

John Benjamins Publishing Company



This is a contribution from *Studies in Language* 36:1
© 2012. John Benjamins Publishing Company

This electronic file may not be altered in any way.

The author(s) of this article is/are permitted to use this PDF file to generate printed copies to be used by way of offprints, for their personal use only.

Permission is granted by the publishers to post this file on a closed server which is accessible to members (students and staff) only of the author's/s' institute, it is not permitted to post this PDF on the open internet.

For any other use of this material prior written permission should be obtained from the publishers or through the Copyright Clearance Center (for USA: www.copyright.com).

Please contact rights@benjamins.nl or consult our website: www.benjamins.com

Tables of Contents, abstracts and guidelines are available at www.benjamins.com

Grammaticalization through inherent variability

The development of a progressive in Spanish

Rena Torres Cacoullous

Penn State University

With the goal of elucidating the diachronic trajectory of a progressive, multivariate analysis is used to track the linguistic factors conditioning variation between the Spanish Progressive and the simple Present, in 13th–15th, 17th, and 19th century texts. The Progressive begins as more of a locative construction, as shown by the early favoring effect of co-occurring locatives. The direction of this co-occurring locative effect is retained over time, but the magnitude weakens relative to aspectual constraints (limited vs. extended duration contexts, dynamic vs. stative verbs), and the Progressive is increasingly disfavored in negatives and interrogatives. Increasing frequency is accompanied by changes in linguistic conditioning. An aspectual opposition arises as, in the course of speakers' recurrent choices between variant forms, the variants develop functional differentiation.

Keywords: variation, grammaticalization, progressive, linguistic conditioning, frequency

1. Inherent variability and grammaticalization

Grammaticalization has remained a profitable research arena since attracting attention some thirty years ago (e.g., Givón 1979; Heine & Reh 1984; Lehmann 1982; Sankoff & Brown 1976). The methods adopted by scholars have ranged from citation of example sentences from historical sources (e.g., Fleischman 1982), to typological surveys of reference grammars (Bybee, Perkins and Pagliuca 1994) or questionnaires (Dahl 1985) and, more recently, statistical analyses of diachronic corpora. At the same time, the probabilistic structure of grammar (Labov 1969) has been newly recognized by more linguists.

Inherent variability in speech is manifested in form-function asymmetry: variation among different forms serving generally similar discourse functions (Labov 1969, 1972; Sankoff & Thibault 1981; Sankoff 1988a). An example is the expression of present progressive meaning by two variants in Spanish illustrated in (1). The simple Present, in the question (1a), and the Progressive Present — *estar* ‘to be (located)’ + Verb-*ndo* (gerund) — in the answer (1b) both express a situation ongoing at the moment of speech.¹ (In the examples, simple and Progressive forms are bolded; the abbreviations PRS and PROG appear in the English translations.)

- (1) a. *dice Moncho: “¿Pero qué **haces?**”*
 says Moncho but what do.PR.S.2SG
 b. “***Estoy ventilando**, digo;*
 be.PR.S.1SG ventilate-GER I say
 a. ‘Moncho says: “But what **are** you **doing** (PRS)?”’
 b. “I **am ventilating** (PROG),” I say;’
 (Marcos Marín 1992, COREC, ACON017A)

However, example (2) from the same corpus of spoken Peninsular (= Spain) Spanish serves to illustrate the aspectual difference between the Progressive and simple Present: the Progressive (2a) has progressive meaning, (‘[now] he is sleeping’), while here the Present (2c) expresses habitual aspect (‘he [customarily] goes to sleep’).

- (2) a. *Pablo se sobrealimenta también. **Está durmie-ndo.***
 Pablo REFL eat_excessively.PR.S.1SG also be.PR.S.3SG sleep-GER
 b. [...] *¿**Está durmiendo** Pablo?!*
 be.PR.S.3SG sleep-GER Pablo
 c. *Sí, Pablo después de come-r se **duerme** [...]*
 yes Pablo after of eat-INF REFL sleep.PR.S.3SG
 d. *¡Que son la siete menos 10!*
 a. ‘Pablo eats too much too. He **is sleeping** (PROG).
 b. [...] Pablo is sleeping?!
 c. Yes, after eating Pablo goes **to sleep** (PRS) [...]
 d. But it’s 10 to seven!’ (Marcos Marín 1992, COREC, ACON023A)

The distinction between progressive and nonprogressive forms is said to be obligatory in some languages. For example, in English, present progressive is overtly marked by Progressive *be* + Verb-*ing*, while the zero-marked Present signals present habitual (*I drink decaf*) (Bybee 1994: 238–239). In Spanish, in contrast, “it is normally possible to replace the Progressive by other forms, without implying nonprogressive meaning, so that corresponding to English *John is singing*

Spanish may have either *Juan está cantando* (Progressive) or *Juan canta* [simple Present]...” (Comrie 1976:33; but see Blansitt 1975:5–6). Other language varieties, including Old English and Latin, lack progressive morphemes altogether.

How do grammars get progressives? In usage-based theory, grammatical morphemes arise through grammaticalization (Bybee 2010, Ch. 6), the set of processes by which a lexically-particular instance of a construction becomes a new grammatical construction. Grammaticalization is said to be initiated by speakers’ choices of new ways of “saying approximately the same thing” (Hopper and Traugott 1993:65). But how do we get from saying the same thing to obligatorily marked grammatical categories such as progressive? In studies employing quantitative argumentation, the standard gauge of change has been shifts in rates: the token frequency of the grammaticalizing construction and/or its frequency relative to an older variant.

Here we examine the relationship between the evolution of a grammaticalizing construction and its patterns of variation with respect to the older morpho-syntactic expression(s) with which we observe that it alternates. The question can then be reformulated as: is increased frequency of a grammaticalizing construction — absolute or relative — accompanied by changes in the patterns of variation with existing expressions? That is, as the grammaticalizing construction increases in frequency, does it generalize proportionally in all contexts? If not, what form does differential spread take?

Using multivariate analysis, we track linguistic factors conditioning variation between the Spanish Progressive and the simple Present. We find that while these forms have been variants in the functional domain of present temporal reference from the beginning, the aspectual opposition between the two has developed over time. The Progressive begins as more of a locative construction, as shown by strong effects of co-occurring locatives in the earliest period. Over time, locative meaning weakens and aspectual meaning (limited vs. extended duration) strengthens. With this, the Progressive is increasingly disfavored with stative predicates and in negative polarity clauses and interrogatives. The changing relative magnitude of effect of the factor groups is a measure of the gradual development of a progressive-nonprogressive opposition. Thus, increasing *rate* of use of a grammaticalizing construction is accompanied by change in its linguistic *conditioning* — the configuration of linguistic constraints on variation with existing variants.

After presenting the corpus of texts, Section 3 shows frequency increases and change in the constituent structure of the Progressive. In Section 4 the envelope of Progressive — simple Present variation is delimited. Section 5 justifies the coding of tokens for a number of factor groups (independent variables), which operationalize hypotheses about grammaticalization and the choice of variants, based on contextual features. Section 6 compares the linguistic conditioning of

the Progressive across time periods through independent multivariate analyses, providing evidence for locative origins and gradual development as an aspectual expression of limited duration. Section 7 shows the rise of a stativity restriction, with separate analyses for stative and dynamic verbs. In Section 8, analysis of the variation within limited duration contexts suggests that negatives and interrogatives may be conservative contexts. In Section 9, I conclude that inherent variability feeds grammaticalization, like other change, which advances via neutralization of functional distinctions in discourse (Sankoff 1988a). In the case of Progressive and simple Present variation, in the course of speakers' recurrent choices between functionally overlapping variant forms, the variants themselves evolve, becoming functionally (here, aspectually) more distinct.

2. Data

The data are from three time periods, the 13th–15th, 17th, and 19th centuries. These correspond to a traditional periodization distinguishing Old Spanish, Golden Age Spanish, and Modern Spanish (but see Company Company 2006: xxiv–xxvii). The stretch of time for each data set is broad, because tokens of the Progressive are not abundant. The approximate word count is 2.5 million for the Old Spanish corpus, 600,000 for the 17th century, and 900,000 for the 19th century.²

The Old Spanish corpus comprises seventeen (sets of) texts spanning the 13th–15th centuries, in prose and verse, of different genres — chronicles (*Estoria de España*, *General Estoria*, *Crónica de los Reyes Católicos*), drama (*Celestina*, works by Juan del Enzina), a sermon (*Corbacho*), didactic stories (*Calila e Dimna*, *Conde Lucanor*), epic poems (*Cid*, *Poema de Fernán González*), and chivalric and romantic novels (*Caballero Zifar*, *Grimalte y Gradissa*). The 17th century corpus is made up of the entire *Quijote* and fourteen plays in verse by six different authors (Cervantes, Lope de Vega, Guillén de Castro, Gaspar de Ávila, Ruiz de Alarcón y Mendoza, Tirso de Molina, and Calderón de la Barca). The 19th century corpus has twenty-five plays and four novels (*Pepita Jiménez*, *Doña Perfecta*, *La Regenta*, *Los Pazos de Ulloa*), by eighteen authors, all in prose. Tokens extracted from plays make up 34% (254/745), 57% (574/1013), and 41% (592/1460) of the Old Spanish, 17th c. and 19th c. data, respectively. The texts are listed chronologically under Corpus, before References.

3. Token frequency and constituency change of the Progressive

In Old Spanish, besides *estar* ‘be (located)’, other verbs combining with a gerund, whose suffix is *-ndo*, were *ir* ‘go’, *andar* ‘go around’, *venir* ‘come’, *salir* ‘go out’, *seer* ‘be’, *quedar* ‘remain, stand still’, as in (3). Thus we may view *estar* + Verb-*ndo* as a particular instance of a general gerund construction in which finite forms of spatial (locational, postural, or movement) verbs took a gerund complement to mean ‘be at/move while Verb-*ing*’ (Bybee and Torres Cacoullós 2009: 199).

- (3) a. *Yua=sse anda-ndo por la carrera que ua al pozo*
 go.IP.FV.3SG=REFL walk-GER by the road that goes to.ART.SG well
 ‘He went walking along the road that goes to the well.’ (GEI, fol. 66v)
 b. [locative-postural-movement]_{verb} + [-*ndo* (gerund)]_{complement}

Table 1 shows increasing token frequencies of Present *estar* + Verb-*ndo* over five time periods. In the earliest Spanish texts of the 13th century, with close to two million words, raw token frequency is 42 and normalized frequency (per 100,000 words) a mere 2. This rises to 10 in 14th–15th century texts. After the Old Spanish period, frequency triples to 30 in 17th and 36 in 19th century texts.

Rates of the Progressive are sensitive to genre or perhaps other extra-linguistic factors (region, education level), as indicated by the disparate 20th century counts (from Clegg n.d. and Cortes-Torres 2005). Among our 19th century texts, we find that the normalized frequency of the Present Progressive is more than twice as high in the plays, at 54 (word count 383,000; N=207), than in the novels, at 22 (word count 507,000; N=110). This difference may be due in part to the distribution of the genres across aspectual contexts. In the plays, of all tokens coded for aspect (both Progressive and simple Present, see 5.4 ahead), more than two-thirds

Table 1. Increasing frequency of Progressive (Present *estar* + Verb-*ndo*)

Century	13th	14th– 15th	17th	19th	20th novels (Clegg n.d.)	20th interviews (Clegg n.d.)	20th inter- views (Cortes 2005)
Word count (thou- sands)	1,838	739	606	890	1,327	1,292	327
N	42	77	180	317	613	1,053	605
<i>estar</i> _{Present} + Verb- <i>ndo</i>							
Frequency per 100k words	2	10	30	36	46	81	185

* For 13th — 19th c. data, see Corpus. 20th c. Clegg (n.d.) ‘interviews’ are from *Habla culta* ‘educated speech’ corpora of several major cities (cf. Lope Blanch 1971); Cortes-Torres (2005) recorded sociolinguistic interviews in Puerto Rico.

(69%, 403/587) occur in limited duration contexts, whereas in the novels a lower proportion does (58%, 244/419). There may also be a genuine genre effect on the rate of the Progressive, since within limited duration contexts the frequency of the Progressive relative to the simple Present is 44% (179/403) in the plays, as opposed to 37% (90/244) in the novels (not quite statistically significant according to a chi-square test, $p=0.07$). Nevertheless, when we compare normalized token frequency within the plays, we still find an increase: from 33 in the 15th c. (word count 99,000; $N=33$) to 57 (word count 222,000; $N=126$) in the 17th and 54 in the 19th century.

As its frequency of use increases, there is change in the constituent structure of *estar* + Verb-*ndo*. The change is from a sequence of two independent lexical items — main verb *estar* ‘be (located)’ and a gerund complement, as in the general gerund construction depicted in (3b) above — to a periphrastic unit — auxiliary *estar* and a gerund as the main verb. This is the new construction depicted in (4), the Progressive (denoted by capitalization).

(4) [*Estar* + Verb-*ndo*]_{Progressive}

Empirical studies have challenged the view that structural reanalysis — the creation of new grammatical structures — is abrupt (cf. Torres Cacoullós 2006; Bybee 2010: 136–150). Evidence that change in the constituent structure of *estar* + Verb-*ndo* is gradual comes from a steady decline of intervening material (for additional indices of grammaticalization, see Torres Cacoullós 2000: 31–55, Bybee and Torres Cacoullós 2009: 201–203). In all time periods, we find examples with material intervening between *estar* and the gerund, as in (5). Intervening elements are locatives (5a), temporal or manner expressions, subjects (5b), objects, and predicate adjectives (5c) (see also examples 10b, 12, 21a, 32). Were we to discount tokens such as these as locative or predicate adjective or some other kind of *estar* construction fortuitously co-occurring with a gerund rather than as ‘true’ cases of *estar* + Verb-*ndo*, we would miss the change.

- (5) a. *que estan a las muelas molie-ndo*
 who be.PRS.3PL to the mills mill-GER
 ‘who are at the mills milling’ (13th c., GEI, fol. 155v)
- b. *les está una dueña hacié-ndo-les la guarda*
 DAT.3PL be.PRS.3SG a duenna make-GER-DAT.3PL the guard
 ‘a duenna is acting as their lookout’ (17th c., Quijote I, Ch. 32)
- c. *¿Estoy despierto o soña-ndo?*
 be.PRS.1SG awake.ADJ.M or dream-GER
 ‘Am I awake or dreaming?’ (19th c., Boda, Act 3, Scene I)

Table 2. Changing constituent structure of Progressive: decline of intervening material in Present *estar* + Verb-*ndo*

Century	13th	14th–15th	17th	19th	20th**
N*	40	70	139	310	605
Adjacent (nothing intervening)	63% (25)	67%(47)	83%(115)	84%(259)	95%(577)
Intervening locative	10% (4)	11% (8)	4% (6)	3% (9)	1% (5)
Intervening other***	28% (11)	21% (15)	13% (18)	14% (42)	4% (23)

*Tokens with gerund preposed (Verb-*ndo* + *estar* order) are excluded: 2 in 13th c., 7 in 14th–15th c., 41 in 17th c. (verse), 7 in 19th c.

** 20th c. speech from Cortes-Torres (2005).

***Intervening material, “other”: 13th–15th c.: subject (14); 17th c.: object (8); 19th c.: subject pronoun (19).

Table 2 shows the proportion of tokens in which Present-tense forms of *estar* and a gerund are adjacent. In the Old Spanish data (13th–15th centuries), adjacent occurrences make up approximately two thirds of all *estar* + Verb-*ndo* tokens, or conversely, occurrences with intervening material (as in (5)) make up approximately one third. We observe a diachronic increase in adjacency, from 63% of all tokens in the 13th century, to over 80% in the 17th and 19th c. data, to 95% in present-day sociolinguistic interviews (Cortes-Torres 2005:47).³ This increasing adjacency correlates with increasing token frequency, as we may corroborate using Kendall’s tau (cf. Hilpert and Gries 2009:390), which has a value of 1 ($p=0.03$) for the proportions of adjacent tokens and the normalized frequencies in the five time periods of Table 1. Particularly sharp is the decline in the proportion of tokens with an intervening locative, displayed in the third row: from one-tenth of all tokens in Old Spanish, to 3–4% in 17th c. and 19th c. data, to virtually absent in present-day sociolinguistic interviews (Cortes-Torres 2005:47).

In summary, *estar* + Verb-*ndo* has increased in frequency and, by the diagnostic of adjacency, has increasingly become a unit. It is thus a good case of grammaticalization, which Bybee (2010:106–107) characterizes as the process by which a specific instance of a construction (*estar* in the gerund construction) gains in frequency and becomes a new autonomous — less analyzable — construction. Since *estar* + Verb-*ndo* has extremely low token frequency in the earliest texts, it is an opportune case in which we are able to track grammaticalization practically from its initiation.

4. The variable context

4.1 The neutralization in discourse hypothesis and the linguistic variable

Early variation is illustrated in the pair of 15th c. examples in (6), where the Progressive *está devaneando* ‘is raving’ in (6a) and the simple Present *devanea* ‘raves’ in (6b) both express a situation in progress at speech time.

- (6) a. *Está devanea-ndo entre sueños.*
 be.PRS.3SG rave-GER between dreams
 ‘He is raving (PROG) in his sleep.’ (Celestina, Act VIII)
- b. *Hijo, déxa=la dezir, que devanea;*
 ‘Son, let.IMP=her talk that rave.PRS.3SG
 ‘Son, let her talk, she is raving (PRS).’ (Celestina, Act IX)

Inherent (structural) variability has been confronted in the cumulative variationist study of language over the last five decades. “The variationist viewpoint on language is determined first by a scientific interest in accounting for grammatical structure *in discourse* [...] and second by a preoccupation with the polyvalence and apparent instability *in discourse* of linguistic form-function relationships” (Sankoff 1988a: 141, italics in original). The working hypothesis is that of neutralization in discourse: while contexts can almost always be found in which different morphosyntactic expressions have different meanings, there are alternations in which the full accompaniment of meaning distinctions is not pertinent either for the speaker or the interlocutor (Sankoff 1988a: 153). Since we have no direct access to speaker intentions, the contrary view — that different forms always imply meaning differences, however subtle — must also remain a hypothesis. Sankoff (1988a: 153) proposed that neutralization of functional distinctions between morphosyntactic expressions in discourse is the “fundamental discursive mechanism of (nonphonological) variation and change”. In a cognitive linguistics framework, a compatible idea is Croft’s (2010: 42) proposal that language change is possible because of “indeterminacy in verbalizing human experience”.

The theoretical notion of the variationist framework is the *linguistic variable*, a set of *variants* between which speakers alternate in expressing a grammatical function (Labov 1969). From this follows the methodological principle of accountable reporting, which requires that we count not only occurrences, but also non-occurrences (Labov 1972: 72). In the case at hand, this means accounting for tokens of the Progressive as well as for cases where it could have occurred but the simple Present did instead, as in (6b) above.

4.2 Non-occurrences: Extracting a sample of the more frequent variant

Progressive tokens were exhaustively extracted. Given the extremely low frequency of the Progressive, especially in the earlier texts, we consider only a sample of the much more frequent variant.⁴ The simple Present sample was taken by extracting two tokens in the vicinity of each Progressive token, usually the one immediately preceding and following (cf. Harvie 1998), and all Present occurrences of lexical types appearing in the Progressive in a given text.⁵

The variable context, the broadest domain in which speakers have a choice between these two forms, is defined here as present temporal reference (cf. Walker 2001: 14–16). Thus, excluded from the analysis of variation were tokens of the simple Present with future (7a) or past (7b) reference as well as in modal periphrases with *deber* (*de*) ‘should’, *haber de* ‘ought to’, *necesitar* ‘need to’, *poder* ‘can’, *tener* (*de, que*) ‘have to’ (7c). Also excluded was quoted material (8a) and proverbs (8b), which may have frozen expressions.

- (7) a. *verá usted cómo todo se dispone*
‘you’ll see how everything is arranged’ (CN, Act II, Scene II)
- b. *y estaba [...] para montar a caballo [...], cuando oigo ¡tras tris, tras tras!*
‘and I was [...] about to get on the horse, when I hear tras tris, tras tras!’
(Pazos, Ch. XXI)
- c. *no lo puedo llevar el estar aquí sepultado en vida*
‘I can’t bear being buried alive here’ (Quijote II, Ch. 55)
- (8) a. *el cántico que comienza: «Vosotros, cielos, oíd agora lo que yo fablo»*
‘the hymn that begins: “You, skies, hear now what I say”’
(Corbacho, Part IV, I, p. 248)
- b. *ojos que no veen, corazón que no quiebra;*
‘eyes that don’t see, heart that doesn’t break’ (Quijote II, Ch. 67)

A thorny task was excluding collocations with high frequency verbs *decir* ‘say’, *ver* ‘see’, *saber* ‘know’, and *creer* ‘believe’, such as *creo* (*que*) ‘I think (that)’, (*como*) *dicen* / *se dice* ‘they say’ to introduce a proverb, Reflexive + *ver* + predicative ‘find oneself X’ (e.g., *me veo desolado* ‘I find myself desolate’). Most are first or second person discourse routines, as listed in (9a–9c). Also excluded are *ser* ‘be’ constructions such as *es que* ‘it’s that’ (9d). Other particular constructions that were set aside are *se* + Dative + *entender* (e.g., *se me entiende* ‘I am understood’), *esperar* (+ infinitive or finite clause) ‘hope, expect’, *hacer bien/mal* ‘be (un)justified in’ (e.g. *hago mal en no ir* ‘it’s wrong of me not to go’), *hace* + time expressions (e.g., *hace cuatro días* ‘four days ago’). These appear invariably with the simple Present.

- (9) a. *¿qué dices?* ‘what are you saying?’;⁶ parenthetical *digo (que)* ‘I say’; *lo que (yo) digo* ‘what I say’; *como digo* ‘as I say’
 b. *ya (lo) ve(s)* ‘you see’; *no ve(s)* ‘don’t you see’; *ve(s) (aquí/como)* ‘you see (here/how), behold’; *bien veo, ya veo* ‘(now) I see’; *qué veo* ‘what do I see’
 c. *ya sé que* ‘I know’, *sabe(s) que* ‘you know’, *¿no sabes?* ‘don’t you know?’
 d. *(si) es que* ‘it’s that’, *es porque* ‘it’s because’, *es decir* ‘that is to say’, *es imposible* ‘it’s impossible’, *es verdad que, lo cierto es* ‘it’s true that’, *es (eso) lo que* ‘that is what’, *lo que es lo mismo* ‘which is the same as’

4.3 Distribution of the variants in the sample

These protocols yielded 745 Old Spanish, 1013 17th century, and 1460 19th century tokens, with frequencies of the Progressive relative to the simple Present at 16%, 18%, and 22% (Table 3).

Table 3. Rate (artificial) of the Progressive with respect to the simple Present in the three data sets (N = 3218)

13th–15th c. (Old Spanish)	17th c.	19th c.
16% (119/745)	18% (180/1013)	22% (317/1460)

The rate of the Progressive is similar in the three time periods as an artifact of the simple Present sampling.⁷ Nevertheless, while these overall rates are artificial, we can compare the frequency of the variants in sets of linguistic sub-contexts, the hypothesized *constraints* on the variable, to obtain a multivariate model of contextual effects on speakers’ choices (Labov 2004).

5. Hypotheses about the grammaticalization of progressives and coding of tokens

To determine the patterns of speakers’ choice of variant (against the null hypothesis that the variation is random, or “free”), we code tokens for contextual features. These contextual features are operationalized as factors, to test hypotheses about constraints on variant choice. For each factor (particular linguistic environment), we predict an increase — favoring effect — or decrease — disfavoring effect — in the relative frequency of the Progressive with respect to the simple Present. The factors compose factor groups (predictor, or independent, variables). The factor groups considered here are co-occurrence of locative and temporal expressions, lexical stativity, aspect, and polarity-sentence type.

- b. *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 there with the mounts’ (Perfecta, Ch. I)

How do locatives become progressives? It has been proposed that progressive aspect is an implication of the locative construction; thus the mechanism of change need not involve space > time metaphor (as has been proposed e.g. in Claudi & Heine 1986) (Bybee et al. 1994: 137). In example (10b) above, the speaker’s activity is given together with his location. The overlap of temporal and locative meaning is also illustrated in (12), in which the speaker is reassuring a skittish lover: the location of the ducks implies their ongoing noisy activity. This is also a handy example of variation, in that the Progressive is juxtaposed to a series of ongoing activities expressed in the simple Present.

- (12) *Dice ella: «¡Yuy, amigo, no hayáis miedo, que el gato es que huyó desde que vos vio!»,*
o «La gallina es que tiene pepita y hace ruido»,
 or the chicken is that has pip and make.PRS.3SG noise
o «La mula es que come cebada y hace ruido»,
 or the mule is that eat.PRS.3SG barley and make.PRS.3SG noise
o «Dos anadones son que están en aquel corral chapulla-ndo»,
 or two ducklings are that be.PRS.3PL in that pond splash-GER
o «Mi señora la vieja es que tose»,
 or my lady the old_one is that cough.PRS.3SG

‘She says: “Oh my friend don’t be afraid, it’s the cat that ran off when it saw you!” or “It’s the chicken that has a pip and is **making noise**” (PRS), or “It’s the mule that is **eating** (PRS) barley and is **making noise**” (PRS), or “It’s two ducklings that **are** in that pond **splashing**” (PROG), or “It’s my lady the old one who is **coughing**” (PRS)’ (Corbacho, Part III, IX, p. 223)

An indication of the association between locative and aspectual meaning is the proportion of tokens with a co-occurring locative that have a meaning of limited temporal duration (more on coding for aspect in Section 5.4 below). In the Old Spanish data, the proportion of tokens (both Progressive and simple Present) with limited duration meaning is greater with a co-occurring locative (66%, 25/38) than without (42%, 191/455). The association between locative and limited duration aspectual meaning is especially strong for the Progressive: 25% (16/62) of Progressive tokens in limited duration contexts occur with a locative, whereas only 6% (9/154) of simple Present tokens do.

The grammaticalization of progressives from locative sources has been hypothesized to proceed via gradual loss of locative meaning, in a type of change

known as semantic bleaching or erosion (Lehman 1982, Heine and Reh 1984). As a construction increases in frequency, there is generalization of the contexts in which it is used; bleaching of locative meaning allows the expression of temporal meaning in more contexts (Bybee et al. 1994: 137). From this hypothesis follows the prediction that the favoring effect of co-occurring locatives on *estar* + *Verbo-ndo* should fade over time.⁸

5.2 Temporal expressions: Testing the more specific meaning of the newer construction

It is thought that “grammaticalization occurs to indicate the less usual or more specific meaning” (Bybee 2010: 180). Here we operationalize the hypothesis that new constructions express more specific meanings than those of the existing variant (Bybee et al 1994: 133; García 1990: 146–147) by considering co-occurring temporal expressions. The prediction is that the Progressive should be favored in the presence of temporal expressions, for example ‘at this instant’ (13) or ‘always’ (16a, below), which may specify temporal location or “internal temporal constituency” (Comrie 1976: 3).

- (13) a. *la tumba se está abrie-ndo en este instante*
 the tomb REFL be.PRS.3SG open-GER in this instant
 ‘The tomb is opening (PROG) at this instant’
 (Álvaro, Jornada 5, Scene IX)
- b. –¿Hay alguien contigo?
 –Sí... Pero *duerme en este momento*
 yes but sleep.PRS.3SG at this moment
 ‘–Is someone with you?
 –Yes...But he is sleeping (PRS) at this moment’
 (Serafina, Act 3, Scene I)

Over time, the favoring effect of co-occurring temporal expressions may diminish, as the “punch” of the new expression “weakens the more it is used” (García 1990: 157). (In contrast, according to a space > time metaphor account the Progressive should increasingly co-occur with temporal adverbials; no such prediction follows from the view that progressive meaning is an implication of the locative source construction, such that aspectual (temporal) meaning is what remains as locative meaning is bleached (Section 5.1, above).)

5.3 Stativity: Testing generalization of progressives from dynamic to stative verbs

The hypothesis we test by coding the gerund lexical types for stativity is that locative-source progressive forms are first restricted to dynamic predicates and then generalize to stative verbs.⁹ The proposed grammaticalization path is as follows:

- (14) “be located at” > progressive > present/imperfective (Bybee et al. 1994, ch. 4)

Bybee et al. (1994: 139–148) adduce cross-linguistic evidence for the generalization of erstwhile progressive expressions to use in habitual contexts and with stative verbs as they become general presents or imperfectives. If Spanish *estar* + Verb-*ndo* follows this cross-linguistic path, we should find a disfavoring effect of statives, which should be strongest in the earlier period(s).

There is no unique categorization of stative verbs among linguists. Dynamic predicates are said to involve change, whereas statives describe situations that will continue unchanged unless something happens, for example, *know* (*a language*) (Comrie 1976: 48–50). But Vendlerian classes, including states, have been viewed as potential aspectual construals of predicates rather than inherent aspectual types, where construal depends on contextual features such as co-occurring adverbials and the tense-aspect construction (Croft 2009: 146; cf. Dahl 1985: 26–27). For example, progressive morphology has been said to “destativize” stative verbs (e.g. Bertinetto 1994: 403). Furthermore, we find no one-to-one relationship between the construal of a lexical type as stative or dynamic and the tense-aspect construction it appears in, Progressive or simple Present (for example, (22a, 22b) vs. (23a, 23b), ahead). Finally, the dynamic-stative classification is not the same across languages (Comrie 1976: 35).

Therefore, for the sake of consistency across time periods, I coded for stativity by lexical type (infinitive form) regardless of grammatical context (cf. Poplack & Tagliamonte 2001: 128, 141–142; Walker 2001: 17–19). Since the objective is to track change, a list of 99 statives, while undoubtedly not definitive, was applied in the same way to the three data sets.¹⁰ The most frequent stative lexical type is *ser* ‘be’, which alone makes up close to one-fifth (18%, 161/916) of stative tokens, followed by *pensar* ‘think’ (12%, 110/916). Together with these two, 21 verbs make up four-fifths (81%, 744/916) of the tokens coded as stative: *adorar* ‘adore’, *afligir* ‘afflict’, *callar* ‘be silent’, *cuidar* ‘think’, *desear* ‘desire’, *entender* ‘understand’, *figurar* ‘imagine’, *gozar* ‘enjoy’, *haber* ‘there is’, *holgar(se)* ‘rest, enjoy oneself’, *imaginar* ‘imagine’, *padecer* ‘suffer’, *parecer* ‘seem’, *penar* ‘pain, suffer’, *querer* ‘want, love’, *recordar* ‘remember’, *saber* ‘know’, *sentir* ‘feel’, *sufrir* ‘suffer’, *temer* ‘fear’, *tener* ‘have’.¹¹

5.4 Aspect: Gauging advancement toward an obligatory progressive

The prediction following from the locative > progressive > imperfective grammaticalization path depicted in (14) is that, in variation with the simple Present, *estar* + Verb-*ndo* will be favored in progressive as opposed to habitual aspect contexts, from the beginning. The grammaticalization view is that, as the construction is used more frequently (Table 1), the inference is made that not using it means non-progressive meaning; as this inference is conventionalized, progressive becomes an obligatory category and the simple Present becomes a zero morpheme indicating habitual aspect (cf. García and van Putte 1989, Bybee 1994). We should observe, then, decreasing relative frequency of the simple Present in progressive aspect contexts.

Although diverse, most treatments of aspect have relied on invented or selected examples (but see Walker 2010b). Here I use the time dimension of aspect (cf. Croft 2009: 151) to classify tokens of both the Progressive and simple Present based on the duration of the situation, independently of lexical stativity. I distinguish what I will call extended duration, which subsumes habitual aspect and states without temporal limits, from limited duration, which includes progressive or continuous actions and states circumscribed to speech time.

I retained aspect coding — extended vs. limited duration — only for those tokens that I classified the same way on two separate occasions (separated by about six months).¹² To test reliability, a sample of approximately 10% of the tokens classified as extended or limited duration was recoded by another linguist; we obtained good agreement (86%, 312/363 observed agreements, Cohen's Kappa = 0.707).¹³ In the following two subsections I illustrate how the coding applies to dynamic and to stative verbs.

5.4.1 *Extended vs. limited duration for dynamic verbs: habitual vs. progressive or continuous*

For dynamic verbs the aspectual distinction between progressive and habitual is well established in spontaneous speech (e.g. Walker 2001) and typological questionnaire studies (e.g. Dahl 1985: 90–98). The examples illustrate progressive (15) and habitual (16), in both the Progressive (a) and simple Present (b). Progressive actions occur simultaneously with speech time (or another point of reference), while habitual occurrences are customarily repeated such that they are characteristic of an extended period of time (Comrie 1976: 27–28, Bybee et al. 1994: 136–137).¹⁴ A situation was coded as habitual even if the period it characterizes began recently, as in (17), where the new situation is 'now' the norm.¹⁵

- (15) a. *no sabemos quién está dentro; habla-ndo están*
 not know.PRS.1PL who be.PRS.3SG inside speak-GER be.PRS.3PL
 ‘we don’t know who is inside; they **are talking** (PROG)’
 (Celestina, Act XIV)
- b. *escucha, que hablan quedito*
 listen.IMP that speak.PRS.3PL softly
 ‘listen, they **are speaking** (PRS) softly’ (Celestina, Act XII)
- (16) a. *quando non han con quién hablar, están*
 when not have.PRS.3PL with whom speak-INF be.PRS.3PL
habla-ndo consigo mismas entre sí
 speak-GER with_self same.F.PL among self
 ‘when they don’t have anyone to talk to, they **talk** (PROG) to themselves’
 (Corbacho, Part II, XII, p. 194)¹⁶
- b. *Que a quien más quieren, peor hablan*
 that to whom most love worse speak.PRS.3PL
 ‘To those they love best they **speak** (PRS) worst’ (Celestina, Act VI)
- (17) a. *cosas vié-ndo=se están todos los días que...*
 things see-GER=REFL be.PRS.3PL all the days that
 ‘things **are being seen** (PROG) every day that... (Afán, Act I, Scene VII)
- b. *Y ahora se peina a estilo de Madrid.*
 and now REFL comb.PRS.3SG in style of Madrid
 ‘And now she **does her hair** (PRS) in the Madrid fashion.’
 (Perfecta, Ch. XIII)

A third aspect label listed by Comrie (1976: 33) is continuous, defined as “imperfectivity that is not occasioned by habituality” (though Bybee et al. (1994: 139) do not find a separate formal category). Such dynamic situations may involve repeated occurrences but occur over a limited or specified period, for example ‘a month’ (18). I coded these together with progressive aspect as limited duration.

- (18) *Un mes hace que estoy envia-ndo recados a su casa.*
 one month make that be.PRS.1SG send-GER messages to your house
 ‘I **have been sending** (PROG) messages to your house for a month.’
 (Inseparables, Act I)

The following set of examples with *escribir* ‘write’ summarizes the aspect coding for dynamic verbs. ‘Writing’ (19a) and ‘(performative) writing to you’ (19b) are evidently in progress at reference time. ‘Writing a history of the theater’ (20a) and ‘writing strange letters’ (20b) are not strictly in progress at speech time, but neither are these situations maintained or repeated over a period so extended as to view them as usual customs or characteristic features, unlike habitual ‘always here

writing at night' (21a) and 'writing beautifully' (21b). The direct objects 'a history of the theater' and 'strange letters' in (20) define telic or specified situations, in contrast to the open-ended situations of writing (anything) at night and writing beautifully in (21), where the period of time is, in Comrie's (1976: 27–28) words, "so extended...that the situation...is viewed not as an incidental property of the moment but, precisely, as a characteristic feature of a whole period." Examples (19) (progressive) and (20) (continuous) were coded as limited duration whereas (21) (habitual) was classified as extended duration.

- (19) a. *–Está escribiendo.*
 be.PRS.3SG write-GER
–Pues ya presto habrá de dejarlo, que empieza a anochecer...
 '–She is writing (PROG).
 –Well she'll soon have to stop, it's starting to get dark'
 (Sí, Act I, Scene IX)
- b. *para evitarle con tiempo, te escribo hoy,*
 DAT.2SG write.PRS.1SG today
 'to avoid [a terrible scandal], I write (PRS) to you today,' (Pepita, p. 259)
- (20) a. *ahora estoy escribiendo una apología del teatro*
 now be.PRS.1SG write-GER a history of.the theater
 'now I am writing (PROG) a history of the theater'
 (CN, Act II, Scene VIII)
- b. *Luisito me escribe, hace días, extrañas cartas*
 Luisito DAT.1SG write.PRS.3SG make days strange letters
 'Luisito has been writing (PRS) to me, for days now, strange letters'
 (Pepita, p. 258)
- (21) a. *porque siempre aquí está mi padre escribiendo*
 because always here be.PRS.3SG my father write-GER
mucha parte de la noche
 great part of the night
 'because my father is always here writing (PROG) a good part of the night'
 (Casa con dos puertas, Jornada II)
- b. *tiene una labia y escribe que da gusto...*
 have PRS.3SG an eloquence and write.PRS.3SG that give pleasure
 'he speaks beautifully and he writes (PRS) [in such a way] that it gives pleasure...'
 (Sí, Act II, Scene V)

5.4.2 *Extended vs. limited duration for stative verbs: Without vs. with temporal limits*

Statives are said to be neutral to the aspectual distinction between habitual and progressive (e.g., Bybee 1994:238; but see Bertinetto 1994:394). In quantitative studies of English varieties, small proportions of statives have been assigned a habitual interpretation, but in these data I found very few good cases of habitual statives.¹⁷ Thus I take the pertinent distinction to be between states without temporal limits, which began before speech time and continue indefinitely (cf. “state exists” in Bybee et al. 1994:318), and states with temporal limits, which are circumscribed to a period near reference time (cf. “permanent” vs. “contingent” state in Bertinetto 1994:403). In (22), *be* always silent, understand one’s language, and fear God are states with no temporal limits, compared with being silent on a particular occasion, understanding what someone is saying in a conversation, and fearing the imminent noose in (23), which are temporally limited. The former were coded as extended duration, the latter as limited duration.

- (22) a. *los peces son los huéspedes que siempre están calla-ndo*
 that always be.PRS.3PL be_silent-GER
 ‘the fish are the guests who are always (being) silent’
 (Apolonio, verse 506)
- b. *No entiendo otra lengua que la mía*
 NEG understand.PRS.1SG
 ‘I **understand** no language other than my own’ (Quijote II, Ch. II)
- c. *cómo, siendo el principio de la sabiduría el temor de Dios, tú, que temes más a un lagarto que a Él, sabes tanto*
 who fear.PRS.2SG
 ‘how, fear of God being the beginning of wisdom, you, who **fears** a lizard more than Him, know so much’ (Quijote II, Ch. XX)
- (23) a. *fabla p<er>o bermudo por q<ue> estas calla-ndo*
 why be.PRS.2SG be_silent-GER
 ‘speak Pero Bermudo, why **are** you (being) silent?’ (EE II, fol. 240v)
- b. *No me entiendes, Sancho: no quiero decir sino que [...]*
 NEG DAT.1SG understand.PRS.2SG
 ‘You do not **understand** me, Sancho: all I want to say is that [...]
 (Quijote II, Ch. XXV)
- c. *sería largo de contar, y más en tiempo que estoy temie-ndo*
 at time that be.PRS.1SG fear-GER
que, entre la lengua y la garganta, se ha de atravesar el riguroso cordel que me amenaza

‘it would be a long story, especially now when I **am fearing** that the severe noose threatening me must tighten between my tongue and my throat’
(Quijote II, Ch. LXIII)

5.4.3 *Aspectual indeterminacy*

Aspect turns out to be indeterminate for a non-negligible number of tokens. For example, in (24), the question could be about the current occasion or about every occasion; in the temporal clause in (25), ‘talking’ could be viewed as in progress at the reference time, which is defined by the coming to mind of the sayings, but could also be viewed as customarily repeated; in (26), the speaker’s opinion may have been just formed in the course of the conversation or be a longstanding one (in other words, the speaker’s sudden anger could be occasioned by either a realization or by being reminded of an already formed opinion). For other tokens, as a reader, I simply could not tell the aspectual interpretation. Indeterminate and uncodable aspect tokens make up approximately one fourth of the data. It remains to be seen from future replicable studies of verbal aspect whether this substantial proportion of uncodable cases reflects the methodological decisions of the analyst or true indeterminacy of aspectual distinctions in certain contexts for language users.¹⁸

- (24) *Señor, Dios mío, ¿oyes mi voz, o estoy condenada a rezar eternamente sin ser oída?*
‘Lord, **do you hear/are you hearing** (PRS) my voice, or am I condemned to eternally pray without being heard?’ (Perfecta, Ch. XXIV)
- (25) *viénenseme tantos [refranes] juntos a la boca cuando hablo, que riñen por salir unos con otros*
‘so many of them [sayings] come into my mouth all together when I **talk/am talking** (PRS) that they fight each other to burst out’ (Quijote II, Ch. XLIII)
- (26) *Sr. Ramos -dijo Remedios súbitamente enojada-, se me **figura** que no entiende V. gran cosa en esto de casar a la gente.*
‘Mr. Ramos, said Remedios suddenly angry, it **seems** (PRS) to me that you do not understand much about this matter of marrying people’
(Perfecta, Ch. XXVII)

5.5 Other factors: Negatives and interrogatives

A final factor group is polarity and sentence type, with affirmative declaratives opposed to negative polarity clauses (27, 28b) and interrogatives (28a). The prediction is that these contexts, which have been found to disfavor the English Progressive (Jarmasz 2006 §6.2.9), will also disfavor *estar* + Verb-*ndo*.

- (27) *ni espumo de la boca. [...]*
 nor froth.PRS.1SG from the mouth
ni esto tremie-<n>do co<n> sa<n>na.
 nor be.PRS.1SG tremble-GER with rage
 ‘nor am I frothing (PRS) at the mouth, [...] nor am I trembling (PROG)
 with rage.’ (EE I, fol. 87v)
- (28) a. *¿Qué estás murmura-ndo, Sempronio?*
 what be.PRS.2SG murmur-GER
 b. *No digo nada.*
 NEG say.PRS.1SG nothing
 a. ‘What are you murmuring (PROG), Sempronio?’
 b. ‘I’m not saying (PRS) anything.’ (Celestina, Act I)

Not included in the analyses reported below is clause type, although in English subordinate clauses favor the Progressive (Walker 2001).¹⁹ Also set aside here are factors relating to the subject (animacy, grammatical person).²⁰

6. Tracking the conditioning of variation over time: Progressive becomes more of a progressive

The simultaneous effect of the factor groups (independent variables) was considered in multivariate analysis, in this case, Variable-rule analysis (Sankoff, Tagliamonte & Smith 2005). Table 4 shows three independent analyses of the Old Spanish, 17th century, and 19th century data.²¹ For each time period, the numbers in the first column represent the Probability or Factor Weight that each factor (level of the independent variable) listed on the left contributes to the occurrence of the Progressive: the closer to 1, the more favorable, the closer to 0, the more unfavorable (or, conversely, favorable to the simple Present).²² The second column shows Progressive rates and the third column the total number of tokens in each factor. The percentage of the data made up by the given factor appears between parentheses (for example, tokens with a co-occurring locative make up 7–8% of the data).²³

The direction of effect is instantiated in the order of the factors within a group by Factor Weight, from higher to lower, or from more to less favorable. Relative strength or magnitude of effect of factor groups is assessed by the Range between the highest and lowest Factor Weight in a group.²⁴ Does the linguistic conditioning of the Progressive change over time?

Table 4. Three independent multivariate analyses of factors contributing to the choice of the Progressive (vs. the simple Present) in Old Spanish, 17th century, and 19th century Spanish texts (Favoring effects in bold, non-significant factors within [])*

Aspect	Old Spanish			17th century			19th century			
	N	Prob	%	N (% data)	Prob	%	N (% data)	Prob	%	N (% data)**
Limited duration	119/745	.68	29%	216 (44%)	.72	32%	450 (63%)	.73	42%	647 (64%)
	180/1013	.36	11%	277 (56%)	.17	4%	268 (37%)	.15	5%	359 (36%)
Extended duration		.32				55			58	
Locative co-occurrence		.77	43%	49 (7%)	.74	44%	77 (8%)	.62	37%	98 (7%)
		.48	14%	696 (93%)	.48	15%	927 (92%)	.49	21%	1330 (93%)
		.29			.26				13	
Polarity — Sentence type		.55	18%	582 (82%)	.57	20%	807 (83%)	.58	25%	1127 (81%)
		.31	9%	130 (18%)	.21	6%	161 (17%)	.20	7%	269 (19%)
		.24			.36			.38		
Temporal co-occurrence		.70	24%	103 (14%)	[.54]	21%	136 (14%)	[.57]	27%	189 (13%)
		.47	15%	638 (86%)	[.49]	17%	874 (86%)	[.49]	21%	1253 (87%)
		.23								
Stativity		[.50]	17%	552 (74%)	.56	22%	703 (69%)	.58	27%	1047 (72%)
		[.49]	14%	193 (26%)	.36	9%	310 (31%)	.32	9%	413 (28%)
				.20					26	

** Total N in Tables 4, 6, 7 is not the same for all factor groups due to exclusions (e.g., tokens of indeterminate or uncodable aspect are not included in the aspect factor group).

In Old Spanish, the strong favoring effect of a co-occurring locative, which has the highest Probability (.77), supports the hypothesis of the locative origins of the Spanish Progressive.²⁵ We also see that an aspectual effect is already in place, with limited duration situations favoring the Progressive (.68). Cross-tabulation (Table 8, in Appendix) displays the independence of locative co-occurrence and aspect as constraints: the Progressive is favored in the presence of a locative in each aspectual context and in limited duration contexts regardless of the presence of a locative. The rate of the Progressive is highest in the presence of a locative in limited temporal duration contexts, at 64% (16/25), and lowest in the absence of a locative in extended duration contexts, at 10% (27/264).

It is important that the magnitude of the aspect effect is not greater than that of locative co-occurrence, as indicated by the close Ranges (32 and 29). Also note that, while it is disfavored (.36), the Progressive nevertheless has a substantial rate in extended duration contexts, at 11%, compared with limited duration contexts, at 29%. Stativity has no effect at all, as the rate of the Progressive is virtually the same for dynamic and stative predicates (17% and 14%, respectively).

In the 17th century, while there is no change in the direction of effect — the Progressive is still favored by co-occurring locatives, limited duration, and affirmative declaratives — there are clear shifts in relative magnitude of effect. First, the aspectual effect has evidently become the strongest, with a Range which is 50% greater than the next largest, that of the polarity-sentence type factor group (55:36 = 1.5 times). The Progressive is now more strongly disfavored (.17) in extended duration contexts, at a mere 4% compared with 32% in limited duration contexts. Second, stativity appears for the first time as statistically significant, with stative verbs disfavoring (.36) the Progressive. Third, the locative factor group (Range = 26) is relegated to third position, behind that of the polarity-sentence type (Range = 36).

In the 19th century, the locative co-occurrence effect (Range = 13) has weakened further, dropping to fourth position in relative magnitude, considerably behind stativity (Range = 26). As in the 17th century, the factor group that contributes the most to the choice of the Progressive is aspect (with a Range that is 50% (58:38) greater than that of polarity-sentence type) and now the highest Probability is that of limited duration (.73).

Finally, co-occurring temporal expressions appear to favor the Progressive in the Old Spanish data (.70). This is expected if the new Progressive construction expresses more specific temporal meanings than the existing simple Present (Section 5.2). However, cross-tabulation of temporal co-occurrence with aspect (Table 5) shows that the higher rate of the Progressive in the presence of a temporal expression holds for extended duration contexts, for example in habitual situations with *siempre* 'always' (29). There is no temporal co-occurrence effect

Table 5. Rate of Progressive by temporal expression co-occurrence in extended vs. limited duration aspectual contexts, by century*

		Extended duration	Limited duration
		% Progressive (N/N)	% Progressive (N/N)
OLD SPAN	Temporal present	26% (13/50)	21% (3/14)
	absent	8% (17/225)	29% (59/202)
17TH C.	Temporal present	10% (3/29)	35% (22/62)
	absent	3% (8/238)	31% (122/388)
19TH C.	Temporal present	23% (12/52)	37% (28/76)
	absent	1% (4/306)	43% (240/563)

*In extended duration contexts, differences are significant ($p < .05$) in Old Spanish and 19th century; in limited duration contexts, differences are not significant in any time period (chi-square tests).

within limited duration contexts, which favor the Progressive regardless of temporal expression co-occurrence. Note that in Old Spanish, the rate of the Progressive is three times greater (26 vs. 8%) in the presence of a temporal expression in extended duration contexts. In contrast, in the 19th century, the rate is more than twenty times greater (21 vs. 1%). In other words, in the later period, the Progressive occurs in extended duration (habitual) contexts virtually only when there is a temporal expression such as ‘constantly’, ‘now’, ‘every day’ (30) (and examples in (17) above).²⁶

- (29) a. *por mucho que tengan siempre están llora-ndo*
 for much that have.SUBJ.PRS.3PL always be.PRS.3PL cry-GER
e quejá-ndo=se de pobreza
 and complain-GER=REFL of poverty
 ‘for all they have they are always crying (PROG) and complaining of poverty’
 (Corbacho, Part II, I, p. 148)
- b. *la muger parlera siempre habla de fechos ajenos*
 the woman garrulous always speak.PRS.3SG of affairs_of_others
 ‘the garrulous woman is always talking (PRS) about others’ affairs’
 (Corbacho, Part II, XII, p. 194)
- (30) *piensa que tu padre soy, y que*
 think-IMP that your father be.PRS.1SG and that
de continuo estoy soña-ndo tu bien...
 of continuous be.PRS.1SG dream-GER your good
 ‘keep in mind that I am your father, and that I am constantly yearning for what is best for you’
 (Álvaro, Jornada I)

In summary: The direction of effect for locative co-occurrence remains stable: the presence of a locative favors choice of the Progressive throughout its trajectory, in support of locative origins and as predicted by the retention, or persistence, hypothesis (Bybee & Paglicua 1987; Hopper 1991). Change here is manifested in shifts in relative magnitude of effect. The favoring effect of co-occurring locatives weakens over time, a measure of the loss of specific features of meaning, or semantic bleaching, in the course of grammaticalization (Bybee et al. 1994: 6). The aspectual effect is in place from the beginning, which belies an abrupt metaphorical space > time leap. The strengthening of this aspect effect between Old Spanish and the 17th century shows that the Progressive indicates limited duration (progressive action, circumscribed state) *in growing contrast* to the extended duration meaning of the simple Present (habitual action, indefinitely existing state) (cf. Bybee 1994). This accords with the grammaticalization view that with more frequent use of the Progressive, its absence implies non-progressive meaning, leaving the zero-marked simple Present as the exponent of habitual (Bybee 1994). A simplified depiction of this development appears in Figure 1a.

So, increased token frequency (Table 1) is accompanied by changes in the linguistic conditioning (Table 4) of the Progressive. What have we learned by viewing the patterns of variation of a grammaticalizing construction with its older alternative? In the account of the development of zero morphemes referred to above (Bybee 1994), the simple Present is viewed as being increasingly restricted to habitual/extended duration (or, disfavored in progressive/limited duration contexts) as the Progressive spreads in progressive contexts. We have seen here that it is

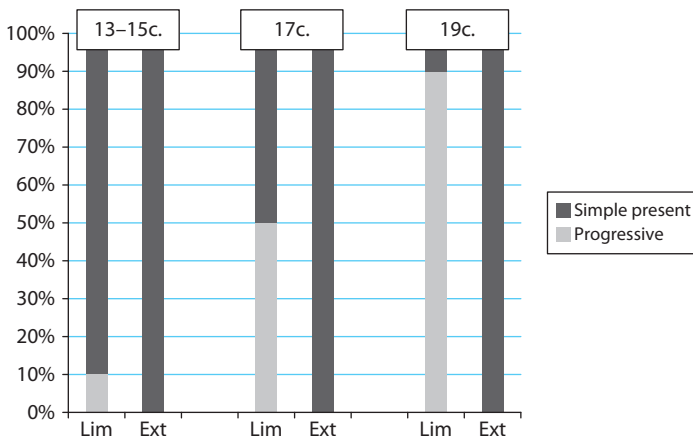


Figure 1a. Hypothesized development of aspectual opposition Progressive (*estar + Verb-ndo*) vs. simple Present (cf. Bybee 1994): the simple Present is gradually restricted to extended duration (Ext) aspectual contexts, as the Progressive increases in frequency in limited duration (Lim) contexts.

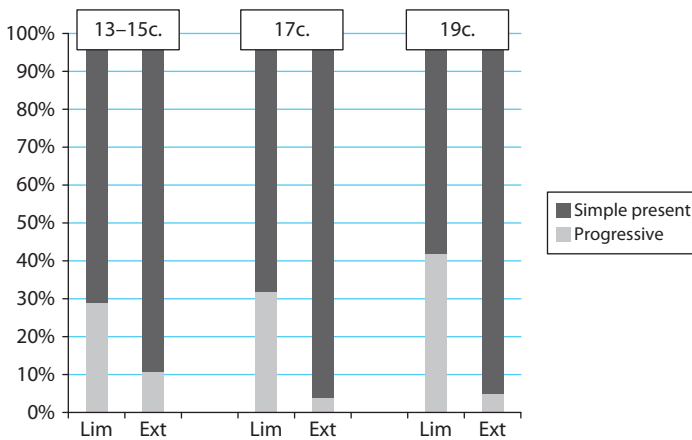


Figure 1b. Development of aspectual opposition Progressive (*estar + Verb-ndo*) vs. simple Present: the simple Present is gradually restricted to extended duration (Ext) aspectual contexts, as the Progressive is favored in limited duration (Lim) and increasingly disfavored in extended duration (Ext) contexts.

not just that the Progressive increases in frequency in progressive contexts, but that, as it loses locative meaning, it loses habitual meaning (or, the Progressive is increasingly disfavored in habitual/extended duration contexts). This is depicted in Figure 1b.

In other words, the Progressive does not come in as a full-blown progressive, which simply increases in token frequency and overall rate with respect to the simple Present, thus ousting it from progressive territory. What we find instead is that, when it started out as more of a locative, the Progressive was also more habitual in aspect (with a substantial rate in extended duration contexts that is two and a half times lower than in limited duration contexts, $29:11 = 2.6$, in Old Spanish). It becomes decreasingly habitual (with a sharply reduced rate in extended duration contexts that is eight times lower than in limited duration contexts, $32:4 = 8$ and $42:5 = 8.4$, in the 17th and 19th centuries).

Thus, the aspectual opposition with the simple Present was not fully in place when *estar + Verb-ndo* was first used. Rather, it has evolved in the course of variation, as the originally more locative construction is increasingly used as an aspectual expression of limited, not extended, duration — or, as it becomes more of a progressive.

7. Stativity and progressivity

The prediction consonant with the hypothesis that locative-source progressives are first restricted to dynamic verbs is that the stativity constraint should be strongest in the earliest period (Section 5.3). This is not borne out, as the stativity effect is significant in the 17th and 19th century but not the Old Spanish data (Table 4). Does the aspectual distinction between extended and limited duration apply to stative verbs? If so, does the Progressive signal it?

We first cross-tabulate the stativity and aspect factor groups to answer the first question. Figure 2, which displays the aspectual contexts of occurrence separately for dynamic and stative verbs, shows that the extended vs. limited duration distinction is pertinent for both dynamic and stative verbs.²⁷ Nevertheless, a greater proportion of stative (34%, 303/881) than dynamic lexical types (22%, 460/2099) is aspectually indeterminate (Section 5.4.3). We also confirm that stative verbs tend to occur in extended duration contexts more than dynamic verbs: approximately half (49%, 1030/2099) of dynamic verb tokens occur in limited duration (progressive, continuous) contexts, whereas only one-third (32%, 283/881) of stative verb situations are circumscribed to speech time.²⁸ This greater association of statives with extended duration is expected under definitions of stativity based on lack of change (Comrie 1976: 48–50; Langacker 1987: 79).

We then conduct separate multivariate analyses of dynamic and stative verbs, shown in Table 6.²⁹ Factor Weights enclosed in square brackets are included to indicate the direction of effect even when the factor group is not statistically significant.³⁰ Three results are important here.

First, comparing the direction of effect within each factor group, the conditioning of variant choice is the same across the two classes of verbs: the Progressive is (decreasingly) favored in the presence of a co-occurring locative and (increasingly) favored in limited duration contexts, for both stative and dynamic

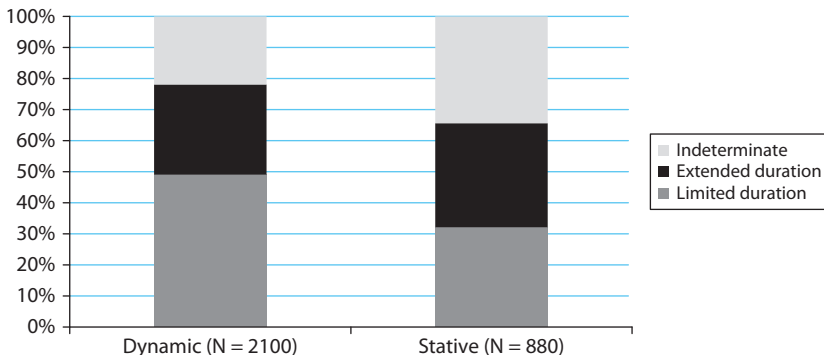


Figure 2. Aspectual context of occurrence by stativity of verb

Table 6. Six independent multivariate analyses of factors contributing to the choice of the Progressive (vs. the simple Present), in Old Spanish, 17th century, and 19th century Spanish texts, for dynamic and stative verbs (Favoring effects in bold, non-significant factors within [])*

Aspect	Old Spanish				17th century				19th century			
	Dynamic		Stative		Dynamic		Stative		Dynamic		Stative	
	Prob	%	Prob	%	Prob	%	Prob	%	Prob	%	Prob	%
N	93/552	29%	26/193	14%	153/703	22%	27/310	9%	282/1047	27%	35/413	9%
Input	.16	(17%)	.15	(14%)	.19	(22%)	.05	(9%)	.25	(27%)	.05	(9%)
Range	33		26		51		68		56		60	
Limited duration	.68	29%	.66	29%	.69	36%	.82	21%	.70	46%	.80	24%
Extended duration	.35	11%	.40	11%	.18	6%	.14	1%	.14	6%	.20	2%
Range	33		26		51		68		56		60	
Locative co-occurrence												
Present	.73	38%	.92	71%	.71	46%	.86	36%	.63	39%	.57	18%
Absent	.48	15%	.48	11%	.48	19%	.48	8%	.49	26%	.50	8%
Range	25		44		23		38		14		14	
Polarity — Sentence type												
Affirmative declarative	.54	19%	.57	21%	.56	24%	.58	10%	.57	31%	.61	10%
Negative, Interrogative	.29	9%	.35	4%	.20	7%	.24	3%	.20	11%	.22	2%
Range	25		36		36		34		37		39	
Temporal co-occurrence												
Present	.69	24%	.72	24%	.54	26%	.59	11%	.52	26%	.83	29%
Absent	.47	16%	.47	12%	.49	21%	.49	9%	.50	27%	.46	7%
Range	22		22		22		22		22		22	

verbs. That is, the aspectual opposition with the simple Present — for states, temporally circumscribed versus existing indefinitely — develops gradually, as *estar* + Verb-*ndo* begins as more of a locative construction with stative, as well as dynamic, verbs, as for example in (31).

- (31) *en tierra está adora-ndo a la más antigua [y] puta tierra*
 in ground be.PRS.3SG adore-GER ACC the most old whore earth
 ‘he is on the ground **adoring** the oldest and most whorish dirt’ (Celestina,
 Act I)

This early more locative construction is not incompatible with meaning “being in a state” such that the gerund was like an adjective (as proposed by Bertinetto 2000: 563–564; Bertinetto et al. 2000: 538–539; see Section 5.3, footnote 9). An adjectival ‘being in an immobile state’ use is illustrated in Old Spanish example (32), where the gerund is conjoined to an adjective and an adjective appears in place of the gerund in following text (for an example of a dynamic-verb gerund conjoined to an adjective, see (5c)).

- (32) *¿Por qué está el león triste et cuida-ndo? [...]*
 for what be.PRS.3SG the lion sad and think-GER
Esto que tú vees estar al león triste et cuidoso [...]
 this that you see be.INF ACC.ART.SG lion sad and pensive
 ‘Why is the lion **sad and thinking?** [...] This that you see, the lion being sad
 and **pensive** [...]’ (Calila, p. 182)

Second, within the 17th and 19th century data sets, the Input values, which indicate the overall tendency for the Progressive to occur, are lower for stative than for dynamic verbs, confirming that statives disfavor progressive forms. This is also the case in American English conversational data, where the rate of the Progressive relative to the simple Present is 22% (95/435) for dynamic verbs vs. 2% (17/871) for stative verbs (Bybee 2010: 178, from Scheibman 2002).

What is noteworthy here is that in the oldest period the Input values are nearly identical (.15–.16). This reiterates the already observed lack of a stativity effect in the Old Spanish data and its appearance in the 17th century, when the limited vs. extended duration aspectual effect has become the strongest (Table 4). In the 17th century the Input for stative (.05) is nearly four times lower than for dynamic (.19), and the difference between stative and dynamic verbs in the rate of the Progressive is even sharper in the 19th c. (Input .05 vs. .25). In other words, the restriction against Progressive use with stative verbs applies not from the beginning but as it becomes more of a progressive.³¹ Similarly, in English the disfavoring effect of statives on *be* + Verb-*ing* may have developed fairly late, in the 18th century (Walker 2001: 25; cf. Nels 1988: 180,185). We may hypothesize, then, that cross-

linguistically the stativity restriction on locative-source progressive forms develops in tandem with their development as expressions of limited duration.

Third, temporal co-occurrence is significant for statives in the 19th century, while locative co-occurrence no longer is. For stative verbs the probability that the Progressive will be chosen is now greatest (.83) in the environment of a temporal expression. Cross-tabulation of temporal co-occurrence with aspect for stative and dynamic verbs separately (Table 10, in Appendix) reveals that, while for dynamic verbs the rate of the Progressive is consistently higher with a co-occurring temporal expression in *extended* duration contexts, for statives, on the other hand, temporal expressions tend to favor the Progressive in *limited* duration contexts, in the 17th and 19th centuries.³² Temporal expressions such as *ahora* ‘now’ in (33) may provide adverbial support for a state circumscribed to speech time encoded in the Progressive, as opposed to an indefinitely existing state in the Present (“contingent” as opposed to “permanent” state (Bertinetto 1994: 403)). This opposition is illustrated in (34), here in the absence of a temporal expression.³³

- (33) *Se me está figura-ndo ahora que no soy*
 REFL DAT.1SG be.PRS.3SG seem-GER now that not be.PRS.1SG
tan malo como yo mismo creo.
 as bad as I myself believe.PRS.1SG
 ‘It seems (PROG) [literally: is seeming] to me now that I am not as bad as I
 myself believe (PRS).’ (Perfecta, Ch. XXVIII)

- (34) a. *Mariquita, yo estoy desea-ndo que nos casemos*
 Mariquita I be.PRS.1SG desire-GER that REFL.1PL marry.SBJV.1PL
 ‘Mariquita, I desire (PROG) [literally: am desiring] that we marry’
 (CN, Act II, Scene II)
- b. *Como el corzo sediento desea y busca*
 as the deer thirsty desire.PRS.3SG and search_for.PRS.3SG
el manantial de las aguas, así mi alma busca
 the spring of the waters so my soul search_for.PRS.3SG
a Dios todavía.
 ACC God still
 ‘Just as the thirsty deer desires (PRS) and searches for the fountain, so
 my soul is still searching for God.’ (Pepita, p. 112)

Even though the rate of the Progressive is much lower for stative than dynamic verbs in the 17th century, and the difference seems to sharpen further in the 19th century, in this later period there appear examples of the construction with prototypical statives, such as ‘be’ + ‘immense’, a time-stable adjective of size (cf. Silva-Corvalán 1994: 92–120) in (35). Thus, the progressive-nonprogressive opposition, which is more robust with dynamic verbs, begins to extend to statives. Evidence for

the extension of the Progressive to the very stative copula *ser* ‘be’ comes from the large on-line *Corpus diacrónico del español* (CORDE) (Real Academia Española), in which the token frequency of Present tense *estar siendo* ‘be being’ (all grammatical persons) relative to all occurrences of the gerund *siendo* leaps fifteen-fold, from 3/13296 in 1600–1650 to 21/6183 in 1850–1900.³⁴ The evolutionary path of the English Progressive may be the same, as it is reported that *be* + predicate adjective/nominal Progressive constructions (*he is being silly*, *he is being a fool*) first appear in the 19th century (Nels 1988: 188).

- (35) *¿Te imaginas que no es, que no está*
 REFL.2SG imagine.PRS.2SG that not be.PRS.3SG that not be.PRS.3SG
sie-ndo, que no será inmenso el sacrificio que hago?
 be-GER that not be.FUT.3G immense the sacrifice which make.PRS.1SG
 ‘Do you imagine that the sacrifice I am making is not, is not being, will not
 be immense?’ (Pepita, 177)

In summary, a limited vs. extended duration distinction does apply to stative verbs, though the proportion of indeterminate aspect tokens is greater for statives than for dynamic verbs. The comparison of the three data sets suggests three stages in the use of the Progressive form with statives. The Progressive has locative origins for both dynamic and stative verbs (evidence: same locative effect in separate analyses of dynamic and stative verbs (Table 6)) and was in fact used early on with statives, at a similar rate as with dynamic predicates (in Old Spanish, near-identical Input values in stative and dynamic verb analyses). It is as the construction’s function as an aspectual expression of limited duration develops that the restriction on Progressive use with stative verbs appears (in the 17th century, drastically lower Input for stative compared with dynamic verbs). The more robust aspectual opposition with the simple Present that applies to dynamic verbs begins to spread to statives (in the 19th century, temporal co-occurrence effect in separate statives analysis; increasing frequency of *estar siendo* ‘be being’).³⁵ This path is depicted in (36).

- (36) Evolution of progressivity and stativity
 more locative, used with both dynamic and stative verbs >
 more aspectual (progressive), largely restricted to dynamic verbs >
 limited vs. extended duration opposition incipiently generalized to statives

8. Conservative contexts: Negatives and interrogatives

We have seen that from its earliest appearance the Progressive is in variation with the simple Present in the domain of present temporal reference, and that the aspectual opposition between these two constructions — limited duration (progressive or continuous action, state circumscribed to speech time) vs. extended duration (habitual action, indefinitely existing state) — develops gradually. But there is variation even within limited duration contexts, as illustrated in the 19th c. examples of the simple Present with limited duration meaning, progressive action in (37) and state circumscribed to speech time in (38). That is, the Progressive is (still) not an obligatory expression of progressive aspect (compare the English translations).

- (37) a. *Parece que hablan de mi función. (Escuchando)*
 seem.PRS.3SG that speak.PRS.3PL of my show
la conversación.
 ‘It looks like they **are talking** [literally: talk] about my show. (Listening to the conversation [stage instructions].)’ (CN, Act I, Scene III)
- b. *¿Pues no lo veis?... Que se abrazan.*
 so NEG it see.PRS.2PL that REFL embrace.PRS.3PL
 ‘Don’t you see?... They’re **embracing** [literally: embrace].’
 (Boda, Act III, Scene XIV)
- c. *— ¿Y madre?*
— Duerme. Se acostó muy tarde.
 sleep.PRS.3SG REFL lay.down.PFV.3SG very late
 ‘—And mother?
 —She is **sleeping** [literally: sleeps]. She went to bed very late.’
 (Regenta I, Ch. XI)
- (38) a. *y... me parece que le estoy oye-ndo...*
 and DAT.1SG seem.PRS.3SG that DAT.3SG be.PRS.3SG hear-GER
me respondió así:
 ‘and...it **seems** as though I’m hearing him...he answered me like this.’
 (Pazos, Ch. 27)
- b. *Un asiento queda en la de Valencia*
 a seat remain.PRS.3SG in the.one of Valencia
 ‘**There is** one remaining seat on the Valencia [coach]’
 (Acertar errando, Act I, Scene II)

We examine the variation within the preferred aspectual context of the Progressive, that is, situations of limited duration, by means of separate multivariate analyses including only tokens of limited duration (Table 7). The structure of the variation has changed over time. In the beginning, the context in which speakers

Table 7. Three independent multivariate analyses of factors contributing to the choice of the Progressive (vs. the simple Present), in Old Spanish, 17th c. and 19th c. Spanish texts, in limited duration aspectual contexts only (Favoring effects in bold, non-significant factors within [])

	Old Spanish		17th century		19th century	
N	62/216		144/450		269/647	
Input	.28 (29%)		.30 (32%)		.40 (42%)	
	Prob	%	Prob	%	Prob	%
	N (% data)		N (% data)		N (% data)	
Locative co-occurrence						
Present	.82	64%	.68	56%	[.57]	48%
Absent	.45	24%	.48	29%	[.49]	41%
	<i>Range</i>		<i>20</i>			
			25 (12%)			64 (10%)
			191 (88%)			569 (90%)
Polarity – Sentence type						
Affirmative declarative	[.54]	33%	.58	38%	.59	49%
Negative, Interrogative	[.37]	18%	.23	10%	.21	15%
	<i>Range</i>		<i>35</i>			<i>38</i>
			156 (76%)			493 (80%)
			49 (24%)			126 (20%)
Stativity						
Dynamic predicate	[.49]	29%	.54	36%	.55	46%
Stative predicate	[.53]	29%	.39	21%	.31	24%
	<i>Range</i>		<i>15</i>			<i>24</i>
			165 (76%)			526 (81%)
			51 (24%)			121 (19%)
Temporal co-occurrence						
Present	[.36]	21%	[.53]	36%	[.41]	37%
Absent	[.51]	29%	[.50]	31%	[.51]	43%
			14 (7%)			76 (12%)
			202 (93%)			563 (88%)

were more likely to choose the Progressive over the simple Present to express limited duration is in the presence of a locative expression, the one significant factor group in Old Spanish. The locative co-occurrence effect weakens in the 17th and disappears in the 19th century, while statives come to disfavor the Progressive and the strongest effect becomes that of polarity — sentence type. Negatives and interrogatives remain as the least favorable context for the Progressive in the 19th century (Probability .21, Range = 38). Example (39) illustrates an affirmative vs. negative pair of clauses with the same verb, the former in the Progressive (39a), the latter in the simple Present (39b). Example (40) illustrates a question about a situation in progress in the simple Present (40a) answered with a Progressive (40b). The few interrogatives we find in the Progressive may serve a “rhetorical function” rather than seeking information (cf. Freed 1994:631–633); half of the eighteen 19th c. tokens are the expressions *¿estoy soñando?* ‘am I dreaming?’, *¿qué (x) estás diciendo?* ‘what are you saying?’, and *¿no está(s) viendo?* ‘don’t you see?’, as in (41), with three tokens each.³⁶

- (39) – [...] *no es posible sino que adivin-ase*
 a. *lo que nos está sucedie-ndo.*
 that which DAT.1PL be.PRS.3SG happen-GER
 b. – *Pero, señora, si no sucede nada.*
 but madam if NEG happen.PRS.3SG nothing
 ‘–It’s not possible but that he must guess
 a. what is **happening** (PROG) to us.
 b. – But, madam, **nothing is happening** (PRS)? [literally: happens]
 (Sí, Act I, Scene V)
- (40) a. – *Y aquella chica, ¿qué hace?*
 and that girl what do.PRS.3SG
 b. – *Está desmenuza-ndo un bizcocho para dar de cenar*
 be.PRS.3SG break.up-GER a cookie
a don Periquito.
 a. ‘– And that girl, what is she **doing** (PRS) [literally: does she do]?
 b. – She’s **taking apart** (PROG) a cookie to give Don Periquito his dinner’
 (Sí, Act I, Scene VI)
- (41) *¿Qué pregunta! Pues, ¿qué! ¿no lo está usted vie-ndo?*
 what question so what NEG it be.PRS.3SG you see-GER
 ‘What a question! What, **don’t** you see? [literally: are you not seeing]’
 (Madrid, Act III, Scene VI)

Why do negated situations in progress and questions about situations in progress tend disproportionately to appear in the simple Present? Negative polarity and interrogative contexts may be conservative, though for different reasons.

On the one hand, Givón (1979: 121–122) proposes that a consequence of the marked presuppositional status of negatives is syntactic conservatism, such that tense-aspect distinctions are first elaborated in the affirmative paradigm. But this argument for the hypothesized conservatism of negation appears circular, since it is not clear in what sense “presuppos[ing] a context in which the affirmative proposition has been asserted or at least entertained” (Horn 1985: 143) is “marked”. Moreover, it has been found that most negative clauses deny something which is “not explicitly present in the conversation”, rather than something that was previously uttered or is presupposed (Thompson 1998: 325; Tottie 1991).

Alternatively, we might think of the aspectual distinction between habitual and progressive as being less applicable under negative polarity, that is, if the situation is not happening, whether it would be of extended or limited duration were it to happen may be less relevant. The disfavoring effect of negation on the Progressive (also in English (Jarmasz 2005)) may be interpreted as being consistent with a greater aspectual-temporal neutrality of negatives. Negation has also been found to disfavor newer forms of the future in Early Modern Greek (Pappas 2001: 83), 17th c. Spanish (Aaron 2010: 10), and still in present-day French (Emirkanian & Sankoff 1985: 195). Studies that consider polarity effects in other languages and constructions can test the hypothesis that negatives are more neutral to tense-aspect distinctions.

On the other hand, interrogatives may appear conservative, perhaps due to a disproportionate number of fixed expressions. In the present data we notice frequent second-person formulas in the simple Present, such as *por qué lloras* ‘why are you crying’, *(por)qué miras* ‘why/what are you looking at’, *de qué te ríes* ‘what are you laughing at’, *qué tienes* ‘what’s troubling you’. In fact, second-person *¿qué hace(is)?* ‘what are you doing?’ alone makes up 8% (11/133) of the 19th c. interrogative data. In English future temporal reference data from Canadian English sociolinguistic interviews, the expression *What am I going to do?* makes up approximately one-fourth of first-person interrogatives and *Is there gonna be...?* and *What’s gonna happen?* just under one-fifth of third-person interrogatives (Torres Cacoullós and Walker 2009: 344–345). It may be the case more generally, then, that interrogatives have a higher proportion of conventionalized collocations compared with declaratives.

What role do we expect frequent collocations to play in grammaticalization? The case of Spanish Progressive — simple Present variation bears on two hypotheses: high frequency collocations of the older construction retard change, while high frequency collocations of the grammaticalizing construction propel it.

First, high frequency routines in the older construction may lag behind other contexts in the generalization of the newer construction, similar to high frequency items which are less likely to undergo analogical change than low frequency ones (the “conserving effect” of token frequency (Bybee 2010:75)). This would be the case of frequent simple Present expressions with progressive meaning such as *¿qué hace(is)?* ‘what are you doing?’. It is important that the opposition between the Progressive and Simple Present here is distributional, not aspectual. Simple Present interrogative *¿qué hace(is)?* simply remains the conventional way to ask about a situation in progress, while the answer increasingly appears in the Progressive.

Second, constructions with particular verbs may take a lead in the semantic-structural changes observed in the grammaticalization of the general construction. Consider *estar + hablando* ‘be talking’, which makes up a relatively high proportion of combinations of *estar + Verb-ndo*. *Hablando* ‘speaking’, usually in some fixed location, is originally harmonious with the locative source meaning of the *estar* construction, whereas gerund *buscando* ‘looking for’, often involving large outdoor expanses in Old Spanish texts, pairs up with motion-verb *andar* ‘go around’ (Torres Cacoullós 2000:120–169). In present-day (Mexican) Spanish varieties, *estar hablando* is the conventional way to express ‘be speaking’, and *andar buscando* the conventional way to ‘be looking for something’, remaining as the tenacious distributional residue of once meaningful associations rather than conveying aspectual or spatial nuances (Torres Cacoullós 2001). Such combinations thus qualify as pre-fabs (Bybee 2010: 34). In Bybee and Torres Cacoullós (2009) we show that in Old Spanish texts *estar hablando* leads in measurable changes such as the proportion of tokens without intervening material and without a co-occurring locative.

Studies of progressive constructions in other languages will help determine whether the disfavoring effect of negation and interrogatives is one of the vagaries of the Spanish Progressive or whether it applies to progressives cross-linguistically or, more generally, to newer (low frequency) constructions.

9. Conclusion: Grammaticalization and changes in linguistic conditioning

In summary, we have seen that increasing frequency and cohesion (Section 3) are accompanied by changes in the linguistic conditioning of a grammaticalizing construction (Section 6). For the Spanish Progressive in variation with the simple Present, while there is stability in direction of effect, which we take as a measure of ‘retention’ (‘persistence’) in grammaticalization, we observed changes in relative magnitude of effect over time. The early strength of the favoring effect of co-occurring locatives supports the locative-source-construction hypothesis for progressives (Bybee et al. 1994:127–133). But comparing variation patterns in between

the beginning and endpoints of grammaticalization gives us a fine view of the locative > progressive grammaticalization path. In particular, the strengthening of the stativity factor group after the Old Spanish period, and analogous reports for English, leads to the hypothesis that the restriction against use of progressive forms with statives is a later development (Section 7). We have also seen that the dominance of the polarity — sentence type constraint within limited duration aspectual contexts in the 17th and 19th century suggests that negatives and interrogatives are conservative contexts (Section 8).

It is important that the development of the Progressive as an aspectual expression of limited duration is gradual. The Progressive does not enter the system as a complete progressive, but is used in habitual contexts in the beginning, when locative meaning is stronger, more than in later periods (Section 6). An aspectual opposition with the simple Present develops gradually, as evaluated by the shift in relative magnitude of effect.

Linguistic change has been said to arise (non-teleologically) as speakers talk in such ways as to express more elaborate nuances (Givón 1979:21) or draw attention (cf. Keller 1990/1994: 101, Haspelmath 1999: 1057, Croft 2000: 74).³⁷ Even among proponents of gradualness, change has been said to involve discrete albeit small micro-steps (Traugott & Trousdale 2010). But in our study of unreflecting language production data, the new construction does not “arise and become part of the variable linguistic system” (Croft 2000: 54) ready-made, nor are even tiny abrupt leaps observable. What linguists think of as ‘innovation’ may well be our posterior depiction of a near-completed or well-advanced process rather than something we can observe, at least in grammaticalization, if the grammatical functions of constructions evolve gradually.

Rather than ‘how do variants arise?’, the question we tackled here was ‘how do variants evolve?’. We answered this empirically, beginning with a broad definition of the variable context as the whole present temporal reference domain, locating constraints on the variable and inferring effects from specific kinds of observed results (Section 5) (rather than ascribing meanings or speaker motivations (as cautioned against by Poplack & Dion 2009)).

Grammatical structure has been argued to emerge from patterns of language use (e.g., Bybee 2010, Hopper 1987). In this study, we have examined the patterns of speakers’ choices between a grammaticalizing construction and the older expression with which it alternates from the beginning. That is, rather than thinking of grammaticalization as resulting in variability (“layering”, per Hopper (1991: 22–24)), we started from the inherent variability of speech (Labov 1969) to test hypothesized grammaticalization processes.

We saw growing aspectual differentiation within the domain of present imperfective, with the Progressive conventionalizing as a progressive (limited duration)

and the simple Present as a nonprogressive (extended duration). This suggests that in the course of the recurrent choices which speakers make in discourse between functionally overlapping variant forms (Sankoff 1988a), the variants themselves may evolve, gradually developing different functions.

Change was measured here by relative strength of effect in the multivariate analyses, while the direction of effect has (so far) remained stable. We know that this is not always the case. In the expression of future time in Spanish, where the newer *ir a* 'go to' + Verb_{Infinitive} periphrasis shows a fifteen-fold rate increase relative to the older synthetic (*-ré*) Future between 17th and 20th c. data sets, Aaron (2010) reports change both in relative magnitude and direction of effect. As with the Progressive, there is weakening of effects traceable to source construction meaning, such as early favoring by motion verbs, but unlike the Progressive, there is reversal of an early disfavoring effect of negative polarity on the future periphrasis in the 20th century. In the Brazilian Portuguese future paradigm, in which four variants have appeared and disappeared over the past five centuries, Poplack and Malvar (2007) also find that direction of effect has not remained constant, for example for adverbial specification. On the other hand, for the French periphrastic and synthetic futures, there is no change in direction of effect between the 19th and 20th century (Poplack & Dion 2009).

The form taken by change in linguistic conditioning thus does not appear to be a question of whether there are two or multiple variants (Poplack & Malvar 2007: 159). It may have to do with the size or speed of frequency increases in the periods examined, which may in turn have to do with whether the change lies in the rise and fall of forms within an existing grammatical category or the rise of a new grammatical category. In the case of the Romance 'go' -based future, change is the gradual replacement of one variant by another. It is important that none of the above-mentioned studies of Romance futures finds evidence for general functional differentiation of the variants — as expressions of future tense — in terms of proximity, certainty, intention, or other oft-ascribed meaning features. In the case of the Spanish progressive, where change is a new grammatical category, gradualness is achieved through increasing functional differentiation of the new and the old — but persisting — variant. More studies on the relationship between changing rates and conditioning (direction and size of effects) will enable us to determine the applicability of these findings to different kinds of linguistic change.

Notes

1. Based on typological findings, it is appropriate to distinguish between similar functions across languages, such as progressive, denoted here by lower case, and the language-specific

constructions encoding them, such as the Progressive *estar* + Verb-*ndo*, denoted by capitalization (Comrie 1976: 10; Croft 2001: 183).

2. Word counts, shown within [], are estimated from the Word Count of electronic texts or from the manual count of words per printed page, with adjustments for extraneous material such as introductory notes, headings, chapter numbers. Unless otherwise indicated, the electronic texts were downloaded from Biblioteca Virtual Miguel de Cervantes, <http://www.cervantesvirtual.com/>.

3. The proportions of adjacent tokens are lower for non-Present tense *estar* + Verb-*ndo* (comparing adjacency figures for all *estar* forms: 36% (13th c.), 50% (15th c.), 67% (17th c.), 78% (19th c.), from Bybee and Torres Cacoullós 2009, Table 4). This suggests that the progressive grammaticalizes in the present before the past tenses.

4. Even in present-day sociolinguistic interviews conducted in Puerto Rico (Cortés Torres 2005), the frequency of the Progressive relative to the simple Present is 5% (22/427). For comparison, the rate of the English Progressive in present temporal reference is 8–14% in a study of Early African American English varieties (Walker 2001: 20) and 9% in American English conversational data (Bybee 2010: 179).

5. Tokens extracted by lexical type make up 69% (429/626), 67% (560/833), and 58% (663/1143) of the simple Present sample, in the Old Spanish, 17th c., and 19th c., respectively.

6. Tokens of *¿qué dice(s)?* (PRS) ‘what are you saying?’ were included from 19th c. texts in which *¿qué está(s) diciendo?* (PROG) appears (e.g., *El estómago*, Act III, Scene IV).

7. Besides the token frequency increase (Table 1), we observe increase in frequency relative to the simple Present among lexical types in which the Progressive appears (i.e., based on the lexical subsamples of the simple Present, after exclusions; see Section 4.2): from 14% (39/282) in the 15th c. *Corbacho* and *Celestina* (not counting *decir* ‘say’ in the *Corbacho*) to 24% (180/744) in the 17th c. and 32% (317/980) in the 19th c. data.

8. Though on an example-by-example basis we find instances of *estar* + Verb-*ndo* with spatial meaning as well as instances with aspectual meaning in any given time period, even overlapping in the same token, degree of grammaticalization is manifested in changing distribution and co-occurrence patterns in the aggregate (Poplack and Tagliamonte 1999; Torres Cacoullós 1999).

9. An alternative hypothesis is that the precursor of (Romance) progressives is a stative construction meaning “being (i.e., finding oneself/itself) in a state” in which the gerund could function as an adjective and which could occur in habitual contexts early on (Bertinetto 2000: 563–564, Bertinetto et al. 2000: 538–539). According to Bertinetto (2000: 566–568), whereas Italian *stare* + gerund has a ‘focalized’ interpretation, the Spanish Progressive represents an earlier ‘durative’ stage, as shown by its compatibility with perfective morphology (for example, *estuvo* (be. PERF.3SG) *leye-ndo* (read-GER) *durante dos horas* ‘he read for two hours’). By this reasoning the ratio of Preterit (perfective) to Imperfect *estar* + Verb-*ndo* should be higher in earlier times, but we find an increase, from 1:10 in Old Spanish texts to 1:2 in a corpus of present-day spoken Spanish (Torres Cacoullós 2000: 153 (Table 34)).

10. Of the 99 stative lexical types, 16 appear in all three periods and make up close to two-thirds of the tokens (63%, 576/916); 24 in two of the three periods, and 10, 12, and 37 only in the Old

Spanish, 17th century, and 19th century data, respectively. More stative lexical types appeared in the 19th c. (75) than in the 17th c. (46) and in Old Spanish (34). Nevertheless, the proportion of stative tokens is between one-fourth and one-third in all three data sets (26% (193/745) in Old Spanish, 31% (310/1013) in the 17th c., and 28% (413/1460) in the 19th c.).

11. Other lexical types coded as stative: *aborrecer* ‘despise’, *acordarse* ‘remember’, *admirar* ‘astonish’, *amar* ‘love’, *aparecer* ‘appear’, *atreverse* ‘dare’, *cabrer* ‘fit’, *carecer* ‘lack’, *cavilar* ‘ponder’, *comprender* ‘comprehend’, *conocer* ‘know’, *creer* ‘believe’, *cuadrar* ‘suit’, *cumplir* ‘satisfy’, *dar ganas* ‘feel like’, *deber* ‘owe’, *desamar* ‘hate’, *desdeñar* ‘scorn’, *doler* ‘pain’, *despreciar* ‘scorn’, *dudar* ‘doubt, hesitate’, *estimar* ‘esteem’, *extrañar* ‘find strange’, *faltar* ‘lack’, *hacer* in weather expressions (e.g., *hace frío* ‘it is cold’), *hallarse* ‘find oneself’, *ignorar* ‘be ignorant of’, *importar* ‘be important’, *llamar(se)* ‘(be) named’, *llevar* ‘have’, *merecer* ‘deserve’, *necesitar* ‘need’, *pecar de* ‘be guilty of’, *pesar* ‘grieve’, *placer* ‘please’, *poseer* ‘possess’, *preocupar* ‘worry’, *presumir* ‘presume’, *pretender* ‘aspire to’, *procurar* ‘endeavor’, *profesar* ‘profess’, *quedar* ‘remain’, *rabiar* ‘be dying to’, *recelar* ‘doubt, suspect’, *representar* ‘represent’, *semejar* ‘seem’, *sentar* ‘suit’, *servir para* ‘serve as’, *sobrar* ‘be in excess’, *sospechar* ‘suspect’, *suponer* ‘suppose’, *traer* ‘have’, *tratarse de* ‘be about’, *valer* ‘be worth’, *vivir* ‘live’.

12. Also excluded from the aspect factor group were narrative clauses or simultaneously reported events (cf. Walker 2001: 11), as in stage instructions (1), and performatives or formulaic speaker stance fragments (Thompson 2002) (2), mostly in the simple Present (under 10% of all tokens).

- (1) a. *stá llamando a TRISTAN*
‘he calls TRISTAN’ (Celestina, Act XIII)
- b. *corre hacia ella*
‘he runs toward her’ (Amor de padre, Act I, Scene III)
- (2) a. *Nada, nada, señora, estoy viendo que a lo mejor esos tunantes asaltan la casa*
‘Nothing, nothing, ma’am, I see that those rogues might attack the house’
(Perfecta, Ch. XXV)
- b. *Vamos, veo que te has enfadado*
‘Well, I see that you’ve gotten angry’ (Perfecta, Ch. IX)

13. Thanks to Joe Bauman for recoding aspect for the reliability test. (Particular works were sampled so that the coder was familiar with the context of the tokens: Old Spanish, *Celestina*; 17th c., *Quijote*; 19th c., *CN*, *Sí*.)

14. Present habitual includes generic sentences (Bybee et al: 1994:126), for example, *el corac'o<n> dell omne siempre esta bulliando* (EE II, fol. 74v) ‘man’s heart always is stirring’, *el grand amor todas las cosas ue<n>ce* (EE II, fol. 75r) ‘great love vanquishes all things’ (only two Progressive tokens).

15. Several scholars ascribe a nuance of transitoriness to Progressive-form habituals, for example, Blansitt’s (1975: 3) “generic progressive”, Comrie’s (1976: 50, n.) “contingent habitual”, Bull’s (1965: 164) notion of deviation from the norm. Torres Cacoullous (2000: 188–204) adduces some evidence for subjectivity in “experiential habitual” uses from the proportion of *estar* + Verb-*ndo* with first or second person subjects or object pronouns and deictic time or place adverbials.

16. This might be an example of a more fine- (vs. coarse-) grained construal of a regularly repeated situation (Croft 2009: 157–158), or a “focalized habitual” as in *Whenever I arrive, he is*

writing (Bertinetto 2000: 570), or “the habitual of a progressive” where “the sum total of all these [individual progressive] occurrences is presented as being habitual” (Comrie 1976: 33).

17. For English, Walker (2001: 24, Table 6) reports 19% (822/4381) habitual statives in the present tense and Jarmasz (2006, Table 31) 7% (161/2204) in Past and Present forms. The following is one of the few examples in these data:

- (1) *magistrados, catedráticos, autoridades, abogados, hasta clérigos, están deseando todo el día, sin darse cuenta, la hora de las tiendas los días que hace bueno y pueden las damas «decorosamente» coger la mantilla y echarse a la calle.* (Regenta I, Ch. IX)
 ‘magistrates, professors, authorities, solicitors, even clergymen, are desiring (PROG) all day, without realizing it, the hour when the shops are open on the days when the weather is good and the ladies can “decorously” take their mantilla and go out onto the street [...]’

18. Consonant with the status of the Present as the default present tense form is that 92% (315/343) of indeterminate aspect tokens are in the Present, in the 19th century data. The proportion of aspectually indeterminate or uncodable tokens was significantly smaller for the Progressive than the simple Present in the 17th (12%, 21/180 vs. 27%, 217/833) and 19th century (9%, 28/317 vs. 28%, 315/1143), though not in the Old Spanish data (19%, 22/114 vs. 25%, 160/631). The smaller proportion of indeterminate Progressive tokens supports the grammaticalization prediction that the meaning of a newer construction is more specific than that of the alternative older one (Bybee et al. 1994: 133).

19. Subordination may be gradient discourse-pragmatically (Thompson 2002). Coding by formal criteria, we find no difference in Progressive rate between main, relative, and subordinate clauses in Old Spanish (16% (52/323), 15% (26/179) and 25% (13/52)) and in the 19th century (24% (175/738), 22% (58/260) and 25% (30/118)); in the 17th century, subordinate showed a significantly higher rate (30%, 21/70) than relative (18%, 45/253) but not main clauses (20%, 70/347).

20. For factors relating to the subject, in prior analyses the third person favored the Progressive in Old Spanish but no longer in 19th century data, nor was grammatical person significant in a Variable-rule analysis of Puerto Rican sociolinguistic interview data (Cortes-Torres 2005: 50). In the present data we find that grammatical person shows skewed distributions with respect to aspect (the proportion of first and second person tokens in limited duration contexts is one and a half to three times greater than for third person) and locatives (third person subjects are two to three times more likely to co-occur with locative expressions (11%, 119/1113) than first and second person (4%, 58/1341)). As to subject animacy, in Old Spanish inanimate subjects with the Progressive tend (3/6) to be metonymically human (e.g., *sus animas* ‘their souls’) and in the 17th century most (8/13) are figuratively personified (such as a favorable wind knocking on the door) or (3/13) co-occur with a human participant in the Dative experiencer construction (*me están bullendo los pies* ‘my feet are itching’, *Quijote* II, Ch. LII). In the 19th century, though most inanimate subjects still co-occur with a human participant (11/26) or are personified (10/26), there are cases of more-inanimate-like subjects (*se está poniendo el sol* ‘the sun is setting’, *Álvaro*, *Jornada* I). The earlier restriction to personified inanimate subjects would be consonant with the nuance of active involvement of the subject in the activity, which has been attributed to (early) progressives (Bybee et al 1994: 133–137).

21. Variable-rule analysis uses logistic regression to perform binomial multivariate analysis for a choice of the “1” variant (here, the Progressive) vs. the “0” variant (the simple Present) (for a short guide to the GoldVarb application, see Walker 2010a: 38–43; on the premises and interpretation of Variable-rule analysis, see Sankoff 1988b). The procedure determines the factor groups 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. When we compare the chosen analysis with the null model (using a chi-square value equal to twice the difference between log likelihoods and a number of degrees of freedom equal to the number of factors minus the number of factor groups) we obtain a χ^2 of 72.18, 190.11, and 380.63, with a df of 4, for the Old Spanish, 17th c., and 19th c., respectively, all with $p < .005$.

22. The ‘input’ indicates the overall likelihood that the Progressive will occur. Since only a sample of the simple Present was taken, comparisons of input values in these analyses are meaningful for subsets within each data set, as in the analyses of dynamic vs. stative verbs in each time period (Table 6), but not across time periods.

23. The distribution of the data by factor is about the same across the three data sets, except for the higher proportion of extended duration contexts in Old Spanish (due to genre).

24. Separate multivariate analyses of the two component subsamples of the Present (Section 4.2) show no difference in direction of effect in any of the factor groups. The subsamples yield a difference in the relative magnitude of the effect of stativity, such that extracting Present tokens by vicinity yields a higher proportion of stative verbs than extracting by lexical type (at least twice as high, about one third of the data) and the stativity effect is magnified. It seems that sampling by vicinity emphasizes lexical (inherent aspect) constraints whereas a simple Present sample of lexical types with which the Progressive actually occurs heightens contextual effects, in particular, that of co-occurring locative expressions. Nevertheless, whatever skewing the sampling procedure entails, it is about the same across the comparison time periods, and over time the relative magnitude of the stativity effect in both samples increases and the co-occurring locative effect in the lexical subsample diminishes (as does the difference between the two subsamples). With respect to the distribution of aspectual contexts, limited duration contexts make up only slightly more of the data in the vicinity subsample (50–72% vs. 46–70%) and the relative magnitude of effect is slightly greater in the lexical sample, but this effect strengthens over time in both cases. There are no distribution or magnitude of effect differences in the remaining factor groups (polarity-sentence type and temporal expression co-occurrence).

25. As a locative, the early Progressive may have a presentative function (cf. Dumont 2006: 295). In Old Spanish, *estar* + Verb-*ndo* is favored by subjects that are full NPs (the form in which new-information referents are typically realized (e.g., Bentivoglio 1993)) and appear postverbally (the position favored by new (first-mention full NP) referents (Bentivoglio & Weber 1986)) (Torres Cacoullos 2009: 219–220).

26. The *siempre* ‘always’ + Progressive habitual may express speaker point of view, such as nuances of disapprobation (cf. Smitterberg 2005, Chapter 7).

27. Figure 2 is based on the pooled data (see Table 9, in Appendix). Distributions are similar across the three periods, although the proportion of extended duration contexts is higher in the Old Spanish data due to genre.

28. In contrast, Walker (2001:24) reports that statives are associated with non-habitual aspect. The discrepancy is due to different coding for aspect: whereas here the extent of time was considered, in that study the criterion was repeated occurrence, such that “states that exist continuously” were coded as “durative” and opposed to habitual (Walker 2001:19).

29. Of the stative lexical types listed in Section 5.3 and footnote 11, the following appear in the Progressive: Old Spanish *adorar* ‘adore’, *callar* ‘be silent’ (6), *comedir* ‘think’ (2), *cuidar* ‘think’ (2), *desear* ‘desire’, *dubdar* ‘waver’ (2) *folgar* ‘rest’ (2), *penar* ‘suffer’, *pensar* ‘think’ (5), *perseverar* ‘persevere’; 17th c., *adorar* (2), *callar*, *desear* (2), *entender* ‘understand’, *gozar* ‘enjoy’ (4), *holgar*, *imaginar* ‘imagine’ (2), *padecer* ‘suffer’, *penar* (3), *pensar* (3), *recelar* ‘fear, suspect’, *sentir* ‘feel’, *temer* ‘fear’ (4), *vivir* ‘live’; 19th c. *afligir* ‘afflict’ (3), *cavilar* ‘ponder’, *desear* (3), *figurar* ‘seem’, *gozar*, *hacer día de verano* ‘be a summer day’, *hacer daño* ‘be harmful’, *hacer falta* ‘lack’, *holgar*, *padecer* (3), *parecer* ‘seem’, *pensar* (7), *recordar* ‘remember’, *sentir*, *ser* ‘be’ (2), *sufrir* ‘suffer’ (5), *temer*.

30. Weights for non-significant factor groups are from the first “stepdown” run in GoldVarb, in which all factors are included in the regression.

31. The restriction against Progressive use with statives may be more one of particular lexical types and their constructions than one of meaning. For example, there seems to be no aspectual difference between ‘knowing’ and ‘reading’ from a face in (a) and (b) (besides one that would be based on the circular argument that the tense-aspect forms signal different meanings).

- (1) a. *En la cara se te conoce que estás mintiendo*
 ‘From your face one knows (PRS) that you are lying’
 (Amor de padre, Act III, Scene VI)
- b. *Os estoy leyendo en el semblante lo que está asando en vuestra alma...*
 ‘I am reading (PROG) on your face what is broiling in your heart’
 (Amor de padre, Act IV, Scene V)

32. Temporal co-occurrence is selected as significant in the analysis of the combined Old Spanish data (Table 4, Section 6), where — for both dynamic and stative verbs — co-occurring temporals favor the Progressive in extended duration contexts, which furthermore make up a greater proportion of the data than in the 17th and 19th centuries.

33. Of course, there is variability, i.e. no one-to-one relation between form and function; a 19th c. stative in the Progressive with habitual meaning is illustrated in footnote 17.

34. The CORDE search was for 1600–1650 and 1850–1900, in all available Peninsular Spanish corpora (*en todos los medios, en ESPAÑA*).

35. In comparison, the *ir* ‘go’ + Verb *-ndo* construction with statives is aspectual from the earliest texts, with a meaning of gradual development, for example, *va sie-ndo* [go.PRS.3SG be-GER] ‘gradually becomes’ (Torres Cacoullós 2000:99–102).

36. Three of the six 19th c. negative polarity Progressive tokens (in three different plays) are the second person interrogative formula *¿no está(s) viendo?* ‘don’t you see?’.

37. The hypothesis of an attention-getting function for new grammaticalizing constructions (cf. Haspelmath’s (1999) “extravagance”) may be operationalized in a linear order effect such that choice of the Progressive is favored in the first of two or more juxtaposed present tense forms, as in the example. I found this effect in 15th c. but not 19th c. data.

- (1) lo puedes echar de ver, Sancho, en que *está partiendo* la capa con el pobre y le *da* la mitad; (Quijote II, Ch. LVIII)
 ‘you can see it, Sancho, in that he is **dividing** (PROG) his cape with the poor man and gives (PRS) him half’

Corpus

Old Spanish

- Cid (circa 1140) = Anonymous. *Cantar de mio Cid. Texto, gramática y vocabulario*, Ramón Menéndez Pidal (ed.), Vol. 3. Madrid: Espasa Calpe, 1944. [30,000]
- Calila (1250) = Anonymous, *Calila e Dimna*, Juan Manuel Cacho Bleuca & María Jesús Lacarra (eds.). Madrid: Castalia, 1984. [86,000]
- FG (1250) = Anonymous, *Poema de Fernán González*. [29,000]
- Apolonio (circa 1260) = Anonymous, *Libro de Apolonio*. [20,000]
- Sacrificio = Gonzalo de Berceo, *Sacrificio de la misa*. [4,000]
- GEI (1260–1280) = Alfonso X, *General estoria. Primera parte*, Antonio G. Solalinde (ed.). Madrid: Centro de Estudios Históricos, 1930. [572,000]
- GEII (1260–1280) = Alfonso X, *General Estoria. Segunda parte*, 2 vols., Antonio Solalinde, Lloyd Kasten & Victor R.B. Oelshläger (eds.). Madrid: CSIC (Consejo Superior de Investigaciones Científicas), 1957. [263,500]
- EE1 (1270–1284); EE2 (1284?-1345?) = Alfonso X, *Estoria de España. The Electronic Texts and Concordances of the Prose Works of Alfonso X, El Sabio*, Lloyd Kasten, John Nitti, and Wilhemina Jonxis-Henkemans (eds.). Hispanic Seminary of Medieval Studies. [279,530; 386,785]
- Leyes (circa 1260) = Alfonso X, *Libro de las leyes. The Electronic Texts and Concordances of the Prose Works of Alfonso X, El Sabio*, Lloyd Kasten et al. (eds.). Hispanic Seminary of Medieval Studies. [165,000]
- Zifar (1310/1320) = Anonymous, *El libro del caballero Zifar*. [34,000]
- LBA (1330–1343) = Arcipreste de Hita, *Libro de Buen Amor*. [57,000]
- Lucanor (1335) = Don Juan Manuel, *El conde Lucanor o Libro de los enxiemplos del conde Lucanor et de Petronio*, José Manuel Bleuca (ed.). Madrid: Castalia, 1971. [79,000]
- Corbacho (1438) = Alfonso Martínez de Toledo, *Corbacho*, Michael Gerli (ed.). Madrid: Cátedra, 1978. [75,000]
- Grimalte (1486) = Juan de Flores, Juan, *Grimalte y Gradissa*, Pamela Waley (ed.). London: Tamesis, 1971. [25,500]
- CRC (1482–1490) = Hernando del Pulgar, *Crónica de los Reyes Católicos*, 2 vols., Juan de Mata Carriazo (ed.). Madrid: Espasa Calpe, 1943. [322,000]
- Celestina (1499) = Fernando de Rojas, Fernando. 1987. *La Celestina*, Dorothy S. Severin, Madrid: Cátedra, 1987 [67,000]; Translation: *La Celestina or The Spanish Bawd: Being the Tragi-Comedy of Calisto and Melibea*, J. M. Cohen, Penguin, 1964.
- Sonetos Íñigo López de Mendoza, marqués de Santillana. [3,000]
- Enzina (1496–1508) = Juan del Enzina, Aucto del repelón; Égloga de Cristino y Febea; Égloga de Fileno, Zambardo y Cardonio; Égloga de las grandes lluvias; Égloga de Mingo, Gil y Pascuala; Égloga de Plácida y Vitoriano; Égloga representada en la misma noche de Navidad; Égloga representada en la noche de la Natividad; Égloga representada en la noche

postrera de Carnal; Égloga representada en requesta de unos amores; Égloga representada la misma noche de Antruejo; Representación a la Pasión y muerte de Nuestro Redentor; Representación a la santísima Resurrección de Cristo; Representación sobre el poder del Amor. [32,000]

17th century

- Quijote = Miguel de Cervantes Saavedra (1547–1616), *Don Quijote de la Mancha*. [384,000]; Translation: Edith Grossman, Harper Collins, 2003.
- Ardenia = Cervantes, *Comedia famosa de La casa de los celos y selvas de Ardenia*. [17,000]
- Dama boba = Lope de Vega (1562–1635), *La dama boba*. [18,000]
- Inocente = Lope de Vega, *Comedia del Príncipe Ynocente*. [15,000]
- Vengadora = Lope de Vega, *La vengadora de las mujeres* [15,000]
- Amor constante = Guillén de Castro (1569–1631), *El amor constante*. [19,000]
- Amistad = Ruiz de Alarcón y Mendoza (1581–1639), *La Amistad castigada*. [14,000]
- Gil = Tirso de Molina (1583?-1648), *Don Gil de las calzas verdes*. [19,000]
- Sótano = Tirso de Molina, *Por el sótano y el torno*. [19,000]
- Discretos = Tirso de Molina, *Amor y celos hacen discretos* [Act I]. [5,000]
- Villana = Tirso de Molina, *La villana de Sagra*. [18,000]
- Boca = Gaspar de Ávila, *La boca y no el corazón o Fingir por conservar*. (1632) [13,000]
- Casa = Calderón de la Barca (1600–1681), *Casa con dos puertas mala es de guardar*. [17,000]
- Dama duende = Calderón de la Barca, *La dama duende*. [18,000]
- Amor, honor y poder = Calderón de la Barca, *Amor honor y poder*. [15,000]

19th century

- CN, Sí (1790–1810) = Fernández de Moratín, Leandro. 1975. *La comedia nueva. El sí de las niñas*, J. Dowling & R. Andioc (eds.). Madrid: Castalia. [34,000]
- Madrid (1828) = Manuel Bretón de los Herreros, *A Madrid me vuelvo*. [13,000]
- Afán (1831) = Carnerero, José María de. *El afán de figurar: comedia en cinco actos, en verso*. [11,000]
- Acertar (1832) = Ventura de la Vega, *Acertar errando, o El cambio de diligencia Comedia en tres actos*. [14,000]
- Inseparables = José de Mariano Larra (1809–1837), *Los inseparables*. [7,000]
- Álvaro (1835) = Duque de Rivas, *Don Álvaro o la fuerza del sino; Drama original en cinco jornadas, y en prosa y verso*. [22,000]
- Boda (1839) = Francisco Martínez de la Rosa. *La boda y el duelo*. [11,000]
- Amor de padre (1849) = Francisco Martínez de la Rosa. *Amor de padre*. [18,000]
- Presupuestos (1852) = Pablo Alonso de la AVECILLA, *Los presupuestos*. [16,000]
- Paloma (1857) = Luis Mariano de Larra (1830–1901). *La paloma y los halcones*. [15,000]
- Piedra (1862) = Luis Mariano de Larra. *La primera piedra. Drama en tres actos y en verso*. [14,000]
- Mal (1868) = Manuel Tamayo y Baus. *No hay mal que por bien no venga*. [15,000]
- Pepita (1870) = Juan Valera, *Pepita Jiménez*. [56,000]
- Perfecta (1876) = Benito Pérez Galdós, *Doña Perfecta*. [65,000]
- Regenta (1870–1880) = Leopoldo Alas “Clarín”, *La Regenta*, vol I & II. [141,000, 162,000]

- Pazos (1886) = Pardo Bazán, Emilia. *Los pazos de Ulloa*. [83,000]
 Siete (1889) = Julio Cuevas, *¡El siete! : juguete cómico en un acto y en prosa /original de Don Julio Cuevas y Don Manuel Labra*. [5,000]
 Enrique Gaspar (1842–1902), *Don Ramón y El señor Ramón* [16,000]; *El estómago* [21,000]; *La gran comedia* [16,000]; *La lengua* [21,000]; *La levita* [17,000]; *Las circunstancias* [15,000]; *Lola* [17,000]; *Los niños grandes* [14,000]; *Problema* [17,000]; *Serafina la devota* [28,000].

References

- Aaron, Jessi Elana. 2010. Pushing the envelope: Looking beyond the variable context. *Language Variation and Change* 22(1). 1–36.
- Bentivoglio, Paola. 1993. Full NPs in spoken Spanish: A discourse profile. In William J. Ashby, Marianne Mithun, Giorgio Perissinotto & Eduardo Raposo (eds.), *Linguistic perspectives on the Romance languages*, 212–224. Amsterdam: John Benjamins.
- Bentivoglio, Paola & Elizabeth Weber. 1986. A functional approach to subject word order in spoken Spanish. In Osvald Jaeggli & Carmen Silva-Corvalán (eds.), *Studies in Romance linguistics: Selected papers from the Fourteenth Linguistics Symposium on Romance Languages*, 23–40. Berlin: Mouton de Gruyter.
- Bertinetto, Pier Marco. 1994. Statives, progressives, and habituals: analogies and differences. *Linguistics* 32(3). 391–423.
- Bertinetto, Pier Marco. 2000. The progressive in Romance, as compared with English. In Östen Dahl (ed.), *Tense and aspect in the languages of Europe*, 559–604. Berlin: Walter De Gruyter.
- Bertinetto, Pier Marco, Karen H. Ebert & Casper de Groot. 2000. The progressive in Europe. In Östen Dahl (ed.), *Tense and aspect in the languages of Europe*, 517–558. Berlin: Walter de Gruyter.
- Blansitt, Edward L., Jr. 1975. Progressive aspect. *Stanford University Working Papers on Language Universals* 18. 1–34.
- Bull, William E. 1965. *Spanish for teachers: Applied linguistics*. New York: The Ronald Press Company.
- Bybee, Joan. 1994. The grammaticization of zero: Asymmetries in tense and aspect systems. In William Paglicua (ed.), *Perspectives on grammaticalization*, 235–254. Amsterdam: John Benjamins.
- Bybee, Joan. 2010. *Language, usage and cognition*. Cambridge: Cambridge University Press.
- Bybee, Joan & William Pagliuca. 1987. The evolution of future meaning. In Anna Giacalone Ramat, Onofrio Carruba & Giuliano Bernini (eds.), *Papers from the Seventh International Conference on Historical Linguistics (ICHL)*, 109–122. Amsterdam: John Benjamins.
- Bybee, Joan & Rena Torres Cacoullous. 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 & Kathleen Wheatley (eds.), *Formulaic language: Volume 1. Distribution and historical change* (Typological Studies in Language (TSL) 82), 187–217. Amsterdam: John Benjamins.
- Bybee, Joan, Revere Perkins & William Pagliuca. 1994. *The evolution of grammar: Tense, aspect, and modality in the languages of the world*. Chicago: University of Chicago Press.
- Claudi, Ulrike & Bernd Heine. 1986. On the metaphorical base of grammar. *Studies in Language* 10(2). 297–335.

- Clegg, J. Halvor. n.d. English syntactic interferences in Puerto Rican Spanish. Brigham Young University, ms.
- Company Company, Concepción (ed.). 2006. *Sintaxis histórica de la lengua española. Primera parte: La frase verbal, Vol. 1*. México: Universidad Nacional Autónoma de México & Fondo de Cultura Económica.
- Comrie, Bernard. 1976. *Aspect*. Cambridge: Cambridge University Press.
- Cortes-Torres, Mayra. 2005. ¿Qué estás haciendo?: La variación de la perífrasis *estar + -ndo* en el español puertorriqueño. In David Eddington (ed.), *Selected Proceedings of the 7th Hispanic Linguistics Symposium*, 42–55. Somerville, MA: Cascadilla Proceedings Project.
- Croft, William. 2000. *Explaining language change*. London: Longman.
- Croft, William. 2001. *Radical construction grammar: Syntactic theory in typological perspective*. Oxford: Oxford University Press.
- Croft, William. 2009. Aspectual and causal structure in event representations. In Virginia Gathercole (ed.), *Routes to language development: In honor of Melissa Bowerman*, 139–166. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Croft, William. 2010. The origins of grammaticalization in the verbalization of experience. *Linguistics* 48(1). 1–48.
- Dahl, Östen. 1985. *Tense and aspect systems*. Oxford: Blackwell.
- Dumont, Jenny. 2006. Full NPs as subjects. In Almeida Jacqueline Toribio & Nuria Sagarra (eds.), *Selected Proceedings of the 9th Hispanic Linguistics Symposium*, 286–296. Somerville, MA: Cascadilla Proceedings Project.
- Emirikian, Louise & David Sankoff. 1985. Le futur simple et le futur périphrastique dans le français parlé. In Monique Lemieux & Henrietta Cedergren (eds.), *Les tendances dynamiques du français parlé à Montréal*, 189–204. Québec: Gouvernement du Québec.
- Fleischman, Suzanne. 1982. *The future in thought and language: Diachronic evidence from Romance*. Cambridge: Cambridge University Press.
- Freed, Alice. 1994. The form and function of questions in informal dyadic conversation. *Journal of Pragmatics* 21. 621–644.
- García, Erica. 1990. Reanalysing actualization and actualizing reanalysis. In Henning Andersen & Konrad Koerner (eds.), *Historical Linguistics 1987: Papers from the 8th International Conference on Historical Linguistics* (8 ICHL), 141–159. Amsterdam: John Benjamins.
- Givón, Talmy. 1979. *On understanding grammar*. New York: Academic Press.
- Harvie, Dawn. 1998. Null subject in English: Wonder if it exists? *Cahiers Linguistiques d'Ottawa* 16. 15–25.
- Haspelmath, Martin. 1999. Why is grammaticalization irreversible? *Linguistics* 37(6). 1043–1068.
- Heine, Bernd & Tania Kuteva. 2002. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.
- Heine, Bernd & Mechtild Reh. 1984. *Grammaticalization and reanalysis in African languages*. Hamburg: Helmut Buske.
- Hilpert, Martin & Stefan Th. Gries. 2009. Assessing frequency changes in multistage diachronic corpora: Applications for historical corpus linguistics and the study of language acquisition. *Literary and Linguistic Computing* 24(4). 385–401.
- Hopper, Paul J. 1991. On some principles of grammaticization. In Elizabeth Closs Traugott & Bernd Heine, *Approaches to grammaticalization*, Vol. 1, 17–35. Amsterdam: John Benjamins.
- Hopper, Paul J. & Elizabeth Closs Traugott. 1993. *Grammaticalization*. Cambridge: Cambridge University Press.

- Jarmasz, Lidia-Gabriela. 2005. "I think I know what I'm thinking": The alternation between simple and progressive forms in stative verbs. Ottawa: University of Ottawa MA thesis.
- Keller, Rudi. 1990/1994. *On language change: The invisible hand in language*. (English version translated by Brigitte Nerlich). London & New York: Routledge.
- Labov, William. 1969. Contraction, deletion, and inherent variability of the English copula. *Language* 45(4). 715–762.
- Labov, William. 1972. *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.
- Labov, William. 2004. Quantitative reasoning in linguistics. In Ulrich Ammon, Norbert Dittmar, Klaus J. Mattheier & Peter Trudgill (eds.), *Sociolinguistics/Soziolinguistik: An international handbook of the science of language and society*, 2nd edn., vol. I, 6–22. Berlin: Mouton de Gruyter.
- Lehmann, Christian. 1982. *Thoughts on grammaticalization: A programmatic sketch*, vol. I. (Arbeiten des Kölner Universalien-Projekts 48). Köln: Universität zu Köln, Institut für Sprachwissenschaft.
- Lope Blanch, Juan M. (ed.). 1971. *El habla de la ciudad de México: Materiales para su estudio*. México: Universidad Nacional Autónoma de México (UNAM).
- Marcos Marín, Francisco (dir.). 1992. *Corpus de Referencia de la Lengua Española Contemporánea: Corpus Oral Peninsular (COREC)*, <http://www.llff.uam.es/~fmarcos/informes/corpus/corpusix.html> (April 2004).
- Pappas, Panayiotis A. 2001. The microcosm of a morphological change: Variation in *thelō* + infinitive futures and *ēthela* + infinitive counterfactuals in Early Modern Greek. *Diachronica* 18. 59–92.
- Poplack, Shana & Nathalie Dion. 2009. Prescription vs. praxis: The evolution of future temporal reference in French. *Language* 85(3). 557–587.
- Poplack, Shana & Elisabete Malvar. 2007. Elucidating the transition period in linguistic change: The expression of the future in Brazilian Portuguese. *Probus* 19. 121–169.
- Poplack, Shana & Sali Tagliamonte. 1999. The grammaticalization of *going to* in (African American) English. *Language Variation and Change* 11(3). 315–342.
- Poplack, Shana & Sali Tagliamonte. 2001. *African American English in the diaspora: Tense and aspect*. Oxford: Blackwell.
- Real Academia Española. Banco de datos (CORDE) [en línea]. *Corpus diacrónico del español*. <<http://www.rae.es>> (5 February 2007).
- Sankoff, David. 1988a. Sociolinguistics and syntactic variation. In Frederick J. Newmeyer (ed.), *Linguistics: The Cambridge survey*, vol. IV, 140–161. Cambridge: Cambridge University Press.
- Sankoff, David. 1988b. Variable rules. In Ulrich Ammon, Norbert Dittmar & Klaus J. Mattheier (eds.), *Sociolinguistics/Soziolinguistik: An international handbook of the science of language and society*, 984–997. Berlin: Walter de Gruyter.
- Sankoff, David & Pierrette Thibault. 1981. Weak complementarity: Tense and aspect in Montreal French. In Brenda B. Johns & David R. Strong (eds.), *Syntactic change: Natural language studies*, eds. Brenda B. Johns & David R. Strong, 205–216. Ann Arbor: Department of Linguistics, University of Michigan.
- Sankoff, David, Sali Tagliamonte & Eric Smith. 2005. GOLDVARB X: A multivariate analysis application for Macintosh and Windows. <http://individual.utoronto.ca/tagliamonte/Goldvarb/GV_index.htm>.
- Sankoff, Gillian & Penelope Brown. 1976. The origins of syntax in discourse: A case study of Tok Pisin relatives. *Language* 52(3). 631–666.

- Scheibman, Joanne. 2002. *Point of view and grammar: Structural patterns of subjectivity in American English conversation*. Amsterdam: John Benjamins.
- Silva-Corvalán, Carmen. 1994. *Language contact and change: Spanish in Los Angeles*. Oxford: Clarendon Press.
- Squartini, Mario. 1998. *Verbal periphrases in Romance: Aspect, actionality, and grammaticalization*. Berlin & New York: Mouton de Gruyter.
- Smitterberg, Erik. 2005. *The progressive in 19th-century English: A process of integration*. Amsterdam: Rodopi.
- Tagliamonte, Sali. 2002. Comparative sociolinguistics. In J. K. Chambers, Peter Trudgill & Natalie Schilling-Estes (eds.), *Handbook of language variation and change*, 729–763. Oxford: Blackwell.
- Thompson, Sandra A. 1998. A discourse explanation for the cross-linguistic differences in the grammar of interrogation and negation. In Anna Siewierska & Jae Jung Song (eds.), *Case, typology and grammar: In honor of Barry J. Blake*, 309–341. Amsterdam: John Benjamins.
- Thompson, Sandra A. 2002. 'Object complements' and conversation. *Studies in Language* 26(1). 125–164.
- Torres Cacoullós, Rena. 1999. Variation and grammaticization in progressives: Spanish *-ndo* constructions. *Studies in Language* 23(1). 25–59.
- Torres Cacoullós, Rena. 2000. *Grammaticization, synchronic variation, and language contact: A study of Spanish progressive -ndo constructions*. Amsterdam: John Benjamins.
- Torres Cacoullós, Rena. 2001. From lexical to grammatical to social meaning. *Language in Society* 30(3). 443–478.
- Torres Cacoullós, 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 (HLS)*, 37–49. Somerville, MA: Cascadilla Proceedings Project.
- Torres Cacoullós, Rena. 2009. Variation and grammaticisation: The emergence of an aspectual opposition. In Stavroula Tsiplakou, Marilena Karyolemou & Pavlos Pavlou (eds.), *Studies in Language Variation: European Perspectives II. (ICLaVE 4)*, 215–224. Amsterdam: John Benjamins.
- Torres Cacoullós, Rena & James A. Walker. 2009. The present of the English future: Grammatical variation and collocations in discourse. *Language* 85(2). 321–354.
- Tottie, Gunnel. 1991. *Negation in English Speech and Writing*. New York: Academic Press.
- Traugott, Elizabeth Closs & Graeme Trousdale. 2010. Gradience, gradualness and grammaticalization: how do they intersect? In Elizabeth Closs Traugott and Graeme Trousdale (eds.), *Gradience, gradualness and grammaticalization*, 19–44. Amsterdam: John Benjamins.
- Vendler, Zeno. 1967. Verbs and times. *Linguistics in philosophy*, 97–121. Ithaca, NY: Cornell University Press.
- Walker, James A. 2001. Using the past to explain the present: Tense and temporal reference in Early African American English. *Language Variation and Change* 13(1). 1–35.
- Walker, James A. 2010a. *Variation in linguistic systems*. London: Routledge.
- Walker, James A. (ed.). 2010b. *Aspect in grammatical variation*. Amsterdam: John Benjamins.
- Weinreich, Uriel, William Labov & Marvin Herzog. 1968. Empirical foundations for a theory of language change. In Winfred P. Lehmann & Yakov Malkiel (eds.), *Directions for historical linguistics*, 95–188. Austin: University of Texas Press.

Appendix

Table 8. Rate of Progressive by locative co-occurrence in extended vs. limited duration aspectual contexts, by century

		Extended duration % Progressive (N/N)	Limited duration % Progressive (N/N)
OLD SPAN	Locative present	23% (3/13)	64% (16/25)
	absent	10% (27/264)	24% (46/191)
17TH C.	Locative present	24% (4/17)	56% (27/48)
	absent	2% (6/248)	29% (116/399)
19TH C.	Locative present	4% (1/25)	48% (31/64)
	absent	4% (13/326)	41% (234/569)

Table 9. Cross-tabulation of Stativity and Aspect*, by time period

		Limited dur	Extended dur	Indeterminate
OLD SPAN	Dynamic	165	196	129
	<i>Stative</i>	51	81	53
17TH C.	Dynamic	339	177	146
	<i>Stative</i>	111	91	92
19TH C.	Dynamic	526	236	185
	<i>Stative</i>	121	123	158

*Narrative and performative uses excluded

Table 10. Rate of Progressive by temporal expression co-occurrence in extended vs. limited duration aspectual contexts, by century, separately for dynamic and stative verbs

		Dynamic		Stative	
		Extended	Limited dur	Extended	Limited dur
		% Prog (N)	% Prog (N)	% Prog (N)	% Prog (N)
OLD SPAN	Temporal present	25% (40)	11% (9)	30% (10)	40% (5)
	absent	7% (154)	29% (156)	8% (71)	28% (46)
17TH C.	Temporal present	17% (18)	37% (49)	0 (11)	31% (13)
	absent	4% (159)	35% (291)	1% (79)	20% (97)
19TH C.	Temporal present	21% (48)	37% (63)	50% (4)	38% (13)
	absent	2% (188)	47% (458)	0 (118)	23% (105)

Author's address

Rena Torres Cacoullos
Spanish, Italian, and Portuguese / Linguistics
Pennsylvania State University
231 Burrowes Building
University Park, PA 16802
rena@psu.edu