

The grammaticalization of epistemicity in Ibero-Romance:

Alike processes, unlike outcomes

Typologically-unexpected overt expletives can be found in a restricted number of non-standard Ibero-Romance null-subject varieties. Historical data suggest that these overt expletives, which in today's varieties show both discourse-oriented and expletive characteristics, have their origin in 15th century impersonal epistemic constructions. This article argues that it is the expletives' epistemic origin which gives rise to, and thus explains, their present-day heterogeneous properties, in particular their function as a marker of epistemicity in a number of varieties. Despite undergoing the same mechanisms of change, the variation in modern Ibero-Romance is understood to be a consequence of the different stages and degrees of grammaticalization reached in each variety.

Keywords: overt expletives; epistemicity; grammaticalization; pragmaticization; Ibero-Romance; dialect syntax

1. Introduction

This article offers a formal analysis of the historical development of so-called “overt expletives” in Ibero-Romance. Expletive subjects, i.e. non-referential “placeholders” used to fill a structural (here, the subject) position as a last-resort mechanism (cf. Svenonius 2002), are attested in a minority of non-standard Ibero-Romance varieties. These Ibero-Romance expletives show discourse-oriented properties — notably, marking epistemic value in many dialects— in addition to their expletive properties (glossed as ELLO, to avoid confusion with Ibero-Romance's homophonous third person referential pronouns):¹

- (1) A: ¿Y tu cre' en brujería?
And you believe.2SG in witchcraft
B: Claro que sí, **ello** existe cosas buenas y cosas malas.
Clear that yes ELLO exist.3SG things good and things bad
'And do you believe in witchcraft? Of course I do, there are good things and bad things' (Dom.Sp.)
- (2) A: Não há hera?
No there.is ivy
B: **Ele** há heras, há.
ELLO there.is ivy there.is
'Isn't there ivy? Yes, there's ivy.' (E.Pg., Vila do Corvo; CORDIAL-SIN CRV38)

¹ The cover term ELLO is used where variation between Ibero-Romance varieties (i.e. the morpholexical items *ele* in Portuguese, *ell* in Catalan and *ello* in Spanish) is not at stake. The article also uses the following abbreviations: (Dom.)Sp. for (Dominican) Spanish; (E.)Pg. for (European) Portuguese; Bal.Cat. for Balearic Catalan; and CONJ as a gloss for the “conjunctive” use of the finite complementizer *que* ‘that’ (cf. Section 3.2; also Alarcos Llorach 1994, Etxepare 2012, Corr 2016b).

Obligatory expletive subjects in non-null subject languages (cf. non-referential yet compulsory English ‘it’, ‘there’ and French *il* in the sentences *il semble qu’il y avait une fête hier soir* ‘it seems that **there** was a party last night’, or *pendant qu’il pleuvait, il est arrivé un homme* ‘whilst **it** was raining, **there** arrived a man’) are the motivation behind Chomsky’s (1982) Extended Projection Principle (EPP) which stipulates that every clause must have a subject (in the canonical subject position, SpecTP). Such elements are strictly prohibited in standard null-subject Ibero-Romance varieties, conforming to Rizzi’s (1982, 1986) theory that the presence of null referential subjects allows for the possibility of null expletive subjects, an implication which is commonly taken as a “ban” on overt expletives in consistent null-subject languages.

However, recent work on the null-subject parameter (Holmberg et al. 2009, Biberauer et al. 2010) and expletives (Biberauer & van der Wal 2012, Biberauer & Cognola 2014) has expanded the respective typologies of both, in the latter case moving beyond a “narrow” definition of the notion of what constitutes a syntactic expletive. Given ELLO’s pragmatic properties, this article appeals to a broader notion of expletive, though, strictly speaking, discourse-sensitive expletives are, according to Biberauer & van der Wal (2012: 4), “forerunners” to what might be referred to as the “narrow” definition of expletives. As such, this article continues to follow the tradition of referring to Ibero-Romance cases as “(overt) expletives” (cf. Uriagereka 1995, 2005; Carrilho 2005, 2008) whilst simultaneously attempting to distance these items from the “expletive” label, associated with “narrow” TP-expletives, by referring to them as ELLO.

In the former case, parametric variation is recognised between different categories of languages exhibiting null subjects, leading to terms such as “partial” null-subject (Holmberg et al. 2009) or “semi” null-subject (Huang 2000) language. Note that Dominican Spanish, whose data will be discussed in the present article, is treated here as a full null-subject language on the view that, although it exhibits an elevated use of overt subjects relative to other varieties of Spanish, Dominican Spanish does not exhibit the specific set of morphosyntactic properties required for classification as a “partial” or “semi” null subject language, or similar terminology (cf. Camacho 2013a: 43-44, 2016 for discussion).

Despite such revisions in the classification of expletives and the null-subject parameter, the occurrence of the Iberian (epistemic) expletive ELLO in otherwise prototypically null-subject varieties is nonetheless considered to be typologically highly anomalous. Accordingly, this article aims to answer two principle questions: i) assuming a “wide” understanding of the term expletive (e.g. one which permits pragmatic values), how do overt expletives arise in a consistent null-subject language in the first place; and ii) how do these elements come to obtain epistemic meaning?

Here, it is argued that there is a link between ELLO’s origin and its present-day role in marking epistemicity —defined by de Haan (2001: 201) as “the degree of confidence the speaker has in his or her statement”—, wherein ELLO is used by the speaker to indicate their (strong) commitment to the truth of their utterance’s proposition (cf. also Palmer 2001). Presenting novel empirical data from historical and modern-day Ibero-Romance, this article presents the case that modern-day ELLO has an unexpected origin in epistemic impersonal constructions from the 15th century onwards, leading to a diversification of ELLO’s value(s) and contexts over time:

- (3) a. Porque, si **ello** es verdad que no hace nada
 [...]

because if ELLO be.3SG truth that not do.3SG nothing

'Because, if it is true that he doesn't do anything...' (Spanish; CORDE, 16th c., San Juan de la Cruz, *Llama*)
- b. **Ell** es veritat que morí en las mans de sos
 ELLO be.3SG truth that die.PST.3SG in the hands of his

enemichs

enemies

'It is true that He died at the hands of His enemies' (Catalan; 1844, Josep Antoni Arnautó, *Meditacions per cada dia del any*)
- c. Oh! **Ele** há frade no caso?
 oh ELLO there.is monk in.the case

'Oh! There's a monk involved?' (Portuguese; CdP; 1846, Almeida Garrett, *Viagens*)

Specifically, ELLO undergoes a particular type of grammaticalization (Roberts & Roussou 2003; Roberts 2011), viz. pragmaticization (Diewald 2011). In this process, ELLO —originally a referential third person (subject) pronoun *ello*— loses its referentiality (undergoing bleaching and a change of formal category), but gains epistemic value via the encoding of contextually-derived epistemic meaning (viz. subjectification) in the pronoun itself. From its epistemic origin (3a-b), ELLO progressively extends via analogy into new environments (3c), eventually giving rise to the diatopic variation in its formal and interpretational properties attested today. Despite a common origin and mechanisms of change, ELLO's present-day variation is thus argued to instantiate the different stages and degrees of grammaticalization reached by the element across Ibero-Romance varieties.

The structure of the article is as follows: Section 2 presents our empirical point of departure, viz. present-day ELLO and its variation; Sections 3 and 4 trace ELLO's historical development, including a brief look at possible examples of ELLO elsewhere in Romance; and Section 5 concludes.

2. Unlike outcomes: properties of modern-day ELLO

Our account starts, perhaps somewhat counterintuitively, by outlining the outcomes of ELLO's diachronic development in Ibero-Romance, using novel data collected from the non-standard regional varieties in which ELLO is best attested today: European Portuguese, Dominican Spanish and, to a lesser degree, Balearic Catalan varieties. Data were obtained from a range of sources, primarily from corpora (detailed in the bibliography) and, in the case of present-day varieties only, native speaker informants via questionnaire and interview/discussion (cf. Corr 2015a for details of the empirical survey). A range of native Ibero-Romance speakers were consulted, but positive confirmation of ELLO was only found in three varieties: European Portuguese from the region of Trás-os-Montes, Dominican Spanish and Balearic Catalan (specifically, *mallorquí*).²

² In the present article, examples without citation are those provided by native speaker consultants.

Given the heterogeneous nature of ELLO across Ibero-Romance (Corr 2015a, 2016a), this article presents ELLO's expletive (Section 2.1) and discourse-oriented properties (Section 2.2) as attested in the surveyed varieties only. Space constraints do not permit the discussion of the microvariation in ELLO's properties across the data discussed in previous work on Ibero-Romance expletives, for which reason the present article's characterisation applies only to the behaviour of ELLO as attested in our empirical survey in the present article. The reader is instead encouraged to consult previous work on Ibero-Romance —in particular, Carrilho (2005, 2008), Silva-Villar (1998), Hinzelin & Kaiser (2006, 2007)— to verify the different behaviours of ELLO across Ibero-Romance, since these do not in all cases coincide with the descriptions given in Section 2 (Corr 2016a: 86-87 provides comment on such cases).³

Here, it is proposed that the variation seen today is an effect of how advanced the grammaticalization of ELLO has been in each variety, following a formal approach to grammaticalization as “reanalysis ‘upwards’ along the functional structure” (Roberts & Roussou 2003: 71), or “up the tree” (van Gelderen 2011: 51). In terms of the degree of grammaticalization, Balearic Catalan ELLO (*ell*) is understood to represent the most advanced variant and Dominican Spanish ELLO (*ello*) the least advanced along this process.

2.1 Expletive characteristics

ELLO is attested in all prototypical expletive contexts (Svenonius 2002), including with existential (4a), extraposition (4b), atmospheric (4c) predicates, and in unaccusative structures involving a postverbal undergoer (4d):

- (4) a. Mamá, mamá! **Ello** hay un búho en el zin!
 mum mum ELLO there.is a owl on the roof
 ‘Mum, mum! There’s an owl on the roof!’ (Dom.Sp.)
- b. **Ele** é sabido que o Mário gosta de dançar.
 ELLO be.3SG know.PST.PTCP that the Mário like.3SG of dance.INF
 ‘It’s known that Mário likes dancing’ (E.Pg, Trás-os-Montes)

³ The present article also leave aside discussion of the parallel between ELLO and expletive-(like) uses of demonstratives in Ibero-Romance (cf. Carrilho 2005, Hinzelin 2009, Camacho 2013b, amongst others):

- (i) **Això** eren tres germanes fadrines, que estaven totes soletes
 that be.PST.3PL three sisters maidens that be.PST.3SG all alone
 ‘Once there were three maiden sisters, who lived all alone’ (Mallorcan Catalan; Hinzelin 2009: 14)

Such cases are disregarded here on the view that there is no *a priori* motivation to assume the phenomena are necessarily *directly* related (although, as an anonymous reviewer observes, personal pronouns can still have demonstrative values). In support of this decision, note that the “expletive”(-like) use of the demonstratives is far more frequent and accepted than that of ELLO; the demonstratives and ELLO do not share a common diatopic or diachronic attestation; and, more crucially, such uses have been shown to have a separate characterisation and distribution from that of ELLO in non-standard European Portuguese (cf. Carrilho 2005):

- (ii) Mas esse já conhecem, {ele/*isto/*isso/*aquilo} já?
 but that already know.3PL ele/isto/isso/aquilo already
 ‘But you already know that, don’t you?’ (E.Pg.; example adapted from Carrilho 2005: 161)

- c. **El** farà més fred que mai aquest hivern.
 ELLO make.3SG more cold than ever this winter
 ‘It’s colder than ever this winter’ (Bal.Cat.)
- d. Eso no e’ na’, **ello** un día nació un niño en la
 that not be.3SG nothing ELLO one day be.born.PST.3SG a child in the
 casa de mi abuela.
 house of my grandmother
 ‘That’s nothing; one day a baby was born in my grandmother’s house’
 (Dom.Sp.)

It is also found in passive (5a), impersonal-*se* (5b) and impersonal copular constructions (5c); and with modal raising verbs only where the non-finite verb in the complement clause is impersonal (cf. the contrast between (5d-e)):

- (5) a. **Ele** são amansadas uma vez por outra
 ELLO be.3PL cultivate.PST.PTCP.F.PL one time by another
 ‘They [zonas ‘the zones’] are cultivated from time to time’ (E.Pg, Cedros, Horta;
 CORDIAL-SIN CDR01-N)
- b. Aqui, **ele** usa-se os trajes novos é pela festa
 here ELLO use.3SG=REFL the suits new be.3SG for.the fiesta
 do Espirito Santo.
 of.the Holy Spirit
 ‘Here, we wear new suits for the festival of the Holy Spirit’ (E.Pg., Cedros,
 Horta; CORDIAL-SIN CDR53-N)
- c. Isso, **ele** são parvoeiras!
 That ELLO be.3PL gobbledeygook
 ‘That, that’s gobbledeygook!’ (E.Pg., Carrapatelo, Évora; CORDIAL-SIN CPT21)
- d. **Ele** deve ser tarde já.
 ELLO ought.3SG be.INF late already
 ‘It must be late by now.’ (E.Pg., Trás-os-Montes)
- e. #**Ele** deve acompanhá-lo para casa
 ELLO ought.3SG accompany.INF=him to home
 (‘S/he ought to take him home.’) (E.Pg., Trás-os-Montes)

Note that (5e) is felicitous if ELLO is taken to be a referential pronoun, but it is deviant with the intended interpretation as an overt expletive. The grammaticality of ELLO in raising constructions appears to be determined by the thematic properties of the lower verb (e.g. *ser* ‘to be’ in (5d), *acompanhar* ‘to accompany’ in (5e)), indicating that the use of ELLO is restricted to a limited subgroup of predicates (i.e. those observed in (4-5)). The appearance of ELLO in an apparent placeholder role and position in the typical environments in which canonical expletive subjects may be found, as well as a number of impersonal environments more generally, gives rise to ELLO’s characterization as an expletive.

2.2 Discourse-oriented properties

However, in addition to its expletive characteristics, ELLO also exhibits many discourse-related properties (as already observed in, *inter alia*, Silva-Villar 1998, Uriagereka

2005, Hinzelin & Kaiser 2007, Hinzelin 2009). This section summarises ELLO's behaviour with respect to its "optionality" (Section 2.2.1); root vs. embedded asymmetry (Section 2.2.2); behaviour in interrogatives (Section 2.2.3); meaning contribution (Section 2.2.4); sentential distribution (Section 2.2.5); and its variation across Ibero-Romance along these dimensions.

2.2.1 Optionality

The use of ELLO is "optional" insofar as it can be omitted from the syntax without producing an ungrammatical sentence, as in (6):

- (6) Acho que (ele) há uma aranha no banho!
 think.1sg that ELLO there.is a spider in.the bathroom
 'I think there's a spider in the bathroom!' (E.Pg., Trás-os-Montes)

However, since the appearance of ELLO entails a change in the sentence's interpretation (cf. Section 2.2.4), ELLO is never really "optional" inasmuch as it is grammatically required for the discourse-oriented effect it has on the sentence's meaning. This contrasts with canonical expletive subjects in non null-subject languages which are compulsory placeholders in the dedicated subject position (viz. SpecTP). The apparent optionality of ELLO relative to canonical non null-subject expletives is in line with the hypothesised grammaticalization path of expletives towards "obligatorization" (cf. Biberauer & van der Wal 2009: 6, and references therein), though we are skeptical about the plausibility of predicting that Ibero-Romance ELLO will develop into a "narrow" TP-expletive on a par with English or French.

2.2.2 Root vs. embedded asymmetry

Cross-linguistically, certain syntactic effects exhibit a root vs. embedded asymmetry, such that some syntactic phenomena are only observed in root clauses or a restricted subset of embedded clauses, known as "main clause phenomena" (Hooper & Thompson 1973; Aelbrecht et al. 2012). Whilst unattested in Balearic Catalan and Dominican Spanish, ELLO is occasionally attested in European Portuguese dialects — and in some Dominican Spanish native-speaker intuitions when elicited — in the embedded complements of verbs of saying (7) and perception (cf. (6) above) and "premise" conditionals (8), all of which are known to be structures which permit main clause phenomena:

- (7) Dicen que **ello** hará mucho frío en este diciembre.
 say.3PL that ELLO make.FUT.3SG much cold in this December
 'They say it'll be very cold this December.' (Dom.Sp.)
- (8) Que préstimo teem agora as talhas, se **êle** não há já vinhas?
 what utility have.3PL now the tubs if ELLO not AUX.3SG still vineyards
 'So what is the utility of the tubs now, if there are no more vineyards?' (from Alentejo, South-Central Portugal in Ribeiro 1927: 28, apud Carrilho 2005: 212)

The possibility of main clause phenomena is often understood in terms of the internal structure of the CP, such that the types of construction permitting main clause phenomena have a larger CP (e.g. Haegeman 2003). Specifically, such constructions

exhibit a CP which projects a Force Phrase, a projection understood to be absent in (structurally smaller) non-assertive clauses. Witness the contrast between the “premise” conditional in (8) and the following (non-assertive) “event” conditional, which is understood to have a “truncated” CP (Haegeman 2003):

- (9) **Vamos ter de esperar, que não sabemos se*
 go.1PL have.INF of wait.INF CONJ not know.1PL if
ele é fácil chegar à noite.
 ELLO be.3SG easy arrive.INF at.the night
 ‘We’re going to have to wait, as we don’t know how easy it’ll be to make our way there at night.’ (E.Pg., Trás-os-Montes)

Since the Force Phrase is understood to encode assertive illocutionary force, the presence of ELLO in (6) appears to be an effect of the syntactic encoding of assertion in the CP, rather than the subject requirement of the EPP.

2.2.3 Behaviour in interrogatives

Another property which suggests that the use of ELLO is determined by illocutionary force rather than the requirements of the EPP is its behaviour in interrogative sentences. In Dominican Spanish and European Portuguese, we observe that ELLO is only grammatical in interrogative sentences which conform to the following three properties (ELLO is ungrammatical in all interrogatives in Balearic Catalan, cf. Corr 2016a). Firstly, the sentence must be a polar interrogative (10), not a wh-interrogative (11):

- (10) **Ele** há carne para comer esta noite?
 ELLO AUX.3SG meat for eat.INF this night
 ‘There’s meat for dinner tonight, right?’ (E.Pg., Trás-os-Montes)
- (11) ***Ele** o que há para comer esta noite?
 ELLO what AUX.3SG for dine.INF this night
 Intended: ‘What’s for dinner tonight?’

Secondly, ELLO must surface in a sentence-initial position (i.e. unlike in prototypical interrogatives in Ibero-Romance, subject-verb inversion is disallowed):

- (12) (**Ele**) há (***ele**) carne para comer esta noite?
 ELLO AUX.3SG ELLO meat for eat.INF this night

Thirdly, at least in the case of European Portuguese, the speaker sets up an expectation that the addressee will confirm the speaker’s query. In (10) the speaker expects an affirmative reply from the addressee (roughly translated as ‘there’s meat for dinner tonight, right?’) and does not make a genuine enquiry about whether or not meat will be on the menu that evening. Thus despite having an interrogative sentence structure, (10) is not a “true” question but is closer in illocutionary value to an assertion (cf. Han 2002).

2.2.4 Interpretation

The expletive ELLO can be broadly characterized as a marker of emphatic epistemic value, though its specific meaning contribution varies between the Ibero-Romance

varieties in which it is used. Thus, in European Portuguese, speakers describe the function of ELLO as conveying speaker commitment to or belief in the truth of the proposition (my informants; Carrilho 2005, 2008). As Carrilho (2005: 203) puts it:

In declarative sentences [...] it is the assertive value of the utterance that is emphasized, the expletive acting as a sort of strong evidentiality marker. In fact, with such expletive constructions the speaker appears to indicate that he strongly stands for the truth of the statement he is making.

However, we characterize ELLO as having an epistemic rather than evidential value on the view that the former marks the speaker's degree of certainty or the necessity/possibility of the truth of the proposition (Faller 2011: 661) whereas evidentiality "indicates the speaker's type of source of information" (ibid.). A similar interpretation is described for Spanish ELLO, characterized at the turn of the 20th century as being synonymous with adverbs like *realmente*, 'really', and *en verdad*, 'in truth' (Henríquez Ureña 1939: 228). Its modern-day pragmatic value is evident in examples such as (13), where the speaker, having mistakenly communicated to her addressee the wrong location of the local bus stop, corrects herself by saying:

- (13) **Ello** hay una parada frente a la Secretaría.
 ELLO there.is a bus.stop opposite to the Secretariat
 'There's a bus stop opposite the secretariat' (Dom.Sp.; *Hoy Digital* newspaper)

The meaning contribution of ELLO in Balearic Catalan is less specific, being described as an exclamatory, "intensifying" element (Veny 1989, Bartra-Kaufmann 2011, Nicolau Dols, p.c.), with an interpretation similar to that of an interjection. If the "epistemic" hypothesis put forward in the remainder of the present article holds, then the prediction would be that Balearic Catalan ELLO has lost its epistemic value and become a more generalised emphatic or exclamative discourse marker. That is, Balearic Catalan is the variety in our data set which exhibits the most advanced stage of grammaticalization of ELLO, a hypothesis borne out in the next section.

2.2.5 Structural position

The empirical data reveal differences between Ibero-Romance varieties with respect to ELLO's distribution within the functional structure, specifically within the left-periphery, an area of the clause identified as, amongst other roles, the locus of discourse-oriented features and interpretation (Benincà 1988, Rizzi 1997, and much subsequent work). On the view taken here that grammaticalization can be understood as reanalysis upwards (or leftwards) along the functional structure, differences in structural position will be taken to correlate with degree(s) of grammaticalization, such that a relatively high position to the left of an element's starting position in the functional structure will reflect a more advanced state of grammaticalization. In contrast, a relatively low position to the left of the starting point will reflect a less advanced state of grammaticalization. For us, the related concept of pragmaticization involves the same mechanisms as grammaticalization, *viz.* discrete reanalysis of structural patterns underlying surface empirical linguistic data, whence the "alike processes" of this article's title. Note though that, as observed in Roberts (2011), it remains to be seen how to reconcile the fact that grammaticalization is understood in terms of feature loss

or semantic bleaching (i.e. the higher along the functional hierarchy, the more logical, or featurally simpler, the content of the functional projections), with the observation that the highest functional projections are typically associated with first-person features (Speas & Tenny 2003, Giorgi 2010, Sigurdsson 2010).

This article assumes that the sentence structure is minimally organized as follows:

(14) $[_{CP} \quad [_{TP} S \text{ AUX } [_{VP} V O (X)]]]$ (adapted from Ledgeway 2013: 289)

However, given the progressive decomposition of the C-domain into a syntactic space with its own internal structure (culminating in the “split” CP of Rizzi 1997), a basic structure for the left-periphery is assumed as follows:

(15) $[_{\text{Force}} [_{\text{Top}} [_{\text{Int}} [_{\text{Foc}} [_{\text{Fin}} [_{\text{TP}} \dots]]]]]]]$

Following Benincà & Poletto (2004: 53), it is assumed that “recursion is not an option” within this domain (and generally), which has led to increasingly complex mappings of the “fine structure” of the left-periphery (see e.g. Rizzi 2004: 242, Ledgeway 2012: 179). Since the finer details of this space are not necessary to the present analysis, we retain use of the simpler functional structures of (14) and (15) here.

In their more typical, and universally-accepted, use as a third person referential pronoun in Ibero-Romance, *ele/ell/ello* may occur in the canonical subject position, SpecTP. In Section 3, we see that it is in this function, viz. as a third person referential pronoun, that today’s ELLO originates. The appearance of ELLO higher than SpecTP will be taken as evidence in support of the analysis that grammaticalization of this element has occurred.

In present-day Ibero-Romance, Balearic Catalan ELLO surfaces in the highest position in the functional hierarchy relative to other Ibero-Romance ELLO variants, occurring exclusively sentence-initially (cf. also Hinzelin 2009: 17), including to the left of imperatives (16a), which are commonly assumed to undergo V-to-C movement, though targeting a higher position (understood here to be ForceP) within this domain than other clause types; wh-exclamatives (16b), and foci (16c):

- (16) a. **El** dóna'm el pa!
 ELLO give.IMP=to.me the bread
 ‘Pass me the bread!’
- b. **El** qué maca que us ha quedat la foto!
 ELLO what pretty that to.you= AUX remained the photo
 ‘That photo turned out so well!’
- c. **El** FLORS creixen al sostre, que estrany!
 ELLO flowers grow.3SG to.the roof what strange
 ‘FLOWERS are growing on the roof, how strange!’

Dominican Spanish ELLO surfaces in the lowest position of the Ibero-Romance varieties under investigation here. Though it always surfaces clause-initially (generally sentence-initially, with some speakers allowing its use in certain embedded structures when tested, as in Section 2.2.2), it cannot precede any left-peripheral elements, as exemplified by its ungrammaticality preceding the wh-exclamative in (17):

(17) *¡**Ello** qué bonita te ha quedado la foto!

ELLO what pretty to.you AUX remained the photo
 ‘That photo turned out so well!’ (wh-exclamative)

Nonetheless, since it can occur sentence-initially in polar interrogatives (usually understood to involve V-to-C movement, though targeting a lower position than imperatives) and entails a pragmatic interpretation (cf. Section 2.2.3-Section 2.2.4), it is assumed that ELLO surfaces in a low left-peripheral position in Dominican Spanish (cf. Corr 2015a for arguments that Dominican Spanish ELLO is located in Fin).

Finally, European Portuguese ELLO can surface in two positions (cf. also Carrilho 2005), one above (18a) and another below (18b-c) topics and foci, but not in between these projections, since it is incompatible with wh-interrogatives ((18d); cf. Section 2.2.3) and wh-exclamatives (18e):

- (18) a. Que **ele**, ainda hoje, há essa tradição, cá.
 That ELLO even Today there.is that tradition here
 ‘Even today we still have that tradition here’
- b. Até a esta altitude, **ele** há gente que vive da terra.
 even at this altitude ELLO aux people who live.3SG of.the earth
 ‘Even at these altitudes, there are people who live off the earth’
- c. TODA A NOITE **ele** choveu!
 all the night ELLO rain.PST.3SG
 ‘It rained all night!’
- d. ***Ele** que fazes aqui?
 ELLO what do.2SG here
 Intended: ‘What are you doing here?’
- e. ??**Ele** (olha) que linda ficou a foto!
 ELLO look.3SG what pretty be.3SG The photo
 ‘(Look how) the photo turned out so well!’

It is thus assumed that European Portuguese ELLO has two dedicated positions: a lower position at the TP/CP border like Dominican Spanish and a higher-peripheral position. Since European Portuguese ELLO cannot precede imperatives (e.g. **Ele dá-me o pão!* ‘Pass me the bread!’), the latter position must be lower than Balearic Catalan ELLO’s position. The availability of two positions for ELLO in European Portuguese is taken to reflect language change in progress. The grammaticalization —indeed, pragmaticization, since we are exclusively dealing with left-peripheral elements— of ELLO is thus most advanced in Balearic Catalan, least advanced in Dominican Spanish, with European Portuguese varieties representing an intermediary developmental stage between the two.

3. Alike processes: origin of ELLO?

This article has proposed that the variation in ELLO’s properties and distribution seen today across Ibero-Romance reflects the different degrees of grammaticalization/pragmaticization undergone by the respective variants (*ello*, *ele*, *el*, etc.) in each variety. Broadly, then, we observe ELLO developing from a T-element, where it affects the propositional content of the sentence, into a C-element, where it contributes not at-issue

content, conveying in many of today's dialects an epistemic evaluation.

This section focuses on the historical circumstances which brought about these changes.

Today's Ibero-Romance referential third person pronouns *ello/ele/ell* are reflexes of the Latin neuter pronoun *ILLUD* (cf. Spanish *él* 'he' < *ILLE*, *ella* 'her' < *ILLA*). According to Hinzelin (2009), it is the *loss* of neuter gender in Ibero-Romance that causes (the reflexes of) *ILLUD* to take on a new function. Conversely, Uriagereka (2005) suggests that Ibero-Romance's *maintenance* of a NEUTER feature from the Latin gender system licenses the overt expletive. However, Adam Ledgeway (p.c.) points out that the non-referential use of *ELLO* under discussion here is *not* a leftover from the Latin gender (nominal declension) system:⁴ accordingly, the neuter characteristics of morphosyntactic items in Latin and modern Ibero-Romance should not be treated as interchangeable.

Other authors (Henríquez Ureña 1939, Bartra-Kaufmann 2011, Gupton & Lowman 2013) have proposed that today's *ELLO* has its origin in a former impersonal expletive usage, but this hypothesis in itself does not bring us any closer to understanding why an overt expletive is available in these prototypically null-subject varieties in the first place, nor does it explain why *ELLO* should develop an epistemic interpretation. On the basis of historical Spanish data, this section advances the hypothesis that *ELLO*'s expletive usage originates in the juxtaposition of an epistemic impersonal construction, *ello es verdad/cierto* ('it is true'), with a separate independent clause headed by the finite complementizer *que*, a paratactic connection which is reanalysed as a subordinating relation. The epistemic interpretation of the sentence becomes codified in the semantics of the pronoun *ello* via the process of pragmatic inference (cf. Section 5). In acquiring an epistemic meaning independent of the original epistemic impersonal context, the pronoun can then extend into new syntactic environments through analogy.

3.1 Textual sources and preliminary evidence

Given the very sparse attestation of *ELLO* in historical Portuguese and Catalan corpora, we restrict our analysis to Spanish historical data, for which the textual record is much more ample. The Spanish data were obtained from Mark Davies' *Corpus de Español* (CdE) and the Real Academia Española's *Corpus diacrónica de español* (CORDE), which feature texts from the medieval period (the *Corpus de Español*'s texts date from the 1200s, whereas *CORDE*'s earliest texts date "from the language's origins") to the 20th century. Both are publically-available annotated corpora, and were searched across all time periods for matches of tokens of constructions involving *ELLO*, which were then manually inspected for the context in which they occurred. Referential uses and repeated tokens of *ELLO* were manually eliminated, with the remaining tokens being retained to form our data set (cf. Table 1 for a breakdown of tokens per construction).

The earliest attestation of an *ELLO* construction in our historical data set is from the late 1400s:

(19) Empero cuenta que jhesu cristo mismo hablo al emperador [...]

⁴ The same applies to the mass/count distinction encoded via "neuter agreement" on adjectives and third person clitic pronouns in Eastern Asturian, which constitutes an innovation in this variety (Fernández-Ordóñez 2012: 82).

Ello es verdad que aparecio el dicho señal
 ELLO be.3SG truth that appear.PST.3SG the said sign
 ‘although it states that Jesus Christ himself spoke to the emperor [...]; it is true
 that the said sign appeared’ (CdE; 1498, Ampíés, *Tratado de Roma*)

The placeholding function of ELLO in (19) contrasts with the referential use of *ello*, as illustrated by its modification by *todo* ‘all’ in (20), during the same period:

(20) Todo ello es bueno para poner debajo tus pies.
 All it be.3SG good for put.INF under your feet
 ‘All [of] that is good to have at your feet.’ (CdE; 1548, Mal Lara, *Recibimiento*)

To justify this article’s proposal that epistemic impersonals are the correct place to begin the reconstruction of the development of ELLO, the following observations are offered for consideration. Firstly, ELLO’s attestation in epistemic constructions predates its appearance in other environments. Whereas the epistemic expressions *ello es verdad/cierto que* (‘it is true that’) are found from the 15-16th centuries onwards in the corpora, analogous extraposition impersonals without epistemic value (e.g. *ello es necesario/visto/fácil que* ‘it is necessary/seen/easy that’) are only attested in our data set from 17–18th centuries, as illustrated in Table 1, which also includes tokens of existential constructions involving ELLO (e.g. *ello hay* ‘there is’; cf. Section 4 for discussion and empirical examples of ELLO in non-epistemic constructions):⁵

Table 1: Tokens of ELLO in CORDE and *Corpus de español* according to century and type of impersonal construction

	1400s	1500s	1600s	1700s	1800s	Total
Epistemic extraposition	1	4	8	9	20	43
Existential			1	1	2	4
Non-epistemic extraposition			2	5	5	13

It does not seem accidental that the earliest and most frequent context in which ELLO is found, predating its use in other environments by some two hundred years in the data set, has a very similar interpretation to the primary semantic-pragmatic value of ELLO today.

Secondly, the attestation of constructions involving ELLO does not conform to the pattern and frequency of their non-ELLO counterpart constructions. In our data set, there are many more examples of “regular” (non-ELLO) existential constructions with *hay* (‘there.is’) than there are tokens of the (non-ELLO) epistemic constructions *es verdad/cierto que*: in the Corpus del Español alone, there are 106,976 tokens of existential *hay* versus 4,603 tokens of *es verdad/cierto que* combined. However, by far the most amply attested ELLO constructions are the epistemic extraposition structures

⁵ This does not exclude the possibility that ELLO could be attested in such environments, or indeed other contexts, at other historical periods (given the accidental nature of historical records and their survival). Further, on the view that language change is abrupt (even if gradual in appearance, cf. Lightfoot 1997), the reanalysis of ELLO constructions is discrete: i.e. it either happens or it does not. The chronology outlined here thus more accurately reflects the diffusion of ELLO across Peninsular Spanish.

ello es verdad/cierto que, of which there are 43 tokens across the data set (cf. Table 1). Existential *haber* constructions with ELLO (viz. *ello hay* ‘ELLO there.is’), on the other hand, are much rarer (four tokens), as illustrated in the comparative Table 2:

Table 2: Tokens of epistemic extraposed and existential constructions in CORDE and *Corpus de español* appearing with or without ELLO

	With ELLO	Without ELLO
Epistemic extraposition	43	4,603
Existential	4	106,976

Existential *haber* constructions involving ELLO thus pattern in frequency and chronology with the attestation of non-epistemic extraposition ELLO structures, of which there are 13 tokens in total with no predicate repeated more than once (e.g. there are two tokens of *ello es preciso que* ‘it is necessary that’ but only one of *ello es indudable que* ‘it’s undoubted that’).

That is, from the rate of incidence of “regular” (i.e. non-ELLO) constructions in the data set, one would predict a high rate of existential structures with ELLO versus a lower rate of extraposition ELLO constructions, contrary to fact. This imbalanced attestation suggests that epistemic constructions were the most common environment in which to find ELLO, which leads us to surmise that this environment would have been the context of the innovation’s origin.

Thirdly, although there are three epistemic expressions with virtually identical meaning in the historical corpora —the prevalent *es verdad que* (‘[it] is true that’), and the rarer constructions *ello es verdad* (‘it [ello] is true’) and *ello es verdad que* (‘it [ello] is true that’)—, the proposed innovation *ello es verdad que* is the last of these constructions to develop. The impersonal biclausal construction *es verdad que* ‘it is true that’ predates the other two structures by an ample margin, being attested from the earliest texts (1200s onwards) in our data set, as in:

- (21) mas es verdad que no niega lo que oyo.
 but be.3sg truth that not deny.3SG that which hear.3SG
 ‘But it’s true that he does not deny what he heard.’ (CdE; 13th c., *Siete Partidas*)

However, the impersonal monoclausal expression *ello es verdad*, in which the pronoun *ello* is referential, in fact is an innovation itself, originating in copular constructions such as (22-23) from the 14th century onwards (on which, cf. Corr 2015b):

- (22) Segund la ujsyon del archo que sse faze en la nuue enel dia que llueue tal era la ujsyon del splandor al derredor. & **ello es la ujsion & color** dela ssemejança de la gloria de adonay
 ‘As the appearance of the rainbow in the clouds on a rainy day, so was the appearance of the surrounding radiance. **Such is the appearance** of the likeness of the glory of the Lord.’ (CdE; 14th c., anon., *Biblia romanceada*)
- (23) el libro dize que ihesu xpisto es fijo de dios [...]. E el dicho sabio moro Respondio & dixo enesto que los xpistianos dizen **ello es asy verdat.**

‘The book says that Jesus Christ is the son of God [...]. And the aforementioned wise Moor replied and said that in this respect what the Christians say is true.’ (CdE; 15th c., anon., *Cancionero castellano*)

3.2 Diachronic development

This section presents the stages of the change which led to the innovation *ello es verdad que*, namely the reanalysis of the parataxis of two independent clauses, the referential epistemic expression *ello es verdad* (itself an innovation, cf. Section 3.1) and a second independent clause headed by the finite complementizer *que*. The diachronic changes proposed are as follows:

- (24) a. Existence of two separate structures, where *ello* is referential:
 [*ello*_{REF} *es verdad*]; [*es verdad que...*]
 (*[*ello*_{REF/NON-REF} *es verdad que...*])
- b. Existence of two independent structures, [*ello*_{REF} *es verdad*], [*que...*], which can co-occur in a pragmatic but not a subordinating relation;
 [*ello*_{REF} *es verdad*] [*que...*]
- c. *Reanalysis* of the pragmatic relation as a subordinating relation:
 [*ello es verdad*] [*que...*]; [[*ello es verdad*][*que...*]]
- d. Existence of three separate structures, where ELLO in the impersonal construction is no longer referential:
 [*ello*_{REF} *es verdad*]; [*es verdad que...*]; [*ello*_{NON-REF} *es verdad que...*]

Note that the independent (matrix) *que*-clauses appealed to in the present article, such as those in the second stage of ELLO’s diachronic development, are attested throughout the history of Ibero-Romance (see Corr 2016b on so-called “illocutionary complementizers”). In particular, widespread and frequent attestation is found both historically and today of the “conjunctive” QUE complementizer (Alarcos Llorach 1994, Etxepare 2012, Corr 2016b), whose function is to “realize a discourse connection between a salient antecedent —either a non-linguistic discourse situation [25a], or an immediately-preceding speech act [25b]— and the conjunctive QUE clause, a connection that the speaker perceives as requiring contextualisation for the benefit of the addressee(s)” (Corr 2016b: 240), whence its traditional association with “speech act” causality:

- (25) a. Context: the addressee is trying to switch on the light.
 Que ya no funciona.
 CONJ already not work.3SG
 ‘Ø It’s no longer working’ (Sp.)
- b. No llores, que yo te perdono
 not cry.SUBJ.2SG CONJ I you=forgive.1SG
 ‘Don’t cry, [for] I forgive you’ (CORDE; 1530, anon., *Los siete sabios de Roma*)

Where the value of the complementizer is ambiguous, creating the conditions for reanalysis —according to Roberts (1993, 2007), surface ambiguity, while not a cause, is

a necessary condition for language change to take place—, we gloss the polyfunctional particle as *que* and translate its “conjunctive” value as ‘for’.

In the initial stage, the two available structures (biclausal, impersonal *es verdad que* and matrix *ello es verdad*, where *ello* has a referential interpretation) occur in complementary distribution:

- (26) a. Es verdad que tienen algún indio
 be.3.SG truth that have.3PL some Indian
 ‘It is true that they speak some Indian’ (CdE; 15th c., Colón)
- b. [...] lo cual si así fuese creído, como **ello**
 the which if thus be.IMPF.SUBJ.3SG believe.PSTP.PTCP as **it**
 es verdad, en todas las edades hubiera habido
 be.3SG truth in all the ages there.would have.PST.PTCP
 pocos malos
 few bad
 ‘[...] which if it were thus believed —as it is true—, there would have been few bad men through the ages.’ (1554, Pérez de Moya, *Philosophía secreta*)

Ello es verdad que is not yet attested. During the second stage, we begin to observe the occurrence of *ello es verdad* immediately preceding a clause headed by the finite complementizer *que* which appears to be interpretationally dependent on the former:

- (27) Un sabio dijo que la memoria es el pulso del amor [...].
 Y así **es ello verdad, que** a poco amor hay poca memoria
 and thus be.3SG ELLO truth que to little love there.is little memory
 ‘A wise man said that memory is the pulse of love [...]. And thus is it true, that/for of little love there is little memory’ (CdE; 1595, de los Ángeles, *Diálogos*)
- (28) no hace sino reír de ver la gracia y liberalidad con que como [...]
Ello es verdad, que anoche uno destos señores pajes [...] me
 ELLO be.3SG truth que last.night one of.these gentlemen pages me=
 dijo que sorbiese una escudilla de caldo
 tell.PST.3SG that sip.IMPF.SUBJ.3SG a bowl of soup
 ‘He just laughs at the grace and bountifulness with which I eat [...]. It is true, that/for last night one of the pages [...] told me to sip a bowl of soup’ (CORDE, 1614, Fernández de Avellaneda, *Don Quixote*)

However, the syntactic status of *ello*, and the dependency relation of the construction as a whole, is ambiguous. Two possible analyses exist: the first is that *ello* refers to the preceding clause, such that the connection between the sentence-initial epistemic clause and the sentence-final *que* clause is pragmatic, and thus the structure has not been reanalyzed (i.e. *que* functions as a “conjunctive” matrix complementizer, translated in (27-28) as ‘for’). The second analysis is that *ello* has no referential value, which results in its being analyzed as the subject of the epistemic clause, in which it functions as a placeholder. In this latter case, the complex expression has undergone reanalysis, syntactically encoding the non-subordinating relation between the epistemic and its following clause as one of complementation (i.e. *que* functions as a subordinating complementizer).

Finally, by the 18th century, we find that the use of ELLO in such environments is no longer ambiguous but consistently entails a non-referential interpretation, a development which is taken here to indicate that the process of reanalysis is complete (cf. fn 5):

- (29) **Ello** es verdad que parece / muy extravagante y raro /
 ELLO be.3SG true that seem.3SG very extravagant and strange
 que el pobre regale al rico
 that the poor gift.SUBJ.3SG to.the rich
 ‘It is true that it seems very extravagant and strange for a poor man to give a gift to a rich man’ (CORDE; 1794, Fernández de Moratín, *Aguinaldo poético*)

4. Extension

From the 17th century, the use of ELLO extends via analogy (cf. Campbell 2013: ch. 4) to other structural environments, viz. impersonal extraposition (30a), raising (30b) and existential *haber* (30c) constructions. These, unlike ELLO’s original context, have no lexical epistemic interpretation (although at least in the first two types, an epistemic interpretation may be derived pragmatically; cf. Section 4.1):

- (30) a. **Ello** es menester que sepamos para qué tanto somos.
 ELLO be.3SG necessary that know.SUBJ.1PL for what so.much be.1PL
 ‘It is necessary for us to know what our purpose is.’ (CdE; 1608, Salas Barbadillo, *La peregrinación sabia*)
- b. **Ello** parece que nuestros amigos los Franceses no
 ELLO seem.3SG that our friends the French not
 son los más delicados del mundo sobre este artículo
 be.3PL the Most delicate of.the world on this topic
 ‘It seems that our friends the French are not the most sensitive in the world on this subject’ (CdE; 1742, Feijoo, *Cartas eruditas y curiosas*)
- c. **Ello** hay virgen, hay clausura
 ELLO there.is virgin there.is monastery
 ‘There is a virgin, there is a monastery’ (CORDE; 1758, Isla, *Fray Gerundio de Campazas*)

That ELLO has a pragmatic-epistemic interpretation in such examples is reinforced by the observation that ELLO only occurs once in (30c), despite the repetition of *hay* (‘there.is’); i.e. were it functioning as a compulsory placeholder of the type observed in non null-subject languages, then ELLO would be expected to occur twice in (30c).

The common property across the above contexts, and the factor which is proposed here to facilitate the analogical spread of ELLO, is their impersonal status. That is, (30a-c) are classic examples of contexts in which a non null-subject language would obligatorily require a phonologically-realized expletive subject, and a consistently null-subject language would disallow one. However, if this is the case, then we must justify why ELLO extends by analogy to some impersonal expressions, but not others, such as impersonal weather predicates, or unaccusative constructions.

One possibility is that speakers perceived a parallel between (the matrix clause of) impersonal ELLO constructions and referential expressions such as *ello/esto/eso* {es

verdad/parece/hay) ('it.REF/this/that {is true/appears so/there is}'), an analogy which cannot extend to weather predicates (**ello/esto/eso llueve* 'it.REF/this/that rains') or unaccusative structures (**ello/esto/eso nació un niño* 'it.REF/this/that was.born a boy'). Such a parallel could have reinforced the initial reanalysis of referential *ello* as non-referential ELLO, insofar as speakers could have inferred a link between, on the one hand, the pragmatic juxtaposition [*ello_{REF} es verdad*][*que...*] and the impersonal subordinate sentence [*es verdad*][*que...*]; and, on the other, the referential monoclausal structure [*ello_{REF} es verdad*]. This analogy would straightforwardly extend to the existential expression (*ello*) *hay* '[ELLO] there.is'; that is, perceiving a link between referential expressions such as *esto/eso hay* ('this/that there is') and *esto/eso es verdad* ('that/that is true') enables the extension of the reanalyzed non-referential *ello es verdad que* ('it.NON-REF is true that') construction to existential constructions, creating non-referential *ello hay* ('ELLO there.is'). Once grammaticalized across a number of impersonal structures (i.e. the extraposed, raising and existential constructions as exemplified in (30a-c)), the perceived syntactic environment for ELLO expressions would extend to impersonal constructions more generally (constituting "conventionalization", in Heine's (2002) terms), eventually producing the expletive characteristics of today's ELLO discussed in Section 3.1.

4.1 The subjectification of epistemicity

This section proposes that ELLO itself comes to obtain epistemic value through pragmatic inference, a process via which meaning is generated inferentially "from the use of a linguistic form in a particular context" (Nicolle 2011: 401; cf. also Traugott & König 1991). Thus, although the use of the impersonal construction *es verdad que* ('it is true that') implies the *speaker's* commitment to the truth of the sentence's proposition, the speaker point-of-view is derived inferentially, not grammatically. In other words, despite being an impersonal construction (with lexical but not grammatical epistemic meaning), the use of *es verdad que* necessarily conveys the speaker's perspective, since by affirming that a proposition is true one ineludibly expresses one's belief in its veracity. A similar implied evaluation is also present in the use of extraposed and raising constructions such as (*ello*) *es importante/parece que* ('it is important that/it seems that'), paraphrasable as "it is my [sc. the speaker's] view that it is true/important that").

It is thus the use of the pronoun *ello* in an epistemic expression (*es verdad que*) with an inferentially-derived speaker-oriented interpretation ('[I believe that] it is true that') which facilitates a process of "conceptual metonymic shift" wherein "a covert shift [involving] the semanticization of an earlier pragmatic implicature or invited inference" (Traugott 2010: 54; cf. also Traugott & Dasher 2002: 27–34; Hopper & Traugott 2003: 87–92) may occur.

Note that the pragmatic inference (i.e. "[it is my [sc. the speaker's] view that] it is true that") is derived from the ELLO expression as a whole, rather than one of its constituent parts. Since it is the interlocutor's subjective attitude which is grammaticalized in ELLO, the process can also be described as one of *subjectification*, understood in broadly Traugottian terms as 'involv[ing] the reanalysis of [pragmatically-inferred] meanings as [grammatically-encoded] meanings in the context

of speaker-hearer negotiation of meaning' (Traugott 2010: 29). For us, such change is intimately associated with grammaticalization and, in particular, pragmaticization.

The historical use of ELLO with pragmatic-epistemic value is illustrated in (31):

- (31) ¡Quántos gramáticos se encuentran que no llegan ni con cien leguas a la dignidad de éstos [...]!
Ello es cierto y muy cierto que todos debieran ser doctos
 ELLO be.3SG true and very true that all ought.3PL be.INF taught
 lo más que pudiesen [...]
 lo more that can.3SG
 'How many grammarians are to be encountered who do not come within even ten leagues of the dignity of these men[...]! **It is very clearly the case that** they should all be as learned as possible [...]' (CORDE Forner, 1776, *Los gramáticos*)

The analysis of ELLO as a marker of epistemicity allows us to explain certain usages of ELLO in modern Ibero-Romance which, without an epistemic origin to account for them, would otherwise be unexpected or inexplicable, as in (32):

- (32) Mas esse já conhecem, **ele** já?
 but that already know.3PL ELLO already
 'But you already know that, don't you?' (E.Pg., Outeiro, Bragança; Carrilho 2005:134)

In (32), the use of ELLO in the positive confirmation-seeking question tag (*ele*) *já*, where *já* functions as a truth-related particle, can be linked to its epistemic origin insofar as it requests the addressee to corroborate (via their knowledge or belief system) the proposition established in the speaker's question. A similar function of ELLO is observed in its (rare) usage as an affirmative/negative particle (Sp. *ello* (*no*); Pg. *ele sim/não* 'yes/(no)') in response to polar questions (e.g. Sp. *¿Vienes a la fiesta esta noche? – Ello*, 'Are you coming to the party tonight?' – 'Yes'.)

Moreover, question tags can involve truth-related expressions in other languages (cf. English *right/correct?*, German *nicht wahr?*, Italian (*nev*)*vero/giusto?*, Spanish *¿verdad?*, and other Ibero-Romance variants). Although such uses extend ELLO's meaning into wider pragmatic domains beyond its original proposed epistemic value, they nonetheless represent a logical expansion of ELLO's function as a marker of epistemicity.

Further, the "interrogative flip" (Speas & Tenny 2003) involved in ELLO's expansion into addressee-oriented structures, which is understood here as a more advanced stage of grammaticalization/pragmaticization, follows Traugott's predicted (unidirectional) cline of subjectification —encoding speaker-centred perspective—preceding *intersubjectification*, the encoding of addressee-centred perspective (cf. Traugott & Dasher 2002: 225, Traugott 2010: 35-36). Whilst ELLO's development beyond its epistemic value is not the focus of this article,⁶ note that the

⁶ Nor do we touch on the development of ELLO outside Europe, due to space constraints. Though Lipski (2002: 3) suggests the phenomenon is "a spontaneous event" in Dominican Spanish, see Corr (2015b) for evidence of ELLO in the textual record from the 1500s to the present day in various Latin-American varieties, a finding which supports the hypothesis that ELLO was already present in the speech of settlers travelling to the New World (Brazilian Portuguese examples are also attested in 19th century texts). The

grammaticalization of ELLO's intersubjective (or addressee-oriented) value could arise from its addressee-oriented use in interrogatives, where it appears in the expected clause-initial position (cf. also (3a)):

- (33) ¿Y conque nos vamos a comer esos víveres o **ello** hay
 and so us= go.1PL to eat.INF those foods or ELLO there.is
 carne para comer?
 meat to eat.INF
 'So are we going to eat this food or is there meat to eat?' (Santo Domingo Este)

4.2 ELLO further afield?

Apparent examples of ELLO exist elsewhere in Romance, such as Corsican *ellu* (34a) and Sardinian *ello* (34b) (see also Sornicola 1996, Ledgeway 2003, 2013 for discussion of Italo-Romance expletive(-like) elements):

- (34) a. Quando **ellu** piuvía
 when ellu rain.PST.3SG
 'When it rained' (Corsican; Remberger & Hinzelin 2009: 9)
 b. **Ello** no sunt parte de s'Unione Europea?
 ello not be.3SG part of the=Union European
 'And are they not a part of the European Union?' (Sardinian; *ibid.*: 9)

Prima facie, the appearance of such elements with similar form and function to ELLO elsewhere in the Romània could lead us to disregard the peninsular epistemic origin hypothesis. However, unlike Ibero-Romance ELLO, Corsican *ellu* is an obligatory subject or expletive pronoun in subordinate contexts (Hinzelin & Remberger 2009), different from the discourse-related function of ELLO. In fact, the peninsular origin hypothesis of ELLO is strengthened by the fact that Sardinian *ello* appears to be a loanword from Balearic Catalan: Remberger & Hinzelin (2009: 10) argue that 'it seems to be very plausible, especially for its functional meaning, but also for its phonological form, to explain Sardinian *ello* as a loanword from the Balearic Catalan [...] *ell'*'. They, alongside Meyer-Lübke (1935), Porru (2002), observe that Sardinian *ello* could not have been inherited directly from Latin, since it has not undergone the expected phonological changes associated with an intervocalic geminate lateral [ll] (rather, one would expect to find the characteristically-Sardinian voiced retroflex geminate plosive [ɖɖ], as in *caddhu* 'horse' < CABALLUS, *biddha* 'city' < VILLA, *cherveddhu* 'brain' < CEREBELLUM). That Sardinian's third person referential pronoun is *iddhu*, and therefore contrasts phonologically —unlike its Ibero-Romance referential counterparts— with ELLO, reinforces Hinzelin & Remberger's (2009) contact account for the Sardinian case.

Further, since Sardinian *ello* is closest in its formal characterization and left-peripheral distribution to the ELLO of today's Balearic Catalan rather than that of non-standard European Portuguese or Dominican Spanish (cf. Section 3.2.5), one can deduce that ELLO would already have been at an advanced state of grammaticalization in the Catalan variety or varieties from which it was borrowed into Sardinian. Noting that

early transfer of ELLO to the New World is also supported by the formal properties of Dominican Spanish ELLO, insofar as it exhibits the syntax associated with an early stage of grammaticalization (cf. Section 2).

Algherese Catalan (in which ELLO is —to the best of my knowledge— unattested historically and today, cf. Pais 1970)⁷ had little impact on neighbouring Sardinian varieties on the island, Remberger & Hinzelin (2009) estimate a period from the 14th until at least the 18th century during which other Catalan varieties would have used *ell*. If the hypothesis put forward in this article can also be extended to Peninsular Ibero-Romance varieties other than Spanish, this timescale would be reduced at the lower end by at least two centuries.

5. Conclusion

This article has traced the development of ELLO from its original function as a referential pronoun in epistemic contexts into its use as a non-referential pragmatic marker of epistemic value, and beyond into further illocutionary functions in present-day Ibero-Romance. Despite ELLO's heterogeneous outcomes in different varieties, the processes in the case of language change under investigation here follow cross-linguistically attested pathways of grammaticalization. The original referential pronoun itself undergoes various processes — a change in formal category, a redistribution of meaning (viz. semantic bleaching and pragmatic enrichment)— commonly associated with grammaticalization. Further, its development from a T-element to a C-element reflects the unidirectionality both of formal models of grammaticalization — namely, upwards progression along the functional structure—, and the Traugottian cline of (inter)subjectification. ELLO's origin as a referential pronoun allows us to explain why we see an apparent expletive in an otherwise fully null-subject language. Likewise, the pragmatic nature of the Iberian expletive would suggest that we should be open to revising our typology of expletives to reflect a more nuanced data set and range of microparametric possibilities.

References

Corpora

- CdE: Davies, Mark. (2002-): *Corpus del Español: 100 million words, 1200s-1900s*. <http://www.corpusdelespanol.org> [consulted 30 September 2016]
- CdP: Davies, Mark & Michael Ferreira. (2006-) *Corpus do Português: 45 million words, 1300s-1900s*. <http://www.corpusdoportugues.org> [consulted 30 September 2016]
- CORDE: Real Academia Española (online): *Corpus diacrónico del español (CORDE)*. <http://www.rae.es>. [consulted 30 September 2016]
- CORDIAL-SIN: Martins, Ana Maria. (coord.) [2000-] (2010): *CORDIAL-SIN: Corpus Dialectal para o Estudo da Sintaxe*. Lisboa, Centro de Linguística da Universidade de Lisboa. <http://www.clul.ul.pt/en/resources/411-cordial-corpus>. [consulted 30 September 2016]

⁷ An anonymous reviewer correctly points out that Pais (1970) is a grammar, and that, since grammars tend not to mirror the use of the spoken language in all its aspects, it cannot consequently be guaranteed that ELLO is not attested Algherese Catalan. References to the ELLO phenomenon do, however, appear in other grammars of Ibero-Romance, so the assumption that ELLO's absence in Pais (1970) indicates its absence more generally in this variety is not without precedent.

Bibliography

- Aelbrecht, Lobke, Liliane Haegeman & Rachel Nye, eds. 2012. *Main clause phenomena: New horizons*. John Benjamins.
- Aijmer, Karin. 1997. *I think* – an English modal particle. *Modality in Germanic languages*, ed. by Toril Swan & Olaf Jansen Westvik, 1–47. Berlin: Mouton de Gruyter.
- Alarcos Llorach, Emilio. 1994. *Gramática de la lengua española*. Madrid: Espasa Calpe.
- Bartra-Kaufmann, Ana. 2011. Recycled neuter expletive pronouns and the nature of the left periphery: *ell* and relatives revisited. *Catalan Journal of Linguistics* 10.185-219.
- Benincà, Paola. 1988. L'ordine degli elementi della frase: Costruzioni con ordine marcato degli elementi. *Grande grammatica italiana di consultazione. Volume 1: La frase. I sintagmi nominale e preposizionale* ed. by Lorenzo Renzi, 129–94. Bologna: Il Mulino.
- Benincà, Paola & Cecilia Poletto. 2004. Topic, focus and V2. In Luigi Rizzi, ed., 52-75.
- Biberauer, Theresa & Jenneke van der Wal. 2012. Expletives beyond English. Handout from *Syntax of the World's Languages V*, Dubrovnik, October 2012.
- Biberauer, Theresa & Federica Cognola. 2014. The expletive-impersonal connection: Mòcheno insights into the typology of null-subject language. Handout from *Forum for Germanic Language Studies* 11, University of Cambridge, 9-11 January.
- Biberauer, Theresa, Andres Holmberg, Ian Roberts & Michelle Sheehan. 2010. *Parametric variation: Null subjects in minimalist theory*. Cambridge University Press.
- Camacho, José. 2013a. *Null subjects*. Cambridge University Press.
- Camacho, José. 2013b. On left-peripheral expletives in Central Colombian Spanish. Ms. Rutgers University.
- Camacho, José. 2016. The null subject parameter revisited. *The Morphosyntax of Portuguese and Spanish in Latin America* ed. by Mary A. Kato & Francisco Ordóñez, 27-48. Oxford University Press.
- Campbell, Lyle. 2013. *Historical linguistics*. Edinburgh University Press
- Carrilho, Ernestina. 2005. *Expletive ele in European Portuguese dialects*. PhD dissertation. Universidade de Lisboa.
- Carrilho, Ernestina. 2008. Beyond doubling: overt expletives in European Portuguese dialects. *Microvariation in Syntactic Doubling* ed. by Sief Barbiers, Olaf Koeneman, Marika Lekakou & Margreet Ham, 301-323. Brill Online.
- Chomsky, Noam. 1982. *Some concepts and consequences of the theory of government and binding*. MIT press.
- Corr, Alice. 2015a. Overt expletives in Ibero-Romance: a diatopic and diachronic perspective. *Revue roumaine de linguistique* 60 (2-3 Special issue on syntactic variation), 205-221.
- Corr, Alice. 2015b. De la sintaxis a la pragmática: el desarrollo del expletivo ibérico desde el español medieval hasta la actualidad. *Armonía y contrastes. Estudios de variación dialectal, histórica y sociolingüística del español* ed. by José María Santos Rovira, 249-263. Editorial Axac.
- Corr, Alice. 2016a. El cambio lingüístico y la interfaz sintaxis-pragmática: evidencias de las lenguas ibero-románicas. *Cuadernos de Lingüística de El Colegio de México* 3:1, 73-118

- Corr, Alice. 2016b. *Ibero-Romance and the syntax of the utterance*. PhD dissertation. University of Cambridge.
- de Haan, Ferdinand. 2001. The relation between modality and evidentiality. *Modalität und Modalverben im Deutschen*, ed. Reimar Müller & Marga Reis, 201-206. Hamburg: Helmut Buske Verlag.
- Diewald, Gabriele. 2011. Grammaticalization and pragmaticalization. In Heiko Narrog & Heine Bernd, ed., 450-461.
- Etxepare, Ricardo. 2012. Subordination and coordination. *Handbook of Hispanic Linguistics*, ed. by José Ignacio Hualde, Antxon Olarrea & Erin O'Rourke, 503-533. Blackwell.
- Faller, Martina T. 2011. A possible worlds semantics for Cuzco Quechua evidentials. *Semantics and Linguistic Theory* 20, 660-683.
- Fernández-Ordóñez, Inés. 2012. Dialect areas and linguistic change: Pronominal paradigms in Ibero-Romance dialects from a cross-linguistic and social typology perspective. *The Dialect Laboratory: Dialects as a Testing Ground for Theories of Language Change*, ed. by Gunther De Vogelaer & Guido Seiler, 73-106. John Benjamins. DOI: 10.1075/slcs.128.04fer.
- Giorgi, Alessandra. 2010. *About the Speaker: Towards a syntax of indexicality*. Oxford University Press.
- Gupton, Timothy & Sarah Lowman. 2013. An F Projection in Cibeño Dominican Spanish. *Selected Proceedings of the 16th Hispanic Linguistics Symposium* ed. by Jennifer Cabrelli Amaro, Gillian Lord, Ana de Prada Pérez & Jessi Elana Aaron, 338-348. Cascadilla Proceedings Project, Somerville, MA.
- Haegeman, Liliane, 2003. Conditional clauses: External and internal syntax. *Mind & Language* 18:4.317-339.
- Han, Chung-Hye, 2002. Interpreting interrogatives as rhetorical questions. *Lingua* 112:3.201-229.
- Heine, Bernd. 2002. On the role of context in grammaticalization. *Typological Studies in Language* 49.83-102.
- Henríquez Ureña, Pedro. 1939. Ello. *Revista de filología hispánica* 1:3.209-29.
- Hinzelin, Marc-Olivier. 2009. Neuter pronouns in Ibero-Romance: Discourse reference, expletives and beyond. *Proceedings of the Workshop «Null-subjects, expletives, and locatives in Romance»* Arbeitspapier 123. Fachbereich Sprachwissenschaft, ed. by Georg Kaiser & Eva-Maria Remberger, 1-25. Universität Konstanz.
- Hinzelin, Marc-Olivier & Georg Kaiser. 2007. El pronombre *ello* en el léxico del español dominicano. *Language Contact and Language Change in the Caribbean and Beyond / Lenguas en contacto y cambio lingüístico en el Caribe y más allá* ed. by Wiltrud Mihatsch & Monika Sokol, 171-188. Frankfurt am Main: Peter Lang.
- Hinzelin, Marc-Olivier & Eva-Maria Remberger. 2009. Expletive vs. discourse marker: Sardinian *ello* and Balearic Catalan *ell*. Handout from 37th Romance Linguistics Seminar, Trinity Hall, Cambridge, 5-6 January.
- Holmberg, Anders, Aarti Nayudu, & Michelle Sheehan. 2009. Three partial null-subject languages: a comparison of Brazilian Portuguese, Finnish and Marathi. *Studia Linguistica* 63:1.59-97.
- Hooper, Joan, & Sandra Thompson. 1973. On the applicability of root transformations. *Linguistic inquiry*, 465-497.

- Hopper, Paul & Elizabeth Traugott. 2003. *Grammaticalization*. Cambridge University Press.
- Huang, Yan. 2000. *Anaphora: A cross-linguistic approach*. Oxford University Press.
- Ledgeway, Adam. 2003. Linguistic Theory and the Mysteries of Italian Dialects. *Multilingualism in Italy: Past and Present*, ed. by Anna Laura Lepschy & Arturo Tosi, 108-140. Oxford: Legenda.
- Ledgeway, Adam. 2012. *From Latin to Romance: Morphosyntactic typology and change*. Oxford University Press.
- Ledgeway, Adam. 2013. Testing linguistic theory and variation to their limits: The case of Romance. *Corpus* 12.271-327.
- Lightfoot, David. 1997. Catastrophic change and learning theory. *Lingua* 100:1.171-192.
- Lipski, John. 2002. Subject pronoun usage among Spanish dialects. Ms. Pennsylvania State University.
- Meyer-Lübke, Wilhelm. 1935. *Romanisches etymologisches Wörterbuch*. 3rd edn. Heidelberg: Winter.
- Michaelis, Laura & Knud Lambrecht. 1996. Toward a construction-based theory of language function: The case of nominal extraposition. *Language* 72:2, 215-247.
- Narrog, Heiko & Bernd Heine. 2011. *The Oxford handbook of grammaticalization*. Oxford University Press.
- Nicolle, Steve. 2011. Pragmatic aspects of grammaticalization. In Heiko Narrog & Bernd Heine, ed., 142-152.
- Palmer, Frank Robert. 2001. *Mood and modality*. Cambridge University Press.
- Pais, Joan. 1970. *Gràmatica algueresa*. Barcelona: Editorial Barcino.
- Porru, Vincenzo. 2002. *Nou dizionari universali sardu-italianu*. Nuoro: Ilisso.
- Remberger, Eva-Maria, & Marc-Olivier Hinzelin. 2009. Expletive vs. discourse marker: Sardinian *ello* and Balearic Catalan *ell*. *38th Romance Linguistics Seminar*, Trinity Hall, Cambridge.
- Rizzi, Luigi. 1997. The fine structure of the left periphery. *Elements of grammar* ed. by Liliane Haegeman, 281-337. Springer Netherlands.
- Rizzi, Luigi. 1982. *Issues in Italian syntax*. Dordrecht: Foris.
- Rizzi, Luigi. 1986. Null objects in Italian and the theory of *pro*. *Linguistic inquiry*, 501-557.
- Rizzi, Luigi, ed. 2004. *The Structure of CP and IP. The Cartography of Syntactic Structures* (vol. 2). Oxford University Press.
- Rizzi, Luigi. 2004. On the cartography of syntactic structures. In Luigi Rizzi, ed., 3-15.
- Roberts, Ian G. 1993. *Verbs and diachronic syntax: A comparative history of English and French*. Dordrecht: Kluwer.
- Roberts, Ian G. 2007. *Diachronic syntax*. Oxford University Press.
- Roberts, Ian G. 2011. Grammaticalisation, the Clausal Hierarchy and Semantic Bleaching. Ms. University of Cambridge.
- Roberts, Ian & Anna Roussou. 2003. *Syntactic change*. Cambridge University Press.
- Sigurðsson, Halldór Á. 2010. On EPP effects. *Studia Linguistica* 64:2.159-189.
- Silva-Villar, Luis. 1998. Subject positions and the roles of CP. *Romance Linguistics. Theoretical Perspectives. Selected Papers from the 27th Linguistic Symposium on Romance Languages (LSRL XXVII), Irvine 20-22 February 1997* ed. by Armin

- Schwegler, Bernard Tranel, & Myriam Uribe-Etxebarria, 247-270. Amsterdam: John Benjamins.
- Speas, Peggy, & Carol Tenny. 2003. Configurational properties of point of view roles. *Asymmetry in grammar* 1.315-345.
- Sornicola, Rosanna. 1996. Alcune strutture con pronome espletivo nei dialetti italiani meridionali. *Italiano e dialetto nel tempo: saggi di grammatica per Giulio C. Lepschy* ed. by Giulio C. Lepschy & Paola Benincà, 323-340. Bulzoni.
- Svenonius, Peter. 2002. *Subjects, Expletives, and the EPP*. Oxford University Press.
- Traugott, Elizabeth C. 2010. (Inter) subjectivity and (inter) subjectification: a reassessment. *Subjectification, intersubjectification and grammaticalization* ed. by Kristin Davidse, Lieven Vandelanotte & Hubert Cuyckens, 29-71. De Gruyter Mouton.
- Traugott, Elizabeth C. & Richard B. Dasher. 2002. *Regularities in semantic change*. Cambridge University Press.
- Traugott, Elizabeth C. & Ekkehard König. 1991. The semantics-pragmatics of grammaticalization revisited. *Approaches to grammaticalization* ed. by Elizabeth C. Traugott & Bernd Heine, 189-218. John Benjamins.
- Uriagereka, Juan. 1995. An F position in western romance. *Discourse configurational languages* ed. by Katalin É Kiss, 79-123. Oxford University Press.
- Uriagereka, Juan. 2005. Iberian Pleonasm. ms. University of Maryland.
- van Gelderen, Elly. 2011. Grammaticalization and Generative Grammar: a difficult liaison. In Heiko Narrog & Heine Bernd, ed., 43-55.
- Veny, Joan. 1989. *Els parlars catalans. Síntesi de dialectologia*. Mallorca: Moll.