Rena Torres Cacoullos, Penn State University

Gradual loss of analyzability: diachronic priming effects

Abstract: In this article, I test accounts of the formation of grammatical units by deploying the facts of variation of the Spanish Progressive. First, unithood and frequency measures support usage-based chunking as more tenable than formal reanalysis as an account of change in constituency. Second, comparison of multivariate models of variation over time reveals that the spread of the Spanish Progressive relative to the simple Present has been differential, as shown in change in the linguistic conditioning of variant choice, in disagreement with an abrupt-reanalysis, constant-rate hypothesis but in support of gradual change in diachrony and inherent variability in synchrony. Third, a priming effect – such that selection of a given construction is favored by previous use of a related construction (here, priming of the *estar* Progressive by non-Progressive *estar* constructions) – is introduced as a measure of internal structure, in particular, of (loss of) analyzability.

1 Introduction

How do grammatical units come about, and how can change in constituency be observed? *Reanalysis* is widely invoked by linguists of otherwise different persuasions as a pivotal mechanism of syntactic change. Reanalysis is understood to change underlying structure, including constituency and syntactic-category labels (Campbell 1998: 284). For example, the English future auxiliary is said to result from reanalysis of the purposive motion construction of main verb *go* with a non-finite clause complement, as represented by rebracketing of some kind: [BE *going* [*to* VERB]] > [BE *going to* VERB] (Hopper and Traugott 1993: 3). A material indication of such reanalysis would be phonetic reduction of *going to* to *gonna*.

Reanalysis has been seen as abrupt, following from the view that each word sequence must have a unique constituent analysis, which in turn follows from the formalist (generative) view that proposed syntactic rules or constraints are categorical and that syntactic categories, for example, main vs. auxiliary verb, are discrete. But the facts of synchronic variation, as between *going to* and *gonna* (e. g., Poplack and Tagliamonte 1999: 328–332), disturb an understanding of grammatical change as abrupt reanalysis.

The probabilistic aspects of grammar (Labov 1969; Cedergren and Sankoff 1974) are now being recognized by more linguists, who are exploring usage-based and emergentist theories of grammar. Bybee (2010) proposes that constituent structure is derivable from domain-general mechanisms in operation as speakers produce and process language. Pertinent here is the fusing of sequential experiences that occurs with repetition or, for language, the *chunking* of frequent word sequences as single processing units (Bybee 2010: 34 and references therein). For example, the vowel of *don't* is more likely to reduce to a schwa in *I don't know* than when the main verb is a less frequent one, as in *I don't inhale*, even though the two expressions are apparently of the same syntactic structure (Scheibman 2001: 114).

In a usage-based view, a consequence of frequent repetition and ensuing chunking of contiguous linguistic units is the *loss of analyzability* of the sequence of (erstwhile) units (Bybee 2010: 44–45; see also Croft and Cruse 2004: 250–253; Langacker 1987: 292). Analyzability is seen as a morphosyntactic parameter that has to do with the degree to which the internal structure is discernable (akin to the morphological "decomposability" of complex words (cf. Hay 2001)), which is not subsumable under a semantic criterion. For example, while *pull strings* has a non-transparent meaning that is not predictable from *pull* and *strings*, it is syntactically analyzable, in that speakers presumably recognize the component parts as individual words and the relation between them, here a verb with its object.

With schematic (productive) constructions that have open classes of items, such as [BE *going to* VERB], loss of analyzability is understood as the weakening of the association between the erstwhile individual components with other instances of the same items. In Bybee's (2003: 618) example, as *going to* reduces to *gonna*, its composite morphemes lose their association with *go*, *to* or *-ing*. But what observations provide evidence for "association" and its loss?

In this paper, the facts of variation are deployed to tackle the question of how grammatical units come about. I use variability in the Spanish Progressive to test gradualness vs. abruptness in the formation of grammatical units and put forward diachronic priming effects as a gauge of analyzability and its erosion over time. After presenting the linguistic variable in Section 2, I begin in Section 3 with a recapitulation of unithood indices and frequency measures, which score an initial point in favor of usage-based chunking as more tenable than formal reanalysis. I then present a multivariate model of variation between the Progressive and simple Present, in Section 4. The shift in the relative importance of aspectual reading and locative co-occurrence scores a further point in favor of gradual change in diachrony and inherent variability in synchrony. In Section 5 I introduce priming effects as a measure of (erosion of) analyzability.

2 Spanish Progressive *ESTAR* + Verb-*ndo*

Latin did not have a dedicated morpheme or construction for progressive aspect, the simple Present serving this function among others (Allen and Greenough 1916: 293, § 465). Probably the most common source for progressives crosslinguistically is locative expressions (Bybee, Perkins and Pagliuca 1994: 127–133; cf. Comrie 1976: 98–105). Beginning from the earliest Spanish texts, we find gerunds (-*ndo* forms) combining with finite forms of spatial (locational, postural or movement) verbs. Besides *estar* 'be (at)', these were usually *ir* 'go', *andar* 'walk, go around', *venir* 'come', *salir* 'go out', *quedar* 'remain, stand still'. Examples are (1a), with *ir*, and (1b), with *venir*.

(1a)   *déxa=me   dezir,   que se   **va**   **hazie-ndo** noche*
        let.IMP=ACC.1SG say.INF that REFL go.PRS.3SG make-GER night
        'let me speak, it **is** [literally: goes] **becoming** night'
        (15th c., Celestina, Act VI)

(1b)   *¿No oyes   lo que   **viene**   **canta-ndo** ese villano?*
        NEG hear.PRS.2SG that.REL come.PRS.3SG sing-GER   that rustic
        'Don't you hear what that rustic **is** [literally: comes] **singing**?'
        (17th c., Quijote II, Ch. IX)

Allen and Greenough (1916: 819, § 507) give a medieval Latin example of this general Spatial Verb + VERB-*ndo* (gerund) construction, *cum una dierum **flendo sedisset**, quidam miles generosus iuxta eam **equitando venit** (Gesta Romanorum,* 66 [58]) 'as one day she **sat weeping**, a certain knight **came riding** by' (*Gesta Romanorum,* 66 [58]).

     In Torres Cacoullos (2000) I adduced evidence for the origins of Spanish Progressive *ESTAR* (< Latin *stare* 'stand') + VERB-*ndo* (gerund) as a locative expression 'be located somewhere VERB-ing' from its early distributions across co-occurring locatives (most frequently with *en* 'in') and verbs (most frequently *hablando* 'talking', other verbs of speech, *esperando* 'waiting', and verbs of body activity). These co-occurring elements are consonant with being stationary in a given place. A 13th century example appears in (2). In contrast, gerund combinations with motion verbs *ir* 'go' and *andar* 'walk (around)' tended to co-occur with other kinds of locatives (*a* 'to', *por* 'along') and verb classes (motion, process, general activity).

(2)   *u<uest>ros hebreos **estan**   aqui **razona-ndo***
                        PRS.3PL here discourse-GER
        'your Hebrews **are** here **conferring**'
        (13th c., General Estoria I, fol. 151r)

The key construct in variation theory is the *linguistic variable* (Labov 1969), a set of variants which "are used interchangeably to refer to the same states of affairs" (Weiner and Labov 1983: 31), i. e. "alternative ways of saying the same thing" (Labov 1982: 22). In the pair of examples from a 19th century play in (3), the "same thing", or grammatical function, is present progressive and the "alternative ways", or *variants*, are the Progressive and simple Present forms. In the English translation, PROG designates the Progressive – *ESTAR* + VERB-*ndo* – as in (3a), PRS the simple Present, as in (3b). Both forms here express a situation in progress at the moment of speech.

(3a)   *EDUARDO. – No me muestres esa compasión. Yo no la merezco.*
        *¿Sabes   tú   con quién **estás**   habla-ndo?*
        know.PRS.2SG you with REL   be.PRS.2SG speak-GER
        'EDUARDO: Don't show me such compassion. I don't deserve it.
        Do you know to whom you **are talking** (PROG)?'
        (19th c., Amor de padre, Act 5, Scene2)

(3b)   *AGENTE. – ¿Cómo tienes valor?*
        *Olvidas   que   **hablas**   con un republicano?*
        forget.PRS.2SG COMP speak PRS.2SG with a   republican
        'AGENT: How do you have the courage?
        Do you forget that you **are talking** (PRS) to a republican?'
        (19th c., Amor de padre, Act 3, Scene VII)

The *variable context* is the sum of contexts where distinctions in grammatical function among different forms may be "neutralized in discourse" (Sankoff 1988a: 153). This is defined here broadly as the domain of present temporal reference, since the Progressive and simple Present also compete as expressions of non-progressive present situations (e. g., (8), below). We circumscribe a variable context in order to adhere to the *principle of accountability*, that not only occurrences but also non-occurrences of a given variant be noted (Labov 1982: 30), here, where the Progressive could have materialized but the simple Present did instead, as in (3b).

Tokens of Present-tense *ESTAR* + VERB-*ndo* were exhaustively extracted from a corpus comprised of 60 texts from three time periods, the 13th–15th, 17th and 19th centuries (traditionally, Old Spanish, Golden Age Spanish, and Modern Spanish; the texts are listed in the Appendix). Tokens of the "non-occurrences", i. e., of the simple Present, were extracted by taking simple Present-tense occurrences of the same lexical types that appear in the Progressive in a given text. From this sample, Present-tense forms with future or past temporal reference were excluded, for example, *estaba [...] para montar a caballo [...], cuando **oigo** ¡tras tris, tras tras!* 'I was [...] about to get on the horse, when **I hear** tras tris, tras tras! (Pazos, Ch. XXI). Also discarded were first- or second-person singular discourse routines (e. g., *digo* 'I say', *ya ve(s)* 'you see') or prefabs involving *ser* 'be' (e. g. *es que* 'it's that').

Following these protocols, a total of 1,656 tokens of the Progressive or simple Present were retained for the analyses of variation. Table 1 depicts the number of texts, word counts and Ns for the three time periods.

Table 1: Data for the study of Progressive – simple Present variation in present temporal reference contexts

	13th–15th century	17th century	19th century
No. texts	17	15	28
Word count	2,500,000	600,000	900,000
N Progressive	119	180	317
N simple Present	429[1]	564	663

All tokens of both forms were coded according to a number of hypotheses about variant choice, operationalized as factors based on the presence or absence of linguistic elements of the context in which the token occurs. Included in the factor groups (independent or predictor variables, or constraints) are co-occurrence of locative adverbials, aspectual reading and priming. The *linguistic conditioning* of variant selection is instantiated in probabilistic associations of forms with contextual elements. A multivariate model of the variation is presented in section 4 ahead, after we first consider evidence from distributional analysis and frequency counts, below.

3 Unithood and frequency

Spanish Progressive *ESTAR* + VERB-*ndo* would seem a good candidate for change via either reanalysis or loss of analyzability. The change in constituent structure would be from a sequence of two independent units – a finite form of main verb *estar* 'to be (located)' with a gerund -*ndo* complement – to a single periphrastic unit, in which *estar* is an auxiliary and the gerund is the main verb (4).

reanalysis / loss of analyzability

(4) $[ESTAR]_{verb}$ + $[VERB\text{-}ndo \text{ (gerund)}]_{complement}$ > $[ESTAR_{aux} + VERB\text{-}ndo_{verb}]_{Progressive}$

---

1 For the 13th–15th century simple Present sample, tokens of lexical types appearing in the Progressive were not extracted from Grima*lte y Gradissa* and *Crónica de los Reyes Católicos*, for which electronic versions were not available (three Progressive tokens each); also omitted were Present tokens of frequent *decir* 'say' in *Corbacho* (of which there was one Progressive token). More on the texts, the simple Present sampling and exclusions is given in Torres Cacoullos (2012).

Whereas in *going to* the items are contiguous, here we have a schematic construction with an intervening slot for the open class of items, the Verb. In the absence of surface phonetic reduction, as with English future *gonna*, what evidence could be assembled for the status of *ESTAR* + Verb-*ndo* as a unitary constituent?

We may take the obverse of analyzability to be *unithood*, operationalized as the proportion of the instances of the construction in which the adjoining items behave as a single unit, i. e. as one word. In previous work (e. g., Torres Cacoullos 2006; Torres Cacoullos and Walker 2011) we developed unithood indices from distributional analysis, which tracks proportions of tokens of an expression across its contexts of occurrence. Increasing unithood of *ESTAR* + Verb-*ndo* has been inferred from a decreasing proportion of occurrences with elements intervening between *estar* and the gerund, with more than one gerund per *estar*, or with the gerund preceding *estar* (Torres Cacoullos 2000: 31–55; Bybee and Torres Cacoullos 2009: 201–203; Torres Cacoullos 2012: 79).

A more direct index of unithood is the positioning of object pronouns (which precede finite verb forms in modern Spanish) (Torres Cacoullos 1999b). In (5a) the object pronoun (underlined) is postposed to the gerund (is telling <u>him</u>), in (5b) it is preposed to *estar* (literally, '<u>it</u> are saying'). The latter configuration, known as "clitic climbing", has been viewed in generative syntax as a restructuring of a series of verbs into a single verbal complex (e. g., Rizzi 1982). In a functionalist view, "clitic climbing" has been seen as a manifestation of the grammaticalization of auxiliaries, as a verb comes to express grammatical (e. g., aspectual, progressive) more than lexical (e. g., spatial, locative) meaning (Myhill 1988).

(5a)   [*ESTAR*] + [Verb-*ndo* + object pronoun/clitic]

Yo voy con tu cordon tan alegre: que se me figura que
esta   dizie-ndo <u>LE</u>   alla   su   coraçon la merced que nos heziste
be.PRS.3SG tell-GER   DAT.3SG there his   heart [...]
'I'll go with your cord so happily, I can almost imagine that his heart there **is telling**
<u>him</u> of the great favor you have done us'
(15th c., Celestina, Act IV, fol.32r)

(5b)   [object pronoun/clitic + *ESTAR* + Verb-*ndo*]

– *Pero que nosotros tampoco les vamos a dar cien días. Vamos a decir lo que nos parezca*
  desde hoy.
– *Ya   <u>LO</u>   **estamos   dicie-ndo**.*
  already ACC.3SG be.PRS.1PL tell-GER
'But we're not going to give them a hundred days. We're going to say what we think
starting today.'
'We [<u>it</u>] **are** already **saying** it'
(20th c., CORLEC, CDEB014A, p215–p216)

In the 15th century example in (5a) *ESTAR* + Verb-*ndo* is compatible with locative meaning, indicated by co-occurring *allá* 'there' in the same clause and the motion verb *voy* 'I go' in a previous clause: the speaker will go to where the person represented metonymically by his heart is located (*está...allá* 'is...there'). There is at the same time aspectual meaning, as conveyed by *se me figura* 'I can imagine': the situation referred to by the gerund is in progress at speech time. In comparison, spatial meaning appears at best attenuated in the 20th century example in (5b), where most prominent is aspectual meaning, indicated by co-occurring temporal adverbial *ya* 'already': the speaker asserts that the verbal situation (*diciendo* 'saying') is in progress.

In Table 2, though the count of all eligible cases is low, , there is a clear trend of increased rates of placement before *estar* (proclisis). Increasing placement of object pronouns before the whole complex (as with single finite verbs) rather than attached to the gerund, can be taken as an indication of enhanced unithood.[2]

Table 2: Increasing unithood of *ESTAR*$_{Present}$ + Verb-*ndo*: placement of object pronouns before *estar* ("clitic climbing")[3]

13th–15th century	17th century	19th century	20th century[4]
71 % (10/14)	72 % (18/25)	89 % (58/65)	97 % (100/103)

Unithood is a theory-neutral measure, compatible with either reanalysis or loss of analyzability. However, the two accounts are distinguished by the expected (non) role of frequency. On the one hand, loss of analyzability attributable to chunking depends on repetition. Applied to the case at hand, with frequent repetition the sequence *ESTAR* + Verb-*ndo* would become a new chunk – more of a fused unit. On the other hand, a theory of syntactic change based on reanalysis in terms of

---

2 The 19th and 20th century rate of preposed object clitics shown in Table 2 is higher than for all tenses of *ESTAR* + Verb-*ndo* (respectively, 70 %, 54/77 in the same texts (reported in Bybee and Torres Cacoullos 2009: 203) and 89 %, 103/115 in Mexico City "habla popular" (UNAM 1976) (reported in Torres Cacoullos 1999b: 146). This is consonant with Progressive grammaticalization advancing in present before past tenses (Torres Cacoullos 2012: 110, n. 3) (whereas habitual markers are said to appear in past before generalizing to present temporal reference contexts (Bybee et al. 1994: chapter 5)).

3 In χ^2 tests, difference between 17th and 19th p < 0.06 (n. s.), between 19th and 20th p < 0. 05. 13th–15th and 17th century totals include object pronouns placed between *estar* and the gerund (e. g., *ell Aguila* **esta** *la* **remira<n>-do**, GEII, fol. 189v) (N = 2, N = 3, respectively).

4 20th century = CORLEC (Marcos Marín 1992) "conversacional" portion (see Table 1 for remaining data)

operations such as "movement and "merge" (e. g., Roberts and Roussou 2003) makes no predictions about frequency of use, under the assumption that usage does not impinge on grammar (e. g., Newmeyer 2003).

Table 3 displays three frequency counts for *ESTAR* + Verb-*ndo*, one absolute, i. e. the token frequency of the construction, and two relative, namely the proportion it constitutes of gerund constructions vis-à-vis other spatial auxiliaries and its rate relative to the simple Present. The first count, in the first row, is straightforward text frequency normalized per 100,000 words (based on the figures given in Table 1 above), by which there is an evident rise (cf. Torres Cacoullos 2012: 77).

Table 3: Frequency increase of *ESTAR*Present + Verb-*ndo*

Century	13th–15th century	17th century	19th century	20th century[5]
Token frequency per 100k words	5	30	35	151
Proportion *estar* gerund constructions with respect to other locative-postural-movement verbs[6]	38 % (45/117)	41 % (54/133)	62 % (93/149)	83 % (364/436)
Rate relative to simple Present occurrences of verbs which appear in *ESTAR* + Verb-*ndo*	14 % (39/282)[7]	24 % (180/744)	32 % (317/980)	

The second frequency measure is the proportion that *ESTAR* + Verb-*ndo* constitutes as an instance of the general gerund construction with finite forms of spatial verbs. Some of these gerund constructions (especially with *ir* and, in some dialects, *andar*) remain robust in modern varieties of Spanish (Torres Cacoullos 1999a). Nevertheless, as Table 2 shows in the second row, the aspectual auxiliary is increasingly more likely to be *estar* than another spatial verb (cf. Torres Cacoullos 2000: 55–60). We can think of this as a measure of string frequency, which may indicate "chunk status" (Brown and Rivas 2011: 42–43).

---

5 20th century = CORLEC (Marcos Marín 1992) "conversacional" portion: word count 241K, N *ESTAR*Pres + Verb-*ndo* 364.

6 Counts from a subset of texts in Table 1: for 3th–15th century from *Calila*, *GE*, *Celestina*, *CRC* (*andar* (17), *ir* (39), *venir* (10)); 17th from *Quijote* (*andar* (12), *ir* (60), *quedar* (5)); 19th century, from *Pepita*, *Perfecta*, *Regenta*, *Pazos* (*andar* (6), *ir* (39), *seguir* (5)); 20th century from CORLEC (*ir* (47), *seguir* (20), *venir* (4)).

7 Count from 15th century *Corbacho* and *Celestina* only, furthermore not counting simple Present *decir* 'say, tell' in the *Corbacho*.

A third frequency measure is the increasing rate of *ESTAR* + Verb-*ndo* relative to the simple Present variant (as in the pair of examples in (3)). Here we count the frequency of Progressive forms relative to simple Present forms of lexical types appearing in the Progressive in the same text (cf. Table 1). Though such a lexically-based count does not provide overall rates, we again observe a rising trend, in the third row of Table 3.

In summary, from enhanced unithood of the construction, measured here by object clitic position, we may infer the increasing absorption of the auxiliary-in-becoming into a periphrastic unit. Coupled with this inference of loss of analyzability is the finding that Present-tense *ESTAR* + Verb-*ndo* sequences increase in frequency, by both absolute and relative frequency measures. We have support, then, for the prediction following from the chunking hypothesis that loss of analyzability of a word sequence is accompanied by frequency increases. In Section 5 I will put forward priming as a more direct measure of analyzability and its erosion. But first, in the next section, evidence for gradualness in the evolution of *ESTAR* + Verb-*ndo* is adduced from the comparison of probabilistic models of its variable use over time.

4 Change as change in linguistic conditioning

4.1 Inherent variability vs. competing grammars: Predictions about contextual effects

The position of Weinreich, Labov and Herzog (1968: 101) is that "command of heterogeneous structures is not a matter of multidialectalism or 'mere' performance, but is part of unilingual linguistic competence". In this view, systematic variation belongs to (a single) grammar (Cedergren and Sankoff 1974: 334).

However, variation and apparent gradualness may be attributed to competing grammars underlying a given form. Abruptness and discreteness may be upheld in the face of observed variation by viewing the aggregate data as reflecting the coexistence of multiple (generative) grammars. Language change is then modeled as modification in the distribution of competing grammars over time (e. g., Yang 2000). For example, the spread of English *do*-support across different sentence types (e. g., negative declaratives and affirmative wh-object questions) is seen as "surface manifestations of a single change in grammar" (such as loss of Verb-to-Infl movement) (Kroch 1989: 199; but see Bybee 2010: chapter 7). Because change is postulated to be a single abrupt change in a parameter setting and the gradual time course is seen as representing a shift from the old

invariant homogenous grammar to the new one, the contexts of the change must be uniform in their effects.

According to this scenario a new structure such as *do*-support may be favored earlier in some contexts than others and begin in those contexts with a higher rate, but the rate of change is constant across contexts; in terms of *linguistic conditioning*, the effects remain fixed in magnitude and direction as the change is propagated (Kroch 1989: 206).

Returning to *ESTAR* + Verb-*ndo*, if the change is abrupt and it is the propagation of change (e. g., across authors and texts) that is gradual, we should observe that the frequency of the newer variant – the Progressive – relative to the older one – the simple Present – increases at a constant rate, uniformly, across linguistic contexts. But if the change itself is a gradual modification of the grammar – one with inherent variability – it is possible that the linguistic contexts and conditions could vary across the course of the change. That is, the rate of occurrence of the Progressive could increase differentially across linguistic contexts.[8] Does it?

4.2 Shifts in relative magnitude of effect

Table 4 shows three independent Variable-rule analyses (Sankoff, Tagliamonte and Smith 2005) of the probability that the Progressive variant will be selected, in 13th–15th century, 17th century and 19th century texts. Variable-rule analysis uses logistic regression to perform binomial multivariate analysis for a choice between two variants, here, the Progressive vs. the simple Present (Sankoff 1988b). There are three lines of evidence in interpreting results of Variable-rule analysis (Poplack and Tagliamonte 2001: 88–95):

1. statistical significance of effect, determining the factor groups (independent variables or constraints) that together account for the largest amount of variation (in terms of stepwise increase of log likelihood, such that the addition of any of the remaining factor groups does not significantly increase the fit to the model);

2. direction of effect, with probabilities (or factor weights, shown in the bordered set of columns) closer to 1 indicating a *favoring*, and closer to 0 a *dis*favoring, effect on *ESTAR* + Verb-*ndo*. That is, the closer to 1 the probability, the greater the likelihood of the Progressive in each of the contexts (factors) listed on the left;[9]

---

9 Factor weights for non-significant groups, from the first "step down" run, in which all groups

3.   relative magnitude of effect, as assessed by the *Range* (shown in italics) between the favoring and the disfavoring probability within each (binary) factor group.

With respect to significance, in Table 4 we see that contributing to variant choice in the 13th–15th century data are Aspect, Locative co-occurrence and Priming (as well as Polarity-Sentence type and Temporal co-occurrence, but not Stativity). In the 19th century, Priming no longer significantly increases the fit to the model (and Stativity has achieved significance). Zeroing in here on Locative co-occurrence and Aspect, the quantitative argumentation that follows will rely on direction of effect, which has generally remained stable, and relative magnitude of effect, which turns out to be the main locus of change (on Priming, see Section 5; on the other effects, see Torres Cacoullos 2012).

For the Locative co-occurrence factor group, tokens of both forms were coded for the presence of a locative in the same clause, as in (6). The hypothesis of retention (Bybee et al. 1994: 15–19) or persistence (Hopper 1991: 28–29), according to which grammaticalizing constructions/morphemes have semantic content deriving from the meaning of their source construction, leads to the prediction that selection of *ESTAR* + Verb-*ndo* will be favored in the presence of a locative (Torres Cacoullos 2012: 83–85).

(6a)   *en la galería*   me   **está**   **esperando**
        In the gallery   DAT.1SG   be.PRS.3SG   wait-GER
        'he **is waiting** (PROG) for me <u>in the</u> <u>gallery</u>'
        (19th c., Perfecta, Ch. X)

(6b)   *ahí*   le   **esperan**   a Vd.   con las caballerías
        there   DAT.3SG   wait.PRS.3PL   ACC you   with the mounts
        'they **are waiting** (PRS) for you <u>there</u> with the mounts'
        (19th c., Perfecta, Ch. I)

For the Aspect factor group, I coded tokens of both the Progressive and simple Present as 'limited duration' if the aspectual reading was one of progressive or continuous (Comrie 1976: 33), as in the pair of examples in (3) and (6), above. Limited duration also applies to stative predicates when the situation is temporally circumscribed, or bound to speech time, again for both variants, as in (7). 'Extended duration', on the other hand, subsumes habitual aspect for dynamic verbs, and states without temporal limits, which exist indefinitely, as in (8) (on coding for aspect, see Torres Cacoullos 2012: 87–91).

---

are included in the regression, are provided within brackets to indicate direction of effect (Poplack and Tagliamonte 2001: 93–94).

Table 4: Three independent Variable rule analyses of linguistic factors contributing to selection of the Progressive[10]

Input	13th–15th century .21 (119/548)			17th century .21 (180/744)			19th century .33 (317/980)		
	Prob	%Prog	N	Prob	%Prog	N	Prob	%Prog	N
Locative co-occurrence									
Present	.90	72 %	21/29	.76	57 %	34/60	.67	51 %	36/70
Absent	.47	19 %	98/519	.47	21 %	142/676	.49	31 %	275/888
Range	43			29			18		
Aspect									
Limited duration	.68	37 %	62/169	.70	41 %	144/355	.71	56 %	269/481
Extended duration	.35	15 %	30/200	.16	6 %	11/182	.12	8 %	16/210
Range	33			54			59		
Priming									
Preceding *estar* construction	.76	56 %	9/16	.69	55 %	13/26	[.49]	30 %	14/46
Preceding 'Other' tenses	.54	21 %	41/194	.53	27 %	70/256	[.56]	38 %	117/310
Preceding simple Present	.46	19 %	55/296	.47	20 %	84/430	[.46]	27 %	133/486
Range	30			22					
Polarity – Sentence type									
Affirmative declarative	.54	24 %	106/440	.57	27 %	164/598	.58	37 %	285/778
Negative, Interrogative	.31	13 %	11/86	.18	8 %	9/116	.18	13 %	20/160
Range	23			39			40		
Temporal co-occurrence									
Present	.76	33 %	25/76	[.54]	28 %	29/103	.60	37 %	50/135
Absent	.45	20 %	93/469	[.49]	24 %	150/639	.48	32 %	266/839
Range	31						12		
Stativity									
Dynamic predicate	[.49]	21 %	93/435	[.52]	26 %	153/593	.51	34 %	282/840
Stative predicate	[.56]	23 %	26/113	[.43]	18 %	27/151	.44	25 %	35/140
Range							7		

---

10 Non-significant factors are shown within square brackets. Ns in some factor groups do not add up to total N because of excluded factors or uncodable tokens.

(7)   Limited duration (in progress, bound to speech time)

(7a)   *fabla   p<er>o bermudo por q<ue> estas   calla-ndo*
      speak. IMP Pero   Bermudo why   be.PRS.2SG be_silent-GER
      'speak Pero Bermudo, why **are** you **(being) silent** (PROG)?'
      (13th c., EE II, fol. 240v)

(7b)   *No   me   entiendes,   Sancho: no quiero decir sino que [...]*
      NEG   DAT.1SG understand.PRS.2SG
      'You do not **understand** (PRS) me, Sancho: all I want to say is that [...]'
      (17th c., Quijote II, Ch. XXV)

(8)   Extended duration (habitual, indefinitely existing)

(8a)   *los peces son´ los huéspedes   que siempre están   calla-ndo*
      the fish be.PRS.3PL the guests   REL   always   be.PRS.3PL be_silent-GER
      'the fish are the guests who **are** always **(being) silent** (PROG)'
      (13th c., Apolonio, verse 506)

(8b)   *No   entiendo   otra   lengua   que la   mía*
      NEG understand. PRS.1SG other language than the mine
      'I **understand** (PRS) no language other than my own'
      (17th c., Quijote II, Ch. II)

With respect to direction of effect, we note that a co-occurring locative consistently favors selection of *ESTAR* + VERB-*ndo* (with probabilities of .90, .76 and .67, in the 13th–15th, 17th and 19th centuries, respectively; in the absence of a co-occurring locative values are close to .50, at .47, .47, .49). Stability in the direction of effect for locative co-occurrence supports the retention hypothesis. Also consistently favoring the Progressive are situations of limited duration. This effect is in place from the beginning (probability .68 in the 13th–15th century). Early favoring of *ESTAR* + VERB-*ndo* in limited duration situations is congruent with the view that progressive aspect is an implication of the locative construction, rather than the result of an abrupt metaphorical space > time leap (Bybee et al. 1994: 137; Torres Cacoullos 2012: 84). By the direction of effect, then, there is continuity in the linguistic conditioning of the Progressive.

Is there change? Yes. The locus of change here is in the relative magnitude of effect. Comparison of the *Range* values for Aspect and Locative co-occurrence within each analysis gives an indication of their relative importance. [11] We see

---

11 This is not a strict mathematical rule. The goal of the stepwise procedure in Variable-rule analysis is to find the set of factor groups which jointly account for the variation and is not primarily meant to order these factor groups according to any criterion. In the analyses shown in Table 4, the order of the factor groups according to the *Range* is consistent with that suggested by

that in the 13th–15th century the *Ranges* are comparable (with a ratio of 33:43 = 0.8) but that in the 17th century the *Range* for Aspect is twice as great (54:29 = 1.9) and in the 19th century it is three times greater (59:19 = 3.1). Here we have the answer to the question of whether the Progressive spreads differentially across linguistic contexts. It does. This contradicts the hypothesis of gradualness in the propagation of change but abruptness in grammatical change itself.

The weakening of the favoring effect of co-occurring locatives, as the probabilities get farther from 1 over time, may be taken as a measure of the loss of source-construction meaning, known as semantic bleaching ("depletion" (Givón 1975: 94)), in the course of grammaticalization. Furthermore, we can note that *ESTAR* + VERB-*ndo* is increasingly disfavored in reference to extended duration (habitual, indefinitely-existing-state) situations, which are becoming more the province of the simple Present, as probability values get closer to zero (at .35, .16 and .12, in the 13th–15th, 17th and 19th centuries, respectively). This means that an aspectual opposition with the simple Present has developed gradually: the originally more locative construction is used more and more as an aspectual expression of limited, as opposed to extended, duration. The developing – but not (yet) obligatory – Progressive – simple Present opposition is illustrated in (9), sleeping in progress (9a) vs. a habitual mode of sleeping (9b). Thus, in the course of speakers' recurrent choices of variants (as in (3), (6)–(8)), among which distinctions in aspectual function can be neutralized in discourse (Sankoff 1988a), the newer and older variant may gradually become aspectually more distinct (Torres Cacoullos 2012).

(9a)   *Coronel ¡Ah!,* **está   durmie-ndo**
         Colonel ah   be.PRS.2SG sleep-GER
         Nada, nada, duerma usted, no se moleste.
         'Colonel ... Ah, you'**re sleeping** (PROG). Never mind, never mind, sleep, don't be bothered'
         (19th c., Serafina la devota, Act 1, Scene VI)

---

the significance of the change in log likelihood, from the first step of the "step down" procedure when the least important group gets "cut": most important in Old Spanish are Aspect, Locative and Temporal (all p = .000), followed by Polarity-Sentence type (p = .001) and Priming (p = .008); in the 17th century Aspect, Locative and Polarity-Sentence type (all p = .000), followed by Priming (p = .032); in the 19th century Aspect and Polarity-Sentence type (both p = .000), followed by Locative (p = .014), Stativity (p = .019), and Temporal (p = .040). The same holds according to the order of selection in the "step up" procedure, another indication of relative magnitude of effect, except for the selection of Aspect before Locative co-occurrence in Old Spanish and of Stativity before Locative co-occurrence in the 19th century.

(9b)   *Dime   cómo **duermes***
       tell.IMP   how   sleep.PRS.2SG
       y te diré quién eres.
       'Tell me how you **sleep** (PRS) and I'll tell you who you are'
       (19th c., Regenta I, Ch. III)

To summarize this section: in comparing the "environmental constraints" (Labov 1982: 20) on the selection of the Progressive over time, we have found shifts in effect strength, or differential increases of the rate of the Progressive by linguistic context. These are compatible not with abrupt reanalysis but with gradual modification of the grammar. The findings here match those in another case of tense-aspect-mood grammaticalization, that of the 'go'-future in Brazilian Portuguese. In tracking future temporal reference over five centuries, Poplack and Malvar (2007: 157–160) found that as a variant receded or advanced, constraints on its selection changed, leading them to conclude that "the transition period in linguistic change is not abrupt, but proceeds as a series of small adjustments, as incoming and outgoing variants jockey for position in the system" (Poplack and Malvar 2007: 121).

     We turn now to the priming effect in the multivariate analysis.

5 Increasing unithood (loss of analyzability): Evidence from priming

Priming (perseveration or persistence) is the repetition of a preceding form or structure. In their study of constraints on the English agentless passive (*The liquor closet was broken into* vs. *They broke into the liquor closet*) in spontaneous discourse, Weiner and Labov (1983: 52) found that the passive variant was most strongly favored by a preceding passive. Priming, defined by Bock and Griffin (2000: 177) as "the unintentional and pragmatically unmotivated tendency to repeat the general syntactic pattern of an utterance" is also robust in psycholinguistic experiments.

     We can first mention that *ESTAR* + VERB-*ndo* is primed by itself. That is, when the verb of the preceding clause is a Progressive *estar* construction (any tense, any grammatical person), as in (10), the rate of the Progressive is higher than average (though Ns are low, at 2/4, 6/6), and 8/10, in the 13th–15th, 17th and 19th century data, respectively). But such self-priming on its own is unrevealing as to unithood/analyzability.

(10)   Progressive to Progressive priming
       Os   *estoy*   *leye-ndo*   *en el semblante*
       DAT.2SG be.PRS.1SG read-GER
       lo que   **está**   *asa-ndo*   *en vuestra alma...*
       that.REL be.PRS.3SG broil-GER
       'I <u>am reading</u> (PROG) on your face what **is broiling** (PROG) in your heart'
       (Amor de padre, Act IV, Scene V)

The priming effect shown in Table 4 is different. Here the question is whether non-Progressive *estar*, i.e., in other than a gerund construction, "triggers" the Progressive. Thus, we consider preceding use of non-Progressive *estar* constructions of the schematic form *ESTAR* + X, including locative (11), predicate adjective (12) and resultative (13) constructions.

Non-Progressive *estar* to Progressive *estar* priming

(11)   *no*   *sabemos*   *quién* <u>*está*</u>   <u>*dentro;*</u> **habla-ndo** **están**.
       NEG know.PRS.1PL who be.PRS.3SG inside speak-GER be.PRS.3PL
       'we don't know who <u>is inside</u>; they **are talking**'
       (Celestina, Act XIV)

(12)   *cuando* *alguno* <u>*está*</u>   <u>*mal*</u> *y*   *al paso de la muerte,*
       when   someone be.PRS.3SG ill and at step from the death
       están   *los expectantes* **roga-ndo**   *a*   *Dios:*
       be.PRS.3PL the expectants beseech-GER ACC God
       'when someone <u>is ill</u> and a few steps from death, there **are** the expectants **beseeching** God'
       (Corbacho IV, II)

(13)   <u>*están*</u>   <u>*cocidas*</u>   *con sus garbanzos, cebollas y tocino,*
       be.PRS.3PL cook.PTCP.F.PL with their garbanzos onion and bacon
       y   *la* *hora de ahora* **están**   **dicie-ndo**: *''¡comé=me! []''*
       and the hour of now be.PRS.3PL say-GER   eat.IMP=me
       'They <u>are cooked</u> with their garbanzos, onions and bacon and now **are saying** "eat me []!"'
       (Quijote II, LIX)

Tokens were coded as having an *estar* construction (as in (11)–(13)), a simple Present, or another finite verb form in the immediately preceding clause. I omitted discourse formulas (such as *digo* 'I say' or ¿qué sé yo? 'what do I know?', *lo mismo da* 'it makes no difference') but included subordinate clauses (complement, concessive, relative, temporal). Finer analysis would consider clause type,

lexical repetition, other preceding periphrastic forms, and the distance at which a preceding *ESTAR* construction has an effect.[12]

I propose that by considering whether other *estar* constructions prime *ESTAR* + VERB-*ndo* we obtain a measure of its analyzability. If *ESTAR* + VERB-*ndo* is "analyzable" – with internal structure and component parts that are recognizable as individual words (section 1), namely a finite form of *estar* and the gerund of another verb – it should also be primed by other constructions composed of *estar* and another unit. If, on the other hand, the Progressive is no longer analyzable, having become, in reanalysis parlance, a single constituent or periphrastic unit (Section 3), no such priming effect should hold.

The multivariate analysis in Table 4 shows that the Progressive is favored by a preceding non-Progressive *estar* construction in the 13th–15th and in the 17th century. This corresponds to Szmrecsanyi's (2005: 139) β-persistence. Such an effect could be taken as lexical, *estar* to *estar*, or as structural, *estar* + X to *estar* + X, priming (where X encompasses various word classes or syntactic roles). Either way, we can think of this kind of priming as based on associations between sub-units of constructions (for example, between English auxiliary *go* in the future construction and lexical verb *go* in various motion constructions) as opposed to priming based on syntactic identity of the entire unit (as when BE *going to* VERB primes itself or, here, when *ESTAR* + VERB-*ndo* primes itself).

The priming by non-Progressive *estar* constructions is as expected, if the Progressive has an analyzable internal structure. It also suggests that, since *estar* has independently increased in frequency to the detriment of copula *ser* 'be' in several constructions (Silva-Corvalán 1994: 94–95 and references therein), priming may be part of the explanation for the advancement of the Progressive in Spanish.

In the 19th century, however, which, as we saw, is when choice of the Progressive is most strongly favored by limited duration aspect (section 4.2), the priming effect is no longer significant (nor is there a discernable direction of effect). This disappearance of the earlier priming effect – *ESTAR* + X no longer triggers *ESTAR*

---

12 On the length of discourse over which priming may operate, see Labov 1994: 567, Travis 2007: 110, 128–129 and references therein. Patterns seemed to be the same when I excluded the first instance of the variable in a speaker turn (in the plays) or in a stretch of discourse attributed to a character (in the novels) as when I included such tokens, if the preceding finite verb appeared in speech directed by the interlocutor to the speaker (for example, *-¿Dónde está? –Está ... **está viajando*** 'Where is she? / She is...is travelling' (Acertar errando, Act 3, Scene XV)) or if the token is separated from the same speaker's preceding finite verb by an interlocutor's turn having no finite verb (for example, – [...] *¿Estás ya contento? – (Va a arrodillarse para besarle la mano.)* ¡Padre mío! – ¿Qué ***haces**, Eduardo?* 'Are you content now? / Dear father! (kneeling to kiss his hand) / What are you doing, Eduardo? (Amor de padre, Act 2, Scene I).

+ Verb-*ndo* – is consonant with diminished analyzability of the whole, i. e. the absorption of *estar* into a periphrastic unit. Together with bleaching of locative meaning and aspectual differentiation from the simple Present (weakened favoring by co-occurring locatives and strengthened disfavoring in extended duration situations), this result is indicative of increasing cohesion of *ESTAR* + Verb-*ndo* as a new unit.

6 Conclusion

The diachronic quantitative patterns of the Spanish Progressive are consistent with gradual loss of analyzability and inherent variability. First, the *ESTAR* + Verb-*ndo* sequence has increased in frequency, on which an abrupt reanalysis account is silent, but which is predicted by an understanding of change in constituency based on chunking that occurs with repetition (cf. Table 3). Second, the spread of the Spanish Progressive relative to the simple Present has been differential, as shown in change in the linguistic conditioning of variant choice, here in the form of shifting relative magnitude of effect, which contradicts an abrupt-reanalysis, constant-rate hypothesis (cf. Table 4).

A measure of internal structure is priming, whereby selection of a given construction is favored by previous use of a different construction with the same verb. Here, favoring of the Progressive over the simple Present when the preceding clause has a different *ESTAR* + X construction indicates a greater degree of analyzability of *ESTAR* + Verb-*ndo* in the 13th–15th and 17th than in the 19th century data, where this priming effect no longer holds (Section 5). The general hypothesis is that structural priming from related constructions is operative in earlier stages of grammaticalization. For example, priming of BE *going to* Verb by lexical *go* should weaken over time.

In "Building on empirical foundations", Labov (1982: 20) wrote:

> Change is the process of replacement, not the outcome of that process. When we study the process directly we are immediately confronted with the heterogeneous character of linguistic systems. Change implies variation; change *is* variation [italics in original]. [...] [The progress of change] is rarely represented by the categorical replacement of one form by another, but normally by changes in the relative frequencies of the variants and *changes in their environmental constraints* [my italics].

Studying change as change in the linguistic conditioning of variant choice makes gradualness tangible. It is hoped that more quantitative diachronic studies of variation, which include tests of priming effects, will elucidate the transition periods between the endpoints of change.

References

Allen, Henry J. and J. B. Greenough (1916): *Allen and Greenough's new Latin grammar for schools and colleges, founded on comparative grammar*. Boston et al.: Ginn.

Bock, Kathryn J. and Zenzi M. Griffin (2000): The persistence of structural priming: Transient activation or implicit learning. *Journal of Experimental Psychology: General* 129(2): 177–192.

Brown, Esther L. and Javier Rivas (2011): Subject-verb order in Spanish interrogatives: a quantitative analysis of Puerto Rican Spanish. *Spanish in Context* 8: 23–49.

Bybee, Joan (2003): Mechanisms of change in grammaticization: the role of frequency. In: Brian D. Joseph and Richard D. Janda (eds.), *The handbook of historical linguistics*, 602–623. Oxford: Blackwell.

Bybee, Joan (2010): *Language, usage and cognition*. Cambridge: Cambridge University Press.

Bybee, Joan, Revere Perkins and William Pagliuca (1994): *The evolution of grammar: Tense, aspect, and modality in the languages of the world*. Chicago: University of Chicago Press.

Bybee, Joan and Rena Torres Cacoullos (2009): The role of prefabs in grammaticization: How the particular and the general interact in language change. In: Roberta L. Corrigan, Edith A. Moravcsik, Hamid Ouali and Kathleen Wheatley (eds.), *Formulaic language: Volume 1. Distribution and historical change*, 187–217. Amsterdam: John Benjamins.

Campbell, Lyle (1998): *Historical linguistics: an introduction*. Cambridge, MA: MIT Press.

Cedergren, Henrietta J. and David Sankoff (1974): Variable rules: performance as a statistical reflection of competence. *Language* 50(2): 333–355.

Comrie, Bernard (1976): *Aspect*. Cambridge: Cambridge University Press.

Croft, William and Alan D. Cruse (2004): *Cognitive linguistics*. Cambridge: Cambridge University Press.

Givón, Talmy (1975): Serial verbs and syntactic change: Niger-Congo. In: Charles N. Li (ed.), *Word order and word order change*, 47–112. Austin: University of Texas Press.

Hay, Jennifer (2001): Lexical frequency in morphology: is everything relative? *Linguistics* 39: 1041–1070.

Hopper, Paul J. (1991): On some principles of grammaticization. In: Elizabeth Closs Traugott and Bernd Heine (eds.), *Approaches to grammaticalization* (Volume 1), 17–35. Amsterdam: John Benjamins.

Hopper, Paul J. and Elizabeth Closs Traugott (1993): *Grammaticalization*. Cambridge: Cambridge University Press.

Kroch, Anthony (1989): Reflexes of grammar in patterns of language change. *Language Variation and Change* 1: 199–244.

Labov, William (1969): Contraction, deletion, and inherent variability of the English copula. *Language* 45: 715–762.

Labov, William (1982): Building on empirical foundations. In: Winfred P. Lehmann and Yakov Malkiel (eds.), *Perspectives on historical linguistics,* 11–92. Amsterdam: John Benjamins.

Labov, William (1994): *Principles of linguistic change: Internal factors* (Volume 1). Oxford: Blackwell.

Langacker, Ronald (1987): *Foundations of cognitive grammar: theoretical prerequisites* (Volume 1). Stanford, CA: Stanford University Press.

Marcos Marín, Francisco (dir.) (1992): Corpus de Referencia de la Lengua Española Contemporánea Peninsular (CORLEC), http://www.lllf.uam.es/ING/Info%20Corlec.html (accessed April 2004).

Myhill, John (1988): The grammaticalization of auxiliaries: Spanish clitic climbing. *Berkeley Linguistics Society* 14: 352–363.

Newmeyer, Frederick (2003): Grammar is grammar and usage is usage. *Language* 79(4): 682–707.

Poplack, Shana and Elisabete Malvar (2007): Elucidating the transition period in linguistic change: The expression of the future in Brazilian Portuguese. *Probus* 19: 121–169.

Poplack, Shana and Sali Tagliamonte (1999): The grammaticization of *going to* in (African American) English. *Language Variation and Change* 11: 315–342.

Poplack, Shana and Sali Tagliamonte (2001): *African American English in the diaspora: tense and aspect.* Oxford: Blackwell.

Rizzi, Luigi (1982): A restructuring rule in Italian syntax. In: Luigi Rizzi (ed.), *Issues in Italian syntax*, 1–48. Dordrecht: Foris.

Roberts, Ian and Anna Roussou (2003): *Syntactic change: a minimalist approach to grammaticalization*. Cambridge: Cambridge University Press.

Sankoff, David (1988a): Sociolinguistics and syntactic variation. In: Frederick J. Newmeyer (ed.), *Linguistics: The Cambridge survey* (Volume IV), 140–161. Cambridge: Cambridge University Press.

Sankoff, David (1988b): Variable rules. In: Ulrich Ammon, Norbert Dittmar and Klaus J. Mattheier (eds.), *Sociolinguistics: An international handbook of the science of language and society*, 984–997. Berlin/New York: Walter de Gruyter.

Sankoff, David, Sali Tagliamonte and Eric Smith (2005): GOLDVARB X: A multivariate analysis application for Macintosh and Windows. <http://individual.utoronto.ca/tagliamonte/ Goldvarb/GV_index.htm>.

Scheibman, Joanne (2000): I dunno but… a usage-based account of the phonological reduction of *don't. Journal of Pragmatics* 32: 105–124.

Silva-Corvalán, Carmen (1994): *Language contact and change. Spanish in Los Angeles.* Oxford: Clarendon Press.

Szmrecsanyi, Benedikt (2005): Language users as creatures of habit: A corpus-based analysis of persistence in spoken English. *Corpus Linguistics and Linguistic Theory* 1(1): 113–150.

Torres Cacoullos, Rena (1999a): Variation and grammaticization in progressives: Spanish *-ndo* constructions. *Studies in Language* 23–1: 25–59.

Torres Cacoullos, Rena (1999b): Construction frequency and reductive change: diachronic and register variation in Spanish clitic climbing. *Language Variation and Change* 11(2): 143–170.

Torres Cacoullos, Rena (2000): *Grammaticization, synchronic variation, and language contact: a study of Spanish progressive -ndo constructions*. Amsterdam: John Benjamins.

Torres Cacoullos, Rena (2006): Relative frequency in the grammaticization of collocations: nominal to concessive *a pesar de*. In: Timothy L. Face and Carol A. Klee (eds.), *Selected proceedings of the 8th Hispanic Linguistics Symposium*, 37–49. Somerville, MA: Cascadilla Proceedings Project.

Torres Cacoullos, Rena (2012): Grammaticalization through inherent variability:

The development of a progressive in Spanish. *Studies in Language* 36(1): 73–122.

Torres Cacoullos, Rena and James A. Walker (2011): Collocations in grammaticalization and variation. In: Bernd Heine and Heiko Narrog (eds.), *Handbook of Grammaticalization*, 225–238. Oxford: Oxford University Press.

Travis, Catherine E. (2007): Genre effects on subject expression in Spanish: Priming in narrative and conversation. *Language Variation and Change* 19(2): 101–135.

UNAM (1976): *El habla popular de la Ciudad de México: Materiales para su estudio*. Mexico City: Centro de Lingüística Hispánica.

Weiner, Judith E. and William Labov (1983): Constraints on the agentless passive. *Journal of Linguistics* 19(1): 29–58.

Weinreich, Uriel, William Labov and Marvin Herzog (1968): Empirical foundations for a theory of language change. In: Winfred P. Lehmann and Yakov Malkiel (eds.), *Directions for historical linguistics*, 95–188. Austin: University of Texas Press.

Yang, Charles D. (2000): Internal and external forces in language change. *Language Variation and Change* 12(3): 231–250.

Appendix: Texts

Author/Title Listed in chronological order	Source/Edition Unless otherwise indicated, electronic texts were downloaded from Biblioteca Virtual Miguel de Cervantes, http://www.cervantesvirtual.com/
Anonymous, *Cantar de mio Cid*	*Texto, gramática y vocabulario*, Ramón Menéndez Pidal (ed.), Vol. 3. Madrid: Espasa Calpe, 1944.
Anonymous, *Calila e Dimna*	Juan Manuel Cacho Blecua & María Jesús Lacarra (eds.). Madrid: Castalia, 1984.
Anonymous, *Poema de Fernán González*	
Anonymous, *Libro de Apolonio*	
Gonzalo de Berceo, *Sacrificio de la misa*	
Alfonso X, *General estoria. Primera parte*	Antonio G. Solalinde (ed.). Madrid: Centro de Estudios Históricos, 1930
Alfonso X, *General Estoria. Segunda parte*	Antonio Solalinde, Lloyd Kasten & Victor R. B. Oelshläger (eds.). Madrid: CSIC (Consejo Superior de Investigaciones Científicas), 1957
Alfonso X, *Estoria de España* (EE1, EE2); *Libro de las leyes*	*The Electronic Texts and Concordances of the Prose Works of Alfonso X, El Sabio*, Lloyd Kasten, John Nitti, and Wilhemina Jonxis-Henkemens (eds.). Hispanic Seminary of Medieval Studies, 1995
Anonymous, *El libro del caballero Zifar*	
Arcipreste de Hita, *Libro de Buen Amor*	
Don Juan Manuel, *El conde Lucanor*	José Manuel Blecua (ed.). Madrid: Castalia, 1971
Alfonso Martínez de Toledo, *Corbacho*	Michael Gerli (ed.). Madrid: Cátedra, 1978
Juan de Flores, Juan, *Grimalte y Gradissa*	Pamela Waley (ed.). London: Tamesis, 1971

Author/Title Listed in chronological order	Source/Edition Unless otherwise indicated, electronic texts were downloaded from Biblioteca Virtual Miguel de Cervantes, http://www.cervantesvirtual.com/
Hernando del Pulgar, *Crónica de los Reyes Católicos*	Juan de Mata Carriazo (ed.). Madrid: Espasa Calpe, 1943
Fernando de Rojas, Fernando, *La Celestina*	Dorothy S. Severin (ed.), Madrid: Cátedra, 1987
Íñigo López de Mendoza, marqués de Santillana, Sonetos	
Juan del Enzina [14 Auctos, Eglogas, Representaciones]	
Miguel de Cervantes Saavedra, *Don Quijote de la Mancha*; *Comedia famosa de La casa de los celos y selvas de Ardenia*	
Lope de Vega, *La dama boba*; *Comedia del Príncipe Ynocente*; *La vengadora de las mujeres*	
Guillén de Castro, *El amor constante*	
Ruiz de Alarcón y Mendoza, *La Amistad castigada*	
Tirso de Molina, *Don Gil de las calzas verdes*; *Por el sótano y el torno*; *Amor y celos hacen discretos* [Act I]; *La villana de Sagra*	
Gaspar de Ávila, *La boca y no el corazón o Fingir por conserver*	
Calderón de la Barca, *Casa con dos puertas mala es de guarder*; *La dama duende*; *Amor honor y poder*	
Leandro Fernández de Moratín, *La comedia nueva*; *El sí de las niñas*	J. Dowling & R. Andioc (eds.). Madrid: Castalia, 1975
Manuel Bretón de los Herreros, *A Madrid me vuelvo*	
José María de Carnerero, *El afán de figurar*	
Ventura de la Vega, *Acertar errando, o El cambio de diligencia*	
José de Mariano Larra, *Los inseparables*	
Duque de Rivas, *Don Álvaro o la fuerza del sino*	
Francisco Martínez de la Rosa, *La boda y el duelo*; *Amor de padre*	
Pablo Alonso de la Avecilla, *Los presupuestos*	
Luis Mariano de Larra, *La paloma y los halcones*; *La primera piedra*	
Manuel Tamayo y Baus. *No hay mal que por bien no venga*	
Juan Valera, *Pepita Jiménez*	

Author/Title Listed in chronological order	Source/Edition Unless otherwise indicated, electronic texts were downloaded from Biblioteca Virtual Miguel de Cervantes, http://www. cervantesvirtual.com/

Benito Pérez Galdós, *Doña Perfecta*

Leopoldo Alas "Clarín", *La Regenta*, vol I & II

Pardo Bazán, Emilia. *Los pazos de Ulloa*

Julio Cuevas, *¡El siete!: juguete cómico en un acto y en prosa* /original de Don Julio Cuevas y Don Manuel Labra

Enrique Gaspar, *Don Ramón y El señor Ramón*; *El estómago*; *La gran comedia*; *La lengua*; *La levita*; *Las circunstancias*; *Lola*; *Los niños grandes*; *Problema*; *Serafina la devota*