

Dirección de la autora

Adriana Patiño Santos
C/ Rosselló, 419, 2-2
08025
Barcelona – España
Tel: (+34)696741578

Variation in the use of *deber* and *deber de* in written and oral materials from Latin America and Spain*

David Eddington and Carmen Silva-Corvalán

According to prescriptive grammars, the uses of the modal verbs *deber* and *deber de* correspond to different meanings. *Deber* is used to indicate obligation or necessity (e.g. *Juan debe presentarse al trabajo mañana*), while *deber de* indicates supposition, probability or conjecture (e.g. *Juan debe de tener unos 30 años*). However, these prescribed normative usages are often not followed; *deber de* and *deber* are used to express both obligation and possibility. We extracted 674 cases of *deber (de)* from the Arthus corpus and submitted them to logistic regression analysis in order to explore a number of factors that play a part in the choice of modal verb. We used the same data to determine what factors correlate with the expression of possibility or obligation.

The analysis indicates that the use of *deber* for obligation and *deber de* for possibility is followed quite often, although more so in Spain; in Latin America, *deber de* is much less frequent. Dynamic verbs and expressions of obligation are more common with *deber*. Stative verbs and expressions of possibility occur with *deber de*. Verbal tense also appears as an influencing factor. The past tense and present perfect are more likely to be used to express possibility rather than obligation. Phonetic factors also play a role. In Spain, *deber de* is more common when *de* is preceded by *-n* or *-s*. Following a vowel, *deber* is more common. This is possibly due to the deletion of /d/ in *de* when the preposition is intervocalic, and its retention following a consonant.

Keywords: *deber (de)*, modal verbs, modality, corpus approach

1. Introduction

According to prescriptive grammars, the uses of the modal verb *deber* and *deber de* followed by an infinitive correspond to different meanings (Cascón Martín 1999; Gómez Torrego 1999; Real Academia Española 1973). *Deber* is used to indicate obligation or necessity, as in (1) and (2).

- (1) Juan debe presentarse al trabajo mañana.
'Juan should/must/needs to show up at work tomorrow.'
- (2) Para sacar una buena nota debes estudiar mucho.
'To get a good grade you should/must/need to study a lot.'

On the other hand, *deber de* indicates supposition, probability or conjecture, as in (3) and (4).

- (3) Juan debe de tener unos 30 años.
'Juan must be about 30 years old.'
- (4) Con esa ropa deben de ser turistas.
'With those clothes they must be/are probably tourists.'

There is, in addition, a transitive verb *deber* 'to owe' (e.g. *Te debo \$1* 'I owe you \$1'), which we do not examine here because, at this synchronic stage of the language, it clearly is a different lexical item from modals *deber* and *deber de*.

One consensus found in the literature is that the prescribed normative usages are often not followed by native speakers of Spanish. That is, *deber de* and *deber* are used to express both obligation and possibility (Aracil 2001; Butt & Benjamin 1988; Cuervo 1994; Real Academia Española 1973). What has been termed confusion between the two forms appears to be typical throughout the Spanish-speaking world rather than a characteristic of a particular variety. It has been documented in Houston, Texas (Fairclough 2000), Valencia (Gómez Molina 2008), Madrid (Gómez Manzano 1992; Ruiz Gurillo 2002; Silva-Corvalán 1995), Las Palmas (Troya 1998), as well as in a number of capital cities in the Americas: Mexico City, San José, Bogotá, Lima, Santiago, Buenos Aires, La Paz, Caracas and San Juan (Samper, Hernández & Troya 1998–1999). The reason for the apparent confusion is quite simple. Historically, the meanings of obligation and possibility have been expressed with both *deber* and *deber de*, which means that the prescribed norms are an innovation that does not necessarily correspond to older usages and is, therefore, an artificial imposition created by modern grammarians (Yllera 1980: 128).

The variation is indeed not surprising given that necessity and possibility are related notions (Silva-Corvalán 1995) located on a semantic continuum rather than being discretely differentiated. Like *must* in English, *deber* and *deber de* can be used in either an epistemic or a deontic sense, as shown further in (5–6). An example such as (5) involves epistemic possibility, while an example of the kind in (6) involves deontic necessity or obligation (see Lyons 1977: 791).

- (5) Ese debe ser / debe de ser Fernando. [Epistemic possibility]
'That must be Fernando.'

- (6) Juan debe / debe de ver médico si tiene fiebre. [Deontic obligation/necessity]
'Juan must see a doctor if he has a fever.'

Modal verbs (e.g., *poder* 'can', *deber* 'must', *solér* 'be used to') convey a range of meanings that reflect a speaker's attitude towards the contents of a proposition: certainty, probability, possibility, belief, obligation, necessity, permission, volition, intention, doubt, etc. For example, the proposition *Juan viene mañana* 'John is coming tomorrow' can be modalized to express obligation (or necessity) or uncertainty about the probability of actualization of the event encoded in the proposition by using the modal verbs *deber* or *deber de*. These modals interact with linguistic and extra-linguistic elements that contribute to producing their contextual or discourse meanings of obligation, necessity or possibility. Every linguistic element contributes a specific meaning to the total, which goes beyond the sum of these meanings because it incorporates pragmatic factors that are context-dependent.

The fact that the modality or semantic meaning of *deber* and *deber de* is not a clear-cut criterion for distinguishing between these modals does not necessarily mean that their uses are in completely free variation. As is to be expected, other factors play a part in determining the choice of modal and its interpretation. For example, some researchers suggest that tense is a factor influencing the choice of *deber* or *deber de* (Aracil 2001; Samper et al. 1998–1999; Silva-Corvalán 1995). The person the verb is inflected for may also influence the choice of whether to include *de* or not (Aracil 2001).

However, this in turn may be due to the phonetic tendency to delete /d/ intervocally in Spanish. Inflections ending in vowels place /d/ in this deletion favoring context. For example, *debe de estar cerca* may be realized phonetically as [deβe:star serka] with no *de* on the surface. In contrast, those ending in a consonant (e.g. -s, -n, *debíamos*, *deben*) would favor the realization of *de*. Some evidence that this phonetic factor influences the variation between *deber* and *deber de* has already been gathered (Gómez Manzano 1992; Gómez Torrego 1999; Samper et al. 1998–1999).

The region of origin of the speaker may also be a factor that governs the variation. For example, in their data from educated speakers from twelve cities, Samper et al. (1998–1999) found that *deber* is used to express probability more often by speakers from Lima, Buenos Aires and La Paz, and less often by speakers from the other nine cities he studied (Mexico City, San José, Costa Rica, Bogotá, Santiago, Caracas, San Juan Puerto Rico, Madrid, Sevilla, Las Palmas).

One difficulty that arises when drawing conclusions based on extant studies is that many are based on a small number of instances of *deber* and *deber de*. Although Samper et al. (1998–1999) gathered a respectable 487 instances of *deber* and 63 of *deber de*, when these are subdivided according to city the data become

quite diluted. For instance, there are only 22 total tokens from Buenos Aires speakers and 45 from Lima. Troya Déniz's (1998) study of Las Palmas is based on 91 total tokens; González Pérez's (2009) on 31 instances from Guadalajara, Spain; Gomez Manzano's (1992) on 146 tokens from Madrid, and although based on a higher number of cases than the preceding studies, Gomez Molina's (2008) 261 tokens from Valencia may still be considered to constitute a small amount of data.

The present study, then, is motivated in part by the scarcity of data in many of the existing treatments of *deber* and *deber de*, and by the need to identify the relative weight of the various factors proposed to favor *deber* or *deber de*. Relative weight could then allow us to establish a hierarchy of probabilistic constraints on the choice of these two modals. To this end, we employ a variationist methodology in the qualitative analysis of the data and apply Goldvarb for the statistical analysis in order to determine the extent to which these modals follow prescribed norms, as well as to uncover the contribution of other factors in the observed variation.

2. Methodology

We obtained the data from the Arthus Corpus compiled at the University of Santiago de Compostela.¹ The corpus includes written material from narratives, newspapers, essays and plays from both Latin American and Peninsular authors. We identified 568 cases of *deber* and 106 cases of *deber de* in the corpus (for a total of 674 tokens). Although this is a much larger data set than in previous studies we recognize up front two limitations of the data. First, about 85 percent of the corpus is from written sources, while the remainder is based on oral materials. Second, only 18 percent of the data is from Latin America² and 82 percent from Spain, and the Latin American data come from only three countries: Argentina 46%, Colombia 11% and Mexico 43%. Of course, conclusions based on these data must keep the limited distribution in mind. We trust that further research will take the speech of other regions into consideration.

Each instance was coded for a number of factors hypothesized to have an effect on the choice of modal. These included tense of the modal, type of modality, final phonetic segment of *deber* and aspect of the main verb. Based on previous findings (Samper et al. 1998–1999; Silva-Corvalán 1995, among others), we coded for three specific tenses: Present indicative, present perfect and past (preterit and imperfect), and grouped the remainder in the 'other tenses' category because the data did not contain a sufficient number of tokens to establish further tense differences. Other tenses included tokens from the past perfect, gerund, infinitive, future, conditional, future perfect, imperfect subjunctive, present perfect subjunctive and past perfect subjunctive tenses.

In 'type of modality' we initially coded for several degrees of obligation: strong and weak obligation (examples 7–8), and also simply suggestion (example 9), taking a number of contextual factors into account to establish these degrees of obligation (see Silva-Corvalán 1995).

- (7) Todos debemos obedecer las leyes del tránsito. [Contextual meaning: strong obligation.]
'We must all obey traffic regulations.'
- (8) Debemos votar cuando hay elecciones. [Contextual meaning: weak obligation]
'We should vote when there are elections.'
- (9) Si te sientes mal debes ver médico. [Contextual meaning: suggestion]
'If you are not feeling well you should/could see a doctor.'

Table 1. Factors influencing the choice of *deber* (versus *deber de*) in Latin America

Corrected Mean	0.983		
Log likelihood	-34.432		
Total N	222		
	Factor weight	%	N
Tense of the modal			
Present indicative	.731	99	96
Other	.465	98	63
Past (preterit and imperfect)	.201	83	63
Present perfect	-		
Range	530		
Type of modality			
Obligation	.637	99	169
Possibility	.139	81	52
Range	498		
Semantic aspect of main verb			
Dynamic	n.s. [.521]	96	190
Stative	n.s. [.372]	87	31
Range	149		
Final phone of inflection of <i>deber</i>: (C (-s, -n) vs. V)			
Consonant	n.s. [.566]	97	72
Vowel	n.s. [.468]	93	149
Range	98		

Table 2. Factors influencing the choice of *deber* (versus *deber de*) in Spain

	Factor weight	%	N
Corrected Mean	0.922		
Log likelihood	-118.864		
Total N	453		
Tense of the modal			
Other	.873	99	95
Present indicative	.511	86	236
Present perfect	.420	63	8
Past (preterit and imperfect)	.158	50	114
Range	715		
Type of modality			
Obligation	.689	94	340
Possibility	.083	34	113
Range	606		
Final phone of inflection of <i>deber</i>: (C (-s, -n) vs. V)			
Vowel	.630	81	314
Consonant	.232	76	139
Range	398		
Semantic aspect of main verb			
Dynamic	.539	85	379
Stative	.309	49	74
Range	230		

The contextual information provided by the data proved not to be sufficient, however, to establish different degrees of obligation. Indeed, only nine instances of a weaker sense of obligation were identified. Therefore, we decided to include only two discrete types of modality for the quantitative analysis: obligation (both weak and strong) versus possibility. With respect to the final phonetic segment of *deber*, the opposition was between final vowel (as in *debe (de)*) and final consonant (*deben (de)*).

Silva-Corvalán (1995) shows that the aspect of the main verb influences the interpretation of the semantics of the modal. A dynamic main verb appears to favor an interpretation of obligation, while a stative main verb favors a reading of possibility. Accordingly, we coded the data by these two verbal aspects.

The data were then submitted to analysis by logistic regression using Gold-Varb (Robinson, Lawrence & Tagliamonte 2001). The first two analyses determine what factors are associated with the overall use of *deber* versus *deber de*. One was carried out on the data from Latin America, and the other on the data from Spain. A third analysis aims at examining the more interesting question of what factors are used to convey the meaning of either obligation or possibility. Factors that do not reach statistical significance in both analyses are indicated as n.s. (not significant) in the tables.

3. Results and discussion with *deber* and *deber de* as the dependent variable

As far as the region of origin of the speaker is concerned, we made a bipartite division between Latin American and Peninsular varieties and this yielded an interesting finding. *Deber de* is highly infrequent in Latin American Spanish; only 12 tokens of *deber de* out of 221 total tokens appeared (i.e. 5% of all tokens; see Appendix). Interestingly, eleven of the twelve examples are in the preterit or the imperfect and ten of these are interpreted to convey possibility. Example (5) in the preterit is interpreted to convey obligation, as is example (12), in the present tense, in the extended context in which it occurs (see examples in the Appendix). This contrasts with Spaniards, who used *deber de* in 21% of the cases (94 tokens of *deber de* out of 453 total tokens from Spain). The data contained some indications of interaction between region of origin (Latin America or Spain) and a number of other variables, which will be discussed below. For this reason, separate analyses were done on the data from each of the two regions.

Our data contain no instances of *deber* or *deber de* from Latin America in the present perfect, which poses no problem for the statistical analysis. However, in the Latin American data all cases in a tense other than the past (preterit or imperfect) or present appear with *deber* and none with *deber de*. This results in a knockout variable, which does not allow logistic regression to be performed. We assume that this may be due to the limitations of our sample, which simply did not register any such instances, rather than to some sort of incompatibility between particular tenses and *deber de*. For example, we found a case of *deber de* in the imperfect subjunctive from a Latin American source in the 20th century section of the *Corpus del Español*³ (*Debieran de existir otras alternativas*). It is not difficult to search Google by country and find thousands of instances of *deber de* in other tenses as well. For this reason we felt justified in adding a fictional token in the 'other tense' category (see Guy & Bayley 1995; Paolillo 2002: 165) in order to eliminate this knockout and allow us to include tense in the analysis.

The factor weights calculated by GoldVarb range from zero to one. Factor weights closer to one strongly favor the use of *deber*, while those closer to zero are least likely to favor *deber* meaning they tend more towards *deber de*. There are two ways of measuring how influential a factor is compared another. One is by comparing the ranges of the factors. The range is the highest weight minus the lowest for a given factor multiplied by 1000. As evident by the ranges in Tables 1 and 2, the most influential factor on the choice of *deber* or *deber de* turns out to be verb tense, thus supporting the effect that has been noted in previous studies (Aracil 2001; Samper et al. 1998–1999; Silva-Corvalán 1995). However, factors with more values tend to have larger ranges which may skew this measure of influence.

A more sensitive measure of factor influence is calculated by comparing the log likelihood of the analysis with all significant factors included with the resulting log likelihood when each factor is removed from the analysis. This is done automatically in the stepping down portion of the GoldVarb analysis. Log likelihoods closer to zero indicate a better fit of the data. Therefore, factors that result in a log likelihood farther from zero when they are removed are more influential factors. The statistical significance of the difference between log likelihoods of the model with all factors, and with all but one factor included, may be calculated.⁴ For example, when verb tense is removed from the Latin American data, the log likelihood moves from -34.432 to -39.870 . This difference is significant at the $p = 0.0043$ level ($\chi^2(2) = 10.876$). However, when type of modality is removed the difference is even more significant ($\chi^2(1) = 9.448, p = 0.0021$) suggesting that type of modality exerts more influence than verb tense. Similar analysis of the data from Spain yields this hierarchy of factor influence in which type of modality and verb tense are equally influential followed by final phone and then semantic aspect.⁵

As the data indicate, in both regions, *deber* occurs most likely in the present and 'other tense' categories, while it is less likely in the past tense. However, an

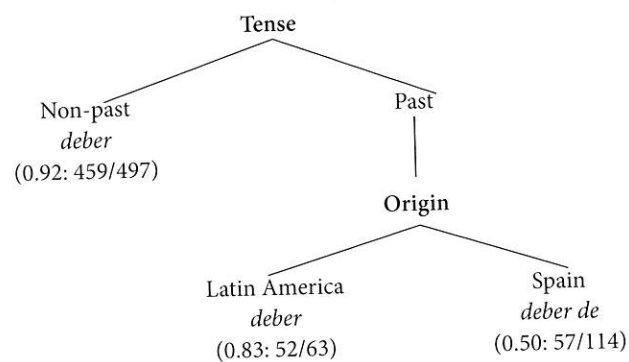


Figure 1. Decision tree of choice between *deber* and *deber de* based on region of origin and tense

interaction between region and tense is evident in the raw percentages. This is illustrated in the tree in Figure 1. *Deber* is the modal of choice for non-past tenses in both regions. However, in past tenses *deber* is favored in Latin America 83% of the time, while in Spain it is an even toss up between *deber* and *deber de*.

Numerous studies have pointed out that the uses of *deber de* and *deber* do not always correspond to possibility and obligation, respectively, as prescriptive grammars would advocate. Although there is some degree of variation in our data, the tendency toward prescriptive norms is indeed significant: *deber* is strongly favored in sentences expressing obligation and disfavored in examples expressing possibility or supposition. Nevertheless, interaction by region is evident here as well. Of the 340 cases from Spain that involve an expression of obligation, 321 (94%) were expressed with *deber* rather than *deber de*. This does not differ much from the Latin American data where obligation was expressed with *deber* rather than *deber de* in 99% of the 169 cases.

When the data from both regions are considered together 96% of cases of obligation (488/509, see Figure 2) are expressed with *deber*. However, Latin Americans also express possibility with *deber* 81% of the time (42 instances of possibility expressed with *deber* over 52 total cases that expressed possibility). This contrasts with Spaniards who express possibility with *deber* in 34% of the 113 cases. To express possibility Spaniards employ *deber de* 66% of the time, and Latin-Americans only 19% of the time. This suggests to us that the prescriptive division of labor between the two modals may be more robust in Peninsular rather than Latin American Spanish.

The phonetic factor was also apparent in the data. In intervocalic position, /ð/ is commonly deleted in Spanish, especially in less formal registers. Inflections of *deber* may end in a vowel, or in the consonants /n/ or /s/. A number of researchers (Gómez Manzano 1992; Gómez Torrego 1999; Samper et al. 1998–1999) suggest

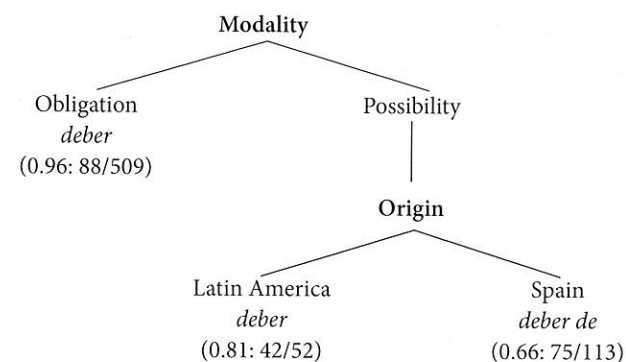


Figure 2. Decision tree of choice between *deber* and *deber de* based on region of origin and modality

that in inflections that place *de* between vowels (e.g. *debía de*, *debe de*) the /ð/ of the preposition may be deleted and the preposition's vowel may be fused with the surrounding vowels. In contrast, deletion is not as probable in cases in which *de* is not intervocalic (e.g. *debían de*, *debes de*). This phonetic factor is statistically significant in the data from Spain where *deber* is indeed favored when the inflection of *deber* ends in a vowel (e.g. *debe de empezar*), while *deber de* is more probable when *deber* ends in a consonant. Although this factor did not reach significance in the Latin American data, the trend is in the same direction, which indicates that this phonetic process plays a part in the use of *deber* versus *deber de*.

A problem with this phonetically motivated deletion is that most of our data come from written sources, which means that a phonetic process is less likely to influence it. This leads us to believe that we are dealing with the historical remnants of a phonetic process in which deletion of *de* in the favorable phonetic contexts throughout the historical development of Spanish eventually became grammaticalized and is now therefore reflected also in the written language.

A small, but significant effect of verb type is also evident from the analysis. In both regions, dynamic verbs (see Vendler 1957) favor *deber* much more than statives. In Spain, however, statives are more likely to appear with *deber de* (Figure 3). This distinction is not significant in Latin America. This same trend is also apparent in how often the statives *ser* and *estar* are used with each modal. Twenty five percent of the cases of *deber de* appear with these stative verbs, while they are used with *deber* only four percent of the time.

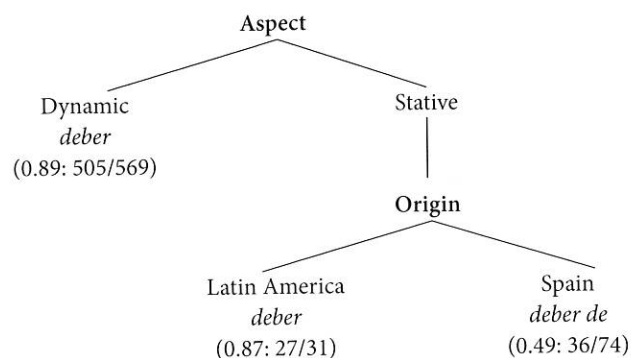


Figure 3. Decision tree of choice between *deber* and *deber de* based on region of origin and the semantic aspect of the modalized verb

4. Analysis with possibility versus obligation as the dependent variable

In the first analysis, we determine what factors influence the use of *deber* and *deber de* without considering the meaning contributed by the modal. Of greater interest, however, is the identification of the factors that correlate with the interpretation of obligation or possibility. We examined this issue by recoding the data so that the dependent variable was obligation versus possibility. The independent variables were verbal tense, modal verb and semantic aspect of the modalized verb.

4.1 Results and discussion of the analysis with possibility versus obligation

Judging by the ranges reported in Table 3, the most influential factor in determining an interpretation of obligation or possibility is the tense of the modal (see also Aracil 2001; Samper et al., 1998–1999; Silva-Corvalán 1995). Comparison of log likelihoods when factors are removed one at a time results in differences that are all highly significant.⁶ Obligation is more often expressed with the tenses in the 'other' category (see Section 2) and to a lesser degree the present tense, while possibility is more likely in the past and present perfect tenses. The modal *deber* is favored to express obligation, and *deber de* to express possibility. This largely coincides with the prescribed usages (Cascón Martín 1999; Gómez Torrego 1999; Real Academia Española 1973).

A new finding of our study is the relationship between the aspect of the main verb (dynamic versus stative) and the choice of modality. Stative verbs highly favor a reading of possibility. Dynamic verbs, in contrast, favor the expression of obligation. This is not surprising since the type of event people feel obligated to carry out or try to oblige others to perform is normally dynamic. On the other hand, suppositions about a possible state of affairs tend to involve stative rather than dynamic situations.⁷ Note the different interpretations of examples (10) and (11) with the modal *deber*. With no further context, example (10) is interpreted to convey possibility with a stative modalized verb and (11) is interpreted to convey obligation with a dynamic verb.

- (10) Pepe debe tener 30 años.
'Pepe must be 30 years old.'
- (11) Pepe debe visitar a su madre.
'Pepe must visit his mother.'

Table 3. Factors favoring a meaning of obligation (versus possibility).

Corrected Mean	0.833		
Log likelihood	-229.960		
Total N	674		
	Factor weight	%	N
Verb tenses			
Other	.790	96	157
Present indicative	.541	82	332
Past (preterit and imperfect)	.212	48	177
Present perfect	.023	13	8
Range	767		
Modal verb			
<i>Deber</i>	.594	86	568
<i>Deber de</i>	.115	20	106
Range	479		
Semantic aspect of main verb			
Dynamic	.581	82	569
Stative	.145	38	105
Range	436		

5. Conclusions

Given the small number of tokens on which previous studies of *deber* and *deber de* were based, our goal was to examine the uses of these two modals by examining a significantly larger number of instances. On the one hand, we concur with the observation of previous researchers that both modals are used to express obligation and possibility. On the other hand, there is a significant tendency for speakers to follow prescribed norms and express obligation with *deber* and possibility with *deber de*. This tendency is much stronger in Peninsular Spanish, reflected in the 21% of tokens with *deber de* in data from Spain, while only 5% of tokens with *deber de* in Latin America.

Since *deber de* is more widely used in Peninsular Spanish, this significantly raises the possibility of using *deber* and *deber de* contrastively. This explains the numerous interactions involving region that appear in the data. Spaniards more often use *deber de* in the past tense and *deber* in non-past tenses; they express obligation with *deber* and possibility with *deber de*; they prefer *deber de* with stative verbs and *deber* with dynamic verbs. The paucity of *deber de* in Latin America

means that *deber* is the modal of general applicability and is not often contrasted with *deber de* to distinguish nuances of usages or meaning.

When taken as a whole, a number of generalizations were found. *Deber* is most likely to be used in the present indicative tense and other non-past tenses. It generally expresses obligation and appears as a modal with dynamic main verbs. In contrast, *deber de* is associated with past and present perfect inflections. It most often expresses possibility, and is used most often with stative verbs. Verbal inflections of *deber* that end in a consonant are more likely to appear followed by *de* while inflections that end in a vowel place *de* in intervocalic position, which favors deletion of /d/, and may explain why *deber* without *de* is favored in these cases.

As is always the case in corpus studies, the results are dependent on the contents of the corpus examined. We made a broad division between speakers from Latin America and those from Spain. In fact, 67% of the instances come from Peninsular speakers. Surely more subtle nuances in usage will be found when data that make finer-grained geographical distinctions are analyzed. In like manner, the large majority of the corpus from which the data were taken (85%) is derived from written sources. Further study into possible written versus spoken usages is clearly warranted.

Notes

* We are grateful to Prof. Guillermo Rojo for allowing us access to the Arthus corpus and for his help with the extraction of the data used in this study. Special thanks go also to Fleur van der Houwen for her help with data coding.

1. Our data were taken from the present Arthus corpus (available from www.sintx.usc.es/Arthus.html) minus the oral data from Buenos Aires, which was added at a later date.
2. The Latin American data come from Argentina (46%), Colombia (11%) and Mexico (43%).
3. www.corpusdelespanol.org
4. The absolute difference between the log likelihood of the two models times two gives the Chi square (Chatterjee & Hadi 2006). The degrees of freedom for one model is total number of values in each factor minus the number of factors in the model. The significance of Chi square is determined by on the basis of the difference between the degrees of freedom in each model.
5. Without verb tense $\chi^2(3) = 34.748, p < 0.0001$; without modality $\chi^2(1) = 81.576, p < 0.0001$; without final phone $\chi^2(1) = 9.764, p = 0.0018$, without semantic aspect $\chi^2(1) = 5.348, p = 0.0207$.
6. Without verb tense $\chi^2(3) = 66.638, p < 0.0001$; without semantic aspect $\chi^2(1) = 50.162, p < 0.0001$; without modal verb $\chi^2(1) = 72.476, p < 0.0001$. Direct comparison of χ^2 at different degrees of freedom is not possible, so verb tense cannot be compared with other factors.

However, the higher χ^2 when modal verb is not included suggest that it is more influential than semantic aspect.

7. Following Comrie (1976), we use 'situation' as a technical term to refer inclusively and indistinctly to actions, processes, events, states, etc.

References

- Aracil, Nadia. 2001. "Deber / deber de". *Panorama de la linguistique hispanique* ed. by Yves Macchi, 195–205. Villeneuve d'Ascq: Université Charles-de-Gaulle-Lille 3.
- Butt, John and Carmen Benjamin. 1988. *A new reference grammar of Modern Spanish*. London: Edward Arnold.
- Cascón Martín, Eugenio. 1999. *Manual del buen uso del español*. Madrid: Editorial Castalia.
- Chatterjee, Samprit and Hadi S. Hadi. 2006. *Regression analysis by example*. Hoboken, New Jersey: Wiley Interscience.
- Comrie, Bernard. 1976. *Aspect*. Cambridge: Cambridge University Press.
- Cuervo, Rufino José. 1994. *Diccionario de construcción y régimen de la lengua castellana*. Bogotá: Instituto Caro y Cuervo.
- Fairclough, Marta. 2000. "Expresiones de modalidad en una situación de contacto: Deber (de) vs. tener que en el español hablado en Houston". *Southwest Journal of Linguistics* 19. 19–30.
- Gómez Manzano, Pilar. 1992. *Perífrasis verbales con infinitivo*. Madrid: Universidad de Educación a Distancia.
- Gómez Molina, José R. 2008. "Estudio sociolingüístico de las perífrasis 'deber+infinitivo' y 'deber de + infinitivo' en el español hablado". *Homenaje al profesor César Hernández*, 865–878. Valladolid: University of Valladolid.
- Gómez Torrego, Leonardo. 1999. "Los verbos auxiliares. Las perífrasis verbales de infinitivo". *Gramática descriptiva de la lengua española*, vol. 2 ed. by Ignacio Bosque and Violeta Demonte, 3323–3389. Madrid: Espasa Calpe.
- González Pérez, M^a Nieves. 2009. "Variación gramatical en el habla de Guadalajara: Alternancia en el uso de las formas que/ de que y deber/ deber de". *Interlingüística* 18. 542–552.
- Guy, Gregory R. and Robert Bayley. 1995. "On the choice of relative pronouns in English". *American Speech* 70. 148–162.
- Lyons, John. 1977. *Semantics*. Cambridge: Cambridge University Press.
- Paolillo, John C. 2002. *Analyzing linguistic variation: Statistical models and methods*. Stanford: CSLI Publications.
- Real Academia Española. 1973. *Esbozo de una nueva gramática de la lengua española*. Madrid: Espasa-Calpe.
- Robinson, John, Helen Lawrence and Sali Tagliamonte. 2001. "GoldVarb 2001: A multivariate analysis application for Windows". Accessed 3 October 2007. Available from www.york.ac.uk/depts/lang/webstuff/goldvarb.
- Ruiz Gurillo, Leonor. 2002. "Algo más sobre el empleo de *deber* y *deber de*". *Español Actual* 77–78. 157–159.
- Samper, Juan Antonio, Clara E. Hernández and Magnolia Troya. 1998–1999. "El uso de *deber* + infinitivo y *deber de* + infinitivo en la norma lingüística culta de América y España". *Boletín de Filología*, 37. 1059–1083.

- Silva-Corvalán, Carmen. 1995. "Contextual conditions for the interpretation of 'poder' and 'deber' in Spanish". *Modality in grammar and discourse* ed. by Joan Bybee & Suzanne Fleischman, 67–105. Amsterdam: John Benjamins.
- Troya Déniz, Magnolia. 1998. *Perífrasis verbales de infinitivo en la norma lingüística culta de Palmas de Gran Canaria*. Madrid: Real Academia Española / Universidad de la Palmas de Gran Canaria.
- Vendler, Zeno. 1957. "Verbs and times". *The Philosophical Review* 66. 143–160.
- Yllera Alicia. 1980. *Sintaxis histórica del verbo español: Las perífrasis medievales*. Zaragoza: Department. de Filología Francesa.

Appendix. Cases of *deber de* from Latin America

- 1) Debió de adivinar mi irritación porque se disculpó.
- 2) Yo estaba tan mal, que debió de creerme.
- 3) La cerradura debía de estar rota, porque una cadena con candado sujetaba las dos hojas de la puerta.
- 4) Debía de figurarse que no me quitaba nada, pues yo me había alejado de Daniela.
- 5) Debí de pensar que ya había hecho bastante de mi parte y que si Massey insistía, no podría negarme por mucho tiempo.
- 6) Debí de perder la conciencia, porque desperté en pleno vuelo, en el momento en que la azafata colocaba las bandejas para la comida.
- 7) Doña Salomé debió de sentirse halagada y agradecida de que su marido se tomara tales cuidados.
- 8) Y sólo una casualidad me permitió rescatar al cabo de cinco años de búsqueda unos 322 pliegos salteados de los más de 500 que debió de tener el sumario.
- 9) Alguna forma, algún aspecto, debía de tener.
- 10) Pensó que las fotografías debieron de tomarse en las primeras horas de la mañana.
- 11) Pero algo debí de vislumbrar en la cara de Bermúdez, que me distrajo de tales consideraciones.
- 12) Deben de tener rodeada toda la iglesia.

Authors' addresses

David Eddington
Department of Linguistics and
English Language
Brigham Young University
4064 JFSB
Provo, UT 84660
eddington@byu.edu

Carmen Silva-Corvalán
Departments of Spanish and Linguistics
University of Southern California
University Park Avenue
Los Angeles, CA 90089-0358
csilva@usc.edu