A Usage-Based Perspective on Spanish Variable Clitic Placement

Pablo E. Requena

Department of Modern Languages and Literatures, University of Texas at San Antonio,
San Antonio, TX 78249, USA; pablo.requena@utsa.edu

Abstract: This study provides a usage-based analysis of Spanish Variable Clitic Placement (VCP). A variationist analysis of VCP in spoken Argentine Spanish indicates that VCP grammar is constrained by lexical (finite verb) and semantic (animacy) factors. Considering the finite effect, the study focuses on usage-based accounts for the gradience attested across finite verb constructions. Grammaticalized meaning and increased frequency tend to account for VCP in general. However, one [*tener que* + infinitive] construction is found exceptional in that it favors enclisis despite its grammaticalized meaning of obligation and its high frequency of use. Data from a larger corpus indicate that the [*tener que* + infinitive] construction lacks unithood, signaling great analyzability of its component elements. Through an exemplar analysis, the [*haber que* 'must' + infinitive] construction that categorically takes enclisis and which is strongly linked to [*tener que* + infinitive] diachronically, semantically, and structurally emerges as a likely analogical model for VCP with *tener que*, pushing *tener que* towards enclisis. This study not only illustrates how usage-based linguistics can capture VCP more generally, but also how this framework provides powerful tools to discover the constraints on VCP in naturalistic use in order to account for individual construction behavior.

Keywords: usage-based linguistics; corpus; variationist sociolinguistics; clitics; Spanish; clitic climbing; clitic placement; finite verb; animacy

1. Introduction

How do usage-based approaches to language account for gradience and variation? In this study, I provide a usage-based perspective on a phenomenon that has attracted linguists' attention for years, namely Spanish Variable Clitic Placement (VCP) in [finite verb + non-finite verb] constructions, as in (1). In modern Spanish, clitics may follow the non-finite verb (enclisis, as in (1a)) or precede the finite verb (proclisis, as in (1b)) in a number of these [finite verb + non-finite verb] contexts.

(1)   a.   *Quiero*   *comprar* = ***lo***   (Enclisis)
          want-prs.1sg   buy-inf = **it-acc-m3sg**
          'I want to buy **it**'

       b.   ***Lo***   *quiero*   *comprar*   (Proclisis)
          it-acc-m3sg   want-prs.1sg buy-inf
          'I want to buy **it**'

Usage-based linguistics considers language a complex adaptive system that exhibits gradience and variation and whose structure emerges from use (Bybee 2010; Givón 1979; DuBois 1985; Hopper 1987). The present study exemplifies how usage-based linguistics accounts for VCP, also drawing on corpus linguistics and variationist sociolinguistics, which provide powerful tools to discover the constraints on VCP in naturalistic use (Tagliamonte 2006).

Previous corpus studies of VCP across dialects of Spanish have identified the finite verb heading the variable construction as the main constraint on the variation (e.g., Myhill 1989; Davies 1995). Other significant factors are of a semantic (animacy of clitic referent) and discourse (referent accessibility, topic persistence)[1] nature (e.g., Aijón Oliva and Nieto 2013; Schwenter and Torres Cacoullos 2014). The first goal of the present study is to extend this literature by presenting the results of a variationist study of VCP in Argentine Spanish, a dialect for which no variationist study exists up to this point. VCP in Argentine Spanish is shown to be constrained by finite verb construction as well as by the animacy of clitic referent. The results indicate that certain finite verb constructions favor enclisis, as do clitics with inanimate referents.

An interesting pattern about the behavior of particular finite verbs with respect to VCP was noted thirty years ago. Myhill (1989) pointed out that verbs whose meanings are aspectual, modal, or auxiliary-like favor proclisis, but verbs with more lexical meanings favor enclisis. His analysis of verbs considered one natural process of language evolution: *grammaticalization*, understood as the process by which a lexical form takes on grammatical functions in specific contexts of use (Bybee and Pagliuca 1987; Givón 1971; Heine and Reh 1984; Hopper and Traugott 2003). Through grammaticalization, for example, a lexical verb (e.g., a verb of motion, such as *ir* 'go') may acquire functional meaning (e.g., as a marker of futurity, as in *ir a* 'go to + infinitive'), becoming more like an auxiliary, modal, or aspectual marker. This process coincides with increased frequency (due to greater productivity of 'go' as future marker) as well as an increase in the degree of *unithood* with the following infinitive ([*ir a* 'go to' + infinitive]) through a process known as *chunking* (Bybee 2010, p. 7)[2]. As a result of the process described above, constructions of the type [finite verb + non-finite verb] that are headed by grammaticalized verbs are assumed to be represented in memory as single units (Goldberg 1995, 2006). During *categorization*, the process by which "words and phrases and their component parts are recognized and matched to stored representations" (Bybee 2010, p. 7), these variable constructions with grammaticalized verbs, now forming a unit that resembles an auxiliary followed by a main verb, may be mapped onto a related clitic construction where a clitic is used with a single finite verb [clitic + finite verb]. This should result in increased proclisis, since the model construction ([clitic + finite verb]) takes proclisis categorically in contemporary Spanish. Synchronic patterns of VCP with different verbs have, therefore, been assumed to result from an ongoing process of grammaticalization (Myhill 1988, 1989; Schwenter and Torres Cacoullos 2014; Torres Cacoullos 1999).

While the data in the corpus study of Argentine Spanish support the grammaticalization account that predicts increased proclisis with grammaticalized and frequent [finite verb + non-finite verb] constructions, one such construction ([*tener que* 'have to' + infinitive]) does not behave as predicted. Despite being a frequent construction headed by a verb of possession that has grammaticalized into obligation, [*tener que* 'have to' + infinitive] clearly favors enclisis. How can usage-based perspectives account for the apparently 'exceptional' behavior of the [*tener que* + infinitive] construction, which favors enclisis in naturalistic use? The second goal of this study is to show how grammaticalization, unithood, and links with other constructions can account for the VCP behavior of [*tener que* + infinitive]. First, I show that, as a relatively recent case of grammaticalization and unlike the other frequent grammaticalized verbs used with VCP, *tener* exhibits weak unithood with infinitives, rendering greater activation of representations of its component parts (*tener*, *que*) as used outside of the variable construction (a property known as *analyzability* (Langacker 1987, p. 292; Bybee 2010, p. 45; Brown and Rivas 2012)). Second, drawing on Bybee's (1985) Network Model of relations between exemplar representations, I show how [*tener que* + infinitive] keeps very strong diachronic, semantic, and structural ties with another construction ([*haber que* 'there be to' + infinitive + clitic]), which constitutes the analogical model pulling [*tener que* + infinitive] towards enclisis.

---

[1]　For some indication of extra-linguistic factors, see Gudmestad (2006).
[2]　In this study, I will also use "strength of association" to refer to "degree of unithood".

　　　While preliminary in nature, this study exemplifies how usage-based linguistics is able to account for gradience and variation in VCP in general, as well as to capture the behavior of one lexical construction that, under other perspectives in the study of language, would be deemed 'exceptional'.

2. VCP and Grammaticalization

　　　The phenomenon under investigation is Spanish direct object (accusative, ACC) clitics. I will define these clitics as phonologically and morphologically dependent elements located at a transitional stage toward affixation, which are used not only to point at a referent that is active in discourse, but also to signal its relative importance or topicality, mainly as a result of animacy considerations. Spanish direct object clitics (*me*, *te*, *lo/la*, *nos*, and *los/las*, in the non-leísta variety of Argentine speakers examined here) occur before a single finite verb (2) or after a non-finite verb (e.g., infinitive or gerund) (3)[3]. However, in a number of complex verb constructions consisting of a [finite verb + a non-finite verb], there are two positions available. Clitics may follow the non-finite verb (enclisis) or precede the finite verb (proclisis), as in (1a and b) above, which I duplicate below as examples (4a) and (4b) for convenience. I refer to this phenomenon as Variable Clitic Placement (VCP), a term that I propose should be used to name the phenomenon, instead of other names that point to theory-specific proposals, such as 'Clitic Climbing' (CC).

(2)　**Lo**　　　　　*compré.*
　　　it-ACC-　　buy-PST.1SG
　　　M3SG
　　　'I bought **it**'

(3)　*Para*　　　　*comprar* = **lo**
　　　to　　　　　buy-INF = **it-ACC-M3SG**
　　　'To buy it'

(4)　a.　*Quiero*　　　　　　*comprar* = **lo**　　　　　　　　(Enclisis)
　　　　　want-PRS.1SG　　　buy-INF = **it-ACC-M3SG**
　　　　　'I want to buy **it**'

　　　b.　**Lo**　　　　　　　*quiero*　　　　*comprar*　　　　(Proclisis)
　　　　　it-ACC-M3SG　　want-PRS.1SG　buy-INF
　　　　　'I want to buy **it**'

　　　From early on, studies of VCP in Romance languages have highlighted the relationship between clitic placement and properties of the finite verb ('matrix verb') in these two-verb constructions or periphrases (Strozer 1976)[4]. Whereas in Romanian, proclisis is obligatory with modal verbs (Gerlach 2002, p. 200)[5], Spanish (among other Romance languages, such as Italian) displays different degrees of what has been described as *optionality* depending on the finite verb. Early studies of clitic placement elicited grammaticality judgments to find out which contexts allow both clitic positions and which contexts allow enclisis only. These studies resulted in lists of finite verbs that allow both clitic positions, such as *querer* 'want', *tratar* 'try', and *soler* 'be in the habit of', and those that only allow enclisis, such as *insistir* 'insist', *soñar* 'dream', and *parecer* 'seem' (Aissen and Perlmutter 1976). Rizzi (1976, 1978) posited that the verbs used in proclisis tend to have meanings similar to those meanings that languages tend to express grammatically, namely modals, aspectuals, and motion verbs. Attention to the context

---

[3]　With imperatives, they occur before negative imperatives (*No lo toques* 'Do not touch it'), but after affirmative imperatives (*Tocalo* 'Touch it'). I do not consider such a context to be a variable context similar to (1a) and (1b).

[4]　I will use *constructions* and *periphrases* to refer to [finite verb+non-finite verb] units here, but I acknowledge that within formal linguistics, only some of these sequences pass certain tests to show their consolidation as periphrases.

[5]　For other languages (or dialects) in which CC seems obligatory, see (Cinque 2006, p. 32, fn. 47).

of a sentence prompted Napoli's (1981) criticism that "no simple list of verbs [. . .] can account for the fact that CC [proclisis] is appropriate with a given verb in one context, but not in another for the same speaker" (p. 878) even within these categories of verbs.

It became clear to some linguists that the position of the finite verb along a continuum from lexical to functional forms was key to understanding clitic placement. Grammaticalization was therefore the process by which linguistic elements move along this continuum towards functional forms (Bybee and Pagliuca 1987; Givón 1971; see also Hopper and Traugott 2003). Myhill (1988) linked patterns of VCP (likelihood of proclisis, in particular) with the grammaticalization of particular verbs by paying attention to changes in meaning, a process known as *desemanticization*, *semantic depletion*, or *bleaching* (e.g., Weinreich [1963] 1966, p. 180f). Along the lines of the observation by Rizzi that resulted in classes of verbs, Myhill found that particular finite verbs heading [finite verb + non-finite verb] constructions that allow both clitic positions appeared in proclisis more frequently if they corresponded to meanings that Bybee (1985) had reported as likely to be represented inflectionally in the world's languages, such as progressive or future meanings as well as epistemic modality. Among the many instances of grammaticalization processes in world languages, of particular interest for VCP was how a finite lexical verb acquires aspectual or modal meaning, leading to a semantic bonding between a semantically weak auxiliary-like finite verb and the non-finite form governing the clitic. Most of the frequent verbs used in VCP contexts have undergone processes of grammaticalization to a greater or lesser extent, resulting in novel uses with more functional meanings (e.g., *ir a* from motion to future). Hence, Spanish VCP has been referred to as " . . . a syntactic–semantic change in progress, which involves the gradual grammaticalization of a number of verbs" (Silva-Corvalán 1994, p. 128). Myhill (1989) considered VCP as a phenomenon that could shed light on synchronic grammaticalization processes in progress, so his quantitative analysis of VCP with frequent finite verbs emphasized the existence of degrees of grammaticalization, as attested in meaning differences that emerge as finite verbs heading VCP constructions behave more auxiliary-like (*auxiliarization*) and less like main verbs.

The result of the 'auxiliarization' of the finite verb in VCP constructions, according to Myhill's view, should be greater use of proclisis because clitic behavior in the construction headed by a grammaticalized finite verb should be analogous to clitic behavior with single finite verbs, which, in Modern Spanish, categorically take proclisis ([clitic + finite verb]). An illustration of this would be the case of [*ir a* 'go to' + infinitive] in sentences where *ir* 'go' marks futurity (a grammaticalized meaning), as in (5a), whereas [*esperar* 'hope'] + [infinitive] would not be functioning as an established periphrasis, therefore constituting neither a semantic nor a syntactical unit (5b). Evidence for this has been provided by Davies (1998, p. 257) when he reports that, in the *Habla Culta* corpus, auxiliary-like verbs such as *ir a* show 86% of proclisis, whereas non-auxiliary verbs, such as *esperar* 'hope', occur with proclisis in 0% of the cases (for more evidence, see also (Davies 1995, p. 375)).

(5)　　a.　　***Lo** [voy a conocer] mañana*
　　　　　　　'I [am going to meet] **him** tomorrow'
　　　　b.　　*??**Lo** [espero] [hacer] mañana.*
　　　　　　　'I [hope] [to do] **it** tomorrow'

Schwenter and Torres Cacoullos (2014) also offer evidence of the link between grammaticalization and verb meaning and its effect on clitic placement. The authors compare constructions that have more grammaticalized meanings, such as [*poder* 'can' + infinitive] (which can mean not only 'ability', but also 'root' or non-epistemic possibility), with constructions with less grammaticalized uses, such as [*querer* 'want' + infinitive]. The reported rates of enclisis in the Mexican data are 13% and 37%, respectively. Furthermore, an important piece of evidence provided in that work shows that two uses of *ir a* 'go to' display different clitic placement patterns. When *ir a* 'go to' is used to convey motion (e.g., I am going to visit you there and then come back), it occurs with enclisis much more often (32%) than when it bears the more grammaticalized meaning of futurity (e.g., I am going to visit you tomorrow) (7%).

Highly frequent lexical elements are particularly susceptible to developing grammatical uses through repetition (Miller and van Gelderen 2017). Therefore, the process of grammaticalization also correlates with increased frequency of the grammaticalized form or sequence of forms as productivity increases (Bybee and Thompson 1997, p. 66; Hopper and Traugott 2003, pp. 126ff., 232). Therefore, the constructions that favor proclisis have very high frequency (e.g., *poder*, *ir*, as mentioned earlier). When high-frequency lexical items or sequences grammaticalize, they acquire new functions as the result of conventionalization. This has been the case, for example, of [BE + *going* + *to* + infinitive] in English, which, through frequency, has consolidated into a construction (unit) used to express future meaning (Bybee 2003). A process within usage-based grammar that can account for these changes that occur through frequent co-occurrence is that of *chunking* (Bybee 2010). Increased unithood is assumed to emerge from frequent co-occurrence of two forms. Bybee (2010) refers to this process as the "chunking of sequential experiences that occurs with repetition", described, among other things, as "the primary mechanism leading to the formation of constructions and constituent structure" (p. 34). So, another way of observing grammaticalization in progress through VCP is to examine the relative frequency of use of a particular finite verb in the [finite verb + non-finite verb] variable construction as indicative of unithood between the two elements in the construction. Torres Cacoullos (1999) operationalizes the degree of grammaticalization of a form by tracking token and type frequencies (Hopper and Traugott 2003, p. 127). She reports that "[a]s *estar* + *-ndo* emerges as a unit in and of itself, the clitic is increasingly preposed to *estar*, as is categorically the case with finite verb forms in present-day Spanish" (p. 153). Similarly, Schwenter and Torres Cacoullos (2014) report that verbs with high token frequency[6] (like *ir a* 'go to', *poder* 'can', *querer* 'want', or *tener que* 'have to' + infinitive, and *estar* 'be' + gerund) favor proclisis, which the authors explain in terms of greater grammaticalizaction into tense–aspect–mood forms. In contrast, verbs with low token frequency (like *andar* 'go/be', *ir* 'go' + gerund, *deber (de)* 'must', *haber de* 'have to', *saber* 'know', *tratar de* 'try', *venir a* 'come to', and *volver a* 'go back to' + infinitive) favor enclisis. Their results not only suggest gradual spread of proclisis from one construction to another as forms grammaticalize, but also that "[d]ifferences between particular constructions are tied to token and relative frequency as measures of unithood (e.g., infinitive constructions with poder vs. querer)" (p. 533).

If grammaticalization accounts are on the right track, grammaticalized meaning (Myhill 1989) and increase in unithood (Torres Cacoullos 1999; Schwenter and Torres Cacoullos 2014) should lead frequent [finite verb + non-finite verb] variable constructions away from enclisis and closer to proclisis. While this seems to be the case for most constructions meeting those criteria of grammaticalized meaning and increased unithood, one VCP construction seems exceptional. I refer here to [*tener que* 'have to' + infinitive], which is not only headed by a verb that has grammaticalized from possession to obligation, but which is also among the most frequent periphrases in corpus studies of VCP (Davies 1995). Despite this, [*tener que* 'have to' + infinitive] favors enclisis.

Davies (1995) was the first large-scale corpus study of clitic placement variation in verbal periphrases such as in (4) in a set of twelve spoken and written corpora (from ten different Spanish-speaking countries). It reported a continuum-like distribution of the 32 verbs analyzed. Table 1 shows the distribution on enclisis with the most frequent verbs in spoken Spanish corpora he examined[7]. The three verbs marked with stars (*) on Table 1 accounted for 70% of all tokens found by Davies in his aggregated corpus of spoken Spanish across dialects. These three verbs have grammaticalized meanings (*ir* > future, *poder* > ability/permission, *tener* > obligation). The prediction

---

6   The authors operationalize high frequency based on observed distribution. This results in verb types that account for >8% of all the extracted tokens in the corpora being considered to be frequent. Under this assumption, *tener que* 'have to' is among the most frequent verbs.

7   Each of these verbs accounted for at least 3% (a minimum of 318 tokens) of all the tokens found in the spoken data (N = 10,626).

from grammaticalization accounts would therefore be that these three verbs should disfavor enclisis. However, for *tener que*, this is not the case, since it is used in enclisis 62% of the time.

Table 1. Distribution of enclisis in the most frequent constructions in spoken corpora (Davies 1995, p. 374).

Construction	% Enclisis
ir + a 'go to' *	14
poder 'can/may' *	40
querer 'want'	53
tener + que 'have to' *	62
deber 'must'	68
haber + que 'must'	100

Table 1 includes the categorical results for the *haber que* 'must' even though this construction falls outside of the envelope of variation given that it categorically selects enclisis. The reason for bringing it to the readers' attention here is twofold. First, this construction is noted by Davies (1995) as a puzzling case where, despite its grammaticalized meaning as a deontic modal, *haber que* exhibits categorical enclisis (pp. 373–74). Second, in Section 4, I will argue that this construction's categorical behavior plays a critical role in the apparently exceptional behavior of *tener que*. Lack of one-to-one correspondence between grammaticalization predictions and particular construction behavior with respect to VCP has been noted in studies of diachronic language change. Torres Cacoullos (1999) proposed that "diachronic increases in CC [proclisis] do not reflect a direct correspondence between the occurrence of CC [proclisis] and the meaning of the periphrastic expression (grammatical vs. lexical) in any given example, but rather indicate the conventionalization of auxiliary + gerund sequences as units whose components are increasingly fused" (p. 147). Therefore, VCP seems to be impacted by the degree to which these two-verb constructions are conventionalized through frequency of use.

To exemplify how usage-based linguistics is able to uncover the constraints that impact VCP, in the next section, I provide a variationist study of VCP in spoken Argentine Spanish, a dialect for which such analysis has not been offered yet. The results indicate which factors, headed by the finite verb construction, constrain clitic placement in this dialect. In addition, this small-scale study provides further evidence of the apparently exceptional VCP behavior of *tener que*, which I will seek to account for in Section 4.

3. Variationist Study of VCP

In his study of 543 tokens of clitics in variable contexts from written texts, (Myhill 1988, p. 360) observed that proclisis was more frequent with some verbs with more grammaticalized meanings, as well as whenever the clitic outranked the subject in the Animacy Hierarchy[8] than in the opposite scenario. The prediction with respect to animacy has been that inanimate referents do not tend to be topicalized as much as animate referents; therefore, the former would favor enclisis and the latter, proclisis. Apart from the effect of the finite verb discussed earlier, the large-scale corpus study of VCP by Davies (1995) also reported a semantic effect of animacy, where clitics with animate referents prefer preverbal position, providing evidence for Myhill's earlier proposal[9].

---

[8] "A proposed hierarchical ordering of noun phrases etc. ranging from personal pronouns, such as I, as maximally 'animate' to forms referring to lifeless objects as minimally 'animate' ... " (Matthews 2007).

[9] Other effects found by Davies (1995) include that proclisis was greater with multiple clitics or reflexives than with single clitics and non-reflexives as well as after the subordinating conjunction *que* 'that' compared with coordinating conjunction *y* 'and'.

More recent studies provided evidence for those effects and uncovered other variable constraints operating on VCP. These studies focused mainly on particular dialects[10]. The dialects examined in later studies include those of Madrid and Bogotá (Sinnott and Smith 2007), Caracas (Gudmestad 2006; Zabalegui 2008), Gran Canaria (Troya Déniz and Pérez Martín 2011), Salamanca (Aijón Oliva and Nieto 2013), Asturias in Nothern Spain (González López 2008, 2013), and Mexico (Schwenter and Torres Cacoullos 2014). Despite their methodological differences, some patterns of VCP use emerge among these studies. In particular, four constraints on VCP are reported as significant: register, finite verb, animacy, and discourse topicality. Specifically, corpus studies indicate that enclisis is favored in written language by certain verbs (discussed below), by inanimate or propositional referents, and by non-persistent referents.

In this section, I add to this literature by reporting on the results of a variationist study of VCP in spoken Argentine Spanish. This study seeks to find out whether the same constraints that impact clitic placement in other dialects (verb, animacy, and topicality) do so in Argentine Spanish as well. In addition, this study seeks to test the apparently exceptional behavior of *tener que*.

3.1. Corpus and Data Extraction

All clear occurrences of third-person singular direct object (3sg DO) clitics (*lo*.MSG, *los*.MPL, *la*.FSG, *las*.FPL) that were adjacent to a verb were manually extracted from the *Habla Culta de la Ciudad de Buenos Aires* corpus (Barrenechea 1987), which consists of 33 transcribed free conversations ranging from 17–55 min in length, each between two and four participants (~260,400 words). The speakers are N = 59 male and female professionals ranging between 26 and 70 years of age, all born in Buenos Aires, so the sample represents educated spontaneous speech or upper-middle socioeconomic level. Even though the interviews for this corpus were conducted in the 1980s and VCP has already been explored in similar *Habla Culta* corpora from other dialects since then, VCP has never been looked at in the *Habla Culta de Buenos Aires* corpus. While this corpus documents educated speech, it was selected because it provided a comparable baseline adult language measure for a larger study on the acquisition of morphosyntactic variation by Argentine children (Requena 2015). It is noteworthy, however, that Schwenter and Torres Cacoullos (2014) found no difference in rates of enclisis between *Habla Culta* 'educated speech' and *Habla Popular* 'popular speech' corpora within the same dialect, namely Mexican Spanish (p. 519). I restricted my study to 3sg DO clitics in an attempt to avoid conflating other persons and syntactic roles (such as reflexives and indirect object clitics). The reason behind this decision is the fact that animacy is a critical factor in variable clitic placement, and I wanted to exclude those clitics that tend to be animate most of the time (i.e., reflexives and indirect object clitics), since this could interfere with the results.

All contexts where clitic placement is categorical in modern Spanish (e.g., [clitic + finite verb] and [non-finite verb + clitic]) were identified. There was a higher frequency of clitics followed by a single finite verb (N = 1262), i.e., [clitic + finite verb], than clitics following infinitives or gerunds (N = 165), i.e., [non-finite verb + clitic]. I, therefore, assume the existence of a [clitic + finite verb] construction in speakers' representation of single finite verbs, which could serve as an analogical model to variable contexts (i.e., [finite verb + non-finite verb]) where the finite verb has grammaticalized into an auxiliary, modal, or aspectual marker, rendering a structure that could be categorized (and stored) as a single verb construction. Given that clitics precede single finite verbs categorically, this analogical link could play an important role in the overall bias toward proclisis that seems to spread gradually from construction to construction in variable contexts (Schwenter and Torres Cacoullos 2014, p. 533). Contexts with categorical clitic placement were excluded from the analysis, as were other contexts with finite verbs that were not sufficiently represented with both clitic positions in the corpus (see

---

[10] With the exception of Déniz (2003) in Troya Déniz and Pérez Martín (2011, p. 15), who examined over twelve dialects included in the Macrocorpus de la norma lingüística de las principales ciudades del mundo hispánico reporting 74% proclisis.

Appendix A for a complete list of exclusions). As the result of data cleaning, I was left with a total of 252 cases of VCP[11].

3.2. Coding of Variable Constraints

The dependent variable (Clitic Placement) was coded categorically as either proclisis or enclisis. Of all the variable cases, 163 (65%) display proclisis, whereas 89 (35%) display enclisis. The prevalence of proclisis overall is a pattern that has been already attested in Modern Spanish regardless of the continuum-like distribution of verbs according to their preference for proclisis (Davies 1995)[12]. However, as shown by previous variationist studies, there are other constraints that are expected to impact VCP in spoken Argentine Spanish as well, namely, the finite verb, the animacy status of the clitic referent, and the accessibility and topicality. In the next section, I address how these independent variables (constraints) were coded, as well as the predictions made for each one. Given that proclisis is pervasive in Modern Spanish, the predictions and results are stated from the perspective of enclisis thereon.
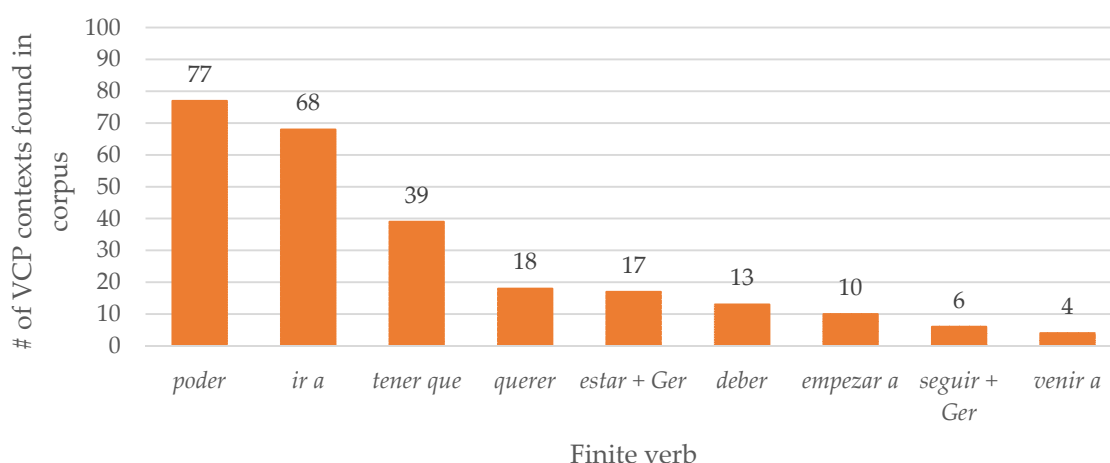
3.2.1. The Finite Verb

The finite verb has been identified as the most influential constraint on VCP (e.g., Davies 1995; Schwenter and Torres Cacoullos 2014). In the present study, I coded for the finite verb heading VCP contexts. The 252 variable contexts found were distributed in constructions headed by nine different finite verbs. Figure 1 shows the number of variable clitic placement contexts found for each verb. The three most frequent finite verb constructions in the Buenos Aires data were *poder* 'can' (6) and *ir a* 'go to' (7); next is *tener que* 'have to' (8). These three verbs have been identified as the most frequent infinitival periphrases (for references, see (Fernández Ulloa 2001, p. 29) cf. (Gómez Manzano 1992)). Accordingly, and following the grammaticalization predictions, I hypothesized that more frequent and grammaticalized verb constructions should disfavor enclisis. This should be the case of *poder* and *ir a* (future meaning), for example. With respect to less grammaticalized verb constructions (e.g., *empezar a* 'begin to'), as well as less grammaticalized uses of certain constructions (e.g., *ir a* with motion meaning), I expect greater enclisis. Even though *tener que* is a frequent and grammaticalized construction, in line with previous studies (Davies 1995), I expected it to favor enclisis, replicating the 'exceptional' behavior by this construction reported elsewhere.

---

[11] Even though this is quite a small data set, the results of this study replicate some previous findings with larger corpora and add to the description of this phenomenon in different dialects, which could be used to establish comparisons across varieties of Spanish, as mentioned by (Gudmestad 2006, pp. 13–14). In addition, in the current manuscript, the corpus study mainly serves the purposes of (a) replicating the effects of the main constraints on VCP in a dialect for which no variationist study has examined this phenomenon, and (b) illustrating the exceptional behavior of *tener que* 'have to' to motivate the theoretical analysis. While a larger dataset would be required for a study that offers proof of a novel effects or where the corpus study is the main focus, I think that the small-scale study contributes to the purposes of the present study ("a" and "b" above).

[12] Formal approaches have also noted how clitic climbing (CC) is possible with some types of verbs. Verbs have, for example, been classified into those that trigger structure simplification and those that do not. Structure simplification would then make the finite and non-finite elements of these periphrases belong to the same clause, thus allowing preverbal clitics. Aissen and Perlmutter (1976) identify "trigger" verbs (such as *querer* 'want', *tratar* 'try', and *soler* 'be in the habit of') and "non-trigger" verbs (like *insistir* 'insist', *soñar* 'dream', and *parecer* 'seem'). The authors support the Clause Union hypothesis (rendering structure simplification) and discard the existence of a specific rule that would account for preverbal clitics. Rizzi (1976), on the other hand, relates proclisis to the verb by proposing a lexically governed rule called 'Restructuring', which allows clitics to appear preverbally. Both Rizzi (for Italian) and Suñer (1980) (for Spanish) agree that modals, aspectuals, and motion verbs allow proclisis in addition to enclisis.

Figure 1. Distribution of all (N = 252) variable clitic placement tokens according to the finite verb in the corpus of Argentine Spanish. Note: In all these variable constructions, the finite verb listed is followed by an infinitive, except in the two cases signaled with "+ Ger" where the non-finite verb is a gerund.

(6)　　Variable use with *poder* 'can':
　　　Proclisis:

a.　*únicamente* **lo** *puedo* *hacer* *en castellano* (XXIV, 185:13)[13]
　　only　　**it-ACC.M.SG** can-PRS.1SG do　　in Spanish
　　'I can only do **it** in Spanish'

　　　Enclisis:

b.　*no* *podés* *hacer* = **lo** *dando* *vuelta* *los* *esquís* (IV, 73:3)
　　NEG　can-PRS.2SG do = **it-ACC.M.SG** turn-GER around the skis
　　'You cannot do **it** turning the skis around.'

(7)　　Variable use with *ir (a)* 'go to':
　　　Proclisis:

a.　**la** *vamos* *a* *destruir* (XXIV, 212:5)
　　it-ACC.F.SG go-PRS.1PL to destroy
　　'We are going to destroy it'

　　　Enclisis:

b.　*vamos* *a* *acortar* = **la** *un* *poquito* (X, 160:3)
　　go-PRS.1PL to shorten = **it-ACC.F.SG** a bit
　　'We are going to shorten **it** a bit'

(8)　　Variable use with *tener que* 'have to':
　　　Proclisis:

a.　**Lo** *tengo* *que* *conversar* *casualmente* (XXVI, 289:28)
　　it-ACC.M.SG have-PRS.1SG to discuss by the way
　　'I have to discuss **it**, by the way'

　　　Enclisis:

b.　*todo* *eso* *tengo* *que* *cuidar* = **lo** *plenamente* (X, 160:3)
　　all that have-PRS.1SG to take care of = **it-ACC.M.SG** completely
　　'All that, I have to take care of **[it]** completely'

3.2.2. Animacy of Clitic Referent

This factor group was included in order to explore Myhill's (1988) hypothesis that inanimate referents would favor enclisis, as they do not tend to be as topical as animate referents. In his characterization of human referents (as well as the more prominent referents in general), Aijón Oliva (2011, p. 28) includes the high frequency of preverbal placement of clitics in variable structures mirroring the subject position in prototypical SVO syntactic order. I coded for animate referents (9), inanimate referents (10), and propositional referents (those clitics whose referent is a whole proposition) (11). I predicted that clitics with inanimate and propositional referents should favor enclisis (Davies 1995; Myhill 1988; Sinnott and Smith 2007).

(9)   Animate Referent: (*lo* 'him' = a man)

lo	*iba*	*a*	*matar* (XXII, p. 99, 10)
it-ACC.M.SG	go-PST.IPFV.3SG	to	kill

'(he) was going to kill **him**'

(10)   Inanimate Referent: (*lo* 'it' = a book)

Voy	*a*	*mirar* = **lo**. (XXIV, p. 166, 19)
go-PRS.1SG	to	look = **it.ACC.M.SG**

'I'm going to look at it'

(11)   Propositional Referent: (*lo* 'it' = the fact that when a player has certain cards s/he should double the bet)

ahora	*que*	*me*	*dijiste*	**lo**	*voy*
now	that-REL	me-DAT	tell-PST.PRF.2SG	**it-ACC.M.SG**	go-PRS.1SG

a	*tener*	*presente*. (XXV, p. 225, 29)
to	have	in mind

'now that you told me, I will have **it** in mind'

3.2.3. Referent Accessibility

Considering the last time that a referent was mentioned (Givón 1995) can provide a measure of how topical the referent is. The hypothesis is that 'the more recently it has been mentioned, the more topical it is' (Myhill 2005, p. 473). I thus coded for whether the referent was immediately accessible (same clause, previous clause) (12), or not immediately accessible (meaning that it was mentioned earlier in discourse, but not in the same or immediately previous clause) (13) (Schwenter and Torres Cacoullos 2014)[14]. I predicted that enclisis should be favored by referents that are not immediately accessible.

---

[13]   Examples from the corpus are followed by references to the source in the format: sample number (in Roman numerals), page number:line number.

[14]   I also coded for the syntactic function of cases of immediate mention, such as Subject, Object, and Other. This was not significant, so the final analysis collapsed all syntactic functions.

(12)   Immediately Accessible Referent[15]:

        *Bueno, <u>lo</u> llama- - -   y el lunes **lo** fui a ver.* (XXIV, p. 202, 18)

Translation:

        Well, he calls <u>him</u>- - - and on Monday I went to see **him-ACC.3SG**

(13)   Not Immediately Accessible Referent:

        Inf. A. ---*Sí... esté... yo creo que sí. Yo he visto pasar por acá todos <u>los colectivos</u> [....] en la línea.*

        Inf. B. ---*Bueno, pero [....] de la línea hasta que oscurezca. Después ya no va a haber más.*

        Inf. A. ---*Pero después empiezan a retirar-**los** porque tienen miedo que pase alguna cosa.* (XXVII, p. 319, 11–13)

Translation:

        Inf. A. ---Yes… I believe that yes. I have seen all <u>the buses</u> [….] of the line go through here.

        Inf. B. ---Well, but […] of the line until the evening. After that, there will be none.

        Inf. A. ---But then they start removing **them-ACC.3PL** because they are afraid that something might happen.

3.2.4. Topic Persistence

Topic Persistence counts how many times the clitic referent is mentioned in upcoming discourse as a cataphoric indicator of topicality (Givón 1983, pp. 14–15; 1992). This could be related to the way Lewis (1979) conceptualizes 'salience' within referential semantics as the most salient entity " . . . in the domain of discourse, according to some contextually determined salience ranking" (p. 348). The author draws on (and critiques) the more philosophical proposal (Russell 1905) according to which a *prominent-objects coordinate* for denoted elements is " . . . determined on a given occasion of utterance of a sentence by mental factors such as the speaker's expectations regarding the things he is likely to bring to the attention of his audience" (Lewis 1970, p. 63) [emphasis is mine].

In an attempt to tap into the speaker's expectations about the ranking of comparative prominence or salience given to the clitic referent, I analyzed the following ten clauses after each token and coded for the number of times that the clitic reference was mentioned (once vs. more than once) and the syntactic function of each of those subsequent mentions (expressed or unexpressed subject, DO, 'other', and in more than one syntactic function)[16]. Syntactic function was not selected as significant, so all uses were collapsed into Persistent uses (mentioned more than once in upcoming discourse) (14) and Non-Persistent uses (either absent or mentioned just once in upcoming discourse) (15). In line with the results of Schwenter and Torres Cacoullos (2014), I predicted enclisis to be favored by non-persistent referents.

---

[15]   The translations corresponding to the examples for each of the discourse factors are provided immediately after the Spanish in more idiomatic form only.

[16]   Unlike Schwenter and Torres Cacoullos (2014), here, I included interlocutor tokens as well as quotative and other discourse formulas.

(14)　Persistent Referent:

　　*No, pero vos **los** podés preparar bien. <u>Los</u> podés preparar bien y <u>los</u> podés preparar mal. Yo <u>los</u> preparé mal. ¡Qué le vas a hacer! [risas]* (XXI, p. 17, 12)

Translation:

　　No, but you can prepare **them-ACC.3PL** well. You can prepare <u>them</u> well and you can prepare <u>them</u> badly. I prepared <u>them</u> badly. What are you going to do! [laugh]

(15)　Not Persistent Referent:

　　Inf. C. ---*...un curso de la Cultura es Lingüística, que puedo dictar = **lo** sin preparar.*
　　Inf. A. ---*No, no, no; porque salís o... o porque...*
　　Inf. B. ---*Y porque tiene cursos [............]*
　　Inf. C. ---*Y el otro es drama que... que... que me interesa, me gusta y lo agarré, ¡qué voy a hacer!, que me exige cuatro o cinco horas de estudio por semana en casa, sábado y domingo.* (XXI, p. 19, 19–22)

Translation:

　　Inf. C. ---...a class at the Cultura is Linguistics, which I can teach **it-ACC.3SG** without preparing.
　　Inf. A. ---No, no, no; because you go out or… or because…
　　Inf. B. ---And because it has courses [………]
　　Inf. C. ---And the other one is Drama that… that… that interests me, I like it and I took it, what am I going to do!, it demands four or five study hours per week at home, on Saturday and Sunday.

　　　I conducted a multiple logistic regression analysis on the data using Rbrul (Johnson 2009) within the R programming environment (R Core Team 2014). This type of analysis allows us to identify the factors (or variables) that, when taken together, affect the application of the selected dependent value (in our case, enclisis) simultaneously. In addition, it enables us to rank the variables according to their relative magnitude of the effect of each variable or factor group (known as 'Range'). In addition, speaker was added to the model as a random effect.

3.3. Multivariate Results

　　　Table 2 shows the results of the multivariate analysis by variable (e.g., Finite Verb, etc.) and levels of a variable (e.g., *tener que* + infinitive, *deber* + infinitive, etc.)[17]. The first column includes the Weight, which indicates the probability that each factor contributes to the application value, namely enclisis. The closer it is to 1, the more likely enclisis is. Taking > 0.5 as a cut-point, the levels with probability weights over 0.5 favor enclisis. The other columns contain enclisis tokens out of all tokens in a level, percentage of enclisis in a level, and percentage of all the data that a level represents, respectively.

17　Two finite verbs (*seguir* and *venir*) did not reach a minimum of 10 tokens each and were initially collapsed into an "Other" category. This, however, resulted in an interaction between finite verb and animacy that not only lacked theoretical motivation in the present study, but also worsened the model fit as measured by the Akaike Information Criterion (AIC). So, we removed the "Other" category from the final model.

Table 2. Variable rule analysis: Factors contributing to speakers' choice of enclitic position. Note: Factors with data in square brackets were not selected as significant.

Application Value: Enclisis	Weight	n	%	% of all Data
Finite Verb ($p. < 4.34 \times 10^{-8}$)				
tener que + Infinitive	0.83	28/39	72	15
empezar a + Infinitive	0.76	6/10	60	4
deber + Infinitive	0.66	7/13	54	5
querer + Infinitive	0.60	8/18	44	7
poder + Infinitive	0.42	24/77	31	31
ir a + Infinitive (movement)	0.39	3/13	23	5
estar + Gerund	0.21	3/17	18	7
ir a + Infinitive (future)	0.14	5/55	9	22
Range	*69*			
Animacy of Referent ($p. < 0.05$)				
Propositional	0.62	17/38	45	16
Inanimate	0.60	57/142	40	59
Animate	0.29	10/62	16	26
Range	*33*			
Referent Accessibility (n.s.)				
Immediately Accessible	[0.52]	49/126	39	52
Not immediately accessible	[0.48]	35/116	30	48
Topic Persistence (10 clauses) (n.s.)				
Not persistent	[0.54]	71/196	36	81
Persistent	[0.46]	13/46	28	19
N = 242, Input Probability = 0.33 (Average rate of enclisis: 35%)				

Here, I will briefly discuss the factors that were significant in the multivariate analysis, since, for the most part, they replicate previous studies. As Table 2 shows, Finite Verb was the most significant predictor, as particular finite verbs favor enclisis more than others (cf. Davies 1995, p. 374). I predicted that the more grammaticalized verb constructions, such as *ir a* (future) and *poder*, would disfavor enclisis. Indeed, the data confirm this. With respect to both uses of *ir a*, the results indicate that, while both uses of *ir a* disfavor enclisis, the more grammaticalized use (i.e., future) exhibits significantly less enclisis (9%) than the less grammaticalized use (i.e., motion) (23%). These rates are similar to those reported by Schwenter and Torres Cacoullos (2014) for Mexican Spanish (7% and 32%, respectively).

As shown by *tener que*, there is not a one-to-one correspondence between more grammaticalized verb constructions and VCP. Whereas two of the most frequent and grammaticalized verbs (*poder* and *ir a* (future)) disfavor enclisis, *tener que* is a frequent grammaticalized verb construction that favors enclisis (71% enclisis). This rate of enclisis with *tener que* falls between the ones reported for this verb in previous studies of written registers (e.g., 85% (Myhill 1989); 87% (Davies 1995)) and spoken registers (e.g., 62% (Davies 1995); 44% (Troya Déniz and Pérez Martín 2011); 30% (Schwenter and Torres Cacoullos 2014)). I believe that the high rate of enclisis with *tener que* in the present dataset is a function of the speakers the corpus represents, namely educated speakers[18]. In any case, the behavior of *tener que* with respect to VCP calls for further exploration, as it goes contrary to the general pattern towards proclisis that grammaticalization accounts predict for modern Spanish.

Bauman (2013) has shown that, since its emergence as a construction in the 12th century until the 20th century, *tener que* has developed from possession to obligation in a process analogous to the one attested for the English 'have to' (Heine and Kuteva 2002). Another interesting finding of Bauman's study is the fact that this gradual (and late, compared to other modal verbs) process of grammaticalization appears to have developed even further in the 20th century through cases of *tener que* used for probability. In the Argentine data, only one token of *tener que* 'have to' used with epistemic meaning of probability was found (see example (16) below). Counter to what the view on grammaticalization followed here would predict, this token appears with enclisis, and was included in

---

18 In Davies (1995), the only dialect that included data from popular (lower socioeconomic status) speakers showed the lowest overall rate of enclisis compared to all the other dialects in that study, for which data came from *Habla Culta* corpora, which contain interviews with educated speakers only. Similarly, Schwenter and Torres Cacoullos (2014) showed lower rates of enclisis in Mexican corpora of *Habla Popular* as well as youth speech than in Habla Culta (educated speech).

a residual category ('Other') for the multivariate analysis. The distribution of enclisis in the rest of the cases of *tener que* 'have to', where this construction conveys the idea of obligation, is 72%, thus strongly favoring enclisis. *Tener que*, therefore, constitutes a very intriguing case, since my study has shown once more its strong preference for enclisis despite its high frequency and its grammaticalized use of obligation.

(16)	*El*	*"pues"*	*tenés*	*que*	*haber* = ***lo***	*dicho*	*mucho* (XXIV, p. 164, 24)
	the	"so"	have-prs.2sg	to	have = **it-ACC.M.SG**	said	a lot

'You must have said "so" a lot'

With respect to Referent Animacy, the results in Table 2 confirm my prediction by showing how propositional and inanimate referents favor enclisis, while referents that rank higher in the animacy scale appear more often in proclisis. The present data clearly show a significant semantic effect of animacy in the expected direction based on Myhill (1988). It is important to note that all the animate referents in the data set were human, as in Schwenter and Torres Cacoullos (2014). Whereas in the present data, clitics with inanimate and propositional referents occur in enclisis twice as often as clitics with animate (human) referents, this was not the case in their study. The rates of enclisis in that study showed only an eight-percentage-point difference between inanimate (21%) and human (29%) referents, which was puzzling based on grammaticalization predictions by Myhill (1988). The authors explain those results as an interaction between animacy and topicality.

Referent accessibility and topic persistence did not reach significance in this model.

3.4. Discussion of Corpus Study

The small-scale study presented in this section examined VCP in a corpus of educated Spanish from Buenos Aires, Argentina from a variationist perspective and showed that:

1. As in previous studies, clitic placement in this variety of Argentine Spanish is constrained by Referent Animacy Clitics since human referents tend to appear preverbally (proclisis).
2. The nature of the finite verb construction emerged as the main factor conditioning clitic position. This result not only extends this known fact to a new dialect of Spanish, but it also provides evidence for the grammaticalization account that predicts a move away from enclisis for frequent periphrases where the finite verb has grammaticalized meaning. In the present study, we find this to be the case between more grammaticalized verbs that disfavor enclisis (e.g., *ir a* (future) and *poder*) as well as with more grammaticalized uses of a verb (e.g., *ir a* (future)) compared to more basic uses (e.g., *ir a* (movement) (see Schwenter and Torres Cacoullos 2014)).
3. *Tener que* constitutes an intriguing case as a frequent grammaticalized verb that favors enclisis. This lack of one-to-one correspondence between grammaticalized verbs and VCP seems exceptional.

In the next section, I show how grammaticalization, unithood, and links with other constructions can account for VCP behavior of [*tener que* + infinitive].

4. Motivating Construction Behavior

4.1. Unithood of [Tener Que + Infinitive]

The usage-based perspective on language followed here posits that language is made up of form–meaning pairings with a sequential structure called *constructions* (Goldberg 1995, 2006). Constructions are assumed to emerge from generalized patterns of use attested in actual utterances and stored via rich memory. When two elements are experienced together with sufficient frequency, they emerge as a unit thanks to processes of chunking and categorization. For example, we have assumed that [clitic + finite verb] must be part of speakers' mental representation, since proclisis occurs categorically with single finite verbs in contemporary Spanish. Following grammaticalization accounts, the proposal has been that the [clitic + finite verb] construction behaves as an analogical model for

VCP in cases where the finite verb and non-finite verb behave like a unit (thus resembling a single verb construction). That may account for the gradual spread of proclisis from construction to construction, as mentioned by Schwenter and Torres Cacoullos (2014, p. 533).

We also saw how one way of looking at the strength of the association between two elements in language is to look at the relative frequency of use together. Schwenter and Torres Cacoullos (2014, p. 527) provide a small-scale analysis that points to an association between finite verb and infinitival continuation that was stronger with *poder* than with *querer*, relative to other uses of these verbs. In what follows, I take the three most frequent verbs in the corpus study of Argentine Spanish (*poder*, *ir a*, and *tener*) and examine their contexts of use in a larger corpus of Spanish (Corpus del Español (CDE), (Davies 2002). I will look at the first-person singular present form[19] across these verbs and establish their relative frequency in the infinitival contexts of [~ + infinitive][20], [~ + direct object][21], and [~ + other continuations (e.g., pause)][22]. I will also look at the strength of the association between these verbs and any type of preverbal clitic [clitic + ~][23]. The analysis was made on N = 1420 tokens of *puedo*, N = 1059 tokens of *voy*, and N = 1511 tokens of *tengo*, and the results are shown in Figure 2. The rates for [~ + infinitive], [~ + lexical direct object], and [~ + other] should add up to 100%. The rates for [clitic + ~] were calculated independently out of 100% of tokens for each verb, are but included in the same figure.
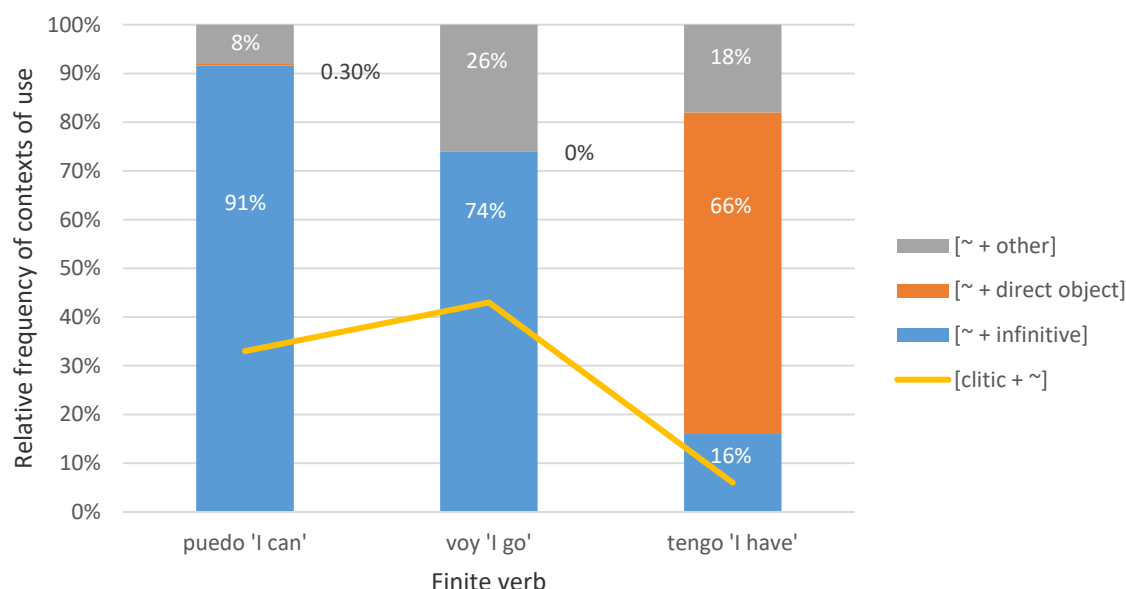
Figure 2. Relative frequency of the 1SG present tense form of each verb followed by an infinitive in Corpus del Español (Davies 2002).

When we consider [~ + infinitive], [~ + lexical direct object], and [~ + other], it is evident how *poder* has the strongest association with infinitives, followed by *ir a*. This association is very weak with

---

[19] I selected 1SG forms because they are the second most frequent singular forms after 3SG. For example, in the oral Corpus del Español (Genre/Historical) the frequencies are: *va* 'he/she/it goes', N = 7096; *voy* 'I go', N = 2661; and *vas* 'you go', N = 862. The advantage of 1SG over 3SG is that the subject is constant and known in the former. With the latter, the subjects can vary in animacy, specificity, and can even be propositional. So, for simplicity, I selected 1SG forms for this section.

[20] An example of [~ + infinitive] is *No te puedo dar una opinión* 'I cannot give you an opinion' (CDE:19-OR, Habla Culta: Lima).

[21] Here, "direct object" refers to a direct object noun phrase or clause that follows the particular verb under consideration. An example of [~ + direct object] *is tengo un amigo que es el director de la revista*. 'I have a friend that is the director of the journal' (CDE:19-OR, Habla Culta Buenos Aires).

[22] [~ + other continuation (e.g., pause)] included continuations that consisted of prepositional phrases, subordinate clauses, adverbials, and pauses. An example of this is *si yo voy a un país* 'if I go to a country' (CDE:19-OR, Habla Culta: Bogotá).

[23] Here, [Cl+V] refers to all clitics regardless of case and person/number (e.g., *me*, *te*, *se*, *lo/s*, *la/s*, *le/s*) that precede the particular verb under consideration.

tener (only 16% of cases, which is similar to the rate of use of this verb with "other" continuations). Therefore, as measured by relative frequency of use, *tener* does not have strong links to infinitives that would suggest a high degree of unithood. In contrast, *tener* is used transitively 66% of the time, a continuation not possible with *poder* or *ir*. In addition, the link between *tener* and a preverbal clitic (measured independently from the other measures of continuation), as a measure of association of *tener* with the [clitic + finite verb] construction, is also much weaker (6%) than with the other two verbs (33% with *poder* and 43% with *ir*). These patterns of use predict that, unlike the variable constructions headed by *poder* or *ir*, which display great unithood and conventionalization, the [*tener que* + infinitive] construction must display strong analyzability and ties to other uses of its components outside of the variable context.

In contrast to the very well-established grammaticalized use of *ir a* with future meaning or *poder*, where clitics tend to occur preverbally as a result of chunking due to frequent repetition (Schwenter and Torres Cacoullos 2014), *tener que* used for obligation constitutes a relatively recent innovation, which could explain the more conservative bias toward enclisis (Blas Arroyo 2018; Blas Arroyo and Schulte 2017). The use of *tener que* to convey obligation comes from a quasi-modal relative construction of old Spanish that was mainly used with verbs of possession. Of those verbs, *tener que* was the most obvious candidate to replace modal constructions with *haber de*, which was disappearing (Olbertz 1998, p. 250ff.). Bauman (2013) examined a number of written corpora from the 18th and 19th centuries, as well as written and spoken 20th century corpora. His results showed that *tener que* has increased its absolute frequency since the 18th century (see Figure 3), taking over not only contexts that had previously belonged to other expressions of obligation (such as *deber (de)* and *haber de*, for example), but probably taking up some features of their preferences of use as well.
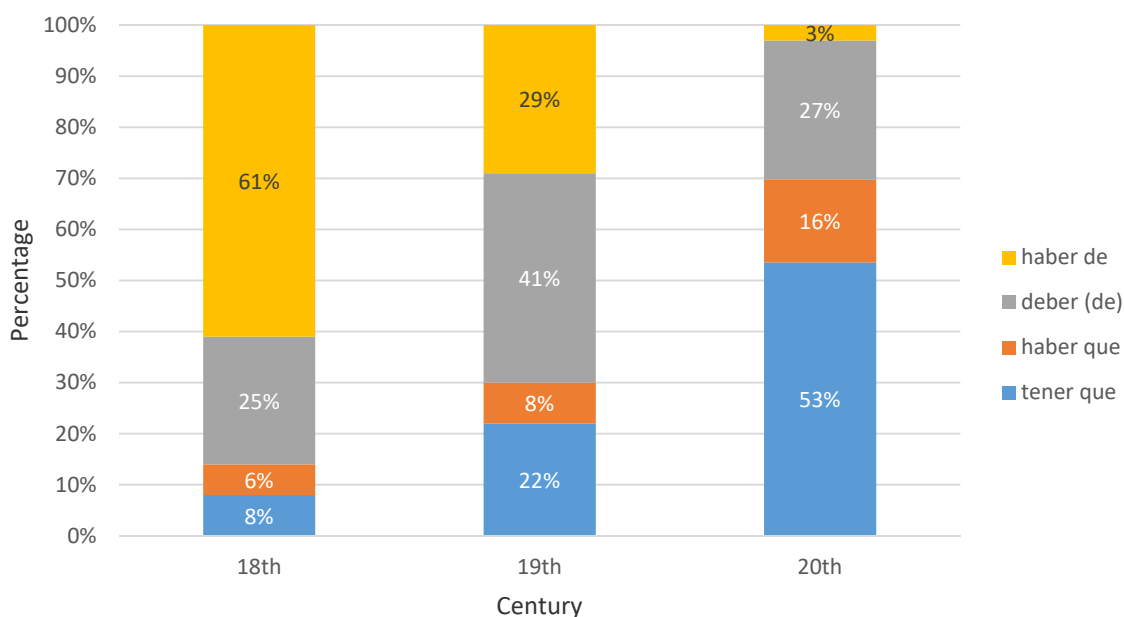
Figure 3. Increase in frequency of *tener que* among modal constructions from the 18th to the 20th century (adapted from (Bauman 2013, p. 147)).

The degree of unithood in emerging periphrases or constructions can be tracked by looking at the possibility of insertion of intervening material. Olbertz (1998) notes a pattern that became common in the use of *tener que* in sentences like *Tenemos pan que comer* 'We have bread to eat'. The antecedent of the infinitival clause [*que comer*] (i.e., *pan* 'bread') started to be implicit most of the time, and this resulted in *que* following the finite verb and preceding the infinitive (i.e., *tenemos ø que comer*), just as prepositions do in other periphrases (p. 252). Despite this conventionalization of *tener que* as a modal periphrasis,

Olbertz acknowledges that *tener que* retains a homonymous non-periphrastic lexical construction. The distinction is clearly seen in (17) below:

(17) a. *No tengo que decir nada.*   (Modal obligation periphrastic construction)
         I don't have to say anything.
      b. *No tengo nada (que decir).*   (Lexical construction)
         I don't have anything (to say).

In connection to the observation that both English deontic *have* and Spanish *tener* 'have' allow both alternative uses shown for Spanish in (17), Giammatteo and Marcovecchio (2009) add: *"Que el auxiliado no sea obligatoriamente adyacente al auxiliary evidencia la no consolidación plena de la perífrasis"*/"The fact that the verb accompanying the auxiliary does not have to be adjacent to it evidences the lack of periphrastic consolidation" (p. 34). This lack of consolidation, probably as a result of the relatively recent grammaticalization of the construction, will be assumed here to have implications in speakers' cognitive representations (see Torres Cacoullos 1999) for this proposal with respect to clitic placement in [finite verb + gerund] constructions).

I have tried to show that *tener* holds weak ties to infinitival continuations in relative frequency of use and alternates as a lexical construction (as in 17b). Thus, [*tener que* + infinitive] displays weak unithood and retains a high degree of analyzability. However, this does not seem to be everything that is drawing *tener* towards enclisis. In the next section, I propose that the greater analyzability of [*tener que* + infinitive] results in strong diachronic, semantic, and structural ties with another construction, which, crucially, only allows enclisis. This construction is [*haber que* 'must' + infinitive + clitic], which may compete (and win) as an analogical model against the conventional model for grammaticalized verbs (namely, [clitic + finite verb]).

4.2. [Haber Que + Infinitive + Clitic] as an Analogical Model for [Tener Que + Infinitive]

Garachana and Hernández (2017) give evidence of how, as the verb *haber* (possessive) became impersonal around the 16th century[24], [*haber que* + infinitive] also changed its meaning towards impersonal obligation[25]. The researchers assume, then, that it was this evolution of [*haber que* + infinitive] towards impersonality that opened the door for another construction to fill the gap in the expression of personal deontic obligation, namely [*tener que* + infinitive]. Therefore, as the expression of obligation was split among very few forms, the authors suggest that "puede aventurarse que [...] la perífrasis *tener que* + *infinitivo* se creó por analogía a partir de *haber que* + *infinitivo* que funcionó como construcción de apoyo."/"It can be assumed that [...] the *tener que* + *infinitive* periphrasis was created by analogy with *haber que* + *infinitive*, which functioned as a grounding construction" (p. 137). They conclude that, since the 18th and 19th centuries, [*tener que* + infinitive] has consolidated as a periphrasis of general obligation in complementary distribution with [*haber que* + infinitive] for impersonal obligation (see Garachana 2017). This would suggest that *tener* and *haber* not only have common origin as verbs of possession, but both also have grammaticalized into obligation, taking over the expression of different types of obligation (Martínez Díaz 2003). Therefore, diachronically and semantically, [*tener que* + infinitive] is very closely linked to [*haber que* + infinitive].

These constructions are also linked in other ways that make it evident that *tener que* finds in *haber que* a very likely analogical model for VCP. In this section, I would like to propose how, under the assumption that the components of a construction can keep ties to their use outside of the construction (Brown and Rivas 2012), the subordinating conjunction *que* can also lead to the conclusion that [*haber*

---

[24] Third-person singular jumped from 50% in the 15th century to 90% in the 19th century, and reached 100% in the 18th century (Garachana and Hernández 2017, p. 133).
[25] [*haber que* + infinitive] was used with personal value 100% until the 15th century and went to 90% impersonal in the 16th century.

que + infinitive] constitutes an analogical model for the *tener que* construction when it comes to VCP. As before, I will draw on frequency data from the Corpus del Español (Davies 2002) and on graphic representations of exemplar models (Bybee 1985, 1988, 2001) in order to show how the internal structure of the *tener que* construction can be analyzable into component sub-units (see Figure 4 below) that still hold tight bonds with their occurrence in other constructions. Below, I add some data to that which has already been presented about *tengo* to show how speakers' previous experience with *que* as used in other constructions can also lead to an enclisis bias[26].
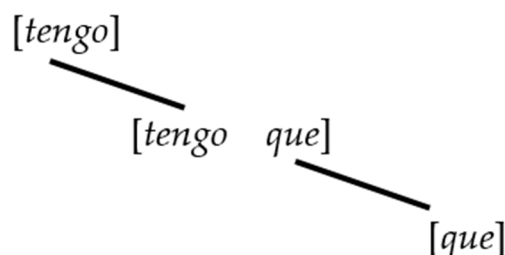

Figure 4. Sub-units of the [*tengo que*] 'have to' construction.

(18) *que* + Infinitive
    Los   poetas tienen   que   escribir   lo   máximo   possible (CDE: 19-OR: Entrevista (ABC))
    the   poets have-PRS.3PL to   write   the   maximum possible
    'Poets have to write as much as possible'

(19) *que* + (Adv) + V
    Sin embargo, la   grabación   que   prefiero   en la   actualidad   es   la
    however,   the   recording   that   prefer-PRS.1SG in the   present   is   the

    "Misa en Si menor"   de   Bach (Entrevista (ABC) CDE: 19-OR)
    "Misa en Si menor"   by   Bach
    "However, the recording that I prefer at present is the '…' by Bach"

(20) *que* + NP/nominal clause/Adv/PP
    ¿Cuándo   descubrió   usted   que   el   piano   lo   era
    when   discover-PST.PRET.2SG you   that   the   piano   it   be-PST.IPFV.3SG

    todo   en   su   vida? (Entrevista (ABC) CDE: 19-OR)
    everything in   your   life
    'When did you discover that the piano was everything for you in life?'

Table 3. The three main linguistic items following *que* (N = 1983).

Following Item	Occurrences	
	Number	**%**
que + (Adv) V	1310	66%
que + NP/nominal clause/Adv/PP	551	28%
que + Infinitive	122	6%

---

26 One variable that Davies introduced to the study of Contemporary Spanish was the nature of the preceding material. He replicated what was true for older stages of Spanish, namely that constructions preceded by the subordinating conjunction *que* 'that' appear in proclitic position more often than with the coordinating conjunction *y* 'and'.

A closer examination of the N = 122 instances of *que* + infinitive shows that only two elements almost exclusively precede this construction (Table 4). They are the verbs *tener* (21) and *haber* (22). This constitutes clear evidence of a very strong connection between *que* + infinitive and only these two verbs.

(21) *tener* + *que* + Infinitive

…*sólo*	*tengo*		*que*	*tapar*	*las*	*cuatro*	*cuerdas*	*con*	*la*
only	have-PRS.1SG		to	cover	the	four	strings	with	the

mano. (Entrevista (ABC) CDE: 19-OR)
hand
'I only have to press the four strings with the hand'

(22) *haber* + *que* + Infinitive

hay		*que*	*tener*	*paciencia* (Entrevista (ABC) CDE: 19-OR)
there be-PRS.3SG		to	have	patience

'one has to have patience'

Table 4. Verbs preceding *que* + infinitive (N = 122).

	Occurrences	
Preceding Verb	**Number**	**%**
tener 'have'	77	63%
haber 'there be'	42	34%
Other[27]	3	3%

In a way that would approximate Bybee's Network Model as extended to grammaticalizing constructions (Bybee 1998, 2010), Figure 5 below displays the data offered so far. As noted in the upper section of the figure, *tengo* shows a very weak association with infinitival continuation, which suggests a low degree of unithood between *tener* and infinitives (as shown earlier). When we now add data on the uses of *que* elsewhere, we note that *que* is usually followed by a finite verb (V) (66%), but in the rare instances when an infinitive follows *que* (only 6% of the tokens in the present dataset), the finite verb preceding *que* is only one of two: *tener* or *haber*. Crucially, *haber*, the one I am arguing here functions as the analogical model for VCP with tener que, only allows enclisis. So, by looking at the associations that the components of the [*tener que* + infinitive] construction hold with other constructions, it becomes clear why enclisis is the preferred position for clitics.

---

[27] Other cases include: *resulta que* + infinitive (CDE:19-OR, Entrevista (ABC)), *Pensaba yo que* + infinitive (CDE:19-OR, Entrevista (ABC)), … *cajón de sastre del que* + infinitive (CDE:19-OR, Entrevista (ABC)).
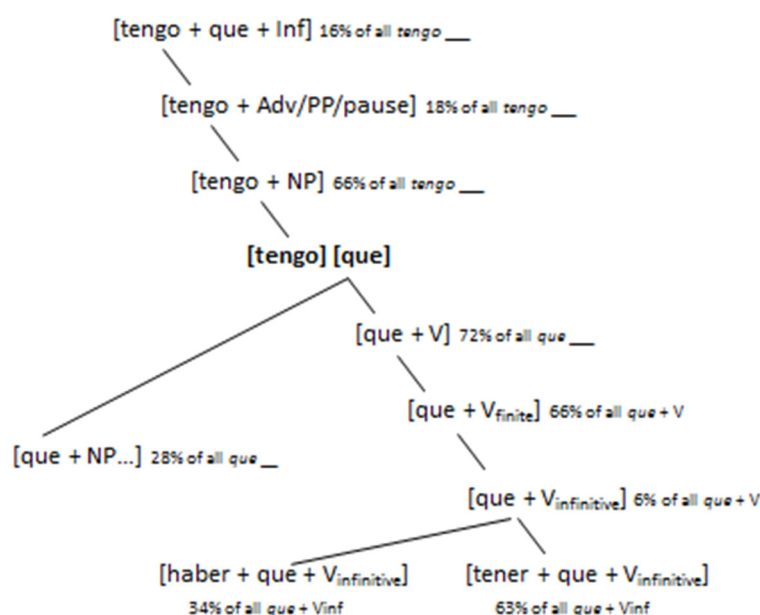
Figure 5. Diagram of links between construction components with their use in their constructions.

In this section, I hope to have shown how [*tener que* + infinitive] is diachronically, semantically, and structurally linked to [*haber que* + infinitive]. I argue that the strength of these links together with the fact that *haber que* selects enclisis categorically (Davies 1995) makes [*haber que* + infinitive + clitic] an analogical model for [*tener que* + infinitive], pulling the latter away from the growing tendency towards proclisis.

5. Conclusions

Variable Clitic Placement (VCP) has been a hot topic within linguistics for years. Here, I have shown how usage-based perspectives of language can account for this phenomenon as speakers experience it in naturalistic language use. This study contributed a variationist analysis of VCP in the variety of Spanish spoken by educated speakers in Buenos Aires, Argentina. The results indicated that VCP grammar is constrained by lexical (finite verb) and semantic (animacy) factors. Considering the finite verb construction as the main constraint on VCP, the study focused on usage-based accounts of the gradience attested across finite verb constructions. Grammaticalized meaning and increased frequency tend to account for VCP in general. However, the [*tener que* + infinitive] construction was found exceptional in that it favors enclisis despite its grammaticalized meaning of obligation and its high frequency of use. Using data from a larger corpus, I then showed that the [*tener que* + infinitive] construction seems to be at an earlier stage of grammaticalization than *ir a* and *poder*. Its lack of unithood of *tener* with infinitives and the possibility of intervening material signaled great analyzability of [*tener que* + infinitive]. Finally, another construction that categorically takes enclisis and which is strongly linked to [*tener que* + infinitive] diachronically, semantically, and structurally was suggested as a likely analogical model for VCP with *tener que*. In this way, I have illustrated how usage-based linguistics can capture this instance of morphosyntactic variation.

This study poses new questions about VCP. For example, it is not clear on the basis of what aspects speakers establish their initial links between a lexical construction and the analogical model. For example, in the case of [*tener que* + infinitive], I tried to show that diachronic information stored in constructions, verb meaning, and construction structure seem to converge to make [*haber que* + infinitive] a likely analogical model. Does all this information work in parallel to map exemplars together during categorization? Or do speakers go by some type of construction information first? How would one test the existence and strength of these links between constructions experimentally? How do children acquire these relations between exemplars in order to arrive at adult-like grammar?

These and other questions would not only expand our understanding of VCP, but also that of the nature of language itself. Here, I hope to have shown how usage-based linguistics provides powerful tools to approach these important questions.

Appendix A. Exclusions as Part of Data Cleaning in Corpus Study

(a) Tokens with (DAT + ACC) clitic clusters such as *me*, *te*, *se*, or *nos* followed by *lo*, *los*, *la*, and *las* ((A1) and (A2)) (N = 198).

(A1)

Me	**lo**	h-ic-e	de	un	color	sepia,	oscur-it-o, (XXVII, 325:20)
me-DAT.1SG	**it-ACC.M.3SG**	do-PRET.1SG	of	a-M.SG	color	sepia,	dark-DIM-M.SG,

'I did it (for myself) on darkish sepia'

(A2)

esa	observación […]	**me**	**la**	han-	hecho	much-ísim-a-s (XXXI, 454:3)
that-F.S	observation […]	**me-DAT.1SG**	**it-ACC.M.3SG**	have-3PL	do-PTCP	many-SUP-F-PL

'many have made that observation to me'

(b) Tokens with *ser* 'be' + adjective + Infinitive, which show invariable enclisis (A3) (N = 3).

(A3)

en	una	mujer... esté...	es	más	fácil	entender-**lo**. (IX, 140:2)
in	a-F	woman… hm…	is	more	easy	understand-**it-ACC.M.3SG**

'It is easier to understand it in a woman'

(c) Tokens with *haber que* 'have to' + Infinitive, which also show invariable enclisis (A4) (N = 46).

(A4)

Hay	que	hacer-**lo**	con	precisión (XIV, 223:1)
have-PRS.3SG	to	do- **it-ACC.M.3SG**	with	precision

'One has to do it with precision'

(d) Tokens of clitics used in juxtaposed or coordinated constructions (A5) (N = 7).

(A5)

puede	adquirir	ese	cuadr-it-o-	y	tenerlo-	y	disfrutarlo (IX, 150:3)
can-3SG	purchase	that-M.SG	painting-DIM-M	and	have-**it-ACC.M.3SG**	and	enjoy-**it-ACC.M.3SG**

'can purchase that little painting and gave it, and enjoy it'

(e) A token that appeared as truncated (A6) (N = 1).

(A6)

porque	**lo**	tiene	que... (XXIX, 383:19)
Because	**it-ACC.M.3SG**	have-3.SG	to…

'because (s)he has to…'

(f) Tokens for which the referent of the clitic could not be identified (N = 7).

(A7) Méjico va a tener-**lo** mucho auténticamente americano nosotros (XXX,
más que 414:6)

Mexico go- to have-**it-** much authentically American us
3.SG **ACC.M.3SG** more than

'Mexico is going to have it more authentically American than us'

(g) Tokens in which there was intervening material between the finite and non-finite verbs (N = 3).

(A8) no podemos… los trabajos ésos ahí en el Ides exponerlos en mitades. (XXI, 17:10)
we can't…the Works those there at the Ides expose-**them.ACC.M.3PL** in halves.

(h) Contexts of a finite verb followed by multiple non-finite forms (N = 25) were identified. Of those, only the instances that allowed only two clitic positions (as in examples (A9a,b)) were included in the analysis (N = 11). Cases in which the clitic could occupy an additional third position (as in examples (A9 c, d)) were excluded (N = 14).

(A9) a. lo tiene que haber matado. (XXXI, 439:1)[28]
it-ACC-M3SG has to have killed
'he/she must have killed it'

b. lo debo haber tenido adentro (XVI, 239:9)
it-ACC-M3SG must have had inside
I must have had it inside'

c. hay que uno jamás las va volver a ver (I, 17:5)
materias
EXIST that one never **them-ACC-F3PL** go to return to see
subjects
'there are subjects that one will never encounter again'

d. no queríamos seguirlo viendo todos los días
NEG want.PRET.1PL keep = **it-ACC-** see.GER all the days
M3SG
'We didn't want to keep seeing it every day.' (XXII, 99:9)

Other cases were extracted, but were excluded for the analysis because they either display no variation in clitic placement, or because very few tokens were found of a given construction. In what follows, we exemplify such cases.

ENCLISIS ONLY

- All constructions consisting of a single non-finite form of a verb that, in present-day Spanish, can only take the clitic postverbally. This also includes imperatives (N = 165).
- [*gustar* 'like' + Infinitive-Clitic] constructions (N = 9)
  (A10) me gustaría hacerlo (XXIV, 210:9)
- [*tratar de* 'try to' + Infinitive-Clitic] constructions (N = 6)
  (A11) El trata de solucionarlos en lo más posible (V, 93:1)
- Expressions with [*vale la pena/da pena* 'it's a pity'+ Infinitive] (N = 2)
  (A12) A mí me da mucha pena- - - dejarlo, abandonarlo (XXIX, 387:17)

---

[28] All examples belong to the Corpus de Habla Culta de la Ciudad de Buenos Aires (Barrenechea 1987) and appear followed by the conversation number (XXXI), then the page number (439), and finally the paragraph number (1).

- [*dejar de* 'stop' + Infinitive-Clitic] constructions (N = 2)
  (A13)　así dejás de— llamarlo. (XXIV, 203:11)
- [*entrar a* 'begin' + Infinitive-Clitic] constructions (N = 2)
  (A14)　entraría a respetarla (X, 157:1)
- [*andar* 'go'+ Gerund-Clitic] (N = 2)
  (A15)　Pero vos sabés que yo ando persiguiéndolo a este señor (XIV, 203:2)
- Constructions occurring only once in the corpus with enclisis: [*estar* 'be' + PP + PP-Clitic] (A15),[haber 'there be' + NP + PP-Clitic] (A16), [*quedar en* 'agree to' + Infinitive-Clitic] (A17), as well as [*acordarse* 'remember' + *haber* 'having'+ Participle-Clitic] (A18), [*animarse a* 'fancy' + Infinitive-Clitic] (A19), [*atreverse a* 'dare' + infinitive-Clitic] (A20), [*comprometerse a* 'commit to' + Infinitive-Clitic] (A21), [*convenir* 'arrange/agree'+ Infinitive-Clitic] (A22), [intentar 'attempt' + Infinitive-Clitic] (A23), [lograr 'achieve' + Infinitive-Clitic] (A24), [*molestar* 'bother' + Infinitive-Clitic] (A25), [*negarse a* 'refuse' + infinitive-Clitic] (A26), [*pensar* 'consider' + infinitive-Clitic] (A27), [*ponerse a* 'start' + Infinitive-Clitic] (A28), [*proponerse* 'intend' + Infinitive-Clitic] (A29), [*pretender* 'pretend' + Infinitive-Clitic] (A30), and [terinar por 'end' + Infinitive-Clitic] (A31).
  (A16)　no estoy en edad de hacerlo tampoco (XIV, 216:5)
  (A17)　no hay forma de aprenderlo (XXII, 88:11)
  (A18)　Quedé en irla a visitar el miércoles. (XXXII, 474:7)
  (A19)　Desde chico me acuerdo haberlo visto (XXVII, 323:3)
  (A20)　¿Te animás a escucharlo de nuevo? (IV, 69:1)
  (A21)　Es que uno no se atreve a dejarlos a ver qué pasa. (IX, 154:7)
  (A22)　Cada cual se comprometió en su país a seguirlos trabajando... (XXI, 17:20)
  (A23)　si ese juicio me conviene conciliarlo o no (XXIII, 114:3)
  (A24)　intentar criarlo allí (VIII, 128:5)
  (A25)　no logro localizarlo (XXIV, 203:2)
  (A26)　nos molestaba hacerla (VII, 114:3)
  (A27)　se negaba a ponerlos en la mesa (IV, 82:9)
  (A28)　así que piensa emplearlo en un... Un empleo comercial
  nomás. (V, 88:2)
  (A29)　un día me pongo a hacerla (V, 93:7)
  (A30)　me propuse- - - hacerlo hasta— un momento determinado... (XXI, 57:19)
  (A31)　como él pretendió hacerlo ver (XXIV, 140:6)
  (A32)　a fuerza de escucharlas terminarán por corearlas junto a nosotros (XIX, 285:1)

PROCLISIS ONLY

- All cases of a clitic used with a single finite verb were excluded, as they only allow the clitic to appear preverbally (N = 1265).
- [Clitic + *hacer* 'make' + Infinitive] constructions (to make someone do something) (N = 12)
  (A33)　El lo hizo correr (XXVIII, 371:12)
- [Clitic + *ir* 'go' + Gerund] constructions (N = 11)
  (A34)　Yo las fui guiando (XI, 169:11)
- [Clitic + *llegar a* 'get to' + Infinitive] (N = 6)
  (A35)　un poco lo llego a dominar y me aburre después (I, 24:18)
- [Clitic + *dejar* 'stop' + Infinitive] (N = 4)
  (A36)　Te está diciendo que lo dejes pensar
- [Clitic + *volver a* 'do . . . again' + Infinitive] (N = 3)
  (A37)　la he vuelto a hacer (XXIV, 212:5)

- [Clitic + *haber* 'there be' *de* + Infinitive] constructions (N = 2)
  (A38)　No la ha de saber manejar (XXII, 102:4)
- [Clitic + *ver* 'see' + Infinitive] (N = 2)
  (A39)　los veo manejarse en coche (XXIII, 131:11)
- [Clitic + *deber de* 'must' + *haber* 'have' + Participle] constructions (N = 2)
  (A40)　Pero él la debe de haber presentido a traves de las sombras (XXIX, 396:1)
- [Clitic + *invitar a* 'invite to' + Infinitive] (N = 2)
  (A41)　de tanto en tanto la invite a salir (X, 158:1)
- Constructions occurring only once in the corpus with proclisis: [Clitic + *dar a entender* 'suggest'] (A41), [Clitic + *obligar a* 'force to' + Infinitive] (A42), [Clitic + *mandar (a)* 'have something done' +Infinitive] (A43), [Clitic + *saber* 'know' + Infinitive] (A44), and [Clitic + *terminar de* 'finish' + Infinitive] (A45).
  (A42)　como lo dan a entender sus títulos (XX, 294:1)
  (A43)　Por eso, los obligó a leer. (XI, 177:12)
  (A44)　No los mandé hacer todavía. (XXIV, 144:7)
  (A45)　no lo sé hacer (XXIV, 158:5)
  (A46)　Todavía no lo terminamos de- - - elaborar (XIV, 213:3)

References

Aijón Oliva, Miguel A. 2011. Variación Sintáctica y creación de estilos: Los clíticos reflexivos en el discurso. In *Variación Variable*. España: Círculo Rojo, pp. 21–56.

Aijón Oliva, Miguel A., and Julio Borrego Nieto. 2013. La variación gramatical como forma y significado: El uso de los clíticos verbales en el español peninsular. *Lingüística* 29: 93–126.

Aissen, Judith L., and David M. Perlmutter. 1976. Clause Reduction in Spanish. In *Studies in Relational Grammar 1*. Edited by David M. Perlmutter. Chicago: University of Chicago Press, pp. 360–404.

Barrenechea, Ana M., ed. 1987. *El habla culta de la ciudad de Buenos Aires: Materiales Para su Estudio*. Buenos Aires: Universidad Nacional de Buenos Aires.

Bauman, Joseph R. 2013. Grammaticalization and Variation in the Modal Domain of Obligation: The Evolution of Spanish [tener que + Infinitive]. Unpublished Ph.D. dissertation, The Pennsylvania State University, State College, PA, USA.

Blas Arroyo, José L. 2018. Comparative variationism for the study of language change: Five centuries of competition amongst Spanish deontic periphrases. *Journal of Historical Sociolinguistics* 4: 177–219. [CrossRef]

Blas Arroyo, José L., and Kim Schulte. 2017. Competing modal periphrases in Spanish between the 16th and the 18th centuries. *Diachronica* 34: 1–39. [CrossRef]

Brown, Esther L., and Javier Rivas. 2012. Grammatical relation probability: How usage patterns shape analogy. *Language Variation and Change* 24: 317–41. [CrossRef]

Bybee, Joan. 1985. *Morphology: A Study of the Relation between Meaning and Form*. Amsterdam: John Benjamins.

Bybee, Joan. 1988. Morphology as lexical organization. In *Theoretical Morphology*. Edited by Michael Hammond and Michael Noonan. San Diego: Academic Press.

Bybee, Joan. 1998. The emergent lexicon. *Chicago Linguistic Society* 34: 421–35.

Bybee, Joan. 2001. *Phonology and language Use*. Cambridge: Cambridge University Press.

Bybee, Joan. 2003. Mechanisms of change in grammaticization: The role of frequency. In *The Handbook of Historical Linguistics*. Edited by Brian D. Joseph and Richard Janda. Oxford: Blackwell, pp. 602–23.

Bybee, Joan. 2010. *Language, Usage and Cognition*. Cambridge: Cambridge University Press.

Bybee, Joan L., and William Pagliuca. 1987. The evolution of future meaning. In *Papers from the 7th International Conference on Historical Linguistics*. Amsterdam and Philadelphia: John Benjamins, pp. 109–22.

Bybee, Joan, and Sandra A. Thompson. 1997. Three frequency effects in syntax. *Berkeley Linguistics Society* 23: 65–85. [CrossRef]

Cinque, Guglielmo. 2006. *Restructuring and Functional Heads*. Oxford: Oxford University Press, vol. 4.

Davies, Mark. 1995. Analyzing Syntactic Variation with Computer-Based Corpora: The Case of Modern Spanish Clitic Climbing. *Hispania* 78: 370–80. [CrossRef]

Davies, Mark. 1998. The evolution of Spanish clitic climbing: A corpus-based approach. *Studia Neophilologica* 69: 251–63. [CrossRef]

Davies, Mark. 2002. Corpus del Español: 100 Million Words, 1200s–1900s. Available online: http://www.corpusdelespanol.org (accessed on 11 July 2020).

DuBois, John W. 1985. Competing motivations. In *Iconicity in Syntax*. Edited by John Haiman. Amsterdam: John Benjamins, pp. 343–65.

Fernández Ulloa, Teresa. 2001. Perífrasis verbales en el castellano de Bermeo (Bizkaia). *Revista Española de Lingüística/Spanish Journal of Linguistics* 30: 1–34.

Garachana, Mar. 2017. Perífrasis formadas en torno a tener en español. Ser tenudo/tenido o/a/de + infinitivo, tener a/de + infinitivo, tener que + infinitivo. In *La Gramática en la Diacronía. La Evolución de las Perífrasis Verbales Modales en Español*. Edited by Mar Garachana. Madrid-Frankfurt: Iberoamericana/Vervuert.

Garachana, Mar, and Áxel Hernández. 2017. La reestructuración del sistema perifrástico en el español decimonónico. El caso de *haber de/tener de* + infinitivo, *haber que/tener que* + infinitivo. In *Herencia e Innovación en el Español del Siglo XIX*. Edited by Elena Carpi and Rosa M. García Jiménez. Pisa: Pisa University Press, pp. 127–46.

Gerlach, Brigit. 2002. *Clitics between Syntax and Lexicon*. Amsterdam and Philadelphia: John Benjamins, vol. 51.

Giammatteo, Graciela M., and Ana M. Marcovecchio. 2009. Perífrasis verbales: Una mirada desde los universales lingüísticos. *Sintagma* 21: 21–38.

Givón, Talmy. 1971. On the verbal origin of the Bantu verb suffixes. *Studies in African linguistics* 2: 145.

Givón, Talmy. 1979. *On Understanding Grammar*. New York: Academics Press.

Givón, Talmy. 1983. Topic continuity in discourse: An introduction. In *Topic Continuity in Discourse: A Quantitative Cross-Language Study*. Edited by Talmy Givón. Amsterdam: John Benjamins, vol. 3, pp. 1–42.

Givón, Talmy. 1992. The grammar of referential coherence as mental processing instructions. *Linguistics* 30: 5–55. [CrossRef]

Givón, Talmy. 1995. Coherence in text vs. coherence in mind. In *Coherence in Spontaneous Text*. Edited by M. A. Gernsbacher and T. Givón. Pittsburgh: John Benjamins, pp. 59–115.

Goldberg, Adele E. 1995. *Construction Grammar*. Chicago: The University of Chicago Press.

Goldberg, Adele E. 2006. *Constructions at Work: The Nature of Generalization in Language*. Oxford: Oxford University Press.

Gómez Manzano, Pilar. 1992. *Perífrasis Verbales Con Infinitivo: Valores y usos en la Lengua Hablada*. Madrid: UNED.

González López, Verónica. 2008. Spanish Clitic Climbing. Unpublished Ph.D. dissertation, The Pennsylvania State University, State College, PA, USA.

González López, Verónica. 2013. Asturian identity reflected in pronoun use: Enclisis and proclisis patterns in Asturian Spanish. In *Selected Proceedings of the 6th International Workshop on Spanish Sociolinguistics*. Edited by Ana M. Carvalho and Sara Beaudrie. Somerville: Cascadilla Proceedings Project, pp. 76–86.

Gudmestad, Aarnes. 2006. Clitic climbing in Caracas Spanish: A sociolinguistic study of ir and querer. *Indiana University Linguistics Club Working Papers* 6: 1–14.

Heine, Bernd, and Tania Kuteva. 2002. *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.

Heine, Bernd, and Mechthild Reh. 1984. *Grammaticalization and Reanalysis in African Languages*. Hamburg: Helmet Buske.

Hopper, Paul J. 1987. Emergent Grammar. *Berkeley Linguistics Society* 13: 139–57. [CrossRef]

Hopper, Paul J., and Elizabeth C. Traugott. 2003. *Grammaticalization*. Cambridge: Cambridge University Press.

Johnson, Daniel E. 2009. Getting off the GoldVarb standard: Introducing Rbrul for mixed-effects variable rule analysis. *Language and Linguistics Compass* 3: 359–83. [CrossRef]

Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar: Descriptive Application*. Stanford: Stanford University Press, vol. 1.

Lewis, David. 1970. General semantics. *Synthese* 22: 18–67. [CrossRef]

Lewis, David. 1979. Scorekeeping in a Language Game. *Journal of Philosophical Logic* 8: 339–59. [CrossRef]

Martínez Díaz, Eva. 2003. La frecuencia de haber y tener en las estructuras perifrásticas de obligación. Algún fenómeno de variación en el español de Cataluña. *Interlingüística* 14: 691–94.

Matthews, Peter H. 2007. Animacy Hierarchy. The Concise Oxford Dictionary of Linguistics. Available online: https://www-oxfordreference-com.libweb.lib.utsa.edu/view/10.1093/acref/9780199202720.001.0001/acref-9780199202720-e-173?rskey=reZLMy&result=1 (accessed on 11 July 2020).

Miller, D. Gary, and Elly van Gelderen. 2017. Grammaticalization. In *Oxford Bibliographies Online: Linguistics*. Available online: https://www.oxfordbibliographies.com/view/document/obo-9780199772810/obo-9780199772810-0019.xml (accessed on 4 September 2020).

Myhill, John. 1988. The Grammaticalization of Auxiliaries: Spanish Clitic Climbing. *Annual Meeting of the Berkeley Linguistics Society* 14: 352–63. [CrossRef]

Myhill, John. 1989. Variation in Spanish clitic climbing. In *Georgetown University Round Table on Languages and Linguistics 1988*. Washington, DC: Georgetown University Press, pp. 227–50.

Myhill, John. 2005. Quantitative methods of discourse analysis. In *Quantitative Linguistics: An International Handbook*. Edited by Reinhard Köhler, Gabriel Altmann and Rajmund Piotrowski. Berlin: Mouton de Gruyter, pp. 471–98.

Napoli, Donna. J. 1981. Semantic Interpretation vs. Lexical Governance: Clitic Climbing in Italian. *Language* 57: 841–87. [CrossRef]

Olbertz, Hella. 1998. *Verbal Periphrases in a Functional Grammar of Spanish*. Berlin: Walter de Gruyter, vol. 22.

R Core Team. 2014. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing. Available online: http://www.r-project.org (accessed on 11 July 2020).

Requena, Pablo E. 2015. Direct Object Clitic Placement Preferences in Argentine Child Spanish. Unpublished Ph.D. dissertation, The Pennsylvania State University, State College, PA, USA.

Rizzi, Luigi. 1976. Ristrutturazione. *Rivista Di Grammatica Generativa Roma* 1: 1–54.

Rizzi, Luigi. 1978. A restructuring rule in Italian syntax. Recent transformational studies in European languages. In *Linguistic Inquiry Monograph*. Edited by Samuel J. Keyser. Cambridge: MIT Press, vol. 3, pp. 113–58.

Russell, B. 1905. On Denoting. *Mind, New Series* 14: 479–93. [CrossRef]

Schwenter, Scott A., and Rena Torres Cacoullos. 2014. Competing constraints on the variable placement of direct object clitics in Mexico City Spanish. *Revista Española de Lingüística Aplicada/Spanish Journal of Applied Linguistics* 27: 514–36. [CrossRef]

Silva-Corvalán, Carmen. 1994. *Language Contact and Change: Spanish in Los Angeles*. New York: Oxford University Press.

Sinnott, Sarah, and Ella Smith. 2007. ¿Subir o no subir? A Look at Clitic Climbing in Spanish. Paper presented at the 36th New Ways of Analyzing Variation, University of Pennsylvania, Philadelphia, PA, USA, October 12.

Strozer, Judith R. 1976. Clitics in Spanish. Ph.D. Dissertation, UCLA, Los Angeles, CA, USA.

Suñer, Margarita. 1980. Clitic Promotion in Spanish Revisited. In *Contemporary Studies in Romance Languages*. Edited by Frank H. Naussel. Bloomington: Indiana University Linguistics Club, pp. 300–30.

Tagliamonte, Sali A. 2006. *Analysing Sociolinguistic Variation*. Cambridge: Cambridge University Press.

Torres Cacoullos, Rena. 1999. Construction frequency and reductive change: Diachronic and register variation in Spanish clitic climbing. *Language Variation and Change* 11: 143–70. [CrossRef]

Troya Déniz, Magnolia, and Ana M. Pérez Martín. 2011. Distribución de clíticos con perífrasis verbales en hablantes universitarios de las Palmas de Gran Canaria. *Lingüística* 26: 9–25.

Weinreich, Uriel. 1966. On the semantic structure of language. In *Universals of Language*, 2nd ed. Report of a Conference Held at Dobbs Ferry, NY, USA, April 13–15. Edited by Joseph H. Greenberg. Cambridge: MIT Press, pp. 114–71. First published 1963.

Zabalegui, Nerea. 2008. La posición de los pronombres átonos con verbos no conjugados en el español actual de Caracas. *Akademos* 10: 83–107.