

No lo he visto ‘masque’ yo? Emergence and properties of a negative polarity item in Peninsular Spanish

Borja Herce

University of Zurich

borja.hercecalleja@uzh.ch



Received: 13-06-21

Accepted: 22-01-22

Published: 18-02-22

How to cite: Borja, Herce. 2022. No lo he visto ‘masque’ yo? Emergence and properties of a negative polarity item in Peninsular Spanish. *Isogloss. Open Journal of Romance Linguistics* 8(1)/1, 1-23.

DOI: <https://doi.org/10.5565/rev/isogloss.148>

Abstract

This paper shows that Spanish ‘más que’ (lit. more than) is much more than a comparative construction synchronically. Phonological, syntactic, and semantic evidence shows that various grammatically different entities hide under this single spelling. The most prominent of these is a (phonologically unstressed) negative polarity item with a meaning “only” or “just”. By means of robust synchronic and diachronic corpus evidence, this paper explores its morphosyntactic properties and geographic distribution in the modern language, as well as when and how a comparative expression with no polarity associations could come to grammaticalize into a negative polarity item.

Keywords: Spanish, negative polarity, corpus, grammaticalization.

1. Introduction

This paper was inspired by a disagreement between a language-learning app's expectations, and the author's native speaker knowledge of Spanish. When asked to produce the Spanish equivalent of 'There are just two doctors in the hospital', the author produced 'No hay más que dos médicos en el hospital' (1), which was not the translation that the app was looking for. On the surface, it seems to be a very poor translation indeed of the original English sentence: the verb is negated (unlike in English), there is no adverb 'just' or 'only', and it appears to contain a comparative construction ('more than') whose presence is completely unexpected judged by the sentence's intended meaning.

- (1) No hay más que dos médicos en el hospital
NEG there.is more than two doctors in the hospital
'There are only two doctors in the hospital'

Despite these differences, the two sentences, that is (1) and its English translation, are complete functional equivalents. A fact about Spanish grammar that appears to have eluded systematic research to date is that under the (orthographic) form 'más que'¹ (literally 'more than/that') hide different constructions. When used for an inequality comparison (2), the 'más' in 'más que' is phonologically stressed and, as expected, imposes no restrictions regarding the polarity of the verb in its matrix clause. Under the 'just/only' meaning, 'más' is phonologically unstressed,² and occurs in negative polarity contexts exclusively.

- (2) a) Habla más que un cura b) No habla más que un cura
 talks more than a priest NEG talks more than a priest
 '(S)he talks more than a priest' '(S)he doesn't talk more than a priest'
- (3) a) *Habla masque un cura b) No habla masque un cura
 talks only a priest NEG talks only a priest
 'It is just a priest talking' 'It is just a priest talking'

A few more example sentences are in order to provide an overview of the grammatical properties of the expression. Examples (4) and (5) show that N-words other than 'no' can also license *masque* as long as they appear preverbally and before it. The examples also illustrate the flexibility of 'masque' regarding the role and type of the constituent to which it can attach/modify, e.g. adjunct PP in (4), patient NP in (5).

¹ 'Más de' is often preferred over 'más que' in the comparative sense, especially before numerals. Although this construction would need to be addressed in any complete account of the Spanish comparatives, it will not be mentioned in the rest of this paper, as its focus is not on comparatives.

² Spanish prescriptive spelling does not distinguish these two kinds of 'más que'. However, because it will be necessary to distinguish them consistently in this paper, the ad-hoc spelling *masque* will be adopted in the remainder of this paper to represent an unstressed realization of 'más que'. This use follows the orthographic convention that already distinguishes other word pairs in the language such *por qué* ('why', stressed) vs *porque* ('because', unstressed), *más* ('more', stressed) vs *mas* ('but', unstressed), *dé* (give.1/3SG.PRS.SBJV, stressed) vs *de* ('of/from', unstressed), *sí* ('yes', stressed) vs *si* 'if/whether', unstressed)...

- (4) a) Nunca bebo masque por la noche *Masque por la noche nunca bebo
 never drink only at the night
 'I never drink except at night'
- (5) a) No vi masque el primer partido *(No) vi el primer partido masque
 NEG saw only the first match
 'I saw only the first match'

Negative polarity items differ with respect to the exact domain where they may occur. Some like 'ever' need not occur in the scope of an N-word but can also be found in questions ('have you ever crashed your car?'), in conditional sentences ('if I ever see him again...'), or after negative predicates ('he denied ever having met her before'). These contexts (see 6-9) seem unable to license Spanish 'masque', which seems to have, thus, very strict negative polarity requirements.

- (6) a) *¿Hablas masque español?
 speak.2SG only Spanish
 'Do you only speak Spanish?' b) ¿No hablas masque español?
 NEG speak.2SG only Spanish
 'Do you only speak Spanish?'
- (7) a) *Si tuviera masque cinco...
 if had.3SG only five
 'If (s)he had only five...' b) Si no tuviera masque cinco...
 if NEG had.3SG only five
 'If (s)he had only five...'
- (8) a) *Rechaza hablar masque contigo b) No rechaza hablar masque contigo
 refuses talk only with.you NEG refuses talk only with.you
 '(S)he only refuses to talk to you' '(S)he only refuses to talk to you'
- (9) a) *Rechaza masque hablar contigo b) No rechaza masque hablar contigo
 refuses only talk with.you NEG refuses only talk with.you
 '(S)he only refuses to talk to you' '(S)he only refuses to talk to you'

In the examples presented so far, 'masque' seems to be straightforwardly translatable or replaceable by an adverb (e.g. Sp. *solo*, Eng. *only*): 'only two doctors' (1), 'only a priest' (3), 'only at night' (4), 'only the first match' (5), 'only Spanish' (6b), etc. Unlike these adverbs, however, 'masque' must occur under the scope of negation, as we have shown in (3-9), it is phonologically unstressed, and it must precede the element it has scope over (see 9b, paraphrasable as 'talking to you is the only thing (s)he refuses to do' vs 9b 'you are the only person (s)he refuses to talk to'). Thus, although its semantic contribution is very similar to 'solo', the syntactic and phonological properties of 'masque' point toward a much more grammaticalized role in the language than the average adverb.

It seems plausible that a process of grammaticalization and semantic change could have led to the emergence of this expression in relatively recent timescales. On the one hand, the expression appears to be subject to considerable dialectal variation, as will be shown later (even within Peninsular Spanish). On the other hand, the orthography seems to still reflect an earlier *état de langue*. Without further research, however, it is unclear what the exact origin is of the expression. The comparative construction (with stressed *más*) is one, but not the only possible source. It is worth

mentioning, in this regard, that an (unstressed) adversative conjunction ‘mas’ meaning ‘but’, although archaic nowadays, was in vigorous use in the past. This conjunction may have simply managed to survive longer in this particular use.³ Adversative conjunction uses of (an unstressed) *masque*, in fact, are still attested in dialectal Spanish.

- (10) Hay dos discotecas masque una está cerrá
 there.is two discos but one is closed
 ‘There are two discos but one is closed’ (Cuenca, COSER)

The synchronic and diachronic challenges of ‘masque’ are, therefore, considerable. The role of this paper will be to explore this idiosyncratic expression in both domains. Section 2 contains a detailed quantitative corpus exploration of the synchronic geographical and morphosyntactic distribution of ‘masque’ and its orthographic twins and siblings. Section 3, in turn, will explore (also from a quantitative corpus perspective) the diachronic origin of the expression and its evolution into the negative polarity item we find today. Section 4 includes a wrap-up discussion and a conclusion to the paper.

2. Synchrony

Introspection and the grammaticality judgments (1-9) that have been presented in Section 1 need to be complemented with other (e.g. corpus) data to obtain a clearer picture of the contemporary situation with regards to ‘masque’ at the community level. At the same time, one needs to delimit the object of study to make data collection a commensurable task. With these factors in mind, it was decided that the optimal strategy was to pursue a quantitative corpus analysis of contemporary Peninsular Spanish in the corpus COSER (Fernández-Ordóñez 2004). This corpus is particularly appropriate for the present research because it spans the whole range of variation of the Spanish spoken in Europe, and (crucially) because it provides access to the audio recordings it is based on. This was a necessary feature to check the (numerous) cases where the stressed vs unstressed nature of ‘más que’ was not clear from the context. At the time it was accessed (February 2021), the size of the transcribed and accessible audio corpus COSER was 270 hours, totalling 4.526.735 words from 200 different rural locations all around the country.⁴

After a preliminary pilot exploration, the following traits were chosen as the potentially most interesting ones to survey and to manually code in the later synchronic corpus analysis:

- a) Type of ‘más que’: phonologically stressed vs unstressed variants, with two subtypes in the later: with vs without *nada* ~ *na* preceding.⁵
- b) Presence of negation (and type of negation) in the matrix verb.

³ For the semantic link between adversative and ‘only’ consider such uses of English ‘but’ as ‘It takes but one second to do this’.

⁴ For more details on the corpus see de Benito Moreno et al. (2016).

⁵ In the preliminary exploration, it was found that the adverb *nada* ‘nothing’, or more often a reduced and unstressed version of it *na*, occurred before unstressed ‘más que’ in a sizable proportion of tokens. In addition, as will be shown later, the morphosyntactic properties of ‘masque’ and ‘namasque’ sometimes differ in important respects. This made advisable a separate exploration of the two subtypes.

- c) Presence of numeral quantifiers in the scope of '(nada) más que'.
- d) Type of constituent that ‘más que’ precedes/modifies: most often a noun phrase, a prepositional phrase, an infinitive, or an adverb.
 - i) In the case this constituent was a noun phrase: definiteness.
- e) Semantic role of the constituent: most often patient, subject, agent, or adjunct. Conjunction (i.e. whole-sentence-linker) uses of ‘más que’ were also found.
- f) Possible semantic contribution of ‘más que’ in context: inequality comparison (i.e. ‘more than’), small amount/insignificance (i.e. ‘only’), contrast (i.e. ‘however’), etc.
- g) Location (geographical coordinates and province) of the informant.

The string “más que” was searched in the corpus, which returned 1978 tokens. Of these 1951 were classified into the types in a), with a (low) number of tokens discarded because of insufficient context or phonological evidence for taxonomization. The total prevalence of the three types in COSER was found to be 642 tokens (32.9%) of the stressed type, 902 (46.2%) of the unstressed ‘namasque’ type, and 407 (20.9%) of the unstressed ‘masque’ type. The geographic distribution of the types was found to be skewed:

Figure 1. Geographical spread of ‘masque’ (Locations with >5 tokens only)

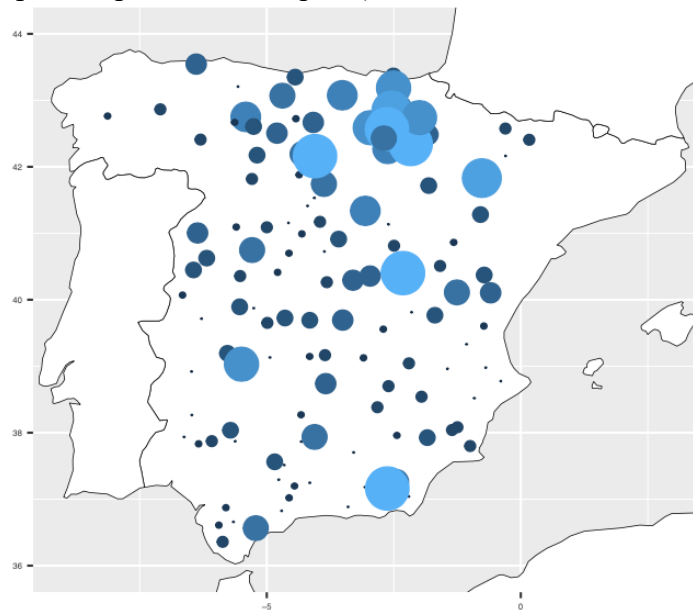


Figure 1 shows the proportion and number of tokens of ‘masque’ in different locations. Lighter blue signals a greater proportion of ‘masque’ relative to the total number of tokens, and circle size corresponds to the number of tokens of ‘masque’ in each location. A moderate northern bias can be observed for the subtype. This distribution is matched by a (much more pronounced) southern bias of the type ‘namasque’.

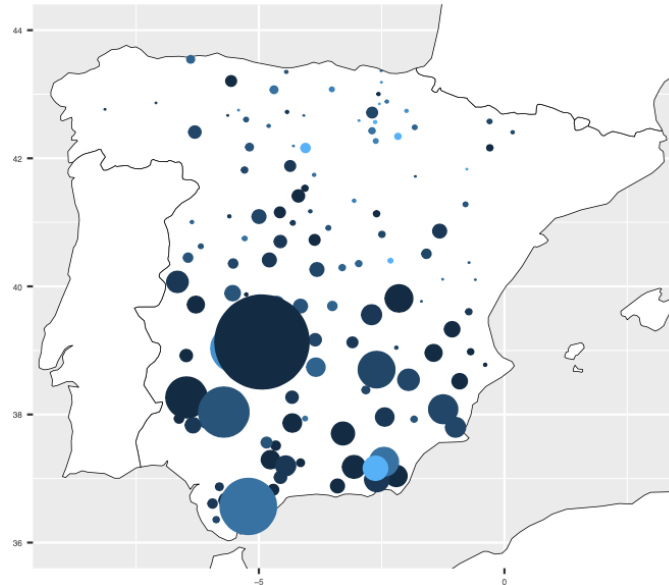
Figure 2. Geographical spread of ‘*namasque*’ (Locations with >5 tokens only)

Figure 2 shows that the type ‘*namasque*’ is only seldom used in northern Iberia. The “complementary distribution” observed for ‘*masque*’ and ‘*namasque*’ in the Spanish geography suggests that the two subtypes are in direct competition and occupy the same functional niche. This does not hold for our third type (phonologically stressed ‘*más que*’), whose geographic distribution was not found to be uneven. Because of the geographical divide illustrated in Figures 1 and 2, it was decided to explore also the possible existence of differences between the ‘*masque*’ tokens from the north and from the south of the Iberian Peninsula. The survey of the geographical distribution of all subtypes was complemented with the coding of the morphosyntactic properties b)-f) outlined before. All tokens of subtype ‘*masque*’, the original focus of this paper, were classified for these variables, as well as 100 tokens of the other types (i.e. ‘*más que*’ and ‘*namasque*’) for comparison.

The functional similarity (of ‘*masque*’ and ‘*namasque*’) notwithstanding, the exploration of their morphosyntactic properties unveiled important differences between them and (more prominently still) with respect to the stressed type ‘*más que*’. All results are displayed in Table 1 in the form of proportions.

The most remarkable one, and the property that inspired this paper, was the obligatory presence of negation in the matrix clause. This property of ‘*masque*’ was confirmed, as 95% of the tokens of this type (N=390) were found to occur in this environment, i.e. with a *no*-negated verb preceding (in 99% of the cases), or, less frequently, with preceding preverbal N-words⁶ *nadie* ‘nobody’, *ninguno/a* ‘none’, or *nunca* ‘never’. Co-occurrence with negation was less frequent in the south (86%), and less frequent still in the subtype ‘*namasque*’ (69%). All these differences are statistically highly significant (e.g. between negation in northern and southern ‘*masque*’, chi-square = 20.3245, p-value <.0001).

⁶ It is well known that these words in Spanish must themselves occur alongside verbal negation when postverbal (e.g. *no vino nadie*, ‘NEG came nobody’), but can contribute the negative meaning all by themselves when they occur preverbally (e.g. *nadie vino* ‘nobody came’). For further details see e.g. Vallduví 1994, Espinal 2000, Poole 2011, Espinal et al. 2016, Giannakidou & Zeijlstra 2017, etc.

Table 1. Types of *más que* and associated properties in COSER

| Variable | | <i>masque</i> (North) ⁷ | <i>masque</i> (South) | <i>masque</i> (All) | <i>namasque</i> | <i>más que</i> |
|---------------------------|----------------------|---------------------------------------|--------------------------|------------------------|-----------------|----------------|
| b) Negated | | 97.9% | 86.4% | 95% | 69% | 3% |
| c) Numeral quantifiers | | 26.7% | 23.3% | 26% | 23% | 11% |
| d) Constituent type | Noun P. | 69% | 64% | 67% | 55% | 70% |
| | i) Definite | 32.7% | 35.1% | 34% | 44% | 81% |
| | Prep. P. | 14.4% | 8.7% | 12% | 20% | 13% |
| | Infinitive | 11.2% | 15.5% | 13% | 7% | 5% |
| | Adverb | 1.4% | 1% | 1% | 4% | 13% |
| e) Semantic role | Agent | 1.8% | 2.9% | 2% | 3% | 29% |
| | Patient | 45.5% | 43.7% | 45% | 33% | 16% |
| | Subject | 27.4% | 31.1% | 30% | 22% | 30% |
| | Adjunct | 21.7% | 13.6% | 20% | 26% | 26% |
| | Conjunc ⁸ | 6.5% | 16.5% | 9% | 14% | 0% |

It seems, thus, that *masque* is aptly described as a negative polarity item in northern peninsular Spanish, but this fact does not extend as clearly to the subtype *na(da)masque*, which, although it still occurs chiefly in this environment, also appears frequently in the absence of verbal negation. It is important to note, however, that presence or absence of verbal negation with this expression is not associated with any semantic difference.

⁷ The classification of geographical locations into North vs South followed the geographic distribution of the types ‘masque’ and ‘namasque’ (Figures 1 and 2), with the following provinces being classified in the southern area: Cáceres, Toledo, Albacete, Alicante and all others to the south, and the rest in the northern one: Salamanca, Avila, Madrid, Cuenca, Valencia and all others to the north. Archipelagos were excluded from either area. This division left 277 tokens of ‘masque’ in the north and 103 in the south.

⁸ This category includes those tokens where ‘más que’ does not obviously belong together with the surrounding phrases, and acts as a sentential linker, often with adversative semantics (see 10). These uses are comparatively few (see Figure 6) and therefore their inclusion/exclusion from Table 1 does not skew the numbers greatly. With respect to the observed prevalence of negation, for example, excluding conjunctive uses increases the numbers somewhat (row b of Table 1 would change to 98.5%, 90.7%, 96.5%, 76.7%, 3%), as this use does not require negation (see 10), but does not alter the observed asymmetries.

- (11) [T]raía a los amigos *namasque* pa que probaran el *ajoblanco* d'aquí
bring ACC the friends only for that they.try the *ajoblanco* from.here
'(S)he brought his friends only for them to try the 'ajoblanco' from here'
(Badajoz, COSER)
- (12) [L]os médicos no mandan *namasque* andar
the doctors NEG command only walk
'Doctors here only command you to walk' (Albacete, COSER)

Example (11) shows an example of *namasque* without verbal negation, and (12) one with negation. In either case, the event denoted by the main predicate (i.e. a bringing event in (11) and a commanding event in (12)) has indeed taken place, whether or not the verb is negated. This suggests that, although the subtype *namasque* falls short of the nearly exceptionless (95%) requirement of negation we observe for *masque*,⁹ its relationship with polarity is certainly not trivial either (as e.g., in the case of stressed *más que*, whose occurrence with verbal negation (3%) corresponds approximately with the overall prevalence of verbal negation in the language).

Other properties that stand out as characteristic of (*na*)*masque* are their frequent use alongside numeral quantifiers (13)-(14), their preference for indefinite noun phrases (15), and their preference for Patient (13), (15), rather than Agent roles. Prototypical uses are thus the following.

- (13) No me preguntaban *namasque* dos cosas
NEG me asked only two things
'They asked me two things only' (Albacete, COSER)
- (14) Si un coche corre más de la cuenta no lo hagáis *masque* a cincuenta
if a car runs more than the count NEG it do.2PL only at fifty
'If a car runs too fast, drive only at fifty (kms/h)' (Cuenca, COSER)
- (15) Pues las gallinas no se le echa *masque* un poquito pienso
well the hens NEG REF them throw only a bit feed
'Well... to hens one throws only a bit of feed' (Burgos, COSER)

The expression, maybe as a result of its contemporary semantics (i.e. its meaning 'only') and its morphosyntactic requirements (i.e. its status as a negative polarity item), seems to disprefer agent roles (which in SVO Spanish cannot always be placed comfortably under the scope of negation), and appears to have specialized for introducing new (rather than given) information, a fact which is concomitantly associated with the greater occurrence of the expression in patient (rather than agent) roles, and with indefinite (rather than definite) noun phrases, which often provide or include a quantification or measure phrase (hence its greater use with numerals).

Another property that distinguishes *masque* and *namasque* is their (continued?) use as sentential linkers, in which function, because of the adversative semantics they usually convey, they also often occur with negation of the predicate in the preceding sentence.

⁹ Given that the form *namasque* includes/has absorbed the N-word *nada* 'nothing', the difference regarding negation with respect to *masque* would appear to be understandable on the basis of the properties that *nada* may have brought to the table (see Footnote 8).

- (16) [L]as sopas no se cuecen, masque se tienen a mojo
 the soups NEG REF boil.3PL instead REF have.3PL in soaking
 ‘Soups are not boiled, instead, they are left soaking’ (La Rioja, COSER)

With regard to the possible meaning contributions of the expression ‘más que’, a qualitative exploration of the tokens in COSER¹⁰ has revealed the following categories:

- a) Excess quantity (i.e. ‘more than’)
- b) Exclusivity/exceptionality (i.e. ‘besides’, ‘except’)
- c) Small amount/insignificance (i.e. ‘only’)
- d) Contrast (i.e. ‘however’, ‘instead’)
- e) Reported speech/behavior (i.e. ‘be/go like...’)

Meanings a), c), and d) should be familiar to the reader by now, as they have appeared in previous examples like 2), 11), and 6) respectively. Meaning b) has also appeared before in example 4). Exploration of the expressions’ uses in COSER has revealed an additional (infrequent) meaning e), see ex. 17, where *masque* is used to introduce reported speech, sounds, or behaviour.

- (17) [Y]o lo veo por esta, que está masque pum, pum, pum, pum...
 I it see by this.F who is like ONOM
 ‘I see it because of her, who is always like “pum, pum, pum, pum...”’

Although these meanings have been reported as discrete entities, this is mostly just a descriptive convenience. In practice, meanings a) and b), b) and c), and c) and d) are intimately associated and very often indistinguishable in a given context. Consider sentences (18) and (19).

- (18) No dejaron más que la piel
 NEG left.3PL more than the skin
 ‘They left nothing but/only the skin’ (Cantabria, COSER)
- (19) No dejaban tener más que cuatro (cabras) cada uno
 NEG allowed.3PL have more than four goats each one
 They didn’t allow you to have more than four goats each / ‘They only allowed you to have 4 goats each’ (Burgos, COSER)

In many contexts, the various meanings a) - c) are largely equivalent. In (18), the skin the crows left was both the only thing left (i.e., no flesh, no bones... meaning b), and a quantity deemed small and insignificant (i.e. meaning c). Turning to example (19), allowing to have *only* four goats (a small amount, meaning c), and not allowing to have more than four goats (i.e. meaning a) become equivalent in a world subject to Gricean (1981) implicatures. Upon learning that ‘They didn’t allow you to have more

¹⁰ It was difficult (and would have been probably also pernicious) to start with a predefined list of meanings to taxonomize actual tokens. Instead, I consider it more empirically responsible to generate this list of categories progressively during data analysis. This approach, inspired by the Autotypologizing method in typology (see Bickel & Nichols 2002) was adopted here.

than four goats each' a hearer would assume that the legal threshold for goat ownership was between 4 (legal) and 5 (illegal), and not elsewhere. That is, meanings a and c (reflected in the alternative English translations given to 19) become, in practice, equivalent in this context, due to the default Gricean assumption that the speaker was maximally informative and owning three or four goats is not also illegal. In this particular example, only knowledge of the world leads one to reject meaning b) as a plausible one, as a law that requires ownership of 4 goats exactly (i.e. not of 3 goats or 5) seems unlikely.

Another example of meaning ambiguity/overlap is presented in (20). In this example (see also 16), *masque*, acting as a sentential conjunction in this case, introduces a proposition that contrasts with the previous one (i.e. meaning d) and also involves doing comparatively little (i.e. meaning c), less than the previous sentence it is in opposition with.¹¹

- (20) Que ahora no los visten masque les ponen una sábana
 that now NEG them dress only them put.3PL a sheet
 'Now they don't dress them up, they only put a sheet over them' / Now they don't dress them up, instead, they put a sheet over them' (León, COSER)

It is unclear, thus, whether one should identify multiple distinct functions (i.e. ambiguity) in the above examples, or whether they should rather be understood as underspecified, that is, as a construction with broad semantics where more specific readings can only emerge, if at all, in particular contexts.

Regarding the association between these meanings and the different phonological *masque* types, the link between a stressed 'más que' and the excess quantity meaning, and between unstressed 'masque' and small amount/insignificance is very strong. In fact, it is the way the two meanings are distinguished where the subtypes are not in complementary distribution, and their distinction does matter for sentence interpretation (see ex. 2b vs 3b). The meaning of exclusivity is intermediate between these two in that it can be conveyed both via a stressed 'más que' (this is the only possibility when 'más' and 'que' are not adjacent, see sentence 21), but also, most often, by means of an unstressed *masque*.

- (21) No vio más solución que el suicidio.
 NEG saw.3SG more solution than the suicide
 '(S)he did not see any other option but to kill him/herself'

The intermediate position of meaning b), between a) and c), suggests it might be understood as the binding context between the two. Similar uses are also found in other languages for the comparative (e.g. in English b: 'This is no more than you deserve' and c: 'He is no more than a mediocre soccer player') although such uses have seemingly not become estranged (via different grammatical properties) from the broader comparative construction in English, unlike in Spanish. Meaning b) might also constitute a step towards meanings d) and e), which have been found in COSER to be always associated with an unstressed *masque*. Table 2 shows the associations found between form and function in the surveyed family of expressions.

¹¹ Consider, similarly, contrastive uses of "only" in English, as in "you can borrow the car, only don't scratch it!", "oh, he was wearing shoes alright, only mine!".

Table 2. Formal/morphological subtypes and associated semantics

| | a) Excess | b) Exclusivity | c) Insignificance | d) Contrast | e) Reported |
|-----------------|-----------|----------------|-------------------|-------------|-------------|
| <i>masque</i> | 19.4% | 85.3% | 90.7% | 8.7% | 1% |
| <i>namasque</i> | 14% | 84% | 86% | 14% | 0% |
| <i>más que</i> | 100% | 3% | 1% | 0% | 0% |

It shows that, maybe as expected (given that at least sometimes phonological stress is the factor which conveys the/a semantic distinction, see examples 2b and 3b, and also 22), the stressed type is always compatible with an ‘excess’ reading, while the unstressed types are so only infrequently. Conversely, unstressed realizations of *masque* are always compatible with (or signal) either the insignificance meaning, or a whole-sentence contrast (with meanings d or e). The exclusivity meaning appears to be very closely associated to the insignificance one in practice. No substantial differences have been found with respect to meaning between the two unstressed types (i.e. *masque* vs *namasque*), which confirms the functional equivalence (i.e. only dialectal difference) reported in Figures 1 and 2.

- (22) Una imagen no vale (más que / masque) mil palabras
 one image NEG be.worth more than only thousand words
 ‘An image isn’t worth more than 1000 words/An image is only worth 1000 words’

Although some evidence (consider the properties of b-meaning *más que* in 21) seems to have emerged in this synchronically-oriented section of the possible diachronic route that led to the emergence of negative polarity *masque* in Spanish (i.e. a shift from the first to the second reading in [orthographically] ambiguous sentence 22), this has only considered one part of the available evidence. As mentioned in the introduction, alternative plausible scenarios may exist that do not rely on the comparative construction and the excess-quantity meaning a) as the source for the negative polarity, unstressed *masque*. The contrastive semantics (meaning d) might also be hypothesized to be the original one from which b) and c) could have derived. The (unstressed) adversative conjunction *mas*, although itself derived from the excess comparative *MAGIS* from Latin, is an old one in Spanish (consider cognates in other Romance languages like Fr. *mais*, It. *ma*). Semantic associations/shifts between adversative conjunctions and exclusivity (meaning b) are also plausible and attested in other languages (e.g. in English ‘but’). More generally, the meanings a) - d) seem to be often collexified or expressed with similar resources in different languages (see e.g. *baino* in Basque, 23-25).

- (23) a) Ni baino altu-agoa da
 I than tall-more is
 ‘(S)he is taller than me’ (Basque)
- (24) Bat baino ez daudat
 one only NEG have.1SG
 ‘I only have one’ (Basque)

- (25) Ez da berria, zaharra baino
 NEG is new old but
 'It's not new but old' (Basque)

In the same geographical area, French also shows a remarkable affinity between these meanings, as the "normal" comparative construction in the language can also express a meaning 'only' under negation (see 19), with interesting phonological and syntactic differences, as in Spanish, between the two uses.¹²

- (26) Il n' en reste (plus) que deux
 it NEG of.them remain more than two
 'There are only two left' (French, Jeremy Pasquereau, p.c.)

It would be most interesting to explore in future work whether this is an areal development, shared by these three languages (and/or others like e.g. Greek, see Von Stechow & Iatridou 2007), and whether these uses emerged at similar times in the history of each language. This, unfortunately, exceeds the goals of this paper, whose focus will be on Spanish alone. The next section will therefore explore the diachrony of this language in search for clues about the directionality between the different uses/meanings identified here so far. Finding out how or why negative polarity became a concomitant to unstressed, insignificance-denoting *masque* is expected to contribute not only to a better understanding of the history of Spanish, but also, more generally, to the research areas of grammaticalization and polarity effects more generally.

3. Diachrony

Some of the synchronic properties of *masque* such as compulsory occurrence with (semantically void) negation, a fixed word order with respect to its modified constituent, unstressed phonological realization, etc. point towards a grammatical (Boye & Harder 2012) status of the expression synchronically. These same properties have been identified as being involved in diachronic processes of grammaticalization generally (Lehmann 1995, Hopper & Traugott 2013), and also in Spanish (Herce 2017a). Furthermore, the type of interpretative enrichment (Grice 1981) that can be found between some of the uses of the expression (see discussion around 15), is particularly suggestive of the diachronic precedence of some of them (e.g. excess quantity a > exclusivity b). This section will explore whether this is the case. It will look for corpus evidence for the origin of the negative polarity *masque* (adversative conjunction *mas* with meaning d, or comparative construction *más que* with meaning a) and will explore whether more of the changes usually associated with grammaticalization (most notably an increase in textual frequency of occurrence, see Bybee 2006) can be identified in *masque* diachrony as well.

Rather than directly embarking into an undirected diachronic corpus exploration from the outset, it is worthwhile to spend some time in less time-consuming strategies

¹² The negator *pas* is needed in French to negate the comparative construction: *il n'en reste pas plus que deux* 'it is not the case that there are more than two left', and only [ply], and not [plys], is possible under the meaning 'only', whereas both are possible in the comparative. (Jérémy Pasquereau, p.c.). For a more thorough discussion of the French construction, see O'Neill (2011).

(and to look further at the synchronic data from section 2) to try and identify the time period(s) most relevant to the construction’s development. Google Ngram Viewer (Mann et al. 2014) offers a convenient way to perform quick explorations and plots of diachronic trends.

Figure 3. Frequency of *_VERB_ más que*, overall (blue), and negated (red)¹³

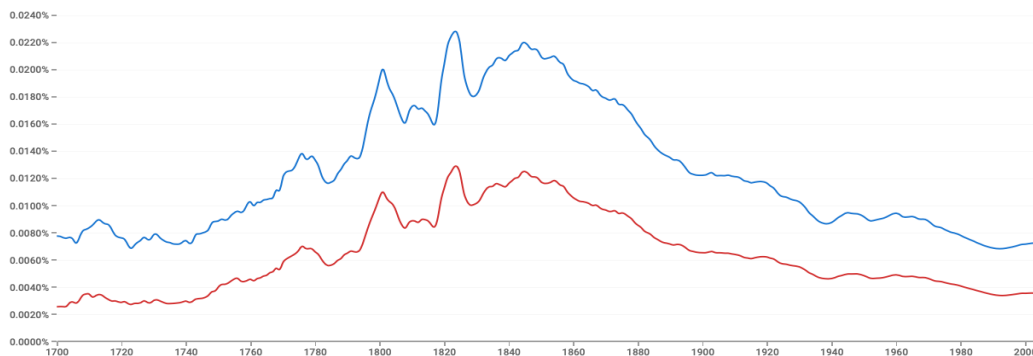
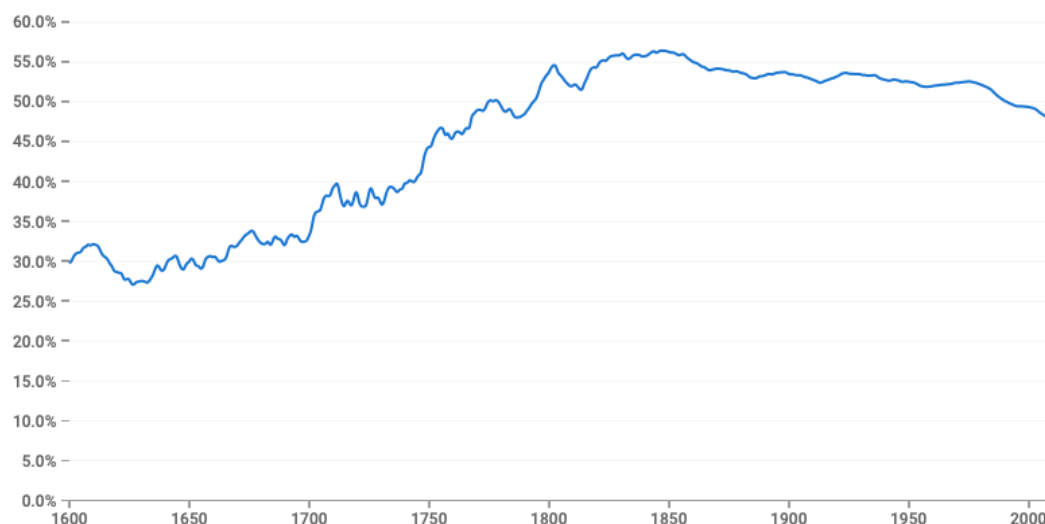


Figure 3 shows the frequency of the expression *más que*¹⁴ immediately after a verb (e.g. ‘tengo más que’) in blue. It shows that the textual frequency of this collocation increased (approximately threefold) in the period 1740-1850. This rise in frequency was driven almost exclusively by a rise of the construction in negative environments (i.e. *no _VERB_ más que*, in red, e.g. ‘no tengo más que’), in which context, its frequency increased 4.5 times in the period (compared to an increase of 2.2 times for the expression in the non-negated environment). It seems, thus, (pending more systematic corpus data collection) that this period may have been the one where the expression grammaticalized into a negative polarity item. It is, however, worth to point out that the negative-to-positive ratio for the expression had been on the rise already before the rise in frequency.

Figure 4. Percentage of negated *_VERB_ más que* as a fraction of all tokens



¹³ A 3-year smoothing has been to this and the rest of the Google Ngram Viewer graphs.

¹⁴ Prescription occurred towards the middle of this period (1850-1900) that resulted in a change from a spelling without an accent (i.e. ‘mas que’) to one with accent (i.e. ‘más que’). Both spelling variants have been added in Figure 1 to eliminate this interference.

Figure 4 shows, thus, that the bias of *más que* towards negative polarity environments had been increasing even before the rise in textual frequency, thus pointing towards it maybe being the cause, rather than the consequence of the increased use. Along with the increase in token frequency until the 1850's, Figure 3 also shows an equally pronounced drop in the century after that. Another Google Ngram plot (Figure 4) may help explain the later drop in frequency of the construction from the mid 19th century onward.

Figure 5. Frequency of *_VERB_ nada más que*, overall (blue), and negated (red)

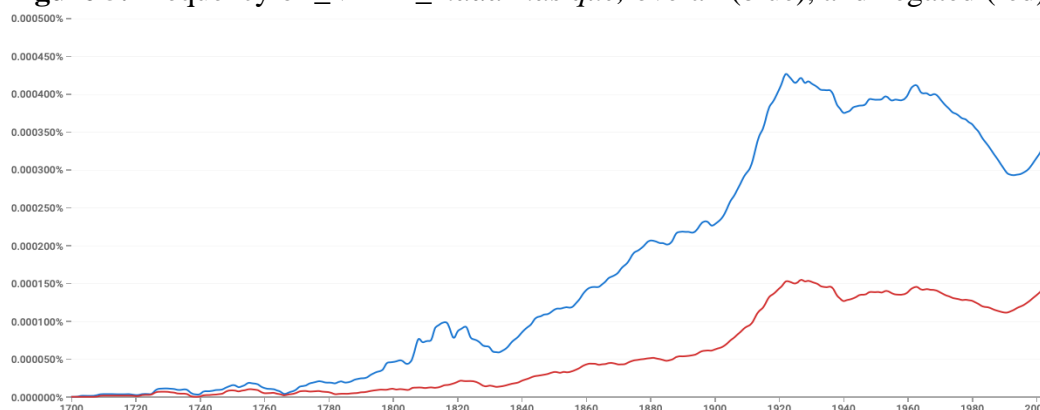


Figure 5 shows that the competing subtype *na(da)masque* which was found in Section 2 to be the preferred one in Southern Spain, appears to have emerged and substantially increased in frequency in a period (1840-1920) that coincides with the decline in textual frequency of *masque*. This would be explained by the fact that, as discussed in Section 2, the two fill the same functional niche and are in direct competition with each other. As observed for the subtype *masque* (Figure 4), the proportion of use of *namasque* in negative environments (in red) appears to have been increasing (from around 30% to 50%) during its increase in frequency. This suggests that the same negative-polarization process has occurred (at different time periods) to both subtypes analyzed in this paper.

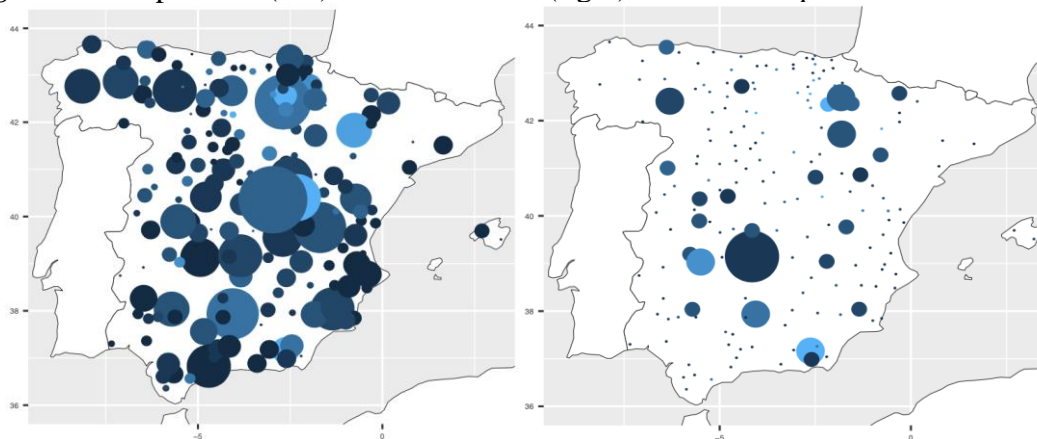
The plots from Google Ngrams allow us to quickly acquire a coarse-grained idea of what might have happened when. In this case, it strongly suggests that the compulsory negative-polarization of the expressions coincided with (or maybe caused) a substantial rise in frequency, a fact which is suggestive of grammaticalization. However, this fact by itself provides little insight as for whether the source of the negative polarity item with the ‘small amount’ meaning is to be found in the ‘excess quantity’ comparative construction, or in the sentential conjunction with contrastive meaning.¹⁵ A finer-grained diachronic corpus approach is needed to ascertain which of these two scenarios is correct, and how other morphosyntactic properties evolved during the construction’s emergence.

Before this, it would be sensible to exhaust the synchronic geographic distribution of the meanings in question as a possible source of information. If it were found, for example, that the contemporary adversative conjunction uses of *masque* are confined to a concrete area, this would suggest that it probably represents a recent

¹⁵ Another limitation of this coarse-grained approach is that Google Ngram Viewer does not allow us to focus exclusively on Peninsular Spanish, like we have done in the rest of the paper, since it conflates all Spanish-language texts regardless of their geographic origin, thus maybe introducing additional noise in our data. This could be addressed in the subsequent exploration in CORDE.

innovation. Unfortunately, conjunctive and comparative uses of *masque* have been found everywhere (see Figure 6). Even though the conjunctive use of *masque* is substandard and much more infrequent, it is nevertheless found across the Peninsula, seemingly even in relic areas (see Andersen 1988, Nichols 2015), which suggests it probably represents a comparatively old feature.

Figure 6. Comparative (left) and adversative (right) uses of *más que* in COSER



On the basis of the insufficient synchronic evidence to decide between the two alternative scenarios, a diachronic corpus analysis was undertaken in CORDE of the crucial period (i.e. mid 17th to mid 19th century, see Figure 4). Spanish texts from Spain were mined from three different five-year periods: the first from the beginning (1660-4) and the last from the end (1860-4) of the (expected) crucial period, as well as from a further intermediate period (1760-4) to provide a sanity check and a control for whether any changes found between the initial and final periods are incremental and meaningful. A total of 100 tokens of *más que* / *mas que*¹⁶ were sampled from each of the three aforementioned periods (also sampling, wherever possible, from different texts and authors) and classified for the same variables identified for the synchronic corpus research in Section 2.

Unlike in synchronic COSER, however, a historical corpus like CORDE cannot possibly provide audio evidence to allow us to identify types on this basis. We cannot know for sure, thus, which tokens of *más que* were stressed and which ones were not. This is particularly complicated if the grammaticalization story reported here is correct, in which case there may have been important changes (e.g. loss of stress) with respect to how non-comparative tokens of *más que* were pronounced. Because the formal types stressed vs unstressed correspond almost perfectly to the different functions of the expression (i.e. ‘excess quantity’ vs ‘exclusivity/insignificance’, see Table 2), these semantic functions have been employed diagnostically to assign tokens to different subtypes. Those cases (see e.g. 27) where the context did not clarify which of the meanings was being conveyed were counted into both types.

¹⁶ Because of the changes in spelling conventions mentioned in Footnote 14, and because spelling choices might conceivably have been affected by the meanings and types (e.g. stressed vs unstressed) that concern us in this paper, the total number of tokens analyzed with and without an accent was arranged to be proportional to the prevalence of each type in a given period. Thus, for example, for the period 1860-1864, texts from Spain in CORDE contain 870 tokens of *más que* (with an accent), and 79 of *mas que* (without an accent). Those proportions were preserved by sampling 92 tokens of the former and 8 of the latter spelling.

- (27) aunque algunos asistieron por espacio de tres años, ninguno aprendió
 although some attended for space of three years nobody learned
 ‘Although some of them attended the class for three years, nobody learned

más que los principios de la aritmética y geometría
 more than the basics of the arithmetic and geometry
 more than/but the basics of arithmetic and geography’ (1761, CORDE)

Example 27 shows that, in certain contexts (see also 18 and 19) and the absence of pronunciation clues, the semantic value of the expression cannot be established with certainty. Cases like 27, thus, have been counted under all their (plausible) meanings/types. These are the number of tokens found of each semantic type in each of the surveyed periods in CORDE:

Table 3. Formal/morphological subtypes and associated semantics

| | a) Excess | b) Exclusivity | c) Insignificance | d) Contrast |
|---------|-----------|----------------|-------------------|-------------|
| 1660-64 | 86% | 35% | 18% | 21% |
| 1760-64 | 71% | 35% | 31% | 26% |
| 1860-64 | 59% | 56% | 51% | 15% |

As Table 3 shows, the proportion of tokens of *más que* that conveys/may convey excess quantity has become progressively lower over time, while the opposite is the case of exclusivity, and, particularly, of insignificance-denoting *más que*. The differences between the initial and the final period are statistically highly significant ($p < 0.01$). These differences confirm the increase in the token frequency (see Figure 3) of the construction (i.e. *masque*) that this paper deals with, and suggests a diachronic route $a > b > c$ with respect to the order in which the new meanings were acquired by the expression. Less clear is the existence of any link between these meanings and the contrastive/adversative meaning d), which appears to be independent of the others and might thus not share the same comparative origin as the other meanings (i.e. it may continue the adversative conjunction *mas*. See, however, footnote 11). No tokens whatsoever have been found of *na(da) más que* in any of these initial periods, which confirms our (grounded, see Figure 5) suspicion that this subtype is the result of a later change, and separate, but probably with the same comparative origin, as *masque*.

Another important (more qualitative) finding is that the unstressed adversative conjunction *mas* (see 6 and Footnote 3) was exceedingly infrequent already by the 1660’s, which suggests it could not possibly have been the source of grammaticalizing, negative polarity *masque*. Although *más que* can indeed be used as a sentence linker with some frequency in these time periods (meaning d in Table 3), this is almost exclusively via the collocation *por más que*,¹⁷ as in (28).

¹⁷ This is a concessive conditionality expression according to RAE-ASALE (2009: 3620-3621).

- (28) Hallarla no he podido, por más que al monte he
 find.her NEG have.1SG been.able for more than to.the mountain have.1SG
 bajado
 descended
 I haven't been able to find her regardless how much I have gone down to the
 mountain (1661, CORDE)

In this context, *más (que)* must be unrelated to the adversative conjunction *mas*: it is stressed (still in the modern language), can be replaced by *mucho* ‘a lot’ (or *poco* ‘little’, or other adjectives or adverbs), and, when used, still conveys a large/excess quantity meaning (reflected in the English translation to 28), with the concomitant concessive semantics being expressed by the abstract construction ‘*por Ad que Sbjv*’ as a whole.¹⁸ The construction in 28, in fact, could even be the origin of the adversative *masque* of examples like 6, although ascertaining this is beyond the purpose of the present paper. Regarding the latter, Table 4 shows the counts and proportions for all the variables surveyed.

One can appreciate that, much as expected according to the preliminary explorations with Google Ngram Viewer (Figure 4), this period witnessed the compulsory negative polarization of the uses of *más que* that denote exclusivity/exactness (b) and/or small amount or insignificance (c). From 1660 to 1860, negative contexts increased from 65.6% to 88.9%. Even the first of these percentages, however, is vastly above the general prevalence of negation in the language or in large/excess quantity (a) *más que*. The reason must be, of course, that the most obvious way in which an expression denoting a large quantity can *in context* denote the opposite is in combination with negation.

Table 4. Types of *más que* and associated properties in CORDE

| Variable | <i>más que</i> All periods | <i>masque</i> 1660-64 | <i>masque</i> 1760-64 | <i>masque</i> 1860-64 |
|-------------|-------------------------------|--------------------------|--------------------------|--------------------------|
| b) Negated | 3.7% | 65.6% | 85.7% | 88.9% |
| c) Numerals | 8.3% | 5.7% | 6% | 3.7% |
| d) | Noun P. | 67.6% | 45.7% | 68.6% |
| | i) Definite | 72.6% | 31.2% | 54.16% |
| | Prep. P. | 11.1% | 17.1% | 20% |
| | Infinitive | 0% | 25.7% | 11.4% |
| e) | Agent | 20.4% | 2.9% | 2.6% |
| | Patient | 11.1% | 45.7% | 37.1% |
| | Subject | 22.2% | 14.3% | 22.9% |
| | Adjunct | 29.6% | 17.1% | 34.3% |

¹⁸ Concessive semantics are still present in the absence of *más*. For example, *por cojo que estés no te voy a dejar mis muletas* ‘Regardless how lame you are I won’t lend you my crouches’.

In the same way as ‘not bad’ means (or can come to mean) ‘good’, ‘no more’ can mean in context either exactness (b, by Gricean implicature, see 19) and, subsequently, insignificance (c). As discussed in Section 2, this occurs in English uses like ‘This is no more than you deserve’ and ‘He is no more than a mediocre soccer player’. The difference with respect to English¹⁹ is, as the numbers in Table 3 have shown, that (Peninsular) Spanish conventionalized the use of NEG + *more than* as the default way of expressing this meaning of insignificance or small amount. The increase in frequency of these uses must have made it more efficient for the underlying syntactic structure and dependencies of the expression to erode (i.e. chunking), and for these new uses to part ways with the original ‘more than’ construction, for example phonologically, with the de-stressing of *más*,²⁰ or syntactically, with the impossibility of separating *que* from *más*, or to modify *más* if the expression is to preserve its meaning “only/except for”.

- (29) Nadie ha sentido (mucho) más que yo la caída del rey de Nápoles
 nobody has felt much more than me the fall of.the king of Naples
 ‘Nobody has lamented (much) more than me the fall of the king of Naples’ (1863, CORDE)
- (30) Nadie ha sentido la caída del rey de Nápoles *masque* yo
 ‘Only I have lamented the fall of the king of Naples’
- (31) Nadie ha sentido más la caída del rey de Nápoles que yo
 ‘Nobody has lamented more than me the fall of the king of Naples’

Example 29 (vs 30) shows that the ‘more than’ reading, and its concomitant stressed pronunciation, are favoured by a location closer to the verb (29), and disfavoured by a location far away from it. The choice of pronunciation (stressed vs unstressed *más*) still manages to convey either meaning, except when the surrounding context discards one of them, e.g. use of ‘mucho’ in 29, or separating *más* from *que* in 31 would make an “only” reading (and an unstressed pronunciation of *más*) impossible.

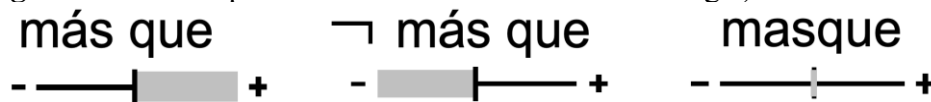
Another trait that makes the “only” reading impossible would be the absence of an N-word (e.g. *nadie*) from the previous examples. Obviously, because, as shown in this paper, unstressed *masque* is a negative polarity item, the absence of negative polarity precludes its use. The obligatorization of negation seems to have occurred during the analyzed time interval (1660-1860, see Table 4), however, this should not hide the fact that negation seems to have been very frequent (65%) in “only”-meaning *más que* from the very beginning. Even more prominently that in the case of negation, data in Table 4 also show that other quantitative tendencies that distinguish *más que* and *masque* synchronically (such as a preference for patient over agent semantic roles, and

¹⁹ See, however, Lee (2015) for a corpus-based research on the emergence of polarity sensitivity in related English expressions like ‘much’, and ‘many’.

²⁰ Although several cases exist in the literature of polarity items that are set apart from other uses by their accentual properties (see e.g. Greek *kanenas* in Giannakidou 1998), it is unclear why this use of *masque* became phonologically unstressed. It may be related with the fact that a functionally similar negative polarity conjunction *sino* is also unstressed in the language (e.g. *No ha trabajado tres días sino dos* ‘NEG has worked three days but two’, see Vicente [2010] for more details on *sino*). Note that, like *masque*, this expression is also transparently formed by a combination of formerly stressed words.

more indefinite nouns, see Table 2) appear to have been already in place from the start, and to have changed little since then. This suggests that they are/were not due to grammaticalization *per se*, but rather due to the conditions imposed by a particular use of *más que*.

Figure 7. Visual representation of the shift from meaning a) ‘more’ to b-c)²¹



Grammaticalization would have thus involved the conventionalization (reflected in an increase in frequency) of the pragmatic enrichment-derived transition from interpretation \neg *más que* in Figure 7 to *masque*. Semantic extensions involving the transition between an interval and an exact point have been found elsewhere (e.g. in the domain of time and space) in relation with grammaticalization processes (Herce 2017b) and might be common as expected from Grice (1981). When a new pragmatically-licensed meaning becomes frequent enough, it increases in autonomy, and thus acquires the possibility to develop (categorical) syntactic rules of its own (e.g. compulsory cooccurrence with preceding preverbal negation, as shown through this paper for *masque*), and/or idiosyncratic pronunciations (unstressed) different from their parent construction. Once this happens, the emerging differences can be used actively (see e.g. 17) to convey the very meaning differences that were initially responsible for the split of a single construction into two.

4 Conclusions

This paper has shown that (Northern) Peninsular Spanish has a negative polarity item *masque* in contemporary use, a fact that to my knowledge has not been noticed/explored to date.²² Because of its semantic, syntactic, and semantic differences with respect to related constructions, it cannot be profitably considered a subtype of any of them, either adversative or comparative (cf. RAE-ASALE’s 2009:862 “comparativas de alteridad”).

The introductory Section 1 presented the basic semantic, syntactic, and phonological traits that characterize this expression and distinguish it from related ones: a meaning “only”, compulsory co-occurrence with a preceding preverbal negation (*no*, *nadie*, *nunca*, *nada*, *ninguno*...), fixed word order with respect to its modified constituent, inseparability of *más* and *que*, and an unstressed phonological realization. It is simply an unfortunate fact about Spanish orthography that the different elements under the sequence *más que* are not distinguished. This fact must be one of the reasons that the expression has not been analyzed independently in the literature so far, and may

²¹ The difference between the meanings b) ‘except’ and c) ‘only’ is one of subjectification. In the terms of Traugott (1995:31) “a pragmatic-semantic process whereby meanings become increasingly based in the speaker’s subjective belief state/attitude toward the proposition”.

²² Various sources exist that investigate the expression *masque* elsewhere (e.g. in the Romance-based creole languages of Asia, where it has a concessive meaning as in Portuguese, see Veiga et al. 2012), as well as the various infrequent, dialectal, or historical uses of Spanish *masque* also as a conjunctive element with adversative or concessive meanings (see e.g. Templin 1929, Brooks 1933).

also be responsible for a few extra learning difficulties to Spanish L2 learners, who are understandably puzzled by the fact that *más que* ‘more than’ behaves nothing like *menos que* ‘less than’ (see Clegg & Seely 1985).

As a way of refining our knowledge of the expression, Section 2 explores, in a quantitative corpus endeavour, the geographic and morphosyntactic distribution of *masque* and related constructions. Based on the analysis of nearly 2000 tokens from oral Peninsular Spanish (COSER), negative polarity *masque* was found to occur more in northern than in southern Spain (see Figure 1), where the related competing alternative *na(da) más que* ‘nothing more than’ has been found to prevail instead (Figure 2). The findings in Table 1, in addition, supported a categorical negative polarity requirement for *masque* (but not so categorical for *namasque*), and identified other morphosyntactic trends (few agents, many patients, preference for indefinite noun phrases) that distinguish the expression from its closest (comparative or adversative) relatives.

To ascertain which of these relatives provided the source for contemporary negative polarity *masque*, Section 3 explored, in a historical corpus of Spanish (CORDE), those time periods (mid-XVII to mid-XIXth century) that appeared (see Figures 3 and 4) to be most critical for the emergence of the construction. It contributes, thus, to our knowledge of how negative polarity items can emerge in language (see e.g. Hoeksema (1994) for verbal sources, and Willis (2011) for indefinite pronouns, etc.). The results confirm that a large increase in frequency (of the ‘only’-denoting *masque*) took place during this period. Over these two centuries, verbal negation became a categorical requirement. These developments, as well as the characteristics that distinguish *masque* from comparative *más que* (i.e. unstressed, fixed word order) point toward the grammaticalization of the latter ‘large amount’-denoting expression into the former ‘only’-denoting one. Comparative *más que* must have therefore undergone a process of divergence (Hopper & Traugott 2003: 118) whereby, in the context/environment of negation, it took on a new meaning (see Figure 7). This use seems to have increased in popularity (and textual frequency, see Table 3), which caused it to achieve a level of independence from its parent construction, and allowed it to undergo independent developments (typical, once again, of grammatical rather than lexical elements) like loss of stress, univerbation (Lehmann 2020) and fixation of word order, that further estranged the new construction from its old source. It must also be mentioned, however, that many of the quantitative morphosyntactic preferences detected for *masque* in Section 2, appear to have been present even before the grammaticalization of the construction took place (see Table 4), which suggests that these were probably just part of the (bridging) context (Larrivé & Kallel 2020) that initially facilitated the innovative and pragmatically-enriched use of the comparative.

All the present findings notwithstanding, much room remains for further research. Because of the highly empirical, data-rich content of this paper, the more theoretical ramifications of *masque* and its history (e.g. with regard to Israel’s 1996, 2001 Scalar Model of Polarity, or with regard to Grammaticalization Theory more generally, Heine 2017) have not been fully explored. Also because of the focus on this concrete construction’s properties and evolution, those of related ones could not be addressed here. The competing *namasque*, for example, seems to have followed (or to be following now) largely the same footsteps as its predecessor. This raises the question of what exactly it is that drives these expressions down that path (i.e. towards polarity sensitivity). Negative-polarity may be the (almost inevitable) result of a certain use achieving independence from neighbouring ones in a context where negation is vastly more frequent than in a language’s grammar as a whole. Context, thus, easily becomes exponent. Future historical corpus explorations on *namasque* and comparable

expressions in other languages (e.g. Basque *baino* in 23-25, French (*plus*) *que* in 26) would be needed to assess to what degree their progression is parallel, or what properties or prerequisites must be present for some pragmatic use (e.g. “only/just”) to become conventionalized and ultimately achieve grammatical independence.

References

- Andersen, Henning. 1988. Center and periphery: adoption, diffusion. In Jacek Fisiak (Ed.), *Historical dialectology: Regional and social*. Trends in Linguistics. Studies and Monographs [TiLSM] 37: 39-83. <https://doi.org/10.1515/9783110848137.39>
- de Benito Moreno, Carlota, Javier Pueyo & Inés Fernández-Ordóñez. 2016. Creating and designing a corpus of rural Spanish. In KONVENS 2016, 20 de septiembre de 2016. Ruhr-Universität Bochum: 78-83.
- Bickel, Balthasar & Johanna Nichols. 2002. Autotypologizing databases and their use in fieldwork. In *Proceedings of the International LREC Workshop on Resources and Tools in Field Linguistics, Las Palmas*, vol. 2627. Nijmegen: MPI for Psycholinguistics.
- Boye, Kasper & Peter Harder. 2012. A usage-based theory of grammatical status and grammaticalization. *Language* 88, 1: 1-44. <https://doi.org/10.1353/lan.2012.0020>
- Brooks, John. 1933. Más que, mas que and mas ¡qué! *Hispania* 16, 1: 23-34. <https://doi.org/10.2307/332588>
- Bybee, Joan. 2006. From usage to grammar: The mind's response to repetition. *Language* 82, 4: 711-733. <https://doi.org/10.1353/lan.2006.0186>
- Clegg, J. Halvor, and Mark Seely. 1985. Set Theory as an Explanation for the Selection Between más que/más de in Spanish. In *Deseret Language and Linguistic Society Symposium* 11, 1: 58-69.
- Espinal, M. Teresa. 2000. On the semantic status of n-words in Catalan and Spanish. *Lingua* 110, 8: 557-580. [https://doi.org/10.1016/s0024-3841\(00\)00002-4](https://doi.org/10.1016/s0024-3841(00)00002-4)
- Espinal, M. Teresa, Susagna Tubau, Joan Borràs-Comes, and Pilar Prieto. 2016. Double negation in Catalan and Spanish. Interaction between syntax and prosody. In Pierre Larrivé, and Chungmin Lee (Eds.), *Negation and polarity: Experimental perspectives*: 145-176. New York: Springer. https://doi.org/10.1007/978-3-319-17464-8_7
- Fernández-Ordóñez, Inés. 2004. Nuevas perspectivas en el estudio de la variación dialectal del español: el Corpus Oral y Sonoro del Español Rural. In David Trotter (Ed.), *Actes du XXIV Congrès de Linguistique et Philologie Romanes*: 29-43. Tübingen : Niemeyer. <https://doi.org/10.1515/9783110923575.29>

- Giannakidou, Anastasia. 1998. Polarity Sensitivity as (non) Veridical Dependency. Amsterdam: John Benjamins. <https://doi.org/10.1075/la.23>
- Giannakidou, Anastasia, and Hedde Zeijlstra. 2017. The Landscape of Negative Dependencies: Negative Concord and N-Words. *The Wiley Blackwell Companion to Syntax, Second Edition*: 1-38. <https://doi.org/10.1002/9781118358733.wbsyncom102>
- Grice, Paul, 1981. Presupposition and Implicature. In Peter Cole (Ed.), *Radical Pragmatics*: 183–199. New York: Academic Press.
- Heine, Bernd. 2017. Grammaticalization. In Brian D. Joseph, and Richard D. Janda (Eds.), *The Handbook of Historical Linguistics*: 573-601. <https://doi.org/10.1002/9781405166201>
- Herce, Borja. 2017a. The diachrony of Spanish haber/hacer + time: A quantitative corpus-based approach to grammaticalization. *Journal of Historical Linguistics* 7, 3: 276-321. <https://doi.org/10.1075/jhl.15018.her>
- Herce, Borja. 2017b. Past–future asymmetries in time adverbials and adpositions: A crosslinguistic and diachronic perspective. *Linguistic Typology* 21, 1: 101-142. <https://doi.org/10.1515/lingty-2017-0003>
- Hoeksema, Jack. 1994. On the grammaticalization of negative polarity items. In *Annual Meeting of the Berkeley Linguistics Society*, 20, 1: 273-282. <https://doi.org/10.3765/bls.v20i1.1458>
- Hopper, Paul J. & Elizabeth C. Traugott. 2003. *Grammaticalization*. Cambridge: Cambridge University Press.
- Israel, Michael. 1996. Polarity sensitivity as lexical semantics. *Linguistics and Philosophy* 19, 6: 619-666. <https://doi.org/10.1007/bf00632710>
- Israel, Michael. 2001. Minimizers, maximizers and the rhetoric of scalar reasoning." *Journal of semantics* 18, 4: 297-331. <https://doi.org/10.1093/jos/18.4.297>
- Larrivéé, Pierre & Amel Kallel. 2020. The empirical reality of bridging contexts: Strong polarity contexts as the transition between NPIs and n-words. *Journal of Historical Linguistics* 10, 3: 427-451. <https://doi.org/10.1075/jhl.00010.lar>
- Lee, Ji Won. 2015. A corpus-based study of change and variation in much, many, far and often as Negative Polarity Items. PhD dissertation. State University of New York at Buffalo.
- Lehmann, Christian. 1995. *Thoughts on grammaticalization*. Munich: LINCOM.
- Lehmann, Christian. 2020. Univerbation. *Folia Linguistica* 54, 1: 205-252. <https://doi.org/10.1515/flih-2020-0007>

- Mann, Jason, David Zhang, Lu Yang, Dipanjan Das, and Slav Petrov. 2014. Enhanced Search with Wildcards and Morphological Inflections in the Google Books Ngram Viewer. In Proceedings of ACL Demonstrations Track. Association for Computational Linguistics, June 2014. <https://doi.org/10.3115/v1/p14-5020>
- Nichols, Johanna. 2015. Types of spread zones. Open and closed, horizontal and vertical. In Rik De Busser & Randy J. LaPolla (Eds.), *Language structure and environment. Social, cultural, and natural factors*. Cognitive Linguistic Studies in Cultural Contexts 6: 261–286. Amsterdam: Benjamins. <https://doi.org/10.1075/clsc.6.10nic>
- O'Neill, Teresa. 2011. The syntax of ne... que exceptives in French. *University of Pennsylvania Working Papers in Linguistics* 17, 1: 175–184.
- Poole, Geoffrey. 2011. Focus and the development of N-words in Spanish. *Romance languages and linguistic theory 2009. Selected papers from 'Going Romance' (Nice 2009)*: 291-303. <https://doi.org/10.1075/rllt.3.17poo>
- RAE-ASALE (= Real Academia Española / Asociación de Academias de la Lengua Española) 2009. *Nueva Gramática de la Lengua Española*. Madrid: Espasa-Calpe. <https://doi.org/10.5944/epos.25.2009.10627>
- Templin, Ernest H. 1929. An Additional Note on mas que. *Hispania* 12, 2: 163-170. <https://doi.org/10.2307/330881>
- Traugott, Elisabeth Closs. 1995. Subjectification in grammaticalisation. In Dieter Stein and Susan Wright, (Eds.), *Subjectivity and subjectivication: Linguistic perspectives*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511554469.003>
- Vallduví, Enric. 1994. *Polarity items, n-words, and minimizers in Catalan and Spanish*. HCRC Publications, University of Edinburgh. <https://doi.org/10.1515/prbs.1994.6.2-3.263>
- Veiga, Nancy Vázquez, and Mauro A. Fernández. 2012. Maskin, maski, masque... in the Spanish and Portuguese creoles of Asia. In Hugo C. Cardoso, Alan N. Baxter, and Mário Pinharanda-Nunes (Eds.), *Ibero-Asian Creoles: Comparative Perspectives* 46: 181-204. <https://doi.org/10.1075/cll.46.07vei>
- Vicente, Luis. 2010. On the syntax of adversative coordination. *Natural Language & Linguistic Theory* 28, 2: 381-415. <https://doi.org/10.1007/s11049-010-9094-0>
- Von Stechow, Kai, and Sabine Iatridou. 2007. Anatomy of a modal construction. *Linguistic Inquiry* 38, 3:445-483. <https://doi.org/10.1162/ling.2007.38.3.445>
- Willis, David. 2011. Negative polarity and the quantifier cycle: Comparative diachronic perspectives from European languages. In Pierre Larrivée, and Richard P. Ingham (Eds.), *The evolution of negation. Beyond the Jespersen cycle*: 285-323. <https://doi.org/10.1515/9783110238617.285>