho de variabilidad y otros factores lingüísticos y comunicativos como los que analizaremos en estas páginas. Algunos se han ensayado en el pasado, pero las muestras obtenidas son a menudo insuficientes para garantizar la representatividad de los resultados. Otros, como los diferentes factores de modalización que en el presente estudio se revelan determinantes, no han sido objeto de atención en el pasado, al menos hasta donde llega nuestro conocimiento. Por otro lado, pocas investigaciones han asumido el reto de evaluar con las herramientas metodológicas apropiadas los contextos que favorecen, o en sentido contrario, frenan la aparición de una determinada variante.

<sup>13</sup> Emilio Náñez, «Sobre dequeísmo», *Revista de Filología Románica*, 2, 1984, págs. 239-248.

<sup>14</sup> En este último sentido, véase también la opinión de Gómez Torrego, *op. cit.*, pág. 189.

<sup>15</sup> Erica C. García, «El fenómeno (de)queísmo desde una perspectiva dinámica del uso comunicativo de la lengua», *Actas del II Congreso Internacional sobre el español de América*, José Moreno de Alba, ed., México, UNAM, 1986, págs. 46-65, pág. 53.

<sup>16</sup> *Vid.* Samper, Hernández y Troya, art. cit., pág. 1.060; José Ramón Gómez Molina, «Estudio sociolingüístico de la perífrasis *deber + infinitivo* y *deber de + infinitivo* en el español hablado», *Lengua viva. Estudios ofrecidos a César Hernández Alonso*, Antonio Álvarez Tejedor, ed., Valladolid, Universidad de Valladolid, 2009, págs. 867-880.

### 3. METODOLOGÍA

Como recuerdan Samper *et al.*<sup>17</sup> en su estudio acerca de este fenómeno de variación en la lengua culta de diversas ciudades del mundo hispánico, las posibilidades de poner a prueba determinadas hipótesis se ven considerablemente limitadas en algunas muestras, ya que el número de frecuencias obtenidas es muy bajo. Así, nada menos que en ocho de las doce urbes analizadas se contabilizan menos de seis ejemplos de la perífrasis *deber de* + infinitivo, lo que lleva a estos autores a afirmar que se trata de cifras «que evidentemente nos impiden conocer el empleo de esta construcción en esas comunidades, si bien es verdad que la pobreza de ejemplos es en sí misma una buena muestra de la escasez de su uso»<sup>18</sup>. Por otro lado, buena parte de estos trabajos empíricos se han centrado en el estudio de los sociolectos más cultivados de la sociedad, lo cual no significa que sus resultados sean necesariamente extensibles a otros segmentos de la comunidad. De hecho, en algún caso en que se ofrece la posibilidad de comparar cifras de unos y otros estratos sociales, llama la atención la escasa coincidencia de los datos. Así, por ejemplo, el reducido porcentaje obtenido en las encuestas de la norma culta de Caracas en el estudio de Samper *et al.*<sup>19</sup> por la perífrasis en incidencia directa (*deber*) para la expresión de la modalidad epistémica (15,9%), difiere considerablemente del que, para el conjunto de esa misma comunidad de habla, han obtenido Galué *et al.*, con frecuencias relativas mucho más elevadas para ese mismo valor (48%) (véase más adelante la tabla 1)<sup>20</sup>.

En la práctica son pocos los trabajos que han analizado las correspondientes comunidades de habla en su conjunto, y aun en algunas de ellas, las cifras que se obtienen son tan limitadas que resulta complicado extraer conclusiones sólidas, más allá de la evidente superioridad cuantitativa de la variante *deber* sobre *deber de*, o del incumplimiento generalizado de la oposición modal epistémico-deontica (véase tabla 1).

En este contexto, el *Macrocorpus sociolingüístico de Castellón y sus comarcas* (a partir de aquí *MCSCS*), del que se extraen los datos de esta investigación, ofrece algunas ventajas evidentes para el estudio de este fenómeno de variación gramatical. Por un lado, abarca todo el espectro social de las comunidades analizadas. Por otro, su magnitud permite la obtención de una muestra de

<sup>17</sup> Samper, Hernández y Troya, art. cit., pág. 1.066.

<sup>18</sup> *Ibíd.*, pág. 1.066.

<sup>19</sup> *Ibíd.*, pág. 1.065.

<sup>20</sup> Daisy Galué, Marcos Suárez y Betsabé Valor, «¿Debe estar ahí... o debe de estar ahí...? Estudio de las perífrasis con el verbo *deber* en el español hablado en Caracas», *Núcleo*, (n.º especial), 2004, págs. 53-77.

datos que en la mayor parte de los casos multiplica por varios enteros las cifras absolutas obtenidas en otras muestras de habla oral. Ello hace que, al margen de algunos contextos en los que las cifras continúan siendo poco representativas (véase más adelante § 4.6.1), en muchos otros se puedan poner a prueba la significación y jerarquía de determinados factores lingüísticos y extralingüísticos. Para ello nos hemos servido del programa estadístico Goldvarb X, habitual en los estudios de variación lingüística, y que sirve para medir los pesos probabilísticos de aquellos factores seleccionados como significativos para explicar la variabilidad, así como su rango explicativo<sup>21</sup>.

La muestra inicial estaba constituida por 156 entrevistas. Sin embargo, las limitaciones ya reseñadas acerca de la recurrencia de este fenómeno en el discurso hicieron que tan solo en 126 encontráramos manifestaciones de al menos una de las variantes analizadas. A esas entrevistas limitamos, pues, nuestro análisis. La muestra definitiva, compuesta por 151 informantes<sup>22</sup>, se halla integrada por diferentes grupos sociales representativos de otros tantos factores sociológicos, como el *sexo* (hombres, mujeres), la *edad* (jóvenes, adultos, mayores), el *nivel de estudios* (primarios, secundarios, superiores); el *estatus social* (bajo, medio-bajo, medio-alto); la *lengua dominante* (castellano, valenciano) y el lugar de *residencia* (Castellón de la Plana, provincia).

Junto a estas variables sociológicas, de las que, sin embargo, no nos ocuparemos en este artículo, cada ocurrencia del presente fenómeno de variación ha sido codificada de acuerdo con diversos factores lingüísticos y comunicativos, como los que se detallarán en el apartado 4. Por último, la extracción de las variantes ofrecidas por el *MCSCS* mediante el programa de concordancias WordSmith 4.0 supuso la obtención de 412 ejemplos, cuya distribución global puede verse en la tabla 1. En ellas se comparan las cifras de la muestra castellonense con las obtenidas en otras comunidades del mundo hispánico.

<sup>21</sup> Así, los valores P. que aparecen en la tabla 3 indican la probabilidad —en una escala de 0 a 1— de que un factor concreto influya en la selección de una determinada variante. Los pesos numéricos próximos a 1 favorecen la elección de la variante elegida (en este caso, la perífrasis con preposición), mientras que, en el extremo opuesto, los más cercanos 0 la desfavorecen. Por el contrario, los situados en cifras intermedias, cercanas a 0,5, ejercen una influencia neutra. Por otro lado, el programa multivariante no solo indica qué factores son significativos, sino también su jerarquía. Esta se obtiene a través de la comparación entre los rangos respectivos, resultantes de calcular la diferencia entre las probabilidades mayor y menor de cada grupo de factores (*vid.* David Sankoff, Sali A. Tagliamonte y Eric Smith, *GOLDVARB X: A multivariate analysis application for Macintosh and Windows*, Toronto/Ottawa, Department of Linguistics (University of Toronto) and Department of Mathematics (University of Ottawa), disponible en <[http://individual.utoronto.ca/Tagliamonte/Goldvarb/GV\\_index.htm](http://individual.utoronto.ca/Tagliamonte/Goldvarb/GV_index.htm)>.

<sup>22</sup> En algunas entrevistas del *MCSCS* participaron tres interlocutores, un entrevistador y dos entrevistados. Ello explica que el número de informantes en la muestra sea superior al de entrevistas. Por otro lado, conviene señalar que buena parte de estas entrevistas figura entre las más espontáneas, lo que, como veremos más adelante (§ 4.3) tiene su relevancia en el fenómeno de variación que nos ocupa.

#### 4. RESULTADOS

##### 4.1. *La oposición modalidad epistémico/deóntica y sus relaciones con la alternancia sintáctica*

Como ha señalado Bybee *et al.*<sup>23</sup>, la tarea de delimitar el alcance de lo que se entiende por modalidad no es empresa fácil, y buena prueba de ello es la larga trayectoria que este tema ha tenido tanto en el ámbito de la filosofía —nada menos que desde Aristóteles—, como más modernamente en la lingüística, sin que los profesionales de esta hayan alcanzado tampoco un consenso definitivo en torno a los límites de dicha categoría. Con todo, suele aceptarse que la modalidad expresa la actitud del hablante hacia el contenido proposicional de sus enunciados. Y aunque el consenso es menor a la hora de definir los tipos de modalidad existentes en una lengua, casi todos aceptan las dos categorías siguientes:

a) la modalidad epistémica está relacionada con el valor veritativo de las proposiciones, en una escala que oscila entre los grados extremos de lo posible y probable, por un lado, y lo no probable y lo no posible, por otro<sup>24</sup>. Por lo general se expresa en términos de probabilidad o conjetura, en el sentido de que lo enunciado puede o no ser inferido como posible.

b) la modalidad deóntica, por el contrario, se halla asociada a contenidos de carácter volitivo, como la expresión de la obligación, el mandato o el deseo.

Como hemos visto en las páginas que preceden, los estudios sobre el fenómeno de variación que nos ocupa se han centrado preferentemente en el análisis de este factor lingüístico para comprobar si las recomendaciones normativas que relegan el uso de cada perífrasis para un tipo determinado de modalidad (deóntica: *deber* + infinitivo; epistémico: *deber de* + infinitivo) se respetan en la lengua real, con conclusiones casi siempre negativas. La tabla 1 ofrece una comparación entre los resultados obtenidos en la región castellanense y otros enclaves del mundo hispánico.

---

<sup>23</sup> Joan Bybee, Revere Perkins y William Pagliuca, *The Evolution of Grammar: Tense, Aspect and Modality in the Languages of the World*, Chicago, University of Chicago Press, 1994, pág. 176.

<sup>24</sup> Cfr. Concepción Otaola, «La modalidad (con especial referencia a la lengua española)», *Revista de Filología Española*, 68, 1-2, 1988, págs. 97-115; Gerhard Bauhr, «Sobre el futuro *cantaré* y la forma compuesta *voy a cantar* en español moderno», *Moderna Sprak*, 86, 1, 1992, pág. 77; Mercedes Sedano, «El futuro morfológico y la expresión *ir a* + infinitivo en el español hablado de Venezuela», *Verba*, 21, pág. 233; Bybee, Perkins y Pagliuca, *op. cit.*, pág. 179. Emilio Ridruejo, «Modo y modalidad. El modo en las subordinadas sustantivas», *Gramática Descriptiva de la Lengua Española*, Ignacio Bosque y Violeta Demonte, dirs., Madrid, Espasa Calpe, 1999, pág. 3.210 y sigs.

## 4.1.1. Distribución global de las variantes

Como puede observarse en la tabla, las dos construcciones no aparecen en nuestro corpus con la misma frecuencia. Como era previsible, la forma sin preposición se usa en muchas más ocasiones (77,7%) que la correspondiente con *de* (22,3%).

Con todo, si extrapolamos los resultados obtenidos en otras investigaciones que se detienen en las capas altas de la sociedad al conjunto de esta —tarea no siempre recomendable, como señalábamos anteriormente— Castellón aparecería en una posición más conservadora, frente a la clara tendencia a la elisión que, con alguna excepción aislada<sup>25</sup>, se advierte en Hispanoamérica. Las cifras castellanenses están más próximas a las de otras poblaciones españolas, como Bermeo, Madrid, Valencia o Las Palmas de Gran Canaria. Hay que notar, sin embargo, que a diferencia de esta última, las demás poblaciones obtienen índices por encima del 30%, superando, pues, los de Castellón. Y ello es especialmente llamativo tras la comparación con la vecina Valencia, cuyo nivel de empleo de la perífrasis en incidencia indirecta es 12 puntos mayor (34,5%) que el correspondiente en Castellón.

	<i>DEBER + INFINITIVO</i>			<i>DEBER DE + INFINITIVO</i>		
	<i>Deóntica</i>	<i>Epistém.</i>	<i>N (%)</i>	<i>Deóntica</i>	<i>Epistém.</i>	<i>N (%)</i>
	%	%		%	%	
Castellón <sup>1</sup>	61	39	320 (77,7)	54,3	45,7	92 (22,3)
Valencia (N.C.) <sup>2</sup>	70,2	29,8	171 (65,5)	50	50	90 (34,5)
Madrid (N.C.) <sup>3</sup>	53,5	45,5	99 (67,8)	53,2	46,8	47 (32,2)
Madrid (N.C.) <sup>4</sup>	81	19	21 (67,7)	40	60	10 (32,3)
Sevilla (N.C.) <sup>4</sup>	91,3	8,7	23 (82,1)	60	40	5 (17,9)
Las Palmas de GC <sup>5</sup> (N.C.)	90	10	70 (77)	100	—	21 (23)
Las Palmas de GC <sup>4</sup> (N.C.)	85,7	14,3	28 (71,8)	100	—	11 (28,1)
Bermeo <sup>6</sup>	61,5	38,5	13 (59)	100	—	9 (41)
Méjico (N.C.) <sup>4</sup>	86,5	13,5	37 (77)	72,7	27,3	11 (21)
Méjico (N.C.) <sup>7</sup>	82	18	388 (75)	62,5	37,5	130 (25)
Puerto Rico (N.C.) <sup>4</sup>	91,7	8,3	36 (76,5)	100	—	10 (23,5)

<sup>25</sup> Mención especial merecen los datos correspondientes a la comunidad hispana de Houston, donde, pese a una cifras absolutas poco concluyentes (tan solo cuarenta y una ocurrencias de la variable en total), es la única ciudad donde se invierten las tornas, siendo aquí claramente dominante la variante prepositiva (70,8%, frente al 39,2% de la otra forma). Aunque Fairclough no formule ninguna interpretación al respecto, no es descartable que la inseguridad lingüística provocada por el aislamiento funcional en que se desarrolla esta variedad norteamericana pueda estar detrás de estas cifras, tan diferentes de las del resto del mundo hispanico.

	DEBER + INFINITIVO			DEBER DE + INFINITIVO		
	Deóntica	Epistém.	N (%)	Deóntica	Epistém.	N (%)
	%	%		%	%	
S. José de Costa R (N.C.) <sup>4</sup>	87,5	12,5	56 (93,3)	75	25	4 (6,7)
Bogotá (N.C.) <sup>4</sup>	97,1	2,9	69 (98,5)	100	—	1 (1,5)
Santiago de Chile (N.C.) <sup>4</sup>	82,6	17,4	23 (92)	100	—	2 (8)
Buenos Aires (N.C.) <sup>4</sup>	52,4	47,6	21 (95,5)	—	100	1 (4,5)
La Paz (N.C.) <sup>4</sup>	75	25	44 (95,6)	50	50	2 (4,4)
Lima (N.C.) <sup>4</sup>	61	39	41(91,1)	75	25	4 (8,9)
Caracas (N.C.) <sup>4</sup>	84,1	15,9	88 (97,7)	—	100	2 (2,3)
Caracas <sup>8</sup>	51,4	48,6	107 (89,1)	89	11	13 (10,9)
Houston <sup>9</sup>	42	58	12 (29,2)	86	14	29 (70,8)

TABLA 1: Distribución de las variantes *deber* + infinitivo y *deber de* + infinitivo para la expresión de los valores deónticos y epistémicos en diversas regiones del mundo hispánico<sup>26</sup>

#### 4.1.2. Distribución global de las variantes perifrásticas por modalidades

4.1.2.1. *Deber* + infinitivo: Al igual que en otras comunidades de habla, en Castellón se utiliza más para la expresión de la obligación (61%) que de la probabilidad (39%). Sin embargo, hay de nuevo diferencias notables entre una regiones hispanas y otras. Y es que Castellón figura esta vez entre las comunidades donde la expresión no canónica de la probabilidad es más alta, con valores similares a los de las clases cultivadas de Lima (39%), y tan solo superada levemente por los mismos sectores de población de Madrid (45,5%), Buenos Aires (47,6%) y Caracas (48,6%). En este sentido, pues, y con excepción de la capital de España, los hablantes castellonenses en su conjunto superan ampliamente en este uso —no canónico, pero menos censurado por la normativa, como vimos— a otras ciudades españolas como Las Palmas de Gran Canaria (Troya: 10%; Samper *et al.*: 14,3%), Sevilla (8,7%), o incluso la vecina Valencia (29,8%).

#### 4.1.2.2. *Deber* + *de* + infinitivo:

Con la excepción de unas pocas urbes, como Valencia, Las Palmas de Gran Canaria Madrid, México y ahora Castellón, la escasez de datos sobre esta perí-

<sup>26</sup> 1 (Blas Arroyo), 2 (Gómez Molina, art. cit.); 3 (Gómez Manzano, *op. cit.*); 4 (Samper *et al.*, art. cit.); 5 (Troya, *op. cit.*); 6 (Teresa Fernández Ulloa, «Perífrasis verbales en el castellano de Bermeo [Vizcaya]», *Revista Española de Lingüística*, 30, 1, 2000, págs. 1-34); 7 (Carmen Valadez, *Perífrasis de infinitivo en la norma lingüística culta de la ciudad de México*, México, UNAM, 1969); 8 (Galué *et al.*, art. cit.); 9 (Marta Fairclough, «Expresiones de modalidad en una situación de contacto: *deber [de]* vs. *tener que* en el español hablado en Houston», *Southwest Journal of Linguistics*, 19, 2, 2000, págs. 19-30).

frasis, hace más difícil todavía realizar generalizaciones válidas. Sea como sea, las cifras disponibles permiten confirmar que la variante se utiliza también tanto para los usos considerados como canónicos por la normativa —esto es, para la expresión de la probabilidad—, como para los no canónicos, cifrados en la expresión de la obligación. En Castellón, estos usos se reparten casi en la misma proporción, con una ligera ventaja para la expresión de la modalidad epistémica (54,3%) que para la deóntica (45,7%). No se cumplen así las predicciones de algunos gramáticos, que dan por sentado el empleo no canónico de *deber* para la expresión de la probabilidad, pero no así los de *deber de* para la obligación.

Llegados a este punto, interesa saber si las frecuencias obtenidas por las dos variantes en cada uno de los valores modales confirman o no el valor predictivo de dicha oposición semántica. Para ello nos servimos del programa estadístico Goldvarb X, mediante el cual es posible averiguar tanto el nivel de significación como la jerarquía de los factores considerados en el estudio. La tabla 2 contiene los resultados de ese análisis multivariante aplicado al corpus y referido a los factores lingüísticos y estilísticos considerados en el presente estudio.

	<i>P.</i>	<i>%</i>	<i>N/Total</i>
Total N (/variante): 92 (/412)			
Media Corregida: 0.17			
MODALIDAD			
Deóntica		[20,4	50/245]
— (Obligación)		[22,5	31/138]
— (Conveniencia)		[17,8	19/107]
Epistémica.		[25,1	42/167]
<i>Rango</i>			
ÉNFASIS (INTENSIFICACIÓN)			
Enfáticos	697	43,5	27/62
Neutros	463	18,6	65/350
<i>Rango</i>	23		
ATENUACIÓN			
Atenuados	381	14,1	25/177
Neutros	591	28,5	67/235
<i>Rango</i>	21		
MODALIDAD ORACIONAL			
Otras	555	24,1	89/370
Interrogativas	257	7,1	3/42
<i>Rango</i>	29		

	<i>P.</i>	<i>%</i>	<i>N/Total</i>
<b>CONTEXTO FONO-SINTÁCTICO</b>			
Intervocálico		[22,3	53/238]
— (e) ( <i>debe (de) estar</i> )		[26,2	27/103]
— (otros) ( <i>deba (de) estar</i> )		[19,3	26/135]
Otros		[22,4	39//174]
— consonante ( <i>deben de estar</i> )		[23,2	36/155]
— cacofónico ( <i>debe (de) decir</i> ).		[15,8	3/19]
<i>Rango</i>			
<b>Nº SÍLABAS (GRUPO VERBAL)</b>			
3 a 5 sílabas		[29,5	46/156]
+ 6 sílabas.		[18,1	45/249]
<i>Rango</i>			
<b>PERSONA Y NÚMERO</b>			
1º pers. sing.		[26,5	9/34]
2º pers. sing.		[23,1	6/26]
3º pers. sing.		[21,4	46/215]
1º pers. plur.		[11,1	4/36]
2º pers. plur.			[—]
3º pers. plur.		[26,7	27/101]
<i>Rango</i>			
<b>SINTAXIS DEL SUJETO</b>			
Explícito		[20,7	29/140]
Omitido		[23,2	63/209]
<i>Rango</i>			
<b>ANIMACIDAD</b>			
Humano		[25	57/228]
No humano		[19	35/184]
<i>Rango</i>			
<b>TIEMPO Y MODO</b>			
Condicional		[18,1	31/171]
Presente		[26,3	45/171]
Otros		[22,9	16/70]
<i>Rango</i>			
<b>V. PRINCIPAL (SIMPLE/COMPUESTO)</b>			
Simple		[23,1	89/386]
Compuesto		[11,5	3/26]
<i>Rango</i>			
<b>MODO DE ACCIÓN</b>			
Estativos		[20	45/235]
Dinámicos		[25,4	47/177]
<i>Rango</i>			

	<i>P.</i>	<i>%</i>	<i>N/Total</i>
<b>ADYACENCIA</b>			
Adyacentes (V. Aux + V. Ppal.)		[22,5	91/405]
No adyacentes		[14,3	1/7]
<i>Rango</i>			
<b>DIÁTESIS</b>			
Activa		[21,8	84/386]
Pasiva refleja		[30,8	8/26]
<i>Rango</i>			
<b>TIPO DE CLÁUSULA</b>			
Subordinadas		[21,8	42/193]
Otras		[22,8	50/219]
<i>Rango</i>			
<b>TIPO DE TEXTO</b>			
Argumentativos		[24,2	58/240]
Otros		[23,2	33/142]
<i>Rango</i>			
<b>TENOR (ENTREVISTA)</b>			
Espontáneas	572	23,4	36/154
No espontáneas	407	17,6	21/119
<i>Rango</i>	17		
<b>FASE DE LA INTERACCIÓN VERBAL</b>			
Iniciales		[20	37/185]
Finales		[26,2	33/126]
<i>Rango</i>			
LOG LIKELIHOOD= -201.028 SIGNIFICACIÓN: 0.039 <sup>27</sup>			

TABLA 2: Contribución de los factores lingüísticos y estilísticos para la selección de la perífrasis *deber de* + infinitivo en las comarcas castellonenses (análisis de regla variable mediante Goldvarb X)

Como puede apreciarse en la tabla, la oposición modal epistémico-deóntica no es significativa, pese a que los porcentajes favorables a la expresión canóni-

<sup>27</sup> En la tabla aparecen tanto los factores que se revelan como significativos, como los no significativos. Los primeros tienen asignado un nivel *P.* de probabilidad, así como un rango, con el que se mide la relevancia explicativa de cada factor: cuanto mayor es este, mayor es también su importancia predictiva. Los factores no seleccionados por el análisis multivariante poseen, lógicamente, menor interés interpretativo, pero en algunos casos muestran tendencias marcadas en las distribuciones porcentuales correspondientes, cuya interpretación es necesario tener en cuenta. En unos casos para descartar que su falta de significación obedezca a la interacción o dependencia; en otros, para vislumbrar tendencias distribucionales posibles a las que, sin embargo, falta el aval estadístico por su escasa representación muestral.

ca de la probabilidad con *deber de* (25,1%) superan ligeramente a los no canónicos de la obligación (20,4%). Por el contrario, sí lo son otros, cuya incidencia en la selección de la variante prepositiva actúa en sentidos opuestos: mientras que algunos representan todavía hoy un factor relevante en el mantenimiento de la preposición, otros aceleran el ya avanzado proceso de elisión. De ellos nos ocupamos en los siguientes apartados.

#### 4.2. *La intensificación expresiva (énfasis)*<sup>28</sup>

Aunque la influencia de este factor no se ha ensayado empíricamente —al menos hasta donde llega nuestro conocimiento— en la bibliografía, encontramos aquí y allá observaciones intuitivas acerca de la potencial contribución del énfasis en el mantenimiento de la preposición, y más concretamente para expresar la obligación. Una buena muestra de estas observaciones es la interpretación formulada hace unos años por Náñez<sup>29</sup> para explicar los usos de *deber de* con sentido deóntico:

[...] al quedar la frase más llena con *de* (*debe de* + infinitivo), indirectamente resulta más expresiva y adecuada para la expresión de la obligación. La idea de la obligación, la actitud del hombre ante la obligación constituye en sí una postura más contundente, activa («agresiva») que la mera enunciación de la probabilidad.

Por su parte, Fernández de Castro<sup>30</sup> comenta también algunos ejemplos extraídos de textos escritos que, en su opinión, son representativos de «lo enfáticamente que se usa con frecuencia la preposición para reafirmar el carácter obligatorio de una circunstancia».

A pesar de las dificultades para definir un criterio como el presente, lo cierto es que sus manifestaciones lingüísticas se dejan ver por doquier en el discurso, y son, por ejemplo, un elemento fundamental en el plano interpersonal de las relaciones comunicativas. Desde un punto de vista pragmático, el énfasis puede ponerse en relación con el concepto más amplio de la intensificación, una macroestrategia de la que se sirven los hablantes para resaltar sus contribuciones, mediante un refuerzo expresivo que en el plano formal supone un incremento fonológico, gramatical y/o léxico<sup>31</sup>.

<sup>28</sup> Pese al carácter impresionista del que se ha tachado a veces un concepto como el énfasis —y en ocasiones no sin razón—, en el presente estudio lo utilizaremos como término sinónimo de intensificación. Por otro lado, y como se señala a continuación, nuestra intención es delimitar el criterio de acuerdo con parámetros objetivos, extraídos del contexto lingüístico y comunicativo.

<sup>29</sup> Náñez, art. cit., pág. 242.

<sup>30</sup> Fernández de Castro, *op. cit.*, pág. 178.

<sup>31</sup> Marta Albelda, *La intensificación en el español coloquial*, Valencia, Universidad de Valencia, 2005.

Con el fin de evaluar de la forma más objetiva posible este criterio, hemos codificado el énfasis a partir de la presencia en el cotexto inmediato a la variable de marcas lingüísticas utilizadas por el hablante para intensificar expresivamente el contenido de sus enunciados. Sin ánimo de ser exhaustivos, algunas marcas de este tipo son, por ejemplo, el empleo de una entonación exclamativa (*¡qué mala!*) y un léxico modalizado (*bicho*), como los que el hablante de (5) emplea para reprender cariñosamente a su interlocutora. Otras veces, los hablantes recurren a mecanismos sintácticos asociados a la intensificación, como la oración consecutiva de (6), o a las estructuras atributivas con un *ser* enfático, como en (7). El empleo de paráfrasis y otros recursos comunicativos, como las pausas estratégicas situadas en la antesala de nuevas unidades remáticas —(8)—, sirven también para introducir contenidos modalizadores. Y tampoco faltan ejemplos como los de (9), en los que el hablante se autocorriga, pasando de un predicado atenuante (*creo que...*) a la expresión reforzada de sus ideas a continuación (*sé que...*)<sup>32</sup>:

- (5) [...] ¡qué mala! ¡tú de pequeña *deberías ser* un bicho! (MCSCS, 119)  
 (6) buenol/ pues: es totalmente diferente/ las travesuras mías, o de mi juventud, o de mi niñez/ con compañía de los otros chicos a lo que se hace hoy en día// hoy en día la juventut como ve tanta televisión/ hay tantas cosas que no *debiera de ver!* y: y- o sea es otro mundo totalmente diferente (MCSCS, 413)  
 (7) [...] yo veo días que hacen muchas excursiones y muchos juegos que no *deberían de hacer!* lo que *deberían de enseñar* más (MCSCS, 2)  
 (8) [...] mira te lo voy a contestar de otra manera/ te lo voy a decir de otra manera// creo que todo el mundo *debería tomarse* unas vacaciones de vez en cuando (MCSCS, 355)  
 (9) [...] podríamos decir que es el mejor amigo que tienen/ entonces creo que- o sé/ que *debería haberse* llevado algún premio más (MCSCS, 42)

Así las cosas, las ocurrencias de la variable en el corpus aparecen divididas en dos grandes grupos. Por un lado, aquellas en las que el hablante recurre a los mecanismos descritos para intensificar el sentido de sus aseveraciones, dudas, deseos, etc. Por otro, las variantes situadas en enunciados donde no aparece ninguna de estas marcas enfáticas. Como era de esperar, estas últimas representan una clara mayoría en el corpus, con 350 casos (85%), frente a los 62 (15%) en los que se dejan ver algunos de los mecanismos de intensificación que hemos reseñado.

Pese a ello, el factor *énfasis* aparece como uno de los más significativos de todos los considerados en el estudio (rango: 23). No en vano, hasta un 43,5 % de todas las variantes del corpus en las que el hablante se sirve de marcas de intensificación, este lo hace mediante el concurso de la perífrasis con preposi-

<sup>32</sup> El lector interesado puede encontrar una relación detallada de otras estrategias de intensificación en Albelda, *op. cit.*, págs. 75 y sigs., especialmente referidas al español coloquial.

ción. Por el contrario, esta frecuencia desciende hasta un 18,6% en los enunciados no enfáticos. De ahí el reparto de probabilidades que arroja el análisis multivariante: mientras que los contextos no enfáticos desfavorecen la preposición (.46), los enfáticos la favorecen (.69).

La relevancia de este factor se ve confirmada mediante la tabulación cruzada con los demás factores considerados en el análisis. En casi todos los subgrupos creados como consecuencia de estos cruces, los avances más significativos en el empleo de la preposición tienen lugar, precisamente, en los contextos enfáticos. Por el contrario, en los neutros o no enfáticos, tales índices descienden significativamente.

#### 4.3. *El nivel de espontaneidad de las interacciones (tenor de la entrevista)*

Otro factor que actúa a favor de la perífrasis en incidencia indirecta podría situarse ahora en el eje estilístico, o si prefiere en el contexto comunicativo en que tienen lugar las interacciones verbales que sirven como marco al presente corpus de habla oral. Se trata del que hemos denominado *nivel de espontaneidad* de las interacciones, medido a través de diversos índices que permiten distinguir distintos tipos de entrevistas sociolingüísticas. En efecto, las muestras de habla que componen el *Macrocorpus sociolingüístico de Castellón y sus comarcas* presentan cierta diversidad en el modo en que se llevaron a cabo las entrevistas semidirigidas. En algunos casos, los entrevistadores participaban activamente en el desarrollo de las conversaciones, por medio de interrupciones, señales de retroalimentación, aportando sus opiniones y/o experiencias personales en diversas fases del diálogo, cuando no, como vimos anteriormente<sup>33</sup>, permitiendo la participación de más de un interlocutor, lo que generalmente facilitaba una considerable mayor distensión comunicativa. En otras ocasiones, por el contrario, el entrevistador se limitaba básicamente a seguir el guión de las preguntas que servían como base para la entrevista, y dejaba que fuera el entrevistado quien llevara el peso de la conversación. Para nuestros actuales intereses, consideramos como más *espontáneas* las entrevistas del primer tipo, y si nuestras hipótesis son correctas, es de esperar que en su desarrollo aparezca más la variante *deber de* que en el segundo tipo de entrevistas, que, a los efectos de la codificación, calificamos como *no espontáneas*<sup>34</sup>.

Como se aprecia en los datos que aparecen en la tabla 2, nuestras intuiciones en torno a este factor se han demostrado realistas. Y es que, en efecto, los

<sup>33</sup> Vid. nota 22.

<sup>34</sup> Para facilitar la comparación, quedan sin codificar en este factor aquellas entrevistas en las que se aprecian signos intermedios entre los dos tipos extremos mencionados.

porcentajes de la variante prepositiva son significativamente mayores en las entrevistas donde reina un mayor grado de espontaneidad entre los interlocutores (23,4%, P. 57) que en las otras (17,6%; P. 40). Aunque el rango (17) no es excesivamente elevado en esta ocasión —lo que habla de un menor peso explicativo que el énfasis—, el factor es seleccionado también como significativo. En el fondo, la naturaleza intrínseca de este factor se halla relacionada con el énfasis, y en cierto modo, forma parte de ese mismo polo de modalización en el que situábamos este último: tanto en uno como en otro nos enfrentamos ante factores que miden un grado de implicación particularmente intenso del hablante en sus enunciados.

#### 4.4. *La atenuación*

Al igual que el énfasis, la estrategia inversa de la atenuación desempeña un papel relevante en el desarrollo de las relaciones interpersonales, hasta el punto de que se ha considerado pieza fundamental en la expresión de la cortesía<sup>35</sup>. Por otro lado, desde un punto de vista formal los llamados *atenuantes* son, en palabras de Haverkate: «[...] una partícula, palabra o expresión que sirve para modificar el significado de un predicado de forma que se indique que ese significado sólo se aplica parcialmente al objeto descrito»<sup>36</sup>.

Ahora bien, precisamente por esos caracteres, las estrategias de atenuación pueden servir al hablante para mitigar su compromiso con la verdad de sus enunciados o, en los casos extremos, añadir matices de duda, incertidumbre, inseguridad, etc. Esto último es, justamente, lo que puede apreciarse en el ejemplo siguiente, donde la hablante que está en el uso de la palabra expresa sus dudas acerca del efecto que en ella pueda tener en el futuro su traslado como maestra a un gran colegio de Castellón, desde su actual destino en una pequeña escuela rural.

(10) [...] bueno, pero pienso que también es una cosa que: a la larga *debe: debe de cansar*, pero no, no lo sé (MCSCS, 210)

La presencia en el discurso de marcadores discursivos (*bueno*), predicados doxásticos (*pienso que...*), alargamientos vocálicos (*que:*), repeticiones (*debe:*

<sup>35</sup> Penelope Brown y Stephen Levinson, *Politeness. Some Universals in Language Use*, Cambridge, Cambridge University Press, 1987; Katherine Kerbrat-Orecchioni, *Les interactions verbales, Vol. II*, Paris, Armand Colin, 1992; Hank Haverkate, *La cortesía verbal*, Madrid, Gredos, 1994; Antonio Briz, «La atenuación en la conversación coloquial. Una categoría pragmática», *El español coloquial. Actas del I Simposio sobre Análisis del discurso oral*, Luis Cortés, ed., Almería, Universidad de Almería, 1995, págs. 101-122.

<sup>36</sup> Haverkate, *op. cit.*, pág. 211.

*debe*) y rutinas conversacionales (*no lo sé*) se dan la mano para completar ese sentido de incertidumbre que la hablante transmite a su interlocutora.

Los recursos formales que traducen en el discurso esta superestrategia pragmática son, pues, diversos, y al igual que hicimos con la intensificación, su presencia en el cotexto inmediato a la variable objeto de estudio nos ha permitido codificar cada una de sus variantes en dos grandes grupos. Por un lado, aquellas ocurrencias que, con independencia de la modalidad epistémica o deóntica que transmiten, no aparecen acompañadas por signos de atenuación como los mencionados. Así sucede en los dos enunciados siguientes, representantes de sendos contenidos modales, epistémicos (11) y deónticos (12), respectivamente:

(11) [...] como ya era tarde y me despierto y yo veo que estaba toda la habitación iluminá(d)a, digo «uhhh» ((risas))// «qué tarde *debe ser*» (MCSCS, 276)

(12) [...] es una realidad y lo acepto/ pero hay otras personas que no les parece bien// entonces *deberían respetar* la opinión de todos (MCSCS, 271)

Por otro lado, se codifican agrupadas en otro bloque aquellas variantes afectadas por la presencia en el discurso inmediato de atenuantes. De nuevo, estos pueden surgir tanto en contextos deónticos como epistémicos. El primero es el caso de (13), donde el hablante mitiga mediante el comprobativo *¿no?* su idea acerca de la necesidad de llenar el tiempo de una manera provechosa, aunque a la gente le toque mucho dinero en la lotería. En el siguiente ejemplo, por el contrario, los alargamientos vocálicos seguidos de pausa (*está://*) así como el marcador discursivo (*pues:*) sirven para vislumbrar el matiz de aproximación con que otro hablante se refiere a la distancia espacial entre dos puntos:

(13) [...] pero, *deberías llenarlo* con otra cosa *¿no?* el tiempo; el tiempo que... (MCSCS, 293)

(14) [...] y nos fuimos al pueblo porque la discoteca *está://* pues *debe estar* a un par de kilómetros del pueblo (MCSCS, 269)

El total de variantes correspondiente a este último grupo es de 177, lo que representa un 43% del total. Por su parte, los contextos neutros en relación con este factor se cifran en 235 (57%).

Al igual que el *énfasis*, este factor se revela como significativo tras el análisis estadístico multivariante, pero esta vez con un efecto inverso. Mientras que, como veíamos en el apartado anterior, los contextos enfáticos favorecen la selección de la variante prepositiva (*deber de*), en esta ocasión, los atenuados la desfavorecen claramente. De este modo se explica que tan solo en el 14,1% de las ocurrencias de la variable en contornos atenuados, los hablantes hayan elegido la variante con preposición, lo que en términos probabilísticos se traduce en un bajo índice del valor P. (.38). Por el contrario, los niveles de la preposición se duplican (28,5%) en los contextos neutros, o no atenuados. En estos, la probabilidad es más favorable a la elección de *deber de* (P. 59).

Por otro lado, estos datos, junto a los que arroja el cruce con la oposición modal epistémico-deóntica, desmienten algunas hipótesis acerca de la funcionalidad de la partícula prepositiva para la expresión del distanciamiento o el grado de evidencialidad<sup>37</sup>. Para el caso del español, por ejemplo, García<sup>38</sup> había observado que la alternancia sintáctica que nos ocupa en estas páginas sería análoga —formal y semánticamente— a la que establecen las formas alternantes *de* y *de que* en los casos de (de)queísmo. De este modo, la obligación, como mensaje imperativo, requiere de la yuxtaposición directa, mientras que la afirmación atenuada, que implica el valor de probabilidad, exige la forma con *de*, es decir, la unión a distancia entre verbo auxiliar y verbo auxiliado. En consecuencia, de ser cierta esta hipótesis, la expresión de la preposición daría lugar a un distanciamiento en el grado de implicación del hablante con respecto a la verdad de su enunciado. Mas aún, se esperaría que este distanciamiento tuviera lugar preferentemente en aquellos contextos en los que no solo se expresa probabilidad —y no obligación—, sino donde, además, esta aparece acompañada de recursos de atenuación, como los advertidos en (10) y (14). Por el contrario, la menor frecuencia de la variante prepositiva se esperaría en los enunciados donde se expresa la obligación pura, sin atenuación alguna, como en (12).

Sin embargo, los datos empíricos resultantes de este análisis arrojan resultados bien diferentes. En primer lugar, y contrariamente a lo esperado según la teoría de la diferenciación funcional, la combinación de probabilidad y atenuación no solo no da lugar a una mayor frecuencia de la preposición, sino todo lo contrario, como puede verse en la tabla 3. Obsérvese como el 18% alcanzado por este contexto se ve ampliamente superado por el 29% de los contextos epistémicos no atenuados. Pero es que, además, los mayores porcentajes en el uso de la preposición se producen no solo en esta última combinación, sino también en la expresión de la obligación, de nuevo sin atenuación (28%).

Modalidad		Contextos atenuados		Contextos neutros	
		N	%	N	%
Deóntica	<i>deber de</i>	15	12,5	35	28
	<i>deber</i>	105	87,5	90	72
Epistémica	<i>deber de</i>	10	18	32	29
	<i>deber</i>	47	82	78	71

TABLA 3: Tabulación cruzada entre la oposición modal deóntico-epistémico y la atenuación

<sup>37</sup> Cfr. Talmi Givon, «Evidentiality and epistemic space», *Studies in Language*, 6, 1982, págs. 23-49; Frank R. Palmer, *Mood and modality*, Cambridge, Cambridge University Press, 1986.

<sup>38</sup> García, art. cit., pág. 64.

#### 4.5. *La modalidad oracional*

La influencia negativa de los contextos atenuados sobre la variante prepositiva puede ponerse en relación con otros resultados derivados de nuestro análisis. Así, al estudiar la incidencia de la modalidad oracional, comprobamos como tan solo las oraciones interrogativas arrojan cifras significativamente diferentes, como puede apreciarse en la tabla 2. En efecto, frente a la modalidad enunciativa, tanto afirmativa como negativa, u otras de aparición en el discurso mucho más anecdótica —exhortativas, imperativas...—, las oraciones interrogativas se erigen como uno de los *loci* menos propicios para la preposición. De hecho, tan solo en un 7,1% de los casos, como en (15), nos encontramos con *deber de* en interrogativas, frente a un porcentaje mucho mayor en el resto de las oraciones (34,1%). Como resultado de ello, y a pesar de los desequilibrios muestrales entre unas y otras<sup>39</sup>, la oposición entre la modalidad interrogativa (P. 25) y el resto (P. 55) surge en el análisis como significativa, y, además, con uno de los rangos más elevados (29):

(15) [...] ¿y en los Oscars hay, hay alguna película que pienses que: que se debería de haber llevado más Oscars?! porque yo sí [...] (MCSCS, 42)

Ahora bien, estos resultados no parecen casuales. Si como hemos comprobado, la preposición disminuye en aquellos contextos en los que el hablante rebaja su compromiso con la verdad contenida en sus enunciados, o muestra signos de inseguridad, no es extraño que ocurra lo mismo con las oraciones interrogativas. No en vano, se trata de la modalidad oracional cuya esencia consiste, precisamente, en el reconocimiento por parte del hablante de un vacío cognitivo propiciado por el desconocimiento o la duda, y para cuya resolución reclama el auxilio del interlocutor.

#### 4.6. *Otros factores lingüísticos*

El resto de factores lingüísticos considerados en el estudio no han resultado significativos tras el análisis estadístico multivariante. Con todo, dentro de estos hay que distinguir entre aquellos que muestran escasas (o nulas) distancias porcentuales entre los grupos correspondientes, y aquellos en los que se advierten diferencias de mayor calado, y para las que no es descartable un mayor grado de significación ante muestras más extensas.

<sup>39</sup> Tan solo hay cuarenta y dos oraciones interrogativas en el corpus analizado, lo que representa apenas un 10% del total.

4.6.1. Entre los primeros, destacan, por ejemplo, la clase sintáctica del sujeto (expreso/omitido), el tipo de cláusula (subordinadas/no subordinadas), el tipo de secuencia textual (argumentativas/no argumentativas) y el modo de acción de los verbos auxiliados (estativos/dinámicos). Sobre algunos de estos cabe realizar, sin embargo, algunas consideraciones adicionales que consideramos de interés, bien porque muestran interacciones relevantes con otros factores lingüísticos, bien porque permiten cotejar nuestros datos con los de otras investigaciones donde han sido considerados previamente.

Este último es el caso del modo de acción verbal de los verbos principales. Siguiendo a Balasch<sup>40</sup>, quien ha estudiado la incidencia de este factor en varios cortes diacrónicos de la historia del español, hemos dividido estos verbos en dos grupos: a) *estativos*, como los de (16), y b) *dinámicos*, como en (17)<sup>41</sup>:

(16) [...] sí/ yo pienso que la muerte más horrorosa *debe ser* la muerte en la que te ahogas (MCSCS, 126)

(17) [...] y luego el otro día estuve discutiendo con una persona sobre si: si los empresarios *deben contratar* para según qué trabajos [...] (MCSCS, 129)

Advierte Balasch que el mencionado factor se revela significativo en el corpus integrado por diversas obras literarias del siglo XIX —no así en el XVII—, con los verbos dinámicos como principales responsables del avance de *deber* en detrimento de *deber de*. En nuestro estudio, sin embargo, los resultados se presentan justamente a la inversa: la preposición aparece levemente por encima cuando *deber* surge en perifrasis con verbos dinámicos (25,4%, frente al 20% con verbos estativos)<sup>42</sup>. Sin embargo, estas diferencias no son significativas, por lo que no puede descartarse el azar en su configuración.

<sup>40</sup> Sonia Balasch, art. cit., págs. 113-118.

<sup>41</sup> Aunque no hay acuerdo en la bibliografía sobre la clase de los verbos estativos, esta autora se basa en las clasificaciones de Comrie (vid. Bernard Comrie, *Aspect*. London, Cambridge University Press, pág. 48-51) y, más específicamente, en el listado que Torres Cacoullós ha ofrecido para el estudio de las perifrasis con *estar* + *gerundio* (vid. Rena Torres Cacoullós, «La evolución de estar + V-ndo hacia la expresión obligatoria del aspecto progresivo en español», *Actas del VII Congreso Internacional de Historia de la Lengua Española*, Concepción Company Company y José G. Moreno de Alba, eds., Madrid, Arco/Libros, 2008, págs. 1.147-1.162). Entre los verbos estativos estudiados por esta autora se encuentran: *adorar, atreverse, carecer, comprender, conocer, cuadrar, creer, deber, desdeñar, desear, despreciar, entender, estimar, faltar, figurarse, haber, holgar, ignorar, llamarse, merecer, necesitar, padecer, parecer, penar, pensar, poder, profesar, querer, saber, sentir, ser, sospechar, temer, tener, tratarse de* y *vivir*. Sin embargo, en nuestro corpus el 96% de estos verbos corresponden al cuarteto formado por *ser, estar, haber* y *tener*, que, por otro lado, figuran entre los verbos más frecuentes en español (Mark Davis, *Corpus del español [100 millones de palabras, siglo XIII-siglo XX]*, 2002, disponible en <<http://www.corpusdelespanol.org>>).

<sup>42</sup> La *Nueva gramática de la lengua española* subraya también la especial profusión de usos de *deber* + infinitivo en contextos propios de la modalidad epistémica, como los verbos impersonales (*haber, hacer*), un dato confirmado también por nosotros (*op. cit.*, pág. 2.143).

Más coincidentes son, sin embargo, nuestros datos con las observaciones realizadas por Gómez Manzano a propósito del verbo estativo más frecuente en el corpus —y en la lengua española en general— el verbo *ser*. En relación con *ser*, señala esta autora que «es frecuente el empleo del verbo *ser* como auxiliado cuando ambas realizaciones (deber y deber de) tienen un valor hipotético o de probabilidad». Por el contrario, en los contextos deónticos, «el uso de *ser* es mucho menor»<sup>43</sup>.

Estos resultados se confirman, efectivamente, en nuestra investigación, y en general son válidos para el conjunto de los verbos de esta naturaleza, como puede apreciarse en la tabla adjunta. A este respecto, obsérvese como el número total de estos verbos en contextos epistémicos es cuatro veces mayor (133) que entre los verbos dinámicos (34). Por el contrario, la modalidad deóntica se halla mucho más equilibrada entre los dos modos de acción verbales, aunque esta vez con una ligera ventaja favorable para los verbos dinámicos (143) sobre los estativos (102).

Modalidad		Estativos		Dinámicos	
		N	%	N	%
Deóntica	<i>deber de</i>	17	17	33	23
	<i>deber</i>	85	83	110	77
Epistémica	<i>deber de</i>	30	23	12	35
	<i>deber</i>	103	77	22	65

TABLA 4: Tabulación cruzada entre la oposición modal deóntico-epistémico y el grado de estatividad de los verbos

Por lo que se refiere al *tipo de secuencia textual*, la tabla 2 muestra como el empleo de la perífrasis con preposición es similar tanto en las secuencias *argumentativas* (24,2%) —aquellas en las que el informante desarrolla sus opiniones acerca diversos temas de actualidad (la jornada semanal de 35 horas, la ampliación del horario escolar para los estudiantes, la creación de guarderías para facilitar la incorporación de la mujer al mercado de trabajo, etc.)—, como en otras que caracterizamos, por oposición, como *no argumentativas* (23,2%), y entre las que sobresalen pasajes narrativos sobre historias personales, fragmentos en estilo directo, episodios dialogales, etc. Los fragmentos de (18) y (19), ambos a cargo del mismo hablante, son representativos de las dos clases textuales, respectivamente:

<sup>43</sup> Pilar Gómez Manzano, *op. cit.*, pág. 162.

(18) E: ¿Qué te parece que el curso escolar para los niños se alargue?

I: *Debería de alargarse//* porque tienen muchas vacaciones y; porque en verdad siempre tenemos que ir un poco/ umm/ deprisa, ¿no? para dar el temario (MCSCS, 196)

(19) [...] yo tenía, pues por decirlo de alguna manera, la posibilidad de: de permitirme desplazar de mi pueblo y marcharme a estudiar a la ciudad// y: bueno/ pues en torno- *debería tener* veinticuatro-veinticinco años [...] (MCSCS, 196)

Pese a ello, la tabulación cruzada con el *énfasis* (véase 4.2) aporta algunos datos de interés, como los que aparecen en la tabla 5. En ella se observa como es en los contextos argumentativos enfáticos donde el incremento en el uso de la preposición se dispara: hasta en un 51% de todas las ocurrencias en este contexto los hablantes seleccionan *deber de*, uno de los porcentajes más amplios obtenidos por esta perífrasis. Por el contrario, el énfasis es menos determinante en las secuencias no argumentativas, alcanzando índices de uso de la preposición más moderados (30%). Estos resultados son lógicos si consideramos que las secuencias argumentativas representan un contexto idóneo para que el hablante se reafirme en sus juicios, y hasta de forma vehemente en ocasiones, como apreciamos en (18).

Tipo de texto		Contextos neutros		Contextos enfáticos	
Argumentativos	<i>deber de</i>	39	17	20	51
	<i>deber</i>	192	83	19	49
Otros	<i>deber de</i>	26	22	7	30
	<i>deber</i>	93	78	16	70

TABLA 5: Tabulación cruzada entre el tipo de secuencia textual y el énfasis expresivo

4.6.2. Como señalábamos al comienzo de este apartado, hay un segundo grupo de factores cuyas diferencias no resultan tampoco significativas, pero en los que estas resultan de mayor calado, y por tanto, merece la pena detenerse en ellas, si quiera brevemente.

Comencemos indicando que algunos de estos factores se han revelado en un sentido contrario al que habíamos esperado inicialmente. Esto es lo que sucede, por ejemplo, con el *contexto fonosintáctico* que rodea a la variable lingüística, factor analizado también en algunos estudios previos. A este respecto, señala Gómez Torrego que «[e]l fonema /d/ entre vocales tiende a desaparecer»<sup>44</sup>; hi-

<sup>44</sup> Gómez Torrego, *op. cit.*, pág. 3.348.

pótesis confirmada en su investigación sobre el habla culta de Madrid por Gómez Manzano, ya que, si bien las diferencias obtenidas por esta autora no son importantes cuando el auxiliar termina en consonante (*deben de...*) (18 casos con preposición, frente a 15 sin preposición), cuando lo hace «en vocal encuentra un 25% de ejemplos de *deber de* (29 casos) y un 75% de *deber* (84 casos)»<sup>45</sup>. Por su parte, y aunque sin la contundencia de los datos madrileños, Samper *et al.* advierten también que en el habla culta de las doce ciudades analizadas «el porcentaje de uso de la construcción sin preposición asciende cuando el auxiliar acaba en vocal»<sup>46</sup>, descenso que es particularmente acusado entre los 43 hablantes que alternan las dos construcciones sin ajustarse a la norma académica. Del mismo modo, Galué *et al.*<sup>47</sup> observan que en Caracas dicho factor lingüístico aparece como el único seleccionado por Goldvarb. En los contextos epistémicos, la terminación vocálica del verbo *deber* desfavorece (P. 36) la preposición. Y complementariamente, en los deónticos esa misma preposición es altamente propiciada (P. 84) cuando al verbo auxiliar termina en consonante.

Estos resultados parecen encontrar su lógica en algunos principios variacionistas universales. Por un lado, un cierto «principio de eco» permitiría explicar una mayor relación entre segmentos fónicos contiguos. De ser ello cierto, la aparición de una consonante favorecería la presencia de la preposición, y viceversa. Por otro lado, la investigación variacionista sobre el español está plagada de manifestaciones de elisión de la consonante /d/ en contextos intervocálicos<sup>48</sup>.

Sin embargo, otras investigaciones han venido a poner en duda la validez general de estos resultados y el poder explicativo de la fonética sintáctica en este caso. Así, ni Troya en Las Palmas<sup>49</sup>, ni Gómez Molina en Valencia<sup>50</sup> encuentran una relación causal como la descrita más arriba. Por nuestra parte, y en nuestro deseo de ahondar en la relevancia de este factor, decidimos establecer algunos contextos adicionales a los ya contemplados en la bibliografía anterior. Así, junto a los enunciados en los que *deber* termina en consonante, como en (20), y en los que potencialmente sería esperable una mayor presencia de la preposición, establecemos otros tres que, también hipotéticamente, desfavorece-

<sup>45</sup> Gómez Manzano, *op. cit.*, pág. 162.

<sup>46</sup> Samper *et al.*, art. cit., pág. 1.074.

<sup>47</sup> Galué *et al.*, art. cit., pág. 70.

<sup>48</sup> Véanse sendos resúmenes en Francisco Moreno Fernández, «Cambios vivos en el plano fónico del español: variación dialectal y sociolingüística», *Historia de la Lengua Española*, Rafael Cano Aguilar, coord., 2 vols., Barcelona, Ariel, 2004, vol. II, págs. 973-1.009; y José Luis Blas Arroyo, «Competing norms in a bilingual community. Ethnolinguistic and social factors in the reversal of a change in progress in peninsular Spanish», *Sociolinguistic Studies*, 1, 3, 2007, págs. 37-68.

<sup>49</sup> Troya, *op. cit.*, pág. 143.

<sup>50</sup> Gómez Molina, art. cit., pág. 871.

rían dicha variante. Por un lado, casos como los de (21), en los que el verbo principal comienza con la misma consonante /d/, en una combinación cacofónica especialmente favorable, pues, para la ausencia de la preposición. Y en el mismo sentido, operarían los contextos intervocálicos, si bien aquí establecemos dos grados diferentes: por un lado, ejemplos como los de (22), en los que *deber* termina en la vocal misma vocal /e/ de la preposición, y donde sería esperable, por tanto, un mayor efecto de la elisión; y por otro, el resto de los contextos intervocálicos, de los que es una buena muestra el ejemplo (23):

- (20) [...] allí *debían de haber*: pues: el doble de niños de los que cabían realmente en aquel colegio (MCSCS, 18)  
 (21) [...] no sé, no sé en qué, pero me imagino que trescientos millones de pesetas *deben dar* una seguridad extraordinaria (MCSCS, 18)  
 (22) [...] yo me imagino que eso *debe- debe ser* el- debe ser el: el reflejo de la época de- del hambre (MCSCS, 18)  
 (23) [...] al colegio iba con un bote/ con un vaso/ y ya *debía haber* el Cola-Cao porque me suena que: que ponía «Cola Cao» (MCSCS, 18)

El análisis de frecuencias permite comprobar como el contexto en el que se produce una mayor disminución (15,8%) de la variante con preposición *es*, conforme a lo esperado, aquel en el que tiene lugar una incómoda contigüidad de dos sonidos idénticos, lo que sucede con verbos auxiliados como *dar*, *decir*, *detener*, etc. A este respecto, Manuel Seco sostiene que el uso de *deber* + infinitivo, al menos en los contextos epistémicos, encuentra una explicación, siquiera parcial, en el hecho de que con la preposición se incurre en ocasiones en contextos cacofónicos «que preferirán evitarse mediante su omisión»<sup>51</sup>. Con todo, la relevancia estadística de este factor resulta difícil de medir puesto que el número de casos disponible es muy bajo (tan solo diecinueve ejemplos, de los que únicamente tres se construyen con la preposición).

Sin embargo, las cifras que ofrecen los otros contextos analizados son prácticamente idénticas. Más aún, cuando distinguimos en el análisis entre los contextos intervocálicos en los que interviene únicamente la /e/ (*debe*) y los demás (*debía*, *debió*...), nos encontramos con resultados que ponen todavía más en duda la relevancia de este factor. Y es que, lejos de ser el primer contexto intervocálico el que más desfavorece la preposición, los datos empíricos lo sitúan, por el contrario, en cabeza (26,2%), incluso por encima de los contextos no intervocálicos (23,2%). O dicho de otra manera, es, justamente, donde se esperarían menores frecuencias de la perífrasis *deber de*, donde más la encontramos.

4.6.2.1. Pese a la escasa relevancia del contexto fono-sintáctico en el presente estudio, el análisis de algunos datos adicionales sugiere el potencial relie-

<sup>51</sup> Manuel Seco, *op. cit.*, pág. 135.

ve de otros factores de carácter fonológico. En efecto, como se desprende de las cifras contenidas en la tabla 3, hay una serie de factores que parecen apuntar en una misma dirección explicativa: cuanto mayor es la pesantez del grupo verbal, menores parecen las posibilidades de que el hablante seleccione la preposición, y viceversa: a mayor simplicidad de esas combinaciones verbales, mayor también la frecuencia de la perífrasis en incidencia indirecta.

Ello se observa, por ejemplo, cuando analizamos la eventual influencia que ejerce el número de sílabas de esos compuestos verbales, cifras que en la práctica pueden oscilar entre las tres sílabas de las combinaciones más cortas, como la que vemos en (24), y la decena del ejemplo siguiente, donde el verbo auxiliar se combina con un verbo principal, que es a su vez auxiliar en una segunda perífrasis.

(24) [...] no sé// está cambiando bastante la familia// pero ese *debe ser* un tema bastante: rayero (MCSCS, 119)

(25) [...] pues yo pienso que *debería dejar de trabajar* los meses que se necesita mucho a una madre, los primeros meses, a lo mejor hasta un año (MCSCS, 129)

Dada la atomización de datos que arrojaría este factor, para el análisis estadístico decidimos agrupar en dos grandes bloques sus posibilidades. Por un lado, aparecen las combinaciones verbales cifradas en cinco o menos sílabas, mientras que en un segundo grupo se sitúan las agrupaciones con seis o más sílabas en total.

El recuento de frecuencias apunta inicialmente en la dirección de la hipótesis planteada: los contextos más “pesados”, esto es, aquellos con mayor número de sílabas, tienden a suprimir la preposición (18,1%) en mayor medida que los menos pesados, donde la presencia de la *de* es once puntos superior (29,5%).

Un factor adicional que apuntalaría estas conclusiones es el que permite distinguir entre dos tipos de verbos auxiliados en función de si: a) aparecen bajo una *forma simple*, como en (26); o b) por el contrario, lo hacen mediante una *forma compuesta*, bien se trate de su participación en otra perífrasis verbal, como veíamos anteriormente en (25), bien en otros complejos verbales no perífrásticos, como en (27).

(26) [...] creo que *debería estar* el carácter reivindicativo de ese sábado (MCSCS, 339)

(27) [...] con la cantidad de estrés a los que (*sic*) estamos sometidos hoy en día más que nada/ vacaciones *debe significar romper/ debe significar hacer* diferentes- *debe significar no estar obliga(d)o a hacer nada* (MCSCS, 339)

De ser ciertas nuestras hipótesis, la mayor pesantez de estos últimos complejos verbales justificaría una menor presencia de la preposición, justo lo con-

trario que con las formas simples de la conjugación. Y en efecto, el análisis de frecuencias parece acomodarse a estas conclusiones: los porcentajes de *deber de* en las formas simples doblan (23,1%) a los de las compuestas (11,5%).

Por último, un tercer criterio relacionado con los anteriores permite distinguir entre aquellos contextos en los que verbo auxiliar y el verbo principal aparecen juntos en la cadena hablada, como en (28), y otros en los que el hablante intercala una o más palabras entre ambos, como en (29). No en vano, y como recuerda Gómez Torrego, «el hecho de que los verbos de una perífrasis constituyan un solo núcleo del predicado, no implica que entre ellos no puedan introducirse otros elementos»<sup>52</sup>.

(28) [...] no dan todas las vacaciones que: que se *deberían de disfrutar!* aunque los papeles conste una cosa/ pues a lo mejor la realidad es otra ¿no?  
(MCSCS, 126)

(29) [...] puede ser perfectamente el hombre el que se deje el trabajo y cuide al niño// o sea/ no *debe* por qué *haber* diferencia de sexo en ese sentido  
(MCSCS, 119)

La lógica parece imponer que, incluso aquellos hablantes proclives al empleo de la preposición, se olviden de esta en aquellos contextos en los que verbo auxiliar y verbo principal se hallan separados por diverso material lingüístico. Y, en efecto, así parece de nuevo de acuerdo con las cifras obtenidas en nuestra investigación: el 22,5% de variantes prepositivas cuando verbo principal y auxiliar van juntos, se reducen a un 14,3% cuando no hay adyacencia directa entre ambos.

Ahora bien, pese a que las cifras ofrecidas por estos factores apuntan en una dirección similar, un análisis detenido de los datos invita a una mayor cautela, fundamentalmente por el hecho, ya reseñado, de que ninguno de esos factores es seleccionado como significativo por el análisis de regresión logística. Es posible que a ello contribuya el que los casos disponibles en algunos grupos sean muy escasos. El más extremo es, sin duda, el ofrecido por el factor de la adyacencia, con tan solo siete ocurrencias en todo el corpus, frente a las 405 restantes, lo que hace imposible obtener conclusiones siquiera aproximadas. Y aunque las distancias a propósito de las formas del verbo principal no sean tan abismales, lo cierto es que los ejemplos correspondientes a las formas compuestas son también muy limitados (apenas 26).

4.6.3. Por último, otros factores lingüísticos en los que aparentemente se observan diferencias frecuenciales de cierto relieve no poseen relevancia explicativa *per se*, sino en interacción con otros. Así ocurre, por ejemplo, con la *persona gramatical* del sujeto.

<sup>52</sup> Leonardo Gómez Torrego, art. cit., pág. 3.325.

En ocasiones se ha puesto en relación esta variable independiente con el grado de implicación del sujeto en la verdad de lo enunciado, en la línea argumental defendida por García<sup>53</sup>, a la que anteriormente hemos hecho ya referencia. Así, para Kitova-Vasileva<sup>54</sup>, por ejemplo, cuando el hablante quiere evitar todo tipo de responsabilidad y distanciarse de lo enunciado, emplea estructuras como *Debe de haber dicho la verdad*, en lugar de la correspondiente sin preposición. En este sentido, sería esperable una mayor presencia de la preposición en la primera persona que en el resto, pero especialmente en los contextos epistémicos, aquellos en los que el compromiso del hablante con lo enunciado es menor.

Los resultados de nuestro análisis ofrecen datos poco concluyentes al respecto. Por un lado, el factor no es seleccionado como significativo por el análisis estadístico. Por otro lado, si bien la primera persona aparece, en efecto, con frecuencias en el empleo de la preposición ligeramente por encima de la media (26,5%), en ello se ve igualada por la tercera persona del plural (26,7%), y con cifras muy próximas se sitúa la segunda del singular (23,1%). Pese a ello, el cruce con la modalidad epistémica ofrece algunos resultados en la dirección señalada por Kitova-Vasileva<sup>55</sup>. En efecto, la primera persona ofrece los índices más elevados de uso de la preposición para la expresión de la probabilidad (44%), al mismo tiempo que los más bajos en los correspondientes contextos deónticos (11%). Ninguna otra persona presenta una diferencia tan destacada entre ambas modalidades. Con todo, datos adicionales serían necesarios para apuntalar esta hipótesis de forma más sólida.

De otro tipo es la interacción que afecta al nivel de significación de otro factor lingüístico considerado en el análisis, el grado de *animacidad* del sujeto. En relación con este factor, distinguíamos entre sujetos *humanos* y sujetos *no humanos*, distinción no seleccionada como significativa por el análisis estadístico multivariante, pese a una ligera diferencia favorable al uso de la perífrasis con preposición entre los primeros (25%, frente a un 19% entre los no humanos). Ahora bien, la tabulación cruzada con el énfasis expresivo (§ 4.2) muestra efectos de diferente intensidad sobre cada grupo. Así, un nuevo repaso a la tabla 5 permite apreciar como en los contextos enfáticos y humanos el empleo de la preposición se dispara (53%) hasta superar incluso a la otra variante (47%), un resultado muy pocas veces observado en el presente corpus. Por el contrario, aunque el énfasis hace aumentar también las cifras de la preposición en los contextos no humanos (33%, frente al 16% en los no enfáticos), en estos la ausencia de preposición continúa siendo la norma (67%). En definitiva, una vez

<sup>53</sup> Erica García, *op. cit.*

<sup>54</sup> María Kitova-Vasileva, *La verosimilitud relativa y su expresión en español*, Santiago de Compostela, Universidad de Santiago de Compostela, 2000, pág. 19.

<sup>55</sup> *Ibíd.*, pág. 19.

más un factor (el énfasis) permite explicar las diferencias observadas en otros, como ocurre en este caso con el grado de animacidad del sujeto.

Para concluir, nos ocupamos de las diferencias encontradas en el análisis del *tiempo verbal*. En relación con este factor hay que destacar en primer término algunas restricciones que afectan a determinadas formas de la conjugación. A este respecto, Silva Corvalán ha augurado «que la oposición *deber* (obligación) y *deber de* (posibilidad) está en vías de perderse, *aunque se mantiene más estable en los tiempos no presentes*» (la cursiva es nuestra)<sup>56</sup>. Extrapolando estas conclusiones a la alternancia general entre las dos variantes, ello significaría que los índices de la preposición son potencialmente más bajos en el presente de indicativo que en el resto de los tiempos verbales. En su análisis sobre la norma culta en diversas ciudades del mundo hispánico, Samper *et al.* muestran diferencias frecuenciales que señalan en esa misma dirección. En sus palabras «el presente de indicativo [...] favorece la presencia de *deber* + infinitivo», mientras que «el imperfecto de indicativo y el condicional simple son los que más propician la aparición de la perífrasis con *de*»<sup>57</sup>. Por su parte, la *Nueva gramática de la lengua española* observa en relación con este mismo tema que «el contraste descrito [la oposición funcional entre las dos perífrasis] se desdibuja a menudo, sobre todo en el presente de indicativo»<sup>58</sup>, una cita que, en el contexto en que aparece, sugiere que se trata del tiempo de la conjugación más proclive a la perífrasis en incidencia directa.

Sin embargo, nuestros datos ofrecen un panorama bien diferente. Lejos de ser el tiempo menos propenso al empleo de la variante prepositiva, el presente surge en nuestro corpus como el más favorecedor de dicha forma (26,3%), frente a cifras más modestas del condicional (18,1%) y del resto de los tiempos de la conjugación (22,9%)<sup>59</sup>. Sin embargo, ninguna de esas diferencias es seleccionada como significativa por el análisis multivariante. Y ello por la notable dependencia de este factor independiente con respecto a la modalidad oracional que analizábamos anteriormente (véase 4.6). Recuérdese como la oposición entre las oraciones interrogativas y las enunciativas arrojaba datos significativos, situándose las primeras como uno de los puntos estructurales menos propicios para la selección de la variante prepositiva (P. 25). Pues bien, como puede advertirse en la tabla 6, las diferencias obtenidas entre los diferentes tiempos de la conjugación no son ajenas a esta oposición.

<sup>56</sup> Carmen Silva Corvalán, «Contextual conditions for the interpretation of *poder* and *deber* in Spanish», *Modality in grammar and discourse*, Joan Bybee y Suzanne Fleishman, eds., Amsterdam, John Benjamins, 1995, págs. 88-89.

<sup>57</sup> Samper *et al.*, *op. cit.*, pág. 1.076.

<sup>58</sup> Real Academia Española/Asociación de Academias de la Lengua Española, *op. cit.*, página 2.143.

<sup>59</sup> Tan solo el imperfecto de indicativo y el pasado simple. No hay ocurrencias de otros tiempos.

<i>Tiempo</i>		<i>Enunciativa</i>		<i>Interrogativa</i>	
		<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>
Otros	<i>deber de</i>	60	25	1	25
	<i>deber</i>	177	75	3	75
Condicional	<i>deber de</i>	29	22	2	5
	<i>deber</i>	104	78	36	95
Presente	<i>deber de</i>	45	27	0	0
	<i>deber</i>	123	73	3	100

TABLA 6: Tabulación cruzada entre el tiempo/modo verbal y la modalidad oracional

El hecho de que el condicional aparezca en el corpus como el paradigma menos asociado a la variante prepositiva tiene mucho que ver con el hecho de que ese modo verbal aparezca con alguna frecuencia en las oraciones interrogativas (treinta y ocho casos en el *MCSCS*). Por el contrario, esa misma modalidad oracional apenas se asocia en la muestra analizada con el presente de indicativo (tres únicos casos) o con el resto de los tiempos (tan solo cuatro ejemplos en todo el corpus). Es por tanto, la asociación con la modalidad oracional —un factor significativo (§ 4.5)— lo que permite explicar diferencias que en otros contextos resultan apenas perceptibles, como puede comprobarse al comparar los índices recogidos por los diferentes paradigmas de la conjugación en la modalidad enunciativa.

## 5. REFLEXIONES FINALES

El análisis variacionista emprendido en la presente investigación nos ha permitido constatar el proceso avanzado de desplazamiento de la variante *deber de* a favor de la perífrasis sin preposición (*deber*). Con todo, dicho proceso parece hallarse en una fase más avanzada en el español de Hispanoamérica, mientras que la perífrasis en incidencia indirecta resiste todavía en algunos ámbitos del español peninsular. Las comunidades de habla castellanenses representan, precisamente, uno de esos ámbitos.

El estudio confirma también que la pretendida alternancia entre las dos variantes para la oposición modal entre contenidos epistémicos y deónticos no se sostiene en la realidad, al menos en la lengua oral. En general, los hablantes utilizan indistintamente ambas perífrasis tanto para la expresión de la probabilidad como de la obligación o la necesidad. Sin embargo, hay otros factores relacionados con la modalización que ayudan a explicar la variabilidad sintáctica. Algunos como el grado de espontaneidad o —sobre todo— el énfasis expresivo,

actúan a favor de la variante prepositiva. Tanto en las interacciones verbales más espontáneas como en aquellos contextos en los que el hablante introduce signos de intensificación para resaltar sus ideas, argumentos, dudas, etc. los índices de presencia de la preposición son significativamente más elevados. Otros factores, por el contrario, como la atenuación o la modalidad oracional interrogativa, lo hacen en contra. En este caso, es posible que el principio de economía condicione la elusión de redundancias innecesarias. Si los contenidos menos seguros en el discurso se hallan ya suficientemente asentados, ya sea por la presencia de atenuantes, ya por la propia naturaleza de las oraciones interrogativas, el empleo de un recurso adicional, como la preposición, parecería innecesario.

En el fondo estos datos resultan lógicos pues, no en vano, intensificación y atenuación representan polos opuestos en un continuum de modalización. De hecho, reanalizados ambos factores en uno solo, obtenemos un nuevo factor significativo, y esta vez con el mayor potencial predictivo de todos los considerados en el estudio (rango 38). Como consecuencia de ello, y tal como puede apreciarse en el gráfico adjunto, se obtiene una marcada distribución lineal. En el extremo menos favorable para la preposición se encuentran los enunciados mitigados (13,2% / P.35), mientras que el énfasis favorece la variante de forma clara (43,5% / P.73), quedando el resto de los enunciados en una posición significativamente neutra (23,5% / P. 55).

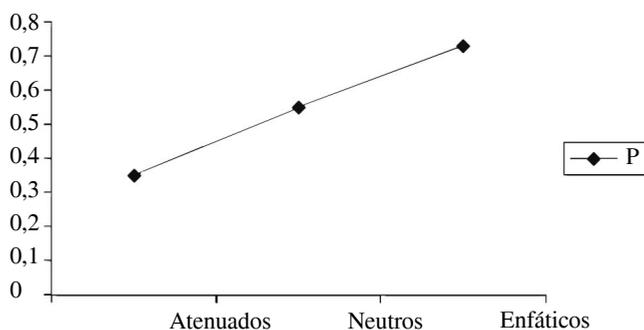


GRÁFICO 1: Probabilidades asociadas a los contextos de la *modalización* en la selección de la variante *deber de* + infinitivo

Por otro lado, hemos comprobado como estos factores —en especial el énfasis— son a menudo responsables de las principales diferencias obtenidas en otros, como sucede con la persona gramatical o el grado de animacidad de los sujetos. Asimismo, la asociación con la modalidad interrogativa permite explicar algunas diferencias entre diversos tiempos de la conjugación.

Ninguno de esos factores permite explicar por sí mismo la variabilidad observada, como tampoco otros que se han considerado en el estudio. Incluso, algunos han arrojado resultados contrarios a lo esperable. Así ocurre, por ejemplo, con el contexto fono-sintáctico, donde al margen de las combinaciones más cacofónicas, que invitan a la supresión de la preposición (*debe [de] decir*), no encontramos apenas diferencias entre los contextos intervocálicos y consonánticos. Con todo, otras diferencias merecen ser estudiadas con más detalle en el futuro, ya que no puede descartarse que una baja participación muestral, como la que se aprecia en algunos grupos especialmente (por ejemplo, la no adyacencia entre verbo auxiliar y auxiliado, o los tiempos compuestos del verbo principal), impida la obtención de resultados más consistentes.

Otras tareas quedan también pendientes para investigaciones futuras. Una de ellas tiene que ver con la distribución sociolectal de esta alternancia sintáctica, es decir, con el modo en que diferentes grupos sociales asimilan los patrones de variación lingüística y estilística descritos en el presente estudio. Al mismo tiempo, sería estimulante comparar los datos de estas investigaciones sincrónicas con los ofrecidos en otros cortes en la historia del español, con el fin de averiguar qué factores han regulado este caso de variabilidad sintáctica en el pasado y si son los mismos que en la actualidad.