

The diachrony of Spanish *haber/hacer* + time

A quantitative corpus-based approach to grammaticalization

Borja Herce

University of the Basque Country and University of Surrey

Using quantitative corpus evidence from different periods, the present article analyzes the emergence and diachronic development of the Spanish time constructions (clausal and adverbial) involving contemporary *hacer* ‘make’ and earlier *haber* ‘have’. The obtained data, as well as cross-linguistic evidence, suggest that the clausal construction must have been the source of the adverbial one. A proposal is presented that could explain that development. The data show, in addition, that the grammatical properties and usage patterns of the clausal and adverbial constructions were very similar until the 16th century but have been diverging ever since. This divergence coincides with an exponential increase in the textual raw frequency of the adverbial construction, where word order fixation, erosion of the inflectional morphology and a change in the possibilities for time adjunction among others are found to occur at around the same time. This points towards a desentencionalization, loss of inner structure and grammaticalization of the adverbial construction in those periods.

Keywords: temporal *hacer*, Spanish, diachrony, grammaticalization, desentencionalization, corpus linguistics, categorization

1. Introduction

In Modern Spanish, using *hace* ‘make.3.sg’ is the most frequent strategy to locate a certain event in the past at a specific distance from the utterance time. Expressions with this same semantic contribution tend to be quite idiosyncratic cross-linguistically (Plank 2011, Herce 2017a). Many of them (e.g. English *ago*, Italian *fa*, Basque *duela*, Maltese *ilu*, etc.) are also deverbal like *hace* and their synchronic status in the language is often controversial. Despite some attempts (Kurzon 2008) to classify these expressions into the traditional grammatical categories of preposition,

adverb, etc., many of these expressions are grammatically idiosyncratic and not amenable to such classifications (Culicover 1999: 71–74) unless, of course, precedence is arbitrarily given to some of their grammatical traits over others.

The Spanish temporal constructions with *hacer* + time-NP are also quite idiosyncratic and have accordingly been the object of abundant research (e.g. Rasmussen 1981; Pérez Toral 1992; Rigau 1999; Howe 2011; Brucart 2015; Fábregas 2016). They are challenging because of the existence of two variants of that construction, the so-called clausal (1) and adverbial (2) constructions, and because of their ability to express different time relations, so-called distance-past (1)–(2) and up-to-now (3) meanings:

- (1) Clausal construction; distance-past
 Hace dos años que te vi
 make.3.SG two years that you see.PST.1.SG
 ‘It has been two years since I last saw you.’
- (2) Adverbial construction; distance-past
 Te vi hace dos años
 you see.PST.1.SG make.3.SG two years
 ‘I saw you two years ago.’
- (3) Clausal construction; up-to-now
 Hace dos años que trabaja aquí
 make.3.SG two years that work.3.SG here
 ‘(S)he has been working here for two years.’

As illustrated by the previous examples, *hacer* appears to be the main verb in the clausal construction, whereas in the adverbial construction *hacer* introduces a syntactically optional constituent. There are many disagreements concerning the status (i.e. grammatical category) of *hacer* in these constructions, especially in the adverbial (RAE-ASALE 2009: 1837), as well as the synchronic relation, if any, between the two constructions (see e.g. Rigau 1999 and Brucart 2015).

It is my contention that the two constructions should better be regarded as largely autonomous synchronically, since their grammatical properties (i.e. allowed verbal morphology, word order possibilities, time adjunction, negation) are quite divergent (Herce 2017b). Concerning categorization, my proposal is that adverbial construction *hace* is better understood as being located somewhere along the verb-adposition continuum (see e.g. Haspelmath 1998: 330) since it has both verbal and adpositional characteristics. It is hardly surprising, therefore, that equivalent deverbal expressions like French *il y a* were included by Hagège (2010: 301–302) in his monograph about adpositions.

An exhaustive analysis of the synchronic grammatical properties of time constructions with *hacer* is beyond the scope of the present article, however. Here the goal will be to provide a picture of their diachronic properties and use, which will in turn provide information about the contemporary status of these constructions in the language. Section 2 will address the origin of the time constructions analyzed in this article. It will present the time relations which these constructions can express (distance-past and up-to-now) and the two variants in which the construction can appear (adverbial and clausal constructions) and will argue in support of the chronological precedence of the up-to-now meaning and of the clausal construction. Section 3 will explore the diachronic change from clausal to adverbial construction cross-linguistically and present a proposal for how that change might have taken place in Spanish. Section 4 will present a corpus-based diachronic analysis of the constructions. The results, based on the analysis of 31,965 verb forms and 1,275 time constructions, show an increase in frequency of the adverbial construction, the fixation of its word order and the erosion of the morphological possibilities of the verb, all of which argue in favor of a grammaticalization process à la Lehmann (1995) having taken place from the 17th to the 19th century. Section 5 will present further discussion of the findings and propose some explanations while Section 6 will conclude by summarizing the results and suggesting possibilities for future research.

2. In search of the origins

2.1 The origin of temporal *haber*

The Spanish construction *hacer* + time is a relatively recent one. With its first documented instances in the 16th century, it was not until the 19th century that *hacer* + time replaced the earlier construction containing *haber* ‘have’. The latter had been present in the language since the first writings and appears to be relatively ancient, since it can be found in other Romance languages as well (e.g. Portuguese, French and Sicilian). It may have been present, therefore, even in Late Latin (4) but was not present in Classical Latin (5), where the verb *esse* ‘be’ was used instead of *habere* ‘have’:

(4) Late Latin, 3rd century AD

Pater eius Appollonius,	ex	quo	hinc	profectus est,	habet
father her Appollonius	from	whom	where	left	be.3.SG have.3.SG
annos XIV					
years 14					

‘It is fourteen years since her father Appollonius set out from here.’

(*Historia Apollonii Regis Tyri* 31.7; cited in Howe & Ranson 2010: 204)

- (5) Classical Latin, 2nd century BC

Hanc domum iam multos annos est cum possideo
 this house already many years be.3.SG that inhabit.1SG
 ‘It is many years that I have been occupying this house.’

(Plautus, *Aulularia*, 3–4; cited in Howe & Ranson 2010: 204)

As explained e.g. by Pérez Toral (1992: 151), there was still a great overlap in the uses and constructions allowed by the two verbs in the earliest writings in Old Spanish,¹ which may explain the replacement of *esse* by *habere* and the loss of the subject in the latter. Some isolated examples do appear, however, which still preserve traces of the verb’s earlier ability to take a subject:

- (6)
- dize que avie Mahomat XL annos de su hedat e IX que*
-
- say.3.SG that have.IPF.3.SG Mahomet 40 years of his/her age and 9 that
-
- regnava

reign.IPF.3.SG

‘It is said that Mahomet was 40 years of age and had been reigning for nine years.’

(14th century, CORDE)

- (7)
- Esso deve de aver más de dos mill y quinientos años*
-
- that must have more than 2000 and 500 years

‘It must be more than 2,500 years since that.’

(16th century, Pérez Toral 1992: 61)

Along with *haber* + time, there were, of course, other constructions for expressing the same time relations (distance-past and up-to-now). We find, for example:

- (8)
- fue tornado a senyorio de cristianos 700 annos aca*
-
- was turned.into an estate of christians 700 years here
-
- ‘[It] was turned into a Christian estate 700 years ago.’

(14th century, CORDE)

- (9)
- no se les ha pagado de tres annos aesta parte*
-
- NEG REFL them have paid from three years to.this part

‘They haven’t been paid for three years.’

(c. 1500, CORDE)

1. In Pérez Toral (1992: 51) it is mentioned that the impersonal existential *habere* (e.g. *Hubo guerras en España*) could have emerged out of the mixture of two different constructions: the possessive (*España hubo guerras*) and the existential (*Fueron guerras en España*). As evidence of the diffuse borders between these, Pérez Toral mentions early examples such as: *Grand alegría es entre todos essos christianos* vs. *El pueblo e la villa hovo grant alegría*; *Quantos que allí ha* vs. *Quantos que y son*. It is well known that there is indeed a semantically motivated cross-linguistic affinity between existence, location and possession (see e.g. Langacker 2004 and Heine 2006).

(8) illustrates the expression of a distance-past time relation by means of a construction different from the one analyzed here. Similarly, (9) shows that the up-to-now time semantics can also be expressed by means other than the verb *haber*. All in all, however, the impersonal *haber* + time construction appears to be well established in Spanish by the time of its earliest attestations in its clausal and adverbial variants, for both the up-to-now and the distance-past meanings.

2.2 The durative and the punctual meanings

As has just been mentioned, the punctual (10) and the durative (11) meanings of the *haber* + time constructions appear in Spanish from the earliest writings:

- (10) ¿Quanto i a que la vistes?
 how.much there have.3.SG that her see.PST.2.SG
 ‘How long has it been since you saw her?’
 (12th century, Rasmussen 1981: 13)

- (11) XLVII annyos ha, al mi cuidar, que de ti no oí
 47 years have.3.SG to.the my mind that from you NEG hear.PST.1SG
 hablar
 talk
 ‘I haven’t heard from you for 47 years, I think.’ (11th century, CORDE)

However, as also argued by Howe & Ranson (2010), the durative up-to-now meaning must have preceded the punctual distance-past. On the one hand, many of the less-grammaticalized temporal clausal constructions, like the ones with *llevar* + time, can only be used with a durative meaning. On the other hand, in the earliest stages of the language, the durative meaning was predominant in *haber* + time. Only later did that proportion decline progressively up to the present, where the punctual is predominant (see Section 4).

Similarly, many equivalent expressions in other languages, which at one point could mark both time relations, have become restricted to distance-past in the modern language:

- (12) Old French
 Trois jours a, ne dormi
 three days ago NEG sleep.1.SG
 ‘I haven’t slept for three days.’ (Díez Itza & Pérez Toral 1991: 49)

- (13) Middle English
 I woot it by myself full yore agon
 I know it by myself many years ago
 ‘I have known that myself for a long time.’ (Chaucer, *The Knight’s Tale*: 1813)

(14) Early Modern Danish

Han er hos Vorherre for snese Aar siden
 he is with our.Lord ago tens year ago
 'He has been dead for decades.'

(Rasmussen 1981: 90)

How, then, can an expression used with an up-to-now interpretation come to acquire the distance-past meaning? I cannot provide a definite answer at this point as to how the change took place, since the construction which is analyzed here already had both interpretations from the first documented stages of Spanish. There are two observations which might be relevant, however:

- (15) Hacer un año que no le veo = Hacer un año que le
 make.3.SG a year that NEG him see.1.SG make.3.SG a year that him
 vi
 see.PST.1SG
 'I haven't seen him for a year.' 'It has been a year since I last saw him.'

If maximal informativeness is assumed for a sentence containing a negation of an up-to-now event (which is indeed the most common use of such an expression), it immediately becomes equivalent to the assertion of the event in distance-past. If the time of the event, rather than the subsequent time interval without the event, wants to be emphasized by the speaker, this may provide a motivation for the change from up-to-now to distance-past. Another link between the two time relations (this time language-specific) may be illustrated by (16):

- (16) Poco tiempo a que es nacida
 little time have.3.SG that be.3.SG born.F
 'She was born a short time ago.' (12th century, Rasmussen 1981: 13)

The verbal tense illustrated by (16) above, much like English present perfect (e.g. *has studied*), has undergone an interesting shift in meaning. From having a present resultative meaning it became a perfect, and cross-linguistically it often comes to develop a perfective meaning. The change in the temporal relation described by this tense means that, while (16) above may have had an up-to-now meaning initially (i.e. 'She is in a state of having been born'; that is, alive and outside the mother's womb), it would not have had the same meaning in later periods. If the time construction with *haber* + time continued to appear with this tense it would have come to have a distance-past meaning here (i.e. 'She was born') which could have acted as a bridge to its further use for distance-past in combination with other past tenses.

2.3 The clausal and the adverbial constructions

Similarly to the durative and punctual meanings of the expressions, the clausal (17) and the adverbial (18) constructions of *haber* + time appear both in writing from the earliest texts of sufficient length (Cantar de mio Cid):

(17) Pocos dias ha, rey, que una lid ha arrancado
 few days have.3.SG king that a dispute has begun
 ‘A dispute broke up a few days ago, my King.’ (12th century, CORDE)

(18) Fuera el rey a San Fagunt aún poco ha
 go.PLUP the king to San Fagunt still little have.3.SG
 ‘The king had gone to San Fagunt only a short time ago.’
 (12th century, CORDE)

It is not too adventurous, however, to assume that the clausal construction must have occurred chronologically prior to the adverbial. There are many reasons to believe so.

First of all, we have the verbal morphology (and therefore the likely verbal origin) of the expression. Despite some proposals for a different, non-verbal origin of *hace* (Elerick 1989), the verbal morphology found in the equivalent, cognate expressions in many other Romance languages clearly seems to support a verbal, clausal origin.

Second, the earliest documented stages of Spanish show a strong quantitative predominance of the clausal over the adverbial construction which only later gives way to the predominance of the adverbial we find in the modern language (see Section 4).

Finally, most, if not all, languages can express a distance-past or an up-to-now meaning by means of a biclausal construction, whereas not all languages have the possibility of encoding those time relations in a single clause (see e.g. Haspelmath 1997: 55, Herce 2017a). All these facts argue for a chronological precedence of biclausal over monoclausal strategies.

2.4 The origin of biclausal constructions

We have concluded that the expressions found in monoclausal constructions for distance-past (e.g. *ago*, *fa*, etc.) are often deverbal and most likely have a (bi)clausal construction as their origin. But which are, in turn, the sources of these biclausal constructions? Biclausal time constructions are very diverse across languages. As argued by Haspelmath (1997: 136–138), most are based upon light, desemantized verbs like ‘make’, ‘exist’, ‘become’, ‘pass’ ... but other possibilities exist:

(19) Russian

Vot uže pjat' let kak ja živu v Pariže

PREST already five year that I live in Paris

'I have been living in Paris for five years.'

(Haspelmath 1997: 138)

(20) Persian

Noh sâl ast ke dar Bamberg zendegi mi-kon-am

nine year be.3.SG that in Bamberg living IPF-do-1.SG

'I have been living in Bamberg for nine years.'

(Haspelmath 1997: 136)

(19), which includes a presentative, shows that not even a verb is necessary in the main clause of biclausal time constructions. (20) shows a biclausal construction involving a copulative verb. A biclausal origin has been proposed here for many monoclausal structures for distance-past. This, of course, does not preclude the possibility that some biclausal constructions are in turn derived from monoclausal constructions. Some, like (20) or the Classical Latin example in (5), could well be derived from monoclausal ones by means of a productive cleft-like syntactic transformation.

Another obvious possible source for biclausal constructions is clause fusion, that is, the gradual combination of two erstwhile independent clauses into one:

(21) Babungo

ŋwá kú. ndwá lùu ŋú'sə bò

he die now be years two

'He died two years ago.' (Lit. 'He died. It's now two years.')

(Haspelmath 1997: 55)

Two independent but semantically connected clauses like these could well combine into a complex biclausal sentence by the subordination or loss of syntactic independence of one of the clauses. These processes have received a fair degree of attention (e.g. Hopper & Traugott 2003: 177, Fischer 2007: 222) and will not be explored further here. I would like to focus instead on a third, less obvious possible source of temporal biclausal constructions.

Clause elaboration (Lehmann 1988, Deutscher 2000), also labeled clause expansion (Heine 2009), is the term used for the gradual acquisition of clausal features like adjuncts, direct objects, subjects, finite verb morphology, etc., by erstwhile non-clausal constituents. A development along these lines has taken place in temporal clausal constructions like the ones with *llevar* 'take' in Spanish, which are semantically equivalent to the clausal constructions with *hacer* with an up-to-now meaning. The initial sparse instances of temporal *llevar* in (22)–(23) differ importantly from the modern usage in (25)–(26):

- (22) en quince años que lleva la fundación ha adelantado poco
 in fifteen years that take.3.SG the foundation has advanced little
 ‘In its 15-year existence, the foundation has achieved little.’ (1803, CORDE)
- (23) en los 22 años que lleva de continua residencia en dicha Iglesia
 in the 22 years that take.3.SG of continuous stay in said.F church
 ‘in the 22 years (s)he has stayed in the aforementioned church.’
 (1828, CORDE)
- (24) Lleva diez años viviendo en Londres
 take.3.SG ten years living in London
 ‘He has been living in London for 10 years.’
- (25) Lleva días que me trae por la calle de la amargura
 take.3.SG days that me bring.3.SG through the street of the sorrow
 ‘(S)he has been troubling me for some days.’ (1984, CREA)

In the present-day examples, (24) and (25), there is semantically a “main event” (*vivir en Londres* and *traer por la calle de la amargura* respectively) which is located in the time axis with an up-to-now meaning. In contrast, the earliest appearances of temporal *llevar*, (22) and (23), do not locate an event in the time axis. The only clause and event are those denoted by *llevar*.

The process of clause elaboration probably proceeded as follows: initially, like any other NP, the time phrase which was the object of *llevar* (e.g. *22 años*) could take a modifier providing further information about the nature of that time extent (e.g. *22 años de sufrimiento* ‘22 years of suffering’); however, as in (23), this was still a monoclausal structure. This began to change when the modifier of the time phrase, by reanalysis, became a modifier of the VP instead (answering to the question ‘how?’) and ceased to form a single constituent with the time NP:

- (26) Lleva [diez años de viaje] → [Lleva diez años] de viaje
 take.3.SG ten years of travel take.3.SG ten years of travel
 ‘He has experienced ten years of travelling.’ ‘He has been travelling for ten years.’

After reanalysis, there was nothing preventing more verbal structures like infinitives, gerunds like in (24), or participles from appearing in these positions, thus giving rise to a biclausal construction. Given that the main predicative content is expressed by the subordinate verb (gerund, infinitive or participle) rather than by *llevar*, it is not surprising that speakers want to provide more information about it. The non-finite verb morphology is thus an obstacle, so some speakers have further elaborated the causality of that phrase to allow finite clauses like in (25). The developments of temporal *llevar* might well provide a parallel for the emergence

of biclausal temporal *haber*. Only time will tell whether the construction with *llevar* + finite clause will spread to become standard, whether it will give rise to a monoclausal strategy or whether it will develop distance-past uses.

3. In search of the diachronic developments

3.1 The emergence of the adverbial construction

As mentioned in Section 2.3, a monoclausal adverbial construction is a frequent diachronic outcome of biclausal time constructions, as was already noticed by Haspelmath (1997: 55). For some languages with deverbal markers for distance-past, proposals have already been made to account for this development. Franco (2012), for example, notes that in Italian, unlike in Spanish, the clausal construction is attested earlier than the adverbial and, thus, he derives the much more grammaticalized *fa* from a clausal construction with *fare* ‘to make’, cognate to Spanish *hacer*:

(27) Old Italian

Oggi fa l'anno che nel ciel salisti
today make.3.SG the.year that to.the heaven go.PST.2.SG

‘It is one year since you went to heaven.’

(Franco 2012: 67)

Similarly, Bourdin (2011: 48) notes briefly that English *ago* might have arisen from a biclausal construction where the Old English verb *āgān* ‘to leave’ was the main verb in the sentence. The following is an illustrative Middle English example of the source construction:

(28) me was told certeyn noght longe agon is that sith þat Crist ne wente ...

‘I was told with certainty not long ago that since Christ didn’t go ...’

(Middle English) (Chaucer, *The Wife of Bath’s Prologue and Tale*: 9–10;
cited in Bourdin 2011: 47)

Note that a sequence like *longe agon is that ...* is structurally ambiguous. On the one hand, it might represent a cleft-like structure in which *longe agon* would be a fronted, focalized constituent. On the other hand, it could represent a clausal structure with a main verb *ago* ‘to leave’ in a structure parallel to that used with *be* in modern English *it has been seven years since ...* Note at this point that a participle like *agon* would have appeared finally in older stages of English and that a verb of motion like *ago* would have taken *be* rather than *have* as an auxiliary. In fact, the change in these two aspects during the Middle English period might have contributed to the reanalysis of the biclausal construction into a cleft from which,

by syntactic back-formation (García-Castillero 2014), the modern adjunctival *ago*-phrases could then start to appear.

In some languages with deverbal expressions for distance-past like Italian or English, therefore, proposals have been made to explain the change from a biclausal to a monoclausal construction. Given the chronological precedence that I have posited for the Spanish biclausal time construction over the adverbial, I will propose a way in which the latter might have arisen from the former.

3.2 Preliminary quantitative observations

A corpus search was conducted in CORDE for the earliest documented stages of Spanish. This was aimed at obtaining some knowledge of the earliest properties and usage patterns of *haber* + time, which might help shed some light into how the adverbial construction emerged from the clausal one. Two strong tendencies were discovered in the earliest instances of the adverbial construction which I believe demand attention.

The first is that, in the earliest adverbial constructions, the phrase headed by *haber* appeared after the main verb in the vast majority of cases (85–90%). This is hardly surprising, since adverbial elements tend to be placed predominantly towards the end of the sentence in any case. This postverbal position, however, becomes more interesting once we see that, in later periods, these clauses started to appear sentence-initially with a progressively greater frequency until this becomes roughly equally frequent to a sentence-final position. I therefore believe that the sentence-final position might constitute the original place in which the adverbial construction first appeared. This is relevant because it renders, for example, the simple omission of the complementizer *que* from the clausal construction an infelicitous explanation for the implementation of the change:

- (29) Ha dos años (que) te vi
 have.3.SG two years that you see.PST.1.SG
 ‘It has been two years since I last saw you.’

As illustrated in (29), such a development would give rise to sentence-initial or at least preverbal adverbial constructions, which would predict a predominance of preposed adverbial constructions in the earliest documented stages.

The second and more interesting tendency is that, in the earliest adverbial constructions, the matrix verb exceedingly frequently appeared embedded into a higher clause, very frequently by means of the complementizer or relative pronoun *que*. I refer to sentences like:

- (30) están perdidas, que non se labraron grant tienpo ha
 be.3.PL lost.F.PL because NEG REFL plowed.3.PL big time have.3.SG
 ‘(The fields) are ruined because they haven’t been plowed for a long time.’
 (14th century, CORDE)
- (31) respuesta avréis la que no vos pudiera dar oy ha
 answer have.FUT.2.PL which NEG you can.PLUP.1.SG give today have.3.SG
 diez días
 ten days
 ‘You will receive the answer I could not give you ten days ago.’
 (14th century, CORDE)
- (32) querie yr ha ayudar al Rey don ferrando que yazie sobre
 want.IPF.3.SG go to help to.the king Mr. Ferrando who lay.IPF.3.SG above
 coynbra siete años auja
 Coimbra seven years have.IPF.3.SG
 ‘He wanted to go to help King Ferrando, who had been in Coimbra for seven
 years.’
 (14th century, CORDE)

In 69.4% ($N = 34$) of the earliest (up to the year 1400) postposed adverbial constructions, the main verb appeared in an embedded clause, subordinated by *que*. This proportion is progressively reduced in later periods, reaching only 32.2% ($N = 19$) in the 19th century, for example. A fraction of that higher frequency in the earliest periods might be explained by the fact that, as mentioned by, for example, Cano-Aguilar (2004: 466) and Cervera Rodríguez (2007), the complementizer *que* enjoyed in Old Spanish a wider array of uses than in later stages of the language. This, however, is insufficient, in my opinion, to explain the strong preference of the adverbial construction for subordinate, *que*-containing contexts. This preference for *que* is, obviously, especially suspicious because the complementizer is the only element overtly distinguishing the adverbial from the clausal construction and, whereas it is expected to occur in the latter, it is not in principle expected above chance levels in the former. I believe this preference of the adverbial construction for *que*-subordination might also be a residue from the grammatical context in which the adverbial construction first arose. Any finely-tuned proposal for the mechanism responsible for the emergence of the adverbial construction must, therefore, be able to account for both the initial predominance of postposed *haber*-phrases and for their early preference for *que*-subordinated matrix clauses. These considerations have guided the proposal I present next.

3.3 Reanalysis of surface structure + syntactic back-formation

The proposal I make concerning how the adverbial construction with *haber* might have emerged from the clausal involves the observation that there existed, in Old Spanish, subordinate topic constructions like the following:

- (33) Rogando al Criador quanto ella meior sabe, que a Mio Çid el
 begging to.the Maker as she best know.3.SG that to Mio Cid the
 Campeador que Dios le curias de mal
 reconnoitrer that God him protect.SBJV.3.SG from misfortune
 ‘Begging the Lord as well as she could to protect the Cid from misfortune.’
 (12th century, CORDE)
- (34) Vio-lo Myo Çid que con los averes que auien tomados,
 see.PST.3.SG-it Mio Cid that with the goods that have.IPF.3.PL taken.M..PL
que si-s pudiessen yr, fer lo yen de grado
 that if-REFL can.SBJV.3.PL go make it would.3.PL gladly
 ‘The Cid saw that if they could leave with the goods they had taken, they
 would do it gladly.’
 (12th century, CORDE)
- (35) Ordenaron assi que los germanos que fincasen en sus tierras
 order.PST.3.PL thus that the Germans that settle.SBJV.3.PL in their lands
 ‘Thus they ordered that the Germans settle on their lands.’
 (13th century, Wanner 2005: 32)

This subordinate topic construction,² which for obvious reasons often receives the name of “complementizer doubling”, was also frequent in other Romance varieties:

- (36) Old Neapolitan
 Le aveva ditto che se sua maistà voleva lo stato suo che
 him/her had.IPF.3.SG told that if his majesty want.IPF.3.SG the state his that
 se llo venesse to fetch with the sword in hand
 him it come.SBJV.3.SG a ppigliare co la spata in mano
 ‘(S)he had told him/her that if the king wanted his/her state he should come
 and take it by force.’
 (Munaro 2015)

As authors like Wanner (2005: 42) mention, the construction was quite popular in Old Spanish but later became more restricted. It is less frequent, but still possible, in the modern language (see e.g. Villa-García 2015):

2. Not to be confused with other similar and possibly not unrelated constructions like the so-called “que de racchocage” (Pusch 2001: 220) or with syntactic prolepsis (Panhuis 1984)

- (37) Le dijeron que si quería que podía ir con ellos
 him/her tell.PST.3.PL that if want.IPF.3.SG that can.IPF.3.SG go with them
 ‘They told him/her that (s)he could go with them if (s)he wanted to.’
- (38) Me dijo que el sábado que no contáramos con él
 me tell.PST.3.SG that the Saturday that NEG count.SBJV.1.PL with him
 ‘He told me not to count on him on Saturday.’

As can be seen, the subordinate topic construction consists of the fronting to the left of a constituent from the subordinate clause and its separation from that clause by a complementizer. The result is that the fronted, topicalized element appears “sandwiched” between two complementizers even if semantically it is part of the subordinate clause.

The proposal I make here for the mechanism responsible for the change from the clausal to the adverbial construction with *haber* capitalizes on the similarity of a *que*-embedded time clausal construction with the subordinate topic construction:

- (39) Viste a[A]braham que ha mas de mjl a[ñ]os que es
 see.PST.2.SG to Abraham that have.3.SG more than 1000 years that be.3.SG
 muerto
 dead
 ‘You saw Abraham, who has been dead for more than 1000 years.’
 (13th century, CORDE)
- (40) Susanna sepas que ha grand tiempo que somos enamorados
 Susanna know.SBJV.3.SG that have.3.SG big time that be.3.PL in.love.PL
 de ti
 of you
 ‘You should know, Susanna, that we have been in love with you for a long
 time.’
 (13th century, CORDE)
- (41) Commo quier que tiempo auja que sanna les tenja
 as want that time have.IPF.3.SG that hatred them have.IPF.3.SG
 ‘As he had hated them for a long time.’
 (14th century, CORDE)
- (42) decia que habia quince años que se habia perdido
 say.IPF.3.SG that have.IPF.3.SG 15 years that REFL have.IPF.3.SG lost
 alli
 there
 ‘He was saying that he had gotten lost there 15 years ago.’
 (16th century, CORDE)

- (43) Las barbas de plata, que no ha veinte años que fueron de oro
 the beards of silver that NEG have.3.SG twenty years that be.PST.3.PL of gold
 ‘A silver-like beard that was gold-like less than 20 years ago.’
 (17th century, Cervantes, *Novelas Ejemplares*)

Given the superficial similarity and semantic compatibility of these sentences with the subordinate topic constructions, these sequences may have been reanalyzed by some speakers as true instances of the subordinate topic construction. As discussed e.g. by Langacker (1987: 110), the ideal, transparent situation in language would be when a given surface string can be assigned a consistent function. In this case, the surface manifestation *que ... que ...*, when the second *que* is neither completive nor relative, would ideally have a single associated meaning, in this case topic. Thus, the assimilation of those strings to subordinate topic constructions would be an attempt towards transparency and would make (39)–(43) examples of bridging contexts between the clausal and adverbial constructions.

Note that in subordinate topic constructions, the element which appears between the complementizers is a member of the subordinate clause which has been fronted. If some speakers came to analyze these sentences (39)–(43) as subordinate topic constructions, they would immediately have had the possibility of producing novel utterances like (44) in which the phrase headed by *haber* is not topicalized:

- (44) Decía que se había perdido allí había quince años
 say.IPF.3.SG that REFL have.IPF.3.SG lost there have.IPF.3.SG 15 years
 ‘He was saying that he had gotten lost there 15 years ago.’ (Adapted from 42)

The present proposal thus hypothesizes a reanalysis of the syntactic structure of *que*-subordinated clausal constructions with *haber*, which were equated to subordinate topic constructions. This is in line with De Smet’s (2009) observation that surface ambiguity is not a sufficient explanation for why a certain syntactic reanalysis takes place, since ambiguity logically only exists as a consequence and not as a cause of the reanalysis itself. Here, it is the analogy to the common subordinate topic construction that makes reanalysis possible. This reanalysis (shown in Figure 1) would have enabled speakers to produce the corresponding non-topicalized construction, thus giving rise to the adverbial construction as we know it:

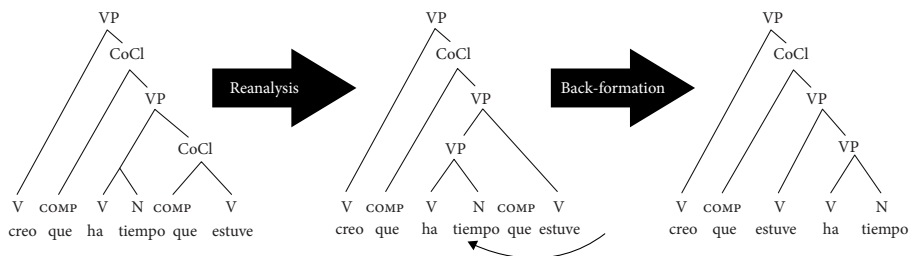


Figure 1. Proposal for the emergence of the adverbial construction: Reanalysis + Back-formation

The present proposal thus involves these stages:

1. Reanalysis of the *que*-subordinated clausal construction as a subordinate topic construction.
2. Syntactic back-formation of the derived subordinate topic construction into the corresponding underived construction.
3. Actualization: Extension of the new construction to other syntactic contexts.

The process of actualization would have later spread the new structure to other contexts like subordinate sentences without *que* or to main clauses. Note, thus, that for some time or some speakers, the clausal and the adverbial constructions would have actually been derivationally related under this hypothesis, something which some formal linguists (Rigau 1999, Fábregas 2016) propose in synchrony. However, it is my contention that, when in the late Middle Ages the subordinate topic construction started to lose ground, the two constructions eventually parted ways since the synchronic connection between them was no longer evident to speakers.

To conclude, it has to be mentioned that this proposal accommodates smoothly the initial observations in Section 3.2. It can explain specifically why it is that adverbial constructions initially occur most often in clauses subordinated by *que*, since, in clausal constructions which were not embedded, reanalysis did not take place because no analogy could be established there to the subordinate topic constructions.

3.4 Discussion

There is still an important point, independent of the actual implementation of the change, which merits attention. It is remarkable that this change has taken place in the same direction in a number of different languages. In both English and Italian, for example, what were originally main sentences containing time information became optional constituents within a clause while the originally subordinate clauses

were promoted to a main clause status by the deletion of the complementizer. The cross-linguistic persistence of that development seems to indicate that there is a strong pressure favoring that change and hence that this cannot be explained merely by an “accident” like an isolated punctual reanalysis. That would also fail to explain why the adverbial construction tends to prevail over the clausal once the two constructions are synchronically present in the language.³ I believe, therefore, that there has to be some pressure inherently favoring the emergence of adverbial constructions and their posterior prevalence over the clausal.

In this respect, it has to be noted that in the clausal construction, the main sentence exclusively provides time information, while the subordinate is the one providing the information about the main event. That is, arguably, cognitively quite anomalous and opposed to the principle of functional stability (see e.g. Heine & Reh 1984: 28, Hengeveld & Mackenzie 2008: 286) whereby constituents with the same specification tend to be placed in the same structural position. Constituents bearing time information tend to be optional elements in their clause. It is therefore not unexpected that the main and subordinate roles of these biclausal constructions tend to reverse in order to make the clause containing the main event the matrix clause and to demote time information to an adjunct position.⁴ The main obstacle for that development is of course the complementizers (*that, che, que ...*) which in the clausal construction overtly signal the subordination of the *que*-clause. Any development which manages to remove the complementizer may thus be strongly favored.

3. Note that in English and (standard) Italian the original time clausal source constructions (containing the verbs *ago* and *fare* respectively) have disappeared while the adverbial constructions (containing adpositional *ago* and *fa*) have survived. Similarly, in Spanish, where the two constructions have long been in competition, the frequency of the adverbial construction has been constantly increasing in comparison with that of the clausal (see Section 4).

4. This development is quite parallel to that studied by Thompson & Mulac (1991) and Brinton (2008) for English expressions like *I think* or *I guess*. In sentences like *I think (that) he came*, they initially constituted the matrix clause, which hosted a complement clause. The main propositional content of those sentences was to be found in the complement clauses, however, while the main clause would have only provided epistemic information. Unsurprisingly, there has been a change of syntactic roles here as well which promoted the erstwhile complement clause to a main clause status while the earlier main clause was demoted to an epistemic modifier of the clause. The complementizer has logically been a victim of the process, since the expressions most often occur now without it and frequently as parentheticals in positions typical of adverbials: *He came, I think*.

4. A corpus-based quantitative analysis of the diachronic changes in time constructions with *haber* and *hacer*

Quantitative corpus studies are becoming more and more popular in linguistic research. This is the result of the greater availability of computer-readable corpora on the one hand and of an increasingly empirical approach to linguistics on the other. Quantitative corpus-based analyses allow for an increased objectivity, extrapolability and replicability of the results. In addition, as mentioned for example by Hilpert & Gries (2016), the increased observational detail that quantitative analyses provide is necessary to understand the internal dynamics of linguistic change, which might in turn provide information about why a given change happened. Moreover, in corpus-driven studies the observed patterns may provide information which the linguist had not necessarily been looking for, as is the case of the data presented earlier in Section 3.2, but which can contribute significantly to our understanding of a given phenomenon or process.

4.1 Methodology

The present corpus study has been carried out in the Corpus Diacrónico del Español (CORDE) of the Real Academia Española. For reasons of homogeneity, the present study is limited to the Spanish of Spain. The focus of this research has been time constructions with *haber* and *hacer* as a whole; however, the search has been limited to the present (*ha, hace*) and the imperfect (*había, hacía*) forms of the verbs. This is due to the fact that, given the lower frequency of other tenses like the past (*hubo, hizo*) or the future (*habrá, hará*), a sufficient (i.e. statistically relevant) number of tokens would have been impossible to obtain. In addition, the most salient distinction, ‘present’ vs. ‘non-present’, can still be captured by comparing present and imperfect.

The mentioned verb forms were tracked in CORDE in five different periods: the 14th century, the 16th century, the 18th century, the second half of the 19th century and the second half of the 20th century. The exact time span which was mined for data depended on the possibility of reaching the previously-established number of tokens (100). Given the lower frequency of the imperfect compared to the present verb forms and given the lower proportion of texts from the earliest periods, the time intervals are usually longer in the earliest periods and for the imperfect forms of the verbs. The exact time intervals have been specified in the appendix. Inside each of those intervals, a search was conducted for each of the two verb forms and, in the periods where temporal *haber* and *hacer* coexisted (the 18th century and the second half of the 19th century), for both verbs (see Table 1).

Table 1. Searched time periods and verb forms and gathered data^a

	14th century	16th century	18th century	1850–1900	1950–2000
Verb form	ha	ha	ha	ha	
	había	había	había	había	
			hace	hace	hace
			hacía	hacía	hacía

a When the goal number (one hundred) was reached for a certain verb form in a certain period, the search was finished (darker shaded cells). In three cases (18th century *había*, 18th century *hacía* and second half of the 19th century *había*) the goal number could not be reached, in which case the search was halted when the period mined for data reached the whole century. These data have been ignored when the achieved number of tokens was below 20 (white cells), since it was deemed that such a small number would not allow for extrapolation.

No concordances or collocations were used in the search. CORDE is not a syntactically tagged corpus and therefore no concordances could be used. In addition, it was decided not to search for specific collocations either (like, e.g., *ha mucho* or *mucho ha*, *hace años* or *años hace*). A search conducted in this manner would have probably introduced certain biases for which it would be difficult to control. Because of this, every single verb form⁵ was “manually” inspected and classified as either temporal or non-temporal. Every temporal instance was in turn classified as either preceded by a preposition or not preceded by a preposition. 100 tokens of the temporal construction not preceded by a preposition were set as the goal number of tokens for each period-verb form combination. Each of those was further classified in terms of these properties:

- Type of construction instantiated (clausal, adverbial preverbal or adverbial postverbal)
- Time relation expressed (durative/up-to-now or punctual/distance-past)
- Negation of the temporal verb (presence or absence of negation in *haber* or *hacer*)
- Negation of the main event (presence or absence of negation on the verb other than the temporal *haber* or *hacer*)
- Time adjunction (presence or absence of time modifiers to temporal *haber* or *hacer*)

5. It has to be noted at this point that, because of the spelling inconsistencies of the earliest periods, *ha* and *había* could appear in writing in many different ways. Thus, *ha* appeared variously as *ha*, *a*, *há* or *á*. More remarkably, *había* could appear in a vast number of different spellings. The most frequent were *había*, *avia*, *avía*, *havía*, *avie*, *auja*, *habia*, *hauja*, *avie*, *hauie*, *abia*, *hauta*, *hauje*, *abia*, *auje* ... The most frequent variants were inspected first and less frequent spellings were only searched if the found number of tokens was insufficient. The various spelling differences have been ignored in this research and subsumed under the verbal forms they represent.

figures that will be presented below. It is therefore possible that the frequency of time constructions has been slightly underestimated as a result.

4.2 Results

The total number of verb forms inspected in the present study was 31,965, of which 16,240 correspond to imperfect and 15,725 to present verb forms. 3.99% of them ($N = 1,275$) appeared in the temporal constructions here studied. Of these, 504 were imperfect verb forms and 771 present verb forms. 89.9% ($N = 1,146$) of these were not preceded by a preposition. Of these, 446 are imperfect and 700 present verb forms. These 1,146 tokens are the ones which have been classified for the six properties presented in the previous section. The most important patterns which have emerged are discussed in Sections 4.2.1 to 4.2.8.

4.2.1 *The replacement of haber by hacer*

One of the aspects which makes the period under study special is that in these years we witness the transition from an older time construction formed with *haber* to a new strategy with *hacer*. This took place most prominently during the 19th century:

Whereas *haber* had been in these time constructions almost three times more frequent⁷ than *hacer* in the 18th century, the situation was reversed by the end of the 19th century, by which time *hacer* had become the dominant choice. As for the reasons for this development, it can be argued that the loss of the lexical uses of *haber* (it was replaced by *tener* in many contexts) and the fact that it ceased to be used as a main verb elsewhere probably favored its replacement by *hacer* in time constructions since the last main verb uses of *haber* (e.g. *ha/había X años que ...*) could have been felt as quite idiosyncratic. The minimal phonological substance of *ha* (i.e. /a/) or its homophony with another very frequent grammatical element like prepositional *a* could also have contributed to the final victory of *hacer* over *haber*.

Concerning the relationship between the two verbs, the emerging data seem to suggest that their behavior and evolution was parallel to a considerable extent

7. CORDE does not provide direct information about frequency. When only a subsection of the corpus is explored, its size is not known. For that reason, it had to be estimated by searching for some of the most common words in Spanish (*de, que, es, la*) in the whole corpus, whose size is known, and by calculating the percentage of the total number of words that these words represented. Assuming that they would constitute a similar proportion of any given subsection of the corpus one can estimate the total size of any subsection. It has to be kept in mind, therefore, that the raw frequencies here provided are estimated. This, in any case, will not affect the relative frequencies of a given period relative to another. The estimated frequencies are provided in the appendix.

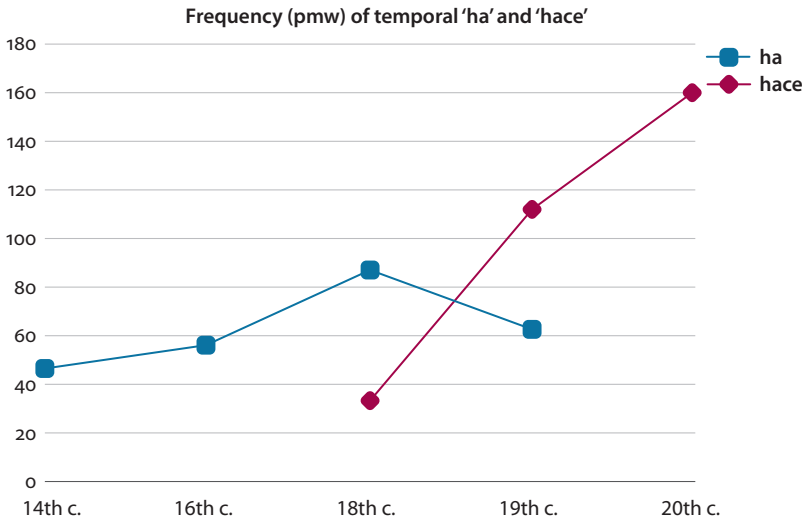


Figure 2. The replacement of *haber* by *hacer*

while they coexisted, which is in agreement with most authors like Díez Itza (1992) and Pérez Toral (1992). In the two periods for which data have been gathered for both verbs, *haber* and *hacer* feature, for example, a similar proportion of clausal, adverbial preverbal and adverbial postverbal constructions:

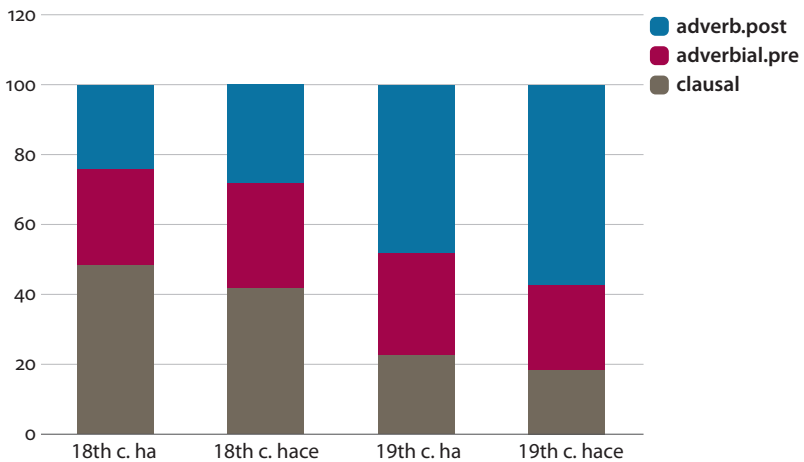


Figure 3. Constructional usage of *ha* and *hace*

There are also important differences, however. With respect to their order relative to the time NP, *ha* and *hace* show similar properties in the clausal construction but not in the adverbial, where *hace* showed a preference for preposing (65.5%, $N = 38$ in the 18th century) and *ha* for postposing (68.6%, $N = 35$ in the 18th century). In

addition, focusing on the 18th century, adverbial *haber* shows a greater proportion of punctual uses (76.6%, $N = 49$) than *hacer* (51.5%, $N = 35$) and a total lack of time adjuncts compared to a 12.1% presence in the case of *hacer*.

In addition, while the two expressions co-existed in the language, some lexemes showed a preference for *haber* and some for *hacer*. *Años* 'years', for instance, seems to occur more frequently with *haber*. Thus, whereas between 1800 and 1860 *hace años* occurs in CORDE (Spain) 19 times, *años ha* occurs 50 times. On the other hand, *poco* 'a little' seems to prefer the construction with *hacer*, since, whereas in the texts dated between 1800 and 1880 temporal *hace poco* occurs 175 times, *ha poco* and *poco ha* combined amount to only 66 instances.

More remarkable still is the preference of different lexemes for different word orders. Between 1800 and 1880, *ha mucho* occurs 93 times and *mucho ha* only 4. On the contrary, *años ha* occurs 73 times, while *ha años* does not occur at all in CORDE between those years. These facts suggest that there may be several sub-patterns and collocational uses within the time constructions as a whole which might be relevant for the diachronic development of the structures here studied. A collostructional analysis of *hacer* + time à la Stefanowitsch & Gries (2003) could be a valuable contribution at this point but exceeds the scope of the present research.

4.2.2 *Durative and punctual meanings*

It has been mentioned in Section 3 that the durative meaning of the time constructions had to precede the punctual, distance-past meaning. The pattern which emerges from the corpus data seems to support that direction for change, since the proportion of distance-past uses of the constructions tends to increase over time. This trend is common to both the clausal and the adverbial construction and to both present and imperfect verb forms. The trend is most clear, however, in the adverbial construction and with present verb forms (see Figure 4).

Distance-past amounted to 60.6% ($N = 20$) of uses in the 14th century, 68% ($N = 79$) in the 18th century, and 94.9% ($N = 74$) in the second half of the 20th century. In the adverbial construction, therefore, the punctual meaning has become virtually compulsory nowadays. Unlike in earlier periods, most contemporary speakers, especially the younger ones, tend to regard the use of *desde* as compulsory in adverbial constructions denoting an interval: *Trabajo aquí (desde) hace dos años*.

4.2.3 *Clausal and adverbial constructions*

The proposal that the clausal construction had chronological precedence over the adverbial has also found quantitative diachronic support. During the analyzed period, for both imperfect and present verb forms, the tendency has been towards a greater use of the adverbial construction compared to the clausal. The greatest

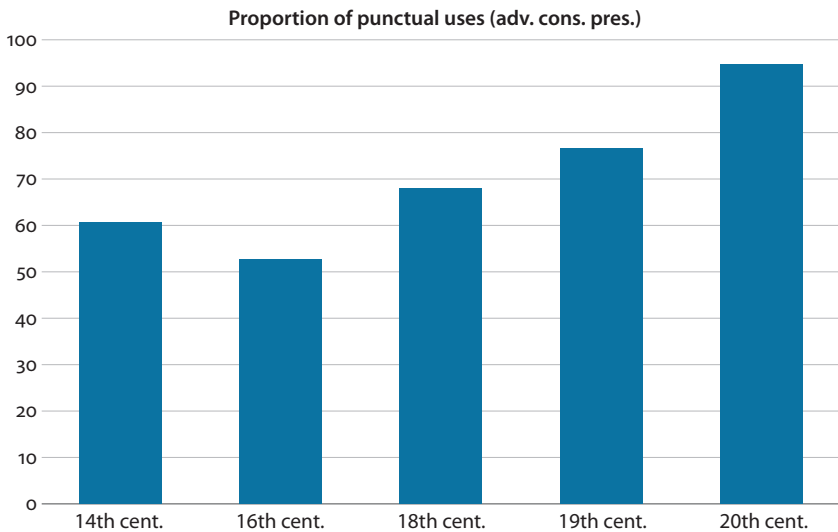


Figure 4. Time semantics of the adverbial construction

change in this respect happened from the 16th century to the second half of the 19th century. In the case of the present forms, the gathered data show that the adverbial construction increased from constituting a mere 20.6% ($N = 21$) in the 16th century to 81% ($N = 179$) in the second half of the 19th century. A chi-square test shows the differences of these intermediate periods to be extremely significant statistically ($\chi^2 = 99.21$; $df = 1$; $p < 0.001$). Even if this tendency runs parallel in the present and imperfect verb forms, it is worth noting that the imperfect always “lags behind” the present forms and shows a lower proportion of the adverbial construction in all time periods:

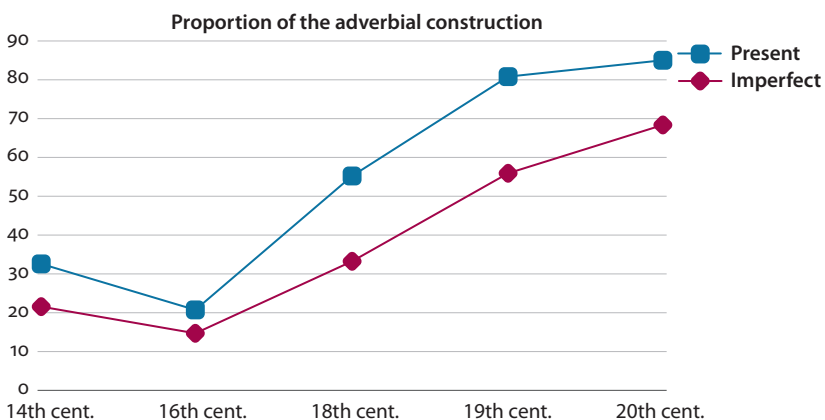


Figure 5. The proportion of the adverbial construction

This difference in the behavior of the present and the imperfect forms is determined by the fact that the verbs in clausal constructions show a more balanced TAM morphology than the verbs in adverbial constructions. Synchronically, therefore, whereas 99% of the adverbial constructions are formed with the present tense form *hace*, clausal constructions show a much more balanced display of the TAM morphological possibilities of *hacer* (see Herce 2017b).

The dramatic change which is observable here in the intermediate periods could potentially be explained either by a decline in the use of the clausal construction or by an increase in the use of the adverbial. Data indicate that it is actually the result of a sharp increase in the frequency of the adverbial construction while the frequency of the clausal construction remains relatively stable:

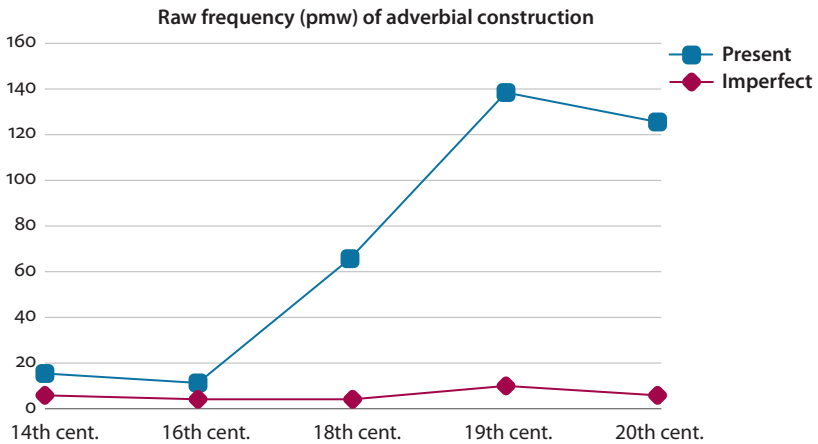


Figure 6. Increase in the frequency of the adverbial construction *ha/hace*

The frequency did not increase equally in all the verb forms, however: while that of the present tense forms soared, that of the imperfect forms remained relatively stable and moderate over time. The change, therefore, involved almost exclusively the raw frequency of the adverbial constructions with *ha* or *hace*, which exploded from slightly above 10 per million words in the 16th century to nearly 140 per million words in the 19th century. This resulted in a great reduction in the morphological variability found in the adverbial construction. Whereas in the earliest periods the present forms were only around three times more frequent than the imperfect ones, in the modern language the present form *hace* is almost compulsory in the adverbial construction.

As mentioned, for example, by Hopper & Traugott (2003: 126), an increase in “textual frequency has long been recognized informally as a concomitant of grammaticalization”. Bybee (2003, 2006) for example, mentions that high frequency may have conserving effects in that it may prevent the application of analogy and

preserve irregular forms and in that, in constructions, it may preserve features otherwise lost. Extremely high frequency, however, results in the erosion of the inner structure and semantics of linguistic expressions.

Grammaticalization and loss of inner structure, in the case of the phrases headed by *ha* or *hace*, implies desentencionalization, a process largely opposed to that of clause elaboration presented in Section 2.4. Lehmann (1988) mentions that, in this event, a subordinate clause is compressed into a nominal or adverbial constituent of the matrix clause. In the process, the clause will lose the semantic components and categories that make up a full-fledged sentence. Lehmann (1988) mentions among others a loss of illocutionary force (capacity of the sentence to be negated, questioned or asserted), of TAM morphology, actants and circumstances and of word order freedom. The rest of the patterns which have arisen in the present corpus research will be explored in order to see if there is evidence of other changes of the sort which usually accompany grammaticalization or desentencionalization.

4.2.4 *Word order of the phrase*

Although in the modern language the verb *hacer* tends to occur overwhelmingly before its accompanying time NP, the verbs *haber* or *hacer* displayed a greater variability in this respect in earlier periods. (47), for example, presents a postposed *hace*:

- (47) Los alzados montañeses eran pocos un mes hace
 the rebels of.mountains were few a month make.3.SG
 ‘The mountain rebels were few a month ago.’ (1867, CORDE)

Regarding the word order of the temporal verb and its accompanying time NP, the overall tendency has been, therefore, toward preposing the verb. In clausal constructions, little difference has been found between the word order of *haber* and of *hacer*, and their progression towards a preposed position is probably parallel to the increasing preference for verbs to precede their objects in the language as a whole. Regarding adverbial constructions, however, more substantial differences are found, since *ha*, as was mentioned earlier, tended to follow the time NP whereas *hace* tended to precede it (see Figure 7).

It is interesting to note that, in its earliest studied periods, for both *haber* and *hacer* the word order found in the clausal and in the adverbial constructions was very similar. It was only later that the two diverged. In the case of adverbial construction *ha*, it tended to preserve a position postposed to its complement despite the word order change found in the clausal construction. In the case of adverbial construction *hace*, it advanced faster than that of the clausal construction toward a rigid position before its complement.

Quite in agreement with previous developments, the intermediate periods (from the 16th to the second half of the 19th century) were the ones which

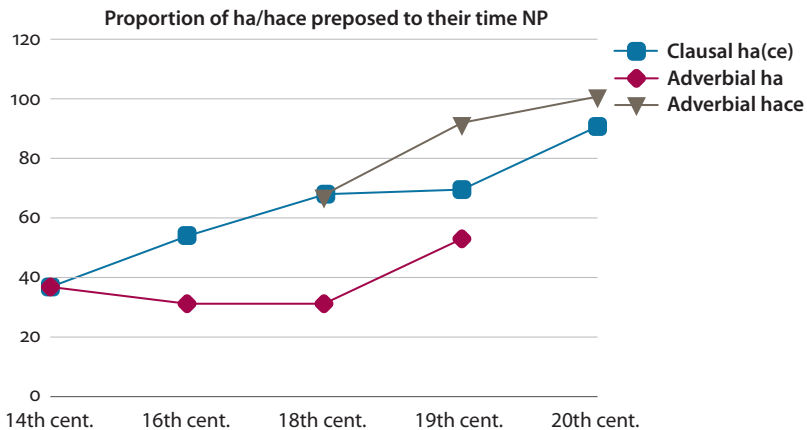


Figure 7. Word order of *ha/hace* and its complement

witnessed the biggest leap towards a rigid preposing, first with the introduction of a predominantly preposed alternative to the earlier postpositional *haber*, then with swift progress towards preposing from the 18th century to the second half of the 19th century and the disappearance of *haber* as a productive strategy. By the second half of the 20th century, *hace* in adverbial constructions had become compulsorily preposed to its complement time NP. This reduction of word order freedom is a typical consequence of grammaticalization. Lehmann (1992: 403), for example, mentions that whereas in German weakly grammaticalized adpositions like *wegen*, *entlang* or *nach* allow both pre- and postposing, more grammaticalized ones (e.g. *von*, *zu*, *in* ...) can only precede their complements. In Lehmann's (1992: 413) words, "Reduction of syntagmatic variability includes fixation of word order. This is why grammaticalization goes hand-in-hand with the loss of word-order freedom".

4.2.5 Time adjuncts

The presence of time adjuncts to *haber* or *hacer* in temporal constructions also shows an interesting, quite parallel evolution. I refer to cases such as (48), where there is a time adjunct to *haber* which overtly specifies its reference time:

- (48) Agora ha un año no hezistes assí
 now have.3.SG one year NEG do.PST.2.SG so
 'You didn't do it like that one year ago.' (1500, CORDE)

In the earliest periods (up to the 16th century), time adjuncts to *haber* occurred with roughly the same frequency in adverbial and in clausal constructions. In the 14th century, for example, we find that time adjuncts occur in clausal constructions in 13.2% ($N = 19$) of cases and in 12.5% ($N = 7$) of cases in the adverbial. This changed later as time adjuncts started to become more infrequent in the adverbial

construction and more frequent in the clausal. For example, in the second half of the 19th century, I found the presence of time adjuncts in adverbial constructions at only 4.7% ($N = 10/213$), as opposed to a 19.5% ($N = 17/87$) presence of time adjuncts in clausal constructions. A chi-square test shows these differences are statistically significant ($\chi^2 = 13.17$; $df = 1$; $p < 0.001$):

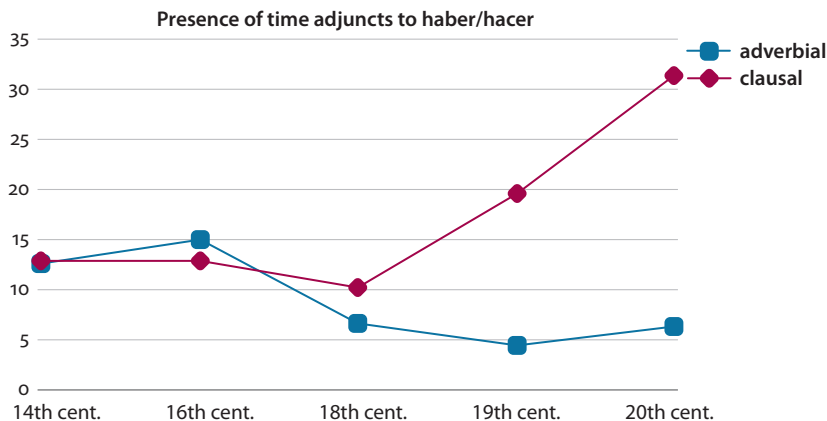


Figure 8. Presence of time adjuncts in clausal and adverbial constructions

This evolution is probably not unrelated to the huge increase in frequency experienced by the adverbial construction (especially with present verb forms of *haber* and *hacer*) between the 16th and the 19th century. In the case of *ha* in adverbial constructions, the drop from the 16th to the 18th century is remarkable: from 15.8% to zero percent in my search. In the case of *hace*, there is also a drop from 12% in the 18th century to 8.6% in the second half of the 19th century. I take these developments to be once more linked to the desententialization of the adverbial construction.

Locational time adjuncts are modifiers of states of affairs. The desemantization and loss of eventivity of *haber* and *hacer* (specifically of their present tense forms) in adverbial constructions meant that there was less motivation to try to locate them in time. The increased use of time adjuncts in the clausal construction could be seen as a side effect of the reduced clausality of the *ha(ce)*-phrases of the adverbial constructions. When a speaker wanted to specify a time reference, possibly different from the utterance time, this would have to be done predominantly by means of a clausal construction, since the adverbial hardly allowed time adjuncts to appear anymore.

4.2.6 Co-occurrence with a preposition

Since finite clauses without a complementizer do not usually combine with adpositions in Spanish, another interesting feature to look at in order to get further

evidence for desententialization is the possibility of the phrases headed by *ha* or *hace* being preceded by a preposition. In Lehmann's (1988: 13) words, "the more a subordinate clause is nominalized, the more easily it combines with adpositions and case affixes". The time phrases headed by *hacer* can and do frequently appear in the modern language preceded by prepositions like *desde* 'since', *de* 'of' or *hasta* 'until'. This was, however, much less frequent in earlier periods:

- (49) (...) *y así los provió fasta agora poco ha*
 and so them regulate.PST.3.SG until now little have.3.SG
 'And they were regulated in this way until short time ago.'

(circa 1500 CORDE)

(49), written by Christopher Columbus, constitutes one of the first sparse instances of a prepositional use of these time phrases. In agreement with what could be expected from previous findings, while in the earliest stages the phrases introduced by *ha* or *hace* could hardly be introduced by a preposition, in later periods this became increasingly frequent:

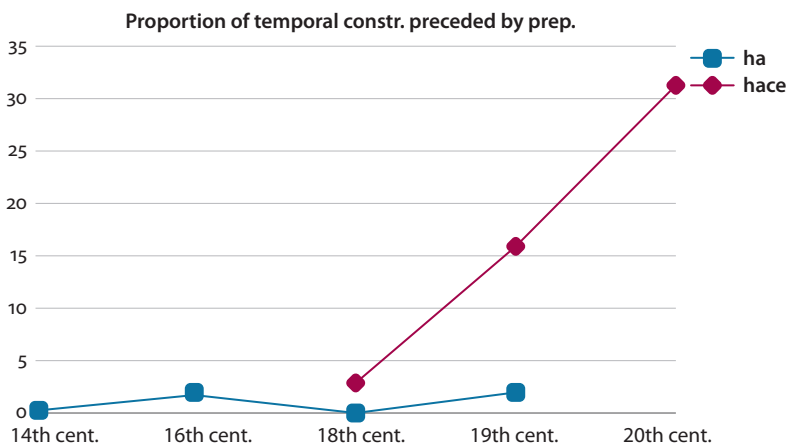


Figure 9. *Ha(ce)*-phrases introduced by prepositions

This, once again, seems to confirm that a major change was effected in the case of the adverbial construction at least, which could not easily function as complements of adpositions in the earliest periods but acquired that property later on. The desententialization of the phrases with *hace* in adverbial position may also be the reason for the present development. When those phrases ceased to be perceived as subordinate finite clauses there was no reason not to extend to them the prepositions that could already occur with other non-clausal adverbial constituents e.g. *hasta ahora* 'until now', *desde la semana pasada* 'since last week', *de ayer* 'from yesterday', etc.

4.2.7 Negation

The picture is more complicated in the case of negation. One reason for this is the existence of the frequent expression *no ha(ce) mucho* which, probably as a result of its high frequency, has managed to survive while most other combinations have lost negation in the adverbial construction. It is revealing that, in the 19th century, for example, out the 121 times that *ha mucho* appears in CORDE, as many as 84 appear negated. By contrast, *ha tiempo* appears 59 times in that period and none of them are negated. This frequent quasi-lexicalized expression *no ha(ce) mucho* increases the proportion of negated uses of *haber/hacer* in the corpus. The result is that diachronic data do not show a clear trend this time. It is worth noting, however, that in the modern language and probably also in earlier stages, negation is barred from occurring in most contexts in the adverbial construction:

- (50) *?Vive usted allí desde no hace 12 días
 live.3.SG you there since NEG make.3.SG 12 days
 ('You haven't lived there for 12 days.')
- (51) *?Era usted de nuestra opinion no hace poco
 be.IPF.3.SG you of our opinion NEG make.3.SG a.little
 ('You agreed with us not a little ago.')
- (52) *?Lo habían maltratado no hacía años
 him have.IPF.3.PL abused NEG make.IPF.3.SG years
 ('They had abused him not years before.')

Note that the previous sentences are all grammatical without negation (e.g. *Vive usted allí desde hace 12 días*) or with negation in the clausal construction (e.g. *No hace 12 días que vive usted allí*). Even in its last stronghold *no hace mucho*, however, negation is slowly retreating, as the alternative word order *hace no mucho* has recently been innovated for its use in the adverbial construction and is becoming increasingly frequent. This development is not unexpected. As Lehmann (1988: 13) points out, "at some stage of strong desentencionalization, the polarity of the subordinate clause is also affected. This usually means that it can no longer be independently negated".

4.2.8 Preverbal vs. postverbal adverbial construction

A pattern which has unexpectedly emerged from the present corpus study concerns the evolution of the word order of adverbial constructions with *ha* or *hace*. The phrases headed by *haber/hacer* (in italics) in these constructions can appear initially before their matrix verb (53) or after it, towards the end of their matrix clause (54):

- (53) Y si Dios me concede lo que [días ha le pido]
 and if God me give.3.SG it that days have.3.SG him ask.1.SG
 ‘And if God grants me what I have been asking him for days.’
 (18th century, CORDE)
- (54) Va cumpliendo lo que [prometió hace pocos años]
 go.3.SG fulfilling it that promise.PST.3.SG make.3.SG few years
 ‘He has been keeping what he promised a few years ago.’
 (18th century, CORDE)

As noted earlier, it was found that these phrases appeared in most cases (81.8%, $N = 27$) postverbally in the earliest period. The proportion of preverbal uses increases gradually until the 18th century, when it had become even more frequent than postverbal position (52.3%, $N = 57$). Later it once more decreased, up to the present, when postverbal position is again predominant (82%, $N = 64$).

This trend has never been spotted before, as far as I know. The figures provided by Pérez Toral (1992: 115), for instance, show rather that the postverbal position was always predominant and that the proportion (around 70% postverbal vs. 30% preverbal) was indeed quite stable from the Middle Ages to the present. There are two aspects, however, which make me trust my numbers over Pérez Toral’s. On the one hand, the present study is based upon a bigger number of analyzed cases. Thus, for the crucial 18th century, she relies on 34 observations whereas the present study has been based on 109. On the other hand, the same development is observed in my data for both *ha* and *hace*, which increases my confidence that the observation is not a product of chance:

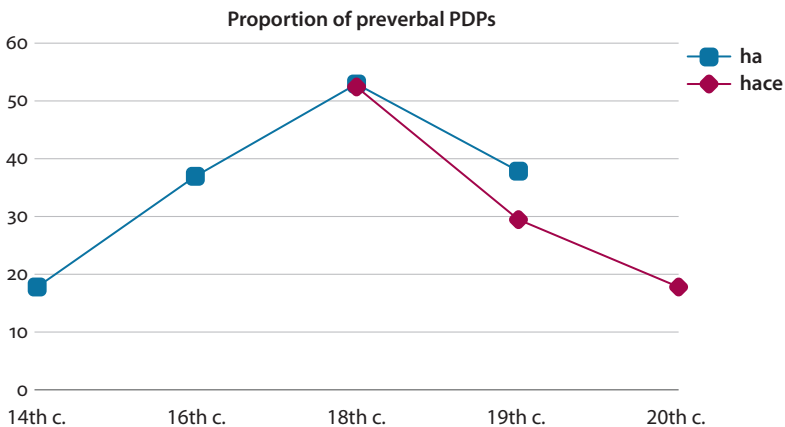


Figure 10. Position of *ha(ce)*-phrases within their matrix clause

A different issue is how to account for the trend itself and what could explain the initial increase of preverbality and its subsequent decline. We see once again that

the turning point is to be found in the intermediate periods, so that it might be again linked somehow to the previously presented diachronic developments. If the predominantly postverbal position of the earliest periods were still a residue of the grammatical context in which the adverbial construction arose in the first place as was discussed in Section 3, we might explain the initial increase of preverbality as a normal adjustment and as part of the actualization after the initial reanalysis, but what would be the pressure for the clauses headed by *haber* or *hacer* to occur initially in so many cases?

Diessel (2008) notes that time adverbial clauses are preferably placed after the main clause and in a chronologically iconic order. The phrases headed by *ha* or *hace*, therefore, would be doubly expected to occur postposed: on the one hand, because of the general preference to place time adverbials in general and time adverbial clauses in particular sentence-finally, and on the other, because the event they present takes place chronologically later than that of the matrix clause:

- (55) Te vi hace dos años
 you see.PST.1.SG make.3.SG two years
 ‘I saw you two years ago.’

In the situation described by (55), the event of me seeing you is chronologically prior to the event of having elapsed two years since that. It is actually the most common word order of the clausal construction which is anti-iconic:

- (56) Hace dos años que te vi
 makes two years that you saw.1.SG
 ‘It has been two years since I saw you.’

The event of elapsing the two years is posterior to that of me seeing you, but is presented first. It might be, therefore, that it is this word order of the clausal construction which has an impact on that of the adverbial, favoring a sentence-initial position by analogy between the two constructions. In the 18th and especially in the 19th century, however, as we have seen, the adverbial construction became extraordinarily frequent and largely lost its clausal status.

As the properties of the two constructions became increasingly different, the analogical link between them may have been loosened, leading to an incipient emancipation of the adverbial construction from the clausal and to a divergence between a verbal *hace* in the clausal construction and a more prepositional-like *hace*⁸ in the adverbial construction. In the latest periods, therefore, because of the

8. As many authors mention, however, (e.g. Bybee 2003: 162) “when grammaticalization is occurring, it may not be possible to uniquely assign elements to particular grammatical categories or structures”.

weaker link to the clausal construction and the greater frequency of the adverbial, the influence of the clausal construction upon the adverbial may have been reduced, which could have contributed to reestablishing the preference for sentence-final position we observed in the data.

5. Discussion

The diachronic developments which have been posited and/or analyzed in the present article can be illustrated by means of the synchronic uses of *hacer*. In (57), *hacer* is a full lexical verb with the meaning of completing something. This can be an amount of time, as in the example, but also a distance or a task. Its temporal meaning is derived, therefore, from the fact that in this use it takes an object NP designating a time interval. In (57) only one event is described, which is that denoted by *hacer*. The phrase *de casados* is an optional modifier specifying a property of the time interval:

- (57) Pedro y María hicieron ayer 25 años de casados
 Peter and Mary make.PST.3.PL yesterday 25 years of married.PL
 ‘Yesterday it was 25 years since Peter and Mary married.’

In later stages, however, this phrase could have emancipated from the object time NP to become a modifier of *hacer* instead. In addition, it became more clausal (e.g. *de casados* > *de estar casados* > *que están casados*), by progressively acquiring the grammatical properties of full (albeit subordinate) clauses like a subject (which will be understood to be the same as that of the main verb), inflectional morphology, negation, some word-order flexibility, etc. This so-called clause expansion was shown in Spanish temporal constructions with *llevar* in Section 2.4. After these developments, the temporal verb might well be described as a so-called ‘raising verb’ (see Boye 2010 for diachronic insights.).

Further diachronic developments (not necessarily in this order) involved the loss of the subject of *hacer*, which became impersonal (58), or the possibility of the new clause to express all sorts of events (i.e. not only states of affairs), thus allowing the biclausal construction as a whole to express both durative (58) and punctual (59) time relations. This is the stage which has been reached by the so-called clausal constructions with *hacer/haber*:

- (58) Pedro y María hizo ayer 25 años que están casados
 Peter and Mary make.PST.3.SG yesterday 25 years that are married.PL
 ‘Yesterday it was 25 years since Peter and Mary married.’

- (59) Pedro y María hizo ayer 25 años que se casaron
 Peter and Mary make.PST.3.SG yesterday 25 years that REFL marry.PST.3.PL
 ‘Yesterday it was 25 years since Peter and Mary married.’

By this stage, we have a main clause, with a very desemanticized verb *hacer*, which provides exclusively time information, and a subordinate clause expressing the main propositional content of the sentence. It is my contention that this maladjustment, which had probably been fuelling the elaboration of the subordinate clause since its origin, also constitutes the explanatory factor for the next diachronic change. At this stage, the same trend of hierarchical promotion of the subordinate clause can only be continued by promoting it to a main clause status by “getting rid” of the complementizer. The erstwhile main clause is, in turn, demoted to a temporal subordinate clause. The main and subordinate roles of the clauses in the sentence are, therefore, reversed, which is the state of affairs illustrated by (60). How exactly this dramatic development could have taken place was explored in Section 3.3.

- (60) Pedro y María se casaron ayer hizo 25 años
 Peter and Mary REFL marry.PST.3.PL yesterday make.PST.3.PL 25 years
 ‘Yesterday it was 25 years since Peter and Mary married.’

The last diachronic development concerns the loss of clausality/desententialization of the phrase headed by *hacer*. After the rise to main clause status of the erstwhile subordinate, the temporal clause became subordinate ad sensum by means of its temporal meaning alone and without adopting any formal marking of subordination. This left the new subordinate very idiosyncratic synchronically. This may have been the ultimate reason why speakers progressively started to eliminate those features most clearly identifying *hacer* in those constructions as a verb (e.g. negation and inflectional morphology) and its phrase as a clause (e.g. word-order flexibility and time adjunction). In addition, by doing this, the phrase has been brought into line with other semantically similar nonclausal time adverbials like *ayer* ‘yesterday’ or *el 14 de julio* ‘July the 14th’ and can now occur in the same syntactic contexts as these (e.g. as the complement of prepositions like *desde*, *hasta* or *de*). These developments were the ones analyzed quantitatively in Section 4. The ultimate result (61) is a non-clausal, syntactically optional phrase headed by an invariable adpositional *hace*. The sentence has become once again monoclausal and contains a single event:

- (61) Pedro y María se casaron hace 25 años
 Peter and Mary REFL marry.PST.3.PL ago 25 years
 ‘Peter and Mary married 25 years ago.’

This stage constitutes the last step toward the complete reversal of the syntactic hierarchical situation in (57) since, after this change, the erstwhile main clause has become an optional nonsentential constituent and an earlier nonsentential optional phrase has become the main and only clause.

The diachronic changes posited here for the time periods preceding the documented developments have been presented and regarded as independent from them. Even if the inversion of the main and subordinate roles of the clauses in the biclausal construction was probably a *sine qua non* for the later changes in the adverbial construction, the two are, in principle, different and logically independent events.

Alternatively, however, it might be possible to regard the whole chain of diachronic developments posited here as an extended grammaticalization process where the original lexical semantics of *haber* ‘have’ or *hacer* ‘make’ constitute the source meaning. When combined with an object denoting a time extent these verbs may have given rise to so-called “untypical contexts” (Diewald 2002). The ambiguous contexts (39)–(43) allowing reanalysis of the underlying structure would be termed “critical” (Diewald 2002) or “bridging” contexts (Heine 2002) under the terminology used in Grammaticalization Theory. The adverbial construction would thus be regarded as a “switch” (Heine 2002) or “isolating” context (Diewald 2002) whose only possible interpretation is the innovative one. All the subsequent changes in the grammatical properties of the adverbial construction which have been presented and discussed here at length (Section 4) would constitute, under this perspective, merely the “conventionalization” (Heine 2002) of the new grammatical target meaning.

Examples (57)–(61), therefore, could potentially illustrate the extended diachronic grammaticalization path that equivalent deverbal adpositions typically follow cross-linguistically in their evolution from full lexical verbs (e.g. Basque *ukan* ‘have’, Old English *āgān* ‘to leave’, Italian *fare* ‘make’, etc.) to adpositions (e.g. Basque *duela*, English *ago*, Italian *fa*).

It must be pointed out that, unlike in the above-mentioned languages, the ultimate diachronic stage (i.e. that with an invariable adpositional *hace*) might not have been yet fully attained in Spanish, since sporadic remnants of verbal morphology, negation or time adjunction continue to appear in contemporary speech and are considered acceptable by prescriptive grammars. However, the synchronic situation of many of these properties is very unclear. The grammatical status of many of the sentences provided by prescriptive grammars as illustrative of these properties is doubtful:

- (62) ?Abandonó la ciudad pronto hará tres meses
 leave.PST.3.SG the city early make.FUT.3.SG three months
 ‘It will be soon three months since he left the city.’ (RAE-ASALE 2009: 1837)

- (63) [?]Se casó no hace ni un mes
 REFL marry.PST.3.SG NEG make.3.SG even one month
 ‘It is not even one month since he got married.’ (RAE-ASALE 2009: 1837)
- (64) [?]Se divorciaron debe de hacer dos años o así
 REFL divorce.PST.3.PL must make.INF two years or so
 ‘It must have been two years or so since they got divorced.’
 (RAE-ASALE 2009: 1837)

A preliminary investigation on the matter shows that speakers of modern Peninsular Spanish find it difficult to classify these sentences as either perfectly grammatical or completely unacceptable (Herce 2017b). There are, in addition, interesting inter-generational differences in some of the judgments which suggest that the grammaticalization process analyzed here might still be underway. This would constitute an excellent opportunity to investigate grammaticalization and desententialization almost in real time. Additional research would be greatly appreciated here.

Regardless of the (un)grammaticality of these sentences, however, this piece of research has shown that the clausal and the adverbial constructions have become more and more different in their usage patterns. This synchronic separation of the two constructions can be tested by checking that focalization can be applied to *hacer* in clausal (65) but not in adverbial (66) constructions:

- (65) Hacer hace una semana que se fue pero llevaba meses
 make make.3.SG a week that left.3.SG but take.IPF.3.SG months
 sin hablar-me
 without talk-to.me
 ‘It has been only a week since he left but he hadn’t talked to me in months.’
- (66) *Se fue hacer hace una semana pero llevaba meses
 leave.PST.3.SG make make.3.SG a week but take.IPF.3.SG months
 sin hablar-me
 without talk-to.me
 (‘It has been only a week since he left but he hadn’t talked to me in months.’)

This suggests that the *hace* in (65) is a lexical element whereas the one in (66) is a grammatical one (Boye & Harder 2012). Other “tests” such as assertion also show the same restriction. In addition, the old lexical and new grammatical meanings can occur within the same clause, which suggests that the two are disjunct and the grammaticalization of *hace* completed (Heine 2002: 85):

- (67) Hace dos días hizo un año que me contrataron
 make.3.SG two days make.PST.3.SG a year that me hire.PST.3.PL
 ‘Two days ago it was one year since they hired me.’

By this it is not suggested that the two constructions are synchronically completely independent from one another. It may be precisely the analogy with the clausal construction that has prevented the complete loss of the verbal properties of *hacer* in the adverbial construction. As mentioned by Hopper (1991), the lexical origin of a grammaticalized feature may remain evident in synchrony and affect its grammatical distribution.

That this might be the case in the presently analyzed time constructions is suggested by some cross-linguistic comparisons. We see, on the one hand, that languages like Spanish and French still have clausal constructions with *hace* and *il y a* respectively where their verbal properties are prominent.⁹ It is these languages that have also managed to preserve some of the verbal characteristics of the expressions in adverbial position, where features like TAM morphology, negation of the verb, time adjuncts etc. continue to appear, albeit infrequently, in the modern language. On the other hand, we have Italian, which has lost the clausal construction with *fare* in most contemporary varieties, or English, which lost *ago* as a verb altogether. Thus, without the moderating influence of the source construction, the grammaticalization process in adverbial constructions with *fare* or *ago* was unrestrained. Maybe as a result, in modern Italian and in English, all traces of verbality (negation, TAM morphology, word order flexibility etc.) have been lost from these expressions.

6. Conclusion and further research

This article has analyzed the diachronic developments of Spanish time constructions with *haber* and *hacer*. The grammatical properties and patterns of use of so-called clausal and adverbial constructions for both *haber* or *hacer* were very similar until the 16th century but have been diverging ever since. More specifically, through the analysis of the different properties which were presented in Section 4, it has been shown that the diachronic development of the time adverbial constructions with *ha* and *hace* shows many of the characteristics typical of grammaticalization.

In the last centuries, for example, adverbial constructions with *hace* have increased their frequency exponentially. In addition to this, we observe a strong erosion of the inflectional morphology of *haber/hacer* which has become in the adverbial construction almost fixed in its earlier present tense forms (integrity). In this construction, *hace* has acquired many adpositional characteristics, advancing on its path from verb to adposition and thus from a major to a minor word class. The

9. Their status as the main predicate in these constructions probably renders them immune to decategorialization (i.e. immune to the loss of their verbal morphosyntactic properties) despite their extremely low semantic contribution to the meaning of the clause as a whole.

expression has become compulsory for conveying a distance-past time relation (paradigmatic variability) and has acquired a rigid word order since *hace* must now occur compulsorily before its time NP in these positions (syntagmatic variability). The time phrase headed by *hacer*, in addition, has lost or severely restricted many of its earlier clausal properties (i.e. has undergone desententialization). This is observed in features like negation, assertion, word order flexibility or time adjunction. None of these developments or the previous ones have taken place in the clausal construction, however, and thus we are faced with a divergence between the adverbial and the clausal construction *hace*. As Hopper & Traugott (2003: 118) mention, the phenomenon known as divergence “is a natural outcome of the process of grammaticalization, which begins as a fixing of a lexical form in a specific potentially grammatical environment where the form takes on a new meaning”.

These diachronic changes (i.e. the deverbalization of *hacer/haber* and the desententialization of its clause) might be just one of the diachronic steps necessary to give rise to these *ago*-like adpositions. One also needs to account for the emergence of the adverbial construction out of the chronologically earlier clausal and for the ultimate origin of these biclausal time constructions themselves. In the case of the presently analyzed Spanish constructions, these other changes are unfortunately not directly observable in the written records, so they can only be hypothesized or reconstructed from the quantitative patterns of the earliest periods and from the attested developments of similar constructions.

Concerning the change from the clausal to adverbial construction in Spanish, a proposal has been presented in this article which could account for the transition and is compatible with the data from the earliest periods. Concerning the ultimate origin of biclausal time constructions, that of Spanish constructions with *llevar* has been presented, which could represent a case parallel to that of temporal *haber*. The overall picture is one where clause expansion can increase the syntactic weight of an erstwhile optional non-clausal constituent giving rise to a biclausal construction. A cognitive preference for giving more relevance to event than to time information could be the driving force behind this tendency and behind the subsequent reversion of the main and subordinate roles of the clauses in biclausal constructions. The later grammaticalization of *hace* to an adposition and the loss of clausality of its phrase would be derived from speakers' tendency to assign the same syntactic position and structure to phrases with the same functional information. There is also a tendency to get rid of constructions which, like these, are badly-aligned with the most frequent grammatical patterns of the language.

Notwithstanding the findings on this article, there is much room for further research concerning these expressions, especially regarding the earliest phases of the diachronic development which has been presented here. A quantitative diachronic analysis of time constructions in their earliest developmental stages (like those with

llevar) could inform about how and why a given construction or lexical item becomes conventionalized for the expression of a given time relation. It could also help us understand better the very understudied phenomenon of clause elaboration which, being opposed to that of desentencionalization, looks difficult to reconcile with unidirectional conceptions of language change. Concerning the deverbalization and loss of clausality of adverbial constructions, a look at languages like English or Italian would allow us to find out whether the findings which have emerged from the present corpus research can be extrapolated to other languages. By providing a contrast to Spanish, a look at these languages would also help us evaluate and quantify whether, as suggested in this article, the coexistence of the source lexical item and the innovative grammatical use may have an influence on the pace of grammatical change or the grammatical properties of these constructions.

Funding

This research was possible thanks to the financial support of the Basque Government (PRE_2015_1_0175).

Acknowledgements

Special thanks are due to Carlos García-Castillero for a thorough reading of the early versions of this article and for helping me improve it in various ways. I also thank Silvia Luraghi and the two anonymous reviewers of JHL for their insightful comments.

Abbreviations

1	First person	NP	Noun phrase
2	Second person	PLUP	Pluperfect
3	Third person	PST	Past tense
COCL	Complement clause	PL	Plural
COMP	Complementizer	PMW	per million words
EVID	Evidential	PREST	Presentative
F	Feminine	REFL	Reflexive
FUT	Future	RES	Resultative
INF	Infinitive	SBJV	Subjunctive
IPF	Imperfect	SG	Singular
IPFV	Imperfective	TAM	Tense, aspect and mood
N	Noun	V	Verb
NEG	Negation	VP	Verbal phrase

References

- Bourdin, Philippe. 2011. Ten Years Ago and Ten Years Since: Competition and Standardization in Early Modern English. *Groninger Arbeiten zur Germanistischen Linguistik* 53:2.45–70.
- Boye, Kasper. 2010. Raising Verbs and Auxiliaries in a Functional Theory of Grammatical Status. *Language Usage and Language Structure* ed. by Kasper Boye & Elisabeth Engberg-Pedersen, 73–104. Berlin: Walter de Gruyter. doi: 10.1515/9783110219180.1.73
- Boye, Kasper & Peter Harder. 2012. A Usage-Based Theory of Grammatical Status and Grammaticalization. *Language* 88:1.1–44. doi: 10.1353/lan.2012.0020
- Brinton, Laurel J. 2008. *The Comment Clause in English: Syntactic Origins and Pragmatic Development*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511551789
- Brucart, José M. 2015. Un análisis derivacional de las construcciones temporales con “hacer”. *Studium Grammaticae: Homenaje al profesor José A. Martínez*, 165–180. Oviedo: Universidad de Oviedo.
- Bybee, Joan. 2003. Mechanisms of Change in Grammaticization: The Role of Frequency. *The Handbook of Historical Linguistics* ed. by Brian D. Joseph & Richard D. Janda, 602–623. Oxford: Blackwell. doi: 10.1002/9780470756393.ch19
- Bybee, Joan. 2006. From Usage to Grammar: The Mind’s Response to Repetition. *Language* 82:4.711–733. doi: 10.1353/lan.2006.0186
- Cano-Aguilar, Rafael. 2004. *Historia de la Lengua Española*. Barcelona: Ariel.
- Cervera Rodríguez, Ángel. 2007. Variedad gramatical y polivalencia pragmática de *que*. *Ex admiratione et amicitia. Homenaje a Ramón Santiago* ed. by Alicia Puigvert Ocal & Inmaculada Delgado Cobos, 205–242. Madrid: Ediciones del Orto.
- Culicover, Peter W. 1999. *Syntactic Nuts: Hard Cases, Syntactic Theory, and Language Acquisition*. Oxford: Oxford University Press.
- De Smet, Hendrik. 2009. Analysing Reanalysis. *Lingua* 119:11.1728–1755. doi: 10.1016/j.lingua.2009.03.001
- Deutscher, Guy. 2000. *Syntactic Change in Akkadian: The Evolution of Sentential Complementation*. Oxford: Oxford University Press.
- Diessel, Holger. 2008. Iconicity of Sequence: A Corpus-Based Analysis of the Positioning of Temporal Adverbial Clauses in English. *Cognitive Linguistics* 19:3.457–482. doi: 10.1515/COGL.2008.018
- Diewald, Gabriele. 2002. A Model for Relevant Types of Contexts in Grammaticalization. In Ilse Wischer & Gabriele Diewald, eds., 103–120.
- Díez Itza, Eliseo. 1992. *Ha, hay, hace* temporales en el siglo de oro. *Actas del II Congreso Internacional de Historia de la Lengua Española (Vol. 1)* ed. by M. Ariza Viguera, 373–380.
- Díez Itza, Eliseo & Marta Pérez Toral. 1991. *Las formas verbales de ciertas expresiones temporales en las lenguas románicas: un estudio de morfosintaxis histórica comparada*. Oviedo: Universidad de Oviedo.
- Elerick, Charles. 1989. Latin *abhinc annos (cum): Spanish *hace anos (que)*. *Hispanic Linguistics* 3:1–2.89–98.
- Fábregas, Antonio. 2016. On the Structure and Variation of ‘*hace*’ as a Temporal Expression. *Stellenbosch Papers in Linguistics* 45.77–108.
- Fischer, Olga. 2007. *Morphosyntactic Change: Functional and Formal Perspectives*. Oxford: Oxford University Press.

- Franco, Ludovico. 2012. Movement Triggers and the Etiology of Grammaticalization: The Case of Italian Postposition *fa*. *Sintagma: Revista De Lingüística* 24.65–83.
- García-Castillero, Carlos. 2014. De-Adjectival Preverbs in the Old Irish Verbal Complex: A Synchronic and Diachronic Study. *Zeitschrift für celtische Philologie* 61:1.89–93.
- Hagège, Claude. 2010. *Adpositions*. Oxford: Oxford University Press.
doi: 10.1093/acprof:oso/9780199575008.001.0001
- Haspelmath, Martin. 1997. *From Space to Time: Temporal Adverbials in the World's Languages*. Munich: LINCOM Europa.
- Haspelmath, Martin. 1998. Does Grammaticalization Need Reanalysis? *Studies in Language* 22:2.315–351. doi: 10.1075/sl.22.2.03has
- Heine, Bernd. 2002. On the Role of Context in Grammaticalization. In Ilse Wischer & Gabriele Diewald, eds., 83–102.
- Heine, Bernd. 2006. *Possession: Cognitive Sources, Forces, and Grammaticalization*. Cambridge: Cambridge University Press.
- Heine, Bernd. 2009. From Nominal to Clausal Morphosyntax: Complexity Via Expansion. *Syntactic Complexity: Diachrony, Acquisition, Neuro-Cognition, Evolution* ed. by Talmy Givón & Masayoshi Shibatani, 23–51. Amsterdam: John Benjamins.
doi: 10.1075/tsl.85.02fro
- Heine, Bernd & Mechthild Reh. 1984. *Grammaticalization and Reanalysis in African Languages*. Hamburg: Buske.
- Hengeveld, Kees & J. Lachlan Mackenzie. 2008. *Functional Discourse Grammar: A Typologically-Based Theory of Language Structure*. Oxford: Oxford University Press.
doi: 10.1093/acprof:oso/9780199278107.001.0001
- Herce, Borja. 2017a. Past-Future Asymmetries in Time Adverbials and Adpositions: A Cross-Linguistic and Diachronic Perspective. *Linguistic Typology* 21:1.101–132.
- Herce, Borja. 2017b. Spanish Time Constructions With *hacer*: Gradient Judgments and Corpus Data to Solve a Syntactic Conundrum. *Belgian Journal of Linguistics* 31.
- Hilpert, Martin & Stephan T. Gries. 2016. Quantitative Approaches to Diachronic Corpus Linguistics. *The Cambridge Handbook of English Historical Linguistics* ed. by Merja Kytö & Päivi Pahta, 36–53. Cambridge: Cambridge University Press.
- Hopper, Paul J. 1991. On Some Principles of Grammaticization. *Approaches to Grammaticalization* ed. by Elizabeth Closs Traugott and Bernd Heine, 17–36. Amsterdam: John Benjamins.
doi: 10.1075/tsl.19.1.04hop
- Hopper, Paul J. & Elizabeth C. Traugott. 2003. *Grammaticalization*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9781139165525
- Howe, Chad. 2011. Structural Autonomy in Grammaticalization: Leveling and Retention with Spanish *hacer* + TIME. *Probus* 23:2.247–282. doi: 10.1515/prbs.2011.007
- Howe, Chad & Diane L. Ranson. 2010. The Evolution of Clausal Temporal Modifiers in Spanish and French. *Romance Philology* 64:2.197–207. doi: 10.1484/J.RPH.3.38
- Kurzon, Denis. 2008. “Ago” and Its Grammatical Status in English and Other Languages. *Adpositions, Pragmatic, Semantic and Syntactic Perspectives* ed. by Denis Kurzon & Silvia Adler, 209–227. (= *Typological Studies in Language*, 74.) Amsterdam: John Benjamins.
doi: 10.1075/tsl.74.11kur
- Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar: Theoretical Prerequisites*. Palo Alto: Stanford University Press.

- Langacker, Ronald W. 2004. Possession, Location and Existence. *Linguagem, Cultura e Cognição. Estudos de Linguística Cognitiva, Vol 1* ed. by Augusto Soares da Silva, Amadeu Torres & Miguel Gonçalves, 85–120. Coimbra: Almedina.
- Lehmann, Christian. 1988. Towards a Typology of Clause Linkage. *Clause Combining in Grammar and Discourse* ed. by John Haiman & Sandra A. Thompson, 181–225. Amsterdam: John Benjamins. doi: 10.1075/tsl.18.09leh
- Lehmann, Christian. 1992. Word Order Change by Grammaticalization. *Internal and External Factors in Syntactic Change* ed. by Marinel Gerritsen & Dieter Stein, 395–416. Berlin: Walter de Gruyter. doi: 10.1515/9783110886047.395
- Lehmann, Christian. 1995. *Thoughts on Grammaticalization*. Munich: LINCOM Europa.
- Munaro, Nicola. 2015. Complementizer Doubling and the Structure of the Topic Field in Italo-Romance. Paper presented at Formal Approaches to Morpho-Syntactic Variation, Vitoria-Gasteiz, 24–25 June 2015.
- Pérez Toral, Marta. 1992. *Sintaxis Histórica Funcional del Español: el verbo 'hacer' como impersonal*. Oviedo: Universidad de Oviedo.
- Panhuis, Dirk. 1984. Prolepsis in Greek as a Discourse Strategy. *Glotta* 62:1/2.26–39.
- Plank, Frans. 2011. Where's Diachrony?. *Linguistic Typology* 15:2.455–471. doi: 10.1515/lity.2011.030
- Pusch, Claus D. 2001. *Morphosyntax, Informationsstruktur und Pragmatik: Präverbale Marker im gaskognischen Okzitanisch und in anderen Sprachen*. Tübingen: Gunter Narr.
- RAE-ASALE. 2009. *Nueva Gramática de la Lengua Española*. (= *Real Academia Española / Asociación de Academias de la Lengua Española*.) Madrid: Espasa-Calpe.
- Rasmussen, Poul. 1981. *El verbo hacer en expresiones temporales: Estudio sintáctico y semántico*. Copenhagen: Akademisk Forlag.
- Real Academia Española. *RDE: Corpus diacrónico del español*. Available at <http://corpus.rae.es/cordenet.html>. Accessed January–August 2015.
- Real Academia Española. *CREA: Corpus de referencia del español actual*. Available at <http://corpus.rae.es/creanet.html>. Accessed June–August 2015.
- Rigau, Gemma. 1999. Temporal Existential Constructions in Romance. *Romance Languages and Linguistic Theory* ed. by Yves d'Hulst, Johan Rooryck & Jan Schroten, 307–333. Amsterdam: John Benjamins.
- Stefanowitsch, Anatol & Stefan T. Gries. 2003. Collostructions: Investigating the Interaction of Words and Constructions. *International Journal of Corpus Linguistics* 8:2.209–243. doi: 10.1075/ijcl.8.2.03ste
- Thompson, Sandra A. & Anthony Mulac. 1991. A Quantitative Perspective on the Grammaticization of Epistemic Parentheticals in English. *Approaches to Grammaticalization, Vol. 2: Focus on Types of Grammatical Markers* ed. by Elizabeth C. Traugott & Bernd Heine, 313–329. Amsterdam: John Benjamins. doi: 10.1075/tsl.19.2.16tho
- Villa-García, Julio. 2015. *The Syntax of Multiple-que Sentences in Spanish*. Amsterdam: John Benjamins.
- Wanner, Dieter. 2005. The Corpus as a Key to Diachronic Explanation. *Romanistische Korpuslinguistik, Vol. 2: Corpora and Diachronic Linguistics* ed. by Claus D. Pusch & Wolfgang Raible, 31–44. Tübingen: Gunter Narr.
- Wischer, Ilse & Gabriele Diewald, eds. 2002. *New Reflections on Grammaticalization*. Amsterdam: John Benjamins. doi: 10.1075/tsl.49

Appendix. Diachronic use data for *hacer/haber* + time. From CORDE (Spain)

Table 2. Inspected items (temporal constructions/tokens inspected)

	14th cent.	16th cent.	18th c. <i>haber</i>	18th c. <i>hacer</i>	19th c. <i>haber</i>	19th c. <i>hacer</i>	20th cent.
Present	100/3569	102/3069	100/2414	103/1884	102/3420	119/717	145/653
Imperfect	100/4050	101/3989	39/3901	–	–	102/1120	153/2680

Because of the greater frequency of *haber* in other (auxiliary) uses, the temporal constructional uses represent a much lower proportion of the total in *haber* than in the case of *hacer*. The temporal use also represents a greater proportion in the present than in the imperfect forms; this gap was small in the earliest periods but was widened progressively.

Table 3. Estimated frequency of the time constructions (per million words)

		14th cent.	16th cent.	18th c. <i>haber</i>	18th c. <i>hacer</i>	19th c. <i>haber</i>	19th c. <i>hacer</i>	20th cent.
Clausal	Present	31.5	45.3	43.1	13.5	14.2	17.9	24.4
	Imperfect	20	22.1	8.7	–	–	7.1	3.1
Adverb	Present	15.5	11.7	44.9	19.5	48.8	94.1	136.6
	Imperfect	7	3.9	4.3	–	–	8.9	6.9

Data show a big increase in the textual frequency of these constructions as a whole; however, the increase is limited to the present tense forms of the adverbial construction.

Table 4. Prepositional uses

Verb form	14th cent.		16th cent.		18th c. <i>haber</i>		18th c. <i>hacer</i>		19th c. <i>haber</i>		19th c. <i>hacer</i>		20th cent.	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%
Present	0	0	2	2	0	0	3	2.9	2	2	19	16	45	31
Imperfect	0	0	1	1	0	0	–	–	–	–	2	2	53	34.6

In the earliest periods, the phrase headed by *haber* or *hacer* could not easily be taken as a complement by a preposition, which is what one would expect from a finite clause. In the latest periods (19th and 20th century), however, there is an important rise in this use. Clausal constructions do not admit this prepositional use: **desde hace tres años que trabajo aquí*.

Table 5. Type of construction and position

Const.	Verb form	14th cent.		16th cent.		18th <i>haber</i>		18th <i>hacer</i>		19th <i>haber</i>		19th <i>hacer</i>		20th cent.		
		N	%	N	%	N	%	N	%	N	%	N	%	N	%	
Clausal	Present	67		81		49		42		23		19		22		
	Imperfect	77		86		26	66.7	–		–		45		48		
Adverbial	Prep.	Present	6		7		27		30		29		24		14	
		Imperfect	1		3		2	5.1	–		–		15		15	
	Postp.	Present	27		12		24		28		48		57		64	
		Imperfect	22		11		11	28.2	–		–		40		37	

In the earliest periods (14th and 16th century), the clausal variant was the most frequent but this has reversed in the latest periods (19th and 20th centuries). The tendency is parallel in the present and the imperfect forms but the clausal construction has always had more weight in the imperfect forms.

Table 6. Time relation (punctual)

Const.	Verb form	14th cent.		16th cent.		18th c. <i>haber</i>		18th c. <i>hacer</i>		19th c. <i>haber</i>		19th c. <i>hacer</i>		20th cent.	
		N	%	N	%	N	%	N	%	N	%	N	%	N	%
Clausal	Present	27	40.3	28	34.6	27	55.1	18	42.9	5	21.7	13	68.4	13	59.1
	Imperfect	26	33.8	28	32.6	9	34.6	–	–	–	–	15	33.3	20	41.7
Adverb.	Present	20	60.6	9	47.4	44	86.3	35	60.3	63	81.8	58	71.6	74	94.9
	Imperfect	12	52.2	5	35.7	5	38.5	–	–	–	–	18	32.7	40	76.9

The present forms tend to express punctual time relations more frequently than the imperfect forms. The adverbial construction with present forms shows a trend of increasingly expressing punctual time relations. The imperfect forms and the clausal construction do not show a clear trend.

Table 7. Negation of the temporal verb (no negation)

Const.	Verb form	14th cent.		16th cent.		18th c. <i>haber</i>		18th c. <i>hacer</i>		19th c. <i>haber</i>		19th c. <i>hacer</i>		20th cent.	
		N	%	N	%	N	%	N	%	N	%	N	%	N	%
Clausal	Present	65	97	74	91.4	42	85.7	40	95.2	22	95.7	17	89.5	20	90.9
	Imperfect	73	94.8	81	94.2	24	92.3	–	–	–	–	40	88.9	42	87.5
Adverb.	Present	32	97	19	100	44	86.3	14	73.7	64	83.1	78	96.3	73	93.6
	Imperfect	22	95.7	14	100	13	100	–	–	–	–	53	96.4	48	92.3

No clear differences exist between the adverbial and the clausal constructions and there is no clear diachronic trend.

Table 8. Negation of the eventive verb (no negation)

Const.	Verb form	14th cent.		16th cent.		18th c. <i>haber</i>		18th c. <i>hacer</i>		19th c. <i>haber</i>		19th c. <i>hacer</i>		20th cent.	
		N	%	N	%	N	%	N	%	N	%	N	%	N	%
Clausal	Present	48	71.6	63	77.8	47	95.9	32	76.2	21	91.3	17	89.5	18	81.8
	Imperfect	68	88.3	68	79.1	23	88.5	–	–	–	–	35	77.8	34	70.8
Adverb.	Present	26	78.8	16	84.2	50	98	56	96.6	75	97.4	78	96.3	76	97.4
	Imperfect	22	95.7	12	85.7	13	100	–	–	–	–	50	90.9	49	94.2

There is a tendency for negation of the event to be more frequent in the clausal construction, *hace tiempo que no ...* than in the adverbial. No clear diachronic trend is found, however.

Table 9. Time adjunction (no adjunction)

Const.	Verb form	14th cent.		16th cent.		18th c. <i>haber</i>		18th c. <i>hacer</i>		19th c. <i>haber</i>		19th c. <i>hacer</i>		20th cent.	
		N	%	N	%	N	%	N	%	N	%	N	%	N	%
Clausal	Present	61	91	74	91.4	46	93.9	32	76.2	21	91.3	14	73.7	16	72.7
	Imperfect	64	83.1	74	86	26	100	–	–	–	–	35	77.8	32	66.7
Adverb.	Present	28	84.8	16	84.2	51	100	51	87.9	77	100	74	91.4	73	93.6
	Imperfect	21	91.3	12	85.7	12	92.3	–	–	–	–	52	94.5	50	96.2

In the 14th and 16th century time adjunction was roughly as frequent in the clausal as in the adverbial construction. This changed in later periods. Nowadays we find time adjunction significantly more frequently in the clausal construction (around 30%) and less frequently in the adverbial (around 5%).

Table 10. Order of *haber/hacer* and time NP (preposed)

Const.	Verb form	14th cent.		16th cent.		18th c. <i>haber</i>		18th c. <i>hacer</i>		19th c. <i>haber</i>		19th c. <i>hacer</i>		20th cent.	
		N	%	N	%	N	%	N	%	N	%	N	%	N	%
Clausal	Present	25	37.3	44	54.3	32	65.3	30	71.4	16	69.6	13	68.4	20	90.9
	Imperfect	57	74	63	73.3	23	88.5	–	–	–	–	35	77.8	41	85.4
Adverb.	Present	12	36.4	6	31.6	16	31.4	38	65.5	41	53.2	74	91.4	78	100
	Imperfect	3	13	4	40	2	18.2	–	–	–	–	47	85.5	50	96.2

There is a strong general trend for *haber* and *hacer* to increasingly appear before their accompanying time NPs: *hace diez años*. That trend is most pronounced, however, in the adverbial construction, where preposing is nowadays the only possibility: **Estuve allí diez años hace*.

There is a remarkable correlation in the evolution of many of the variables analyzed here for the adverbial construction (e.g. frequency, word order, time semantics, time adjunction etc.), which indicates clearly that they are related and somehow part of the same broader phenomenon (grammaticalization):

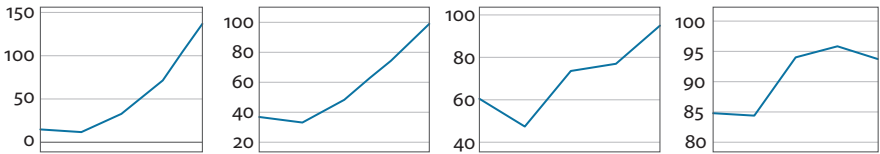


Figure 11. Evolution of present forms of the adverbial construction along four dimensions: frequency (PMW), word order (% preposed), time semantics (% punctual), time adjunction (% no adjunction)

Author's address

Borja Herce
 Faculty of Letters
 University of the Basque Country
 Paseo de la Universidad, 5
 01006 Vitoria-Gasteiz
 Spain
 borja.herce@ehu.eus