

Expletives in Icelandic: a corpus study¹

NB: This is a post-print version reflecting changes made in the peer review and editing process, but not the publisher's PDF. This article appeared in Bridget Drinka (Ed.), *Historical Linguistics 2017: Selected Papers from the 23rd International Conference on Historical Linguistics*, 364–384. Amsterdam: John Benjamins. When citing, please use the page numbers given there. For the published version, see: <https://benjamins.com/catalog/cilt.350.17boo> .

Abstract

Various claims have been made concerning expletives in historical Icelandic (Eythórsson & Sigurðardóttir, 2016; Hróarsdóttir, 1998; Rögnvaldsson, 2002). However, previous studies are limited to certain centuries and/or text types. This paper assesses the status of expletives throughout all stages of attested Icelandic (1150-2008) using the Icelandic Parsed Historical Corpus (IcePaHC). The IcePaHC data indicates that the expletive is already attested in impersonal constructions in Old Icelandic, contra previous accounts. I also claim that *þar* 'there' was historically available as an expletive in

¹ I thank audiences at *ICHL 23* (July 2017), the *22nd International LFG Conference* (July 2017) and the *Manchester Forum in Linguistics* (April 2017) for their useful feedback and comments. I also thank three anonymous reviewers for their valuable insights. The research for this paper was funded by an Arts & Humanities Research Council UK (AHRC) doctoral scholarship.

presentational constructions and argue against a recent claim that the emergence of expletive *það* is a contact phenomenon. Strikingly, postfinite expletives are attested historically, which challenges the standard cross-Germanic account for the diachrony of expletives.

Key words: Syntactic change, historical corpora, Germanic, Icelandic, expletives, cataphora.

1. *Introduction*

The rise of expletives is a diachronic development common to the Germanic languages.² All present-day Germanic languages exhibit expletives (see Vikner, 1995 for an outline). However, in earlier stages of Germanic the expletive is typically absent in contexts where it occurs today. The emergence of expletives has been well researched for many of these languages: e.g. Breivik (1983) and Ingham (2001) on English; Lenerz (1985) and Axel (2007) on German; Falk (1993) on Swedish; Faarlund (1990) and Kinn (2016) on Norwegian. By comparison, the Icelandic diachrony has received less attention, and previous studies have been limited to certain centuries, text

² “Expletive” in this sense is interchangeable with “pleonastic” or “dummy”, i.e. an element which has no semantic content, but is there for syntactic reasons.

genres and/or construction types (Hróarsdóttir, 1998; Rögnvaldsson, 2002; Eythórsson & Sigurðardóttir, 2016). Present-day Icelandic has one main form which functions as an expletive: *það*, formally identical to the 3SG.NT referential pronoun.³ The development whereby expletive *það* emerged in the history of Icelandic is the focus of this paper. I present a quantitative corpus-based study, using the Icelandic Parsed Historical Corpus (IcePaHC, Wallenberg et al., 2011). This allows me to assess the status of *það* in a range of contexts throughout all attested stages of Icelandic (1150-2008), thereby capturing a diachronic breadth and detail not achieved in the literature to date.

The structure of the paper is as follows: in section 2, I outline the background concerning expletives in Icelandic and historical Germanic; section 3 presents the methodology and findings of the corpus study; section 4 discusses these findings; section 5 concludes the paper.

2. *Background*

2.1 *Three types of það*

³ There is also a second form which can function as an expletive in present-day Icelandic: *hann*, formally identical to the 3SG.MASC pronoun, which typically appears in constructions with weather predicates (Eythórsson & Sigurðardóttir, 2016).

For the purposes of this study, I differentiate between three types of *það* in Icelandic:

1. Anaphoric *það* (ANAPH), e.g. (1).
2. Cataphoric *það* (CATPH), e.g. (2).
3. Expletive *það* (EXPL), e.g. (3).

(1) [Þetta andlit]_i, ég hafði einhvern
 DEM.ACC face.ACC I.NOM have.PST one.ACC
 tíma séð **það**: áður...
 time.ACC see.PST.PTCP ANAPH before
 ‘This face, I had seen it one time before...’ (2008, Ofsi.258)⁴

(2) a. **Það**: var samt ljóst [að við þyrftum
 CATPH be.PST still clear COMP we.NOM need.PST.SBJV
 meiri mannskap]_i.
 more crew

‘It was still clear that we would need more crew.’
 (2008, Ofsi.733)

b. **Það**: má segja [að mér hafi
 CATPH may say.INF COMPI.DAT have.PRS.SBJV
 lukkast að fleygja sál minni...].
 succeed.PST.PTCP to throw-away.INF soul my
 ‘One may say that I have succeeded in throwing away my
 soul...’ (1985, Margsgaga.945)

(3) a. **Það** var að vora.
 EXPL be.PST to become-spring.INF
 ‘It was becoming spring.’ (2008, Mamma.1066)

b. **Það** var töluverður snjór yfir öllu.
 EXPL be.PST considerable.NOM snow.NOM over all
 ‘There was a considerable amount of snow over
 everything.’ (2008, Ofsi, 772)

⁴ Unless otherwise stated, all examples are from IcePaHC and referenced: (Year, Text.CorporusID). Further information about the texts and editions used can be found in the IcePaHC documentation: <https://github.com/antonkarl/icecorpus> [last accessed 26.09.2018].

This paper will discuss cataphoric *það* and expletive *það*. Cataphoric *það* occurs in constructions with a clausal argument (“extraposition”, e.g. Thráinsson, 1979), and cataphorically refers to this argument.⁵ By contrast, expletive *það* is neither anaphorically nor cataphorically referential.

For expletive *það*, I distinguish between impersonal constructions which lack an overt subject, e.g. (3-a), and presentational constructions which have a postfinite discourse-new subject, e.g. (3-b).⁶ The impersonal category includes certain predicates which some claim license “quasi-arguments” (non-referential arguments), distinct from “true expletives” (non-referential non-arguments), see Falk (1993, p. 67); Chomsky (1981, pp. 324–327). For authors who recognise quasi-arguments, *það* with the weather predicate in (3-a) would qualify as such (see Vikner 1995, pp. 224–227). However, as I claim in section 2.2, expletive *það* does not qualify as a subject but is a structural placeholder for the prefinite position. As a non-subject, *það* cannot qualify as a quasi-argument.⁷ As such, I classify constructions with weather predicates

⁵ The classic account by Thráinsson (1979, p. 181) claims that, in contexts like (2), prefinite *það* can be referential or expletive, while postfinite *það* is always referential. For this paper, I assume that any *það* which co-occurs with a clausal argument later in the sentence is cataphoric.

⁶ For evidence that the postfinite noun phrase in Icelandic presentational constructions qualifies as a syntactic subject, see Zaenen et al. (2017).

⁷ For a similar view of expletive *það*, see Eythórsson & Sigurðardóttir (2016).

together with other impersonal constructions where *það* is a positionally restricted expletive, such as impersonal passives, e.g. (4).

- (4) **Það** var dansað alla nóttina.
EXPL be.PST dance.PASS.PTCP all night
'It was danced all night.' (Thráinsson, 2007, p. 266)

The presentational category captures both “existentials” with copular *vera* ‘be’, e.g. (3-b), as well as presentational constructions with other verbs, e.g. unaccusative *koma* ‘come’ in (5).

- (5) **Það** kæmi bara tóm þvæla.
EXPL come.PST.SBJV only empty.NOM nonsense.NOM
'There would come only empty nonsense.'
(2008, Mamma.754)

For an overview of subtypes of impersonal and presentational constructions in modern Icelandic, see Thráinsson (2007, pp. 310–311).⁸ A particularly fine-grained classification scheme is presented in Booth (forthcoming), which takes into account both present-day and historical Icelandic data.

2.2 Position of expletive *það*

The positional distribution of expletive *það* has attracted much attention (e.g. Sells, 2005; Sigurðsson, 2007). It is well known for present-day Icelandic that expletive *það* is restricted to the clause-initial prefinite position.

⁸ *Það* also occurs in cleft constructions, but I do not treat clefts in this paper.

Compare the impersonal passive in (6-a) which has a clause-initial *það*, with (6-b), where the clause-initial position is occupied by an adjunct and *það* is ungrammatical.

- (6) a. **Það** var dansað í gær.
 EXPL be.PST dance.PASS.PTCP yesterday
 ‘It was danced yesterday.’
- b. Í gær var (*það) dansað
 yesterday be.PST (*EXPL) danced.PASS.PTCP
 ‘Yesterday it was danced.’

On the basis of the data in (6), I assume that *það* does not qualify as subject, since Icelandic subjects typically invert in fronting contexts like (6-b) (see also Thráinsson, 1979, pp. 480–481; Platzack, 1983; Maling, 1988). Compare (6) with the Swedish data in (7), in which expletive *det* behaves like a subject.

- (7) a. **Det** dansades i går.
 EXPL dance.PST.PASS yesterday
 ‘It was danced yesterday.’
- b. I går dansades **det**.
 yesterday dance.PST.PASS EXPL
 ‘Yesterday it was danced.’

In this paper, I refer to the positionally restricted Icelandic type as a “prefinite expletive”, and the Swedish type as a “subject expletive”.⁹

⁹ Faarlund (1990, pp. 63–65) refers to the positionally restricted type as an “expletive topic”. I prefer “prefinite expletive”, since “expletive topic” implies information structural properties which are not relevant in describing a filler for a certain clause structure position; being non-referential, expletives cannot be topics.

2.3 Historical context

In earlier Icelandic, expletive *það* is typically absent in the impersonal and presentational constructions outlined in section 2.1. The absence of a clause-initial *það* results in verb-initial (V1) structures, e.g. (8); cf. (3) above.¹⁰

- (8) a. Tekur nú að hausta.
begin.PRS now to become-spring.INF
‘It now starts to become autumn.’ (1310, Grettir.48)¹¹
- b. Var þá gleði mikil í kóns höll.
be.PST then joy.NOM great.NOM in king.GEN hall
‘There was then great joy in the king’s hall.’
(1480, Jarlmann.48)

Rögnvaldsson (2002) observed that the first instances of expletive *það* in presentational and impersonal constructions appear in c.1500. Hróarsdóttir (1998) examines 18th and 19th century informal letters and observes an increase in the frequency of *það* in authors born after 1810. However, these two studies are restricted to certain centuries, and therefore offer a snapshot of the diachrony. Moreover, both Rögnvaldsson and Hróarsdóttir only count the number of sentences which feature expletive *það*, without taking into

¹⁰ An anonymous reviewer suggests that the examples in (8) involve so-called “Narrative Inversion”. However, on my understanding, Narrative Inversion refers to V1 declaratives which have a topical subject in the immediately postfinite position (see e.g. Sigurðsson, 1990). Since the examples in (8) are topicless, neither qualify as Narrative Inversion.

¹¹ All IcePaHC texts are normalised to modern orthography, regardless of their date. I follow this normalisation throughout.

account the total contexts in which the expletive could potentially occur, including where it is absent, e.g. (8). As discussed in section 3.1, IcePaHC allows me to take this into account. Concerning cataphoric *það*, e.g. (2), Faarlund (1990, p. 72) and Rögnvaldsson (2002) state that cataphoric *það* is already frequently in Old Norse/Icelandic (1150-1350), a claim which I test in the corpus study in section 3.

As noted in section 1, the development of expletives is common to Germanic. Based on studies of Germanic languages other than Icelandic, the pathway of change in (9) has been proposed, which I refer to as the “Prefinite First Hypothesis”.¹²

- (9) **Prefinite First Hypothesis**
prefinite expletive > subject expletive

The Prefinite First Hypothesis states that positionally restricted prefinite expletives emerge before subject expletives historically. In this connection, it has been claimed that the initial rise of Germanic expletives is not conditioned by subject-related considerations, but by structural considerations concerning the verb-second (V2) constraint (Richards & Biberauer, 2005). In this paper, I test whether the Prefinite First Hypothesis holds for Icelandic.

¹² See Breivik (1983), Ingham (2001) on English; Lernerz (1985) on German; Falk (1993) on Swedish; Faarlund (1990) on Norwegian; more general overviews are Haiman (1974), Richards & Biberauer (2005).

3. *Corpus study*

3.1 *Methodology*

The basis for this study is data from the Icelandic Parsed Historical Corpus (IcePaHC, Wallenberg et al., 2011). IcePaHC contains approximately 1,000,000 words spanning ten centuries (1150-2008), thereby covering all attested stages of Icelandic.¹³ IcePaHC thus allows one to examine change across the centuries which many pre-existing studies contrasting Old Icelandic with present-day data do not capture.¹⁴ IcePaHC follows the Penn treebank format established for historical English (e.g. Kroch & Taylor, 2000; Santorini, 2010) and is compatible with the CorpusSearch query language (Randall, 2005). All content is lemmatised, part-of-speech tagged and annotated for constituency.

IcePaHC treats both cataphoric *það* and expletive *það* as an expletive (“ES”). Moreover, both overt and “null” expletives – identical constructions

¹³ As an anonymous reviewer points out, certain texts dated to 1150 are extant in manuscripts from the 14th century; their dating is an approximation.

¹⁴ IcePaHC has some limitations: the texts included represent a very small sample of attested historical Icelandic, and certain genres are over-represented and others under-represented in individual periods. Some texts are also translations of source texts in other languages. Nevertheless, the advantages offered by the syntactic annotation outweigh such issues.

where *það* is absent, e.g. (8) – are distinctively tagged.¹⁵ I take the tagging at face value; I do not look for additional sentences which qualify as cataphoric/expletive *það* contexts but are not annotated as such. I restrict the study to matrix declaratives.

I isolated the cataphoric/expletive *það* contexts outlined in section 2.1 via CorpusSearch queries. I then manually examined all examples, checking the construction type assignment and assuring for internal category coherence. I excluded 150 examples which are misannotated as cataphoric/expletive *það* contexts. The data extracted by this process is shown in Table 1. I split the IcePaHC diachrony into five periods, see first column.

Period	Cataphoric contexts	Expletive contexts		All
		Impersonal	Presentational	
1150-1350	366	365	90	821
1351-1550	161	237	88	486
1551-1750	172	151	93	416
1751-1900	158	204	100	462
1901-2008	197	225	135	557
All periods	1054	1182	506	2742

Table 1: Cataphoric/expletive *það* contexts in IcePaHC, 1150-2008

¹⁵ For the IcePaHC policy on expletives, see: http://www.linguist.is/icelandic_treebank/Expletives [last accessed 26.09.2018].

All examples in Table 1 were coded for the presence/absence of *það* via the search queries. Additionally, I tagged manually for further properties relevant to the study aims: predicate type; verb position; position of *það* if present.

In this section, I compare the proportion of instances in which *það* is present in the clause-initial prefinite position resulting in V2, e.g. (3), to instances in which *það* is absent, resulting in V1, e.g. (8). I compare the frequency of *það* across the five periods. I conduct χ^2 -tests to examine whether the value for an individual period deviates significantly from the mean value across the whole corpus ($p < 0.05^*$, $p < 0.01^{**}$, $p < 0.001^{***}$; “ns” marks a value which does not significantly deviate from the mean).

3.2 *Cataphoric það*

The IcePaHC findings confirm previous claims (Faarlund, 1990; Rögnvaldsson, 2002) that prefinite cataphoric *það* is already robustly attested in constructions with a clausal argument in Old Icelandic (1150-1350), at a frequency of 77.0%, see Table 2. Cataphoric *það* is frequent through all periods and there is an increase visible in the data for 1901-2008, where *það* occurs at a frequency of 91.6%. The lowest frequencies are observed for the periods 1351-1550 and 1551-1750, coinciding with a genre bias: religious and biographical texts are disproportionately represented in these two periods, compared to the other periods where narrative texts (i.e. sagas/modern fiction)

dominate. This may account for the lower frequencies, but more research on texts beyond IcePaHC would be required to confirm this.

Period	Prefinite <i>það</i> (V2)	No <i>það</i> (V1)	Total	% <i>það</i>	χ^2
1150-1350	124	37	161	77.0%	ns
1351-1550	49	22	71	69.0%	*
1551-1750	52	24	76	68.4%	*
1751-1900	75	18	93	80.6%	ns
1901-2008	109	10	119	91.6%	****
All periods	409	111	520	78.7%	

Table 2: Frequency of prefinite cataphoric *það* in constructions with a clausal argument in IcePaHC, 1150-2008

I show examples with cataphoric *það* from the earliest period (1150-1350) in (10)-(11).

(10) **Það**_i er ráð húsfreyja [að taka vel
CATPH be.PRS advice housewife to receive.INF well
við gestum]_i.
PTCL guests
'It is good advice, housewife, to receive guests well.'
(1310, Grettir.658)

(11) **Það**_i vildi ég [að þú ræddir ekki um]_i.
CATPH wish.PST I.NOM COMP you.NOM talk.PST NEG about
'I wished that you didn't talk about it.' (1275, Morkin.1280)

A particular context in which cataphoric *það* occurs is in constructions like (12), where the main verb in the matrix clause is in present participle form (see Faarlund, 2004, pp. 133–134 for discussion).

- (12) En **það**_i er vitanda [að þá
 but CATPH be.PRS know.PRS.PTCP COMP then
 göfgum vér réttlega postula guðs alla]_i...
 honour.PRS we rightly apostle god.GEN all
 ‘But one is to know that we then rightly honour all of God’s
 apostles...’ (1150, Homiliubok.304)

The majority of constructions like (12) in my dataset are attested in *Homiliubok*. This text is known to have a Latin background, though the exact nature of this connection remains unclear (for discussion, see Weenen, 1993). Weenen (1993, p. 183) discusses constructions like (12) under “Latin construction”. She gives the Latin in (13-a) as the parallel construction for (13-b).¹⁶

- (13) a. Sed quærendum nobis est...
 but seek.PASS.PTCP we.DAT be.PRS
 ‘But we are to seek...’
 b. En **þess**_i er oss leitanda [að...]_i
 but CATPH.GEN be.PRS we.DAT seek.PASS.PTCP COMP
 ‘But we are to seek that...’¹⁷

Strikingly, although there is no cataphoric pronoun in the Latin original (13-a), cataphoric *það* (*þess*) is present in the Icelandic translation in (13-b).

¹⁶ Thanks to Tarrin Wills (p.c) for pointing this out.

¹⁷ The verb *leita* ‘seek’ takes a genitive argument as its object.

The *Ordbog over det norrøne prosasprog* (Dictionary of Old Norse Prose)

lists similar examples from *Homiliubok* with the Latin parallel, e.g. (154).¹⁸

- (14) a. Sciendum est...
 know.PASS.PTCP be.PRS
 ‘It is to be known...’
- b. En það; er vitanda [að...];
 but CATPH be.prs know.PRS.PTCP COMP
 ‘But it is to be known that...’¹⁹

Data like (13) and (14) indicates that the use of cataphoric *það* in such contexts was a robust feature of Old Icelandic, and was employed even in sentences based on a source construction with no such cataphoric pronoun.

3.3 Expletive *það* in impersonal constructions

The literature states that expletive *það* does not occur in impersonal constructions prior to c.1500, at which point *það* is attested in impersonal passives, and with weather verbs in c.1540 (Rögnvaldsson, 2002). However, the IcePaHC data challenges this, see Table 3.

Period	Prefinite <i>það</i> (V2)	No <i>það</i> (V1)	Total	% <i>það</i>	χ^2
--------	------------------------------	-----------------------	-------	--------------	----------

¹⁸ *Ordbog over det norrøne prosasprog*. [Dictionary of Old Norse Prose.] (n.d.) University of Copenhagen, <https://onp.ku.dk> [last accessed 31.07.2018].

¹⁹ <http://onpweb.nfi.sc.ku.dk/webart/v/vi/88016731326cartpnfr.htm> [last accessed 26.09.2018].

1150-1350	6	157	163	3.7%	***
1351-1550	4	110	114	3.5%	***
1551-1750	2	64	66	3.0%	**
1751-1900	17	72	89	19.1%	ns
1901-2008	68	18	86	79.1%	***
All periods	97	190	518	18.7%	

Table 3: Frequency of prefinite expletive *það* in impersonal constructions in IcePaHC, 1150-2008

While expletive *það* is relatively infrequent in the early data, there are six examples with the expletive from different texts in Old Icelandic (1150-1350), contra the literature. I show examples in (15)-(17).

- (15) **Það** er nú að segja frá Hrafni. Hann
EXPL be.PRS now to say.INF from Hrafn he.nom
kom á Eyri og...
come.pst to Eyrir and
‘Now one speaks of Hrafn. He came to Eyrir and...’
(1350, Finnbogi.1394)²⁰
- (16) **Það** var snemma um morgun.
EXPL be.PST early about morning
‘It was early in the morning.’ (1350, BandamennM.1196)
- (17) **Það** er mælt um sakir þær allar
EXPL be.PRS speak.PASS.PTCP about things DEM all
sem hér eru taldar, um frumhlaup
REL here be.PRS tell.PASS.PTCP about personal-assault
og um sár og um víg og
and about injury and about manslaughter and
lagalöstu alla...
law-evasions all

²⁰ The example initiates a new discourse and so it is reasonable to rule out the possibility that *það* is anaphoric.

‘One speaks about all those things which are told here, about personal assault and about injury and about manslaughter and all evasions of the law...’ (1270, Gragas.334)²¹

Contra previous claims, the examples in (15)-(17) indicate that expletive *það* could occur in impersonal constructions in Old Icelandic, although the overall preference at this stage was for there to be no expletive.

Examples with expletive *það* in the period 1351-1550 are shown in (18)-(20).

(18) **Það** er nú sagt af einum ríkum manni
EXPL be.PRS now say.PASS.PTCP of one rich man
og mikilhæfum... Hann tók sótt hættliga.
and talented he.nom take.pst sickness dangerous
‘One now speaks of a rich and talented man... He caught a
dangerous illness.’
(1475, Ævintyri.477)²²

(19) **Það** skall ekki svo ganga ef ég kann hitta hana.
EXPL shall NEG so go.INF if I.NOM can meet.INF she.ACC
‘Things will not go that way if I can meet her.’
(1475, Ævintyri.655)

(20) **Það** er þvottdagur.
EXPL be.PRS Saturday
‘It is Saturday.’ (1540, ntjohn.193.331)

Strikingly, the early examples with the expletive are restricted to ‘say’-type predicates, e.g. (15), (17)-(18), and inherently subjectless predicates (i.e.

²¹ This sentence introduces a new chapter, and so one can rule out the possibility that the clause-initial *það* has anaphoric reference to something in the preceding discourse.

²² This sentence begins a new chapter, see footnote 21.

those which do not select a subject argument), e.g. (16), (19)-(20).²³ Only in the late 19th/early 20th century does expletive *það* appear in impersonal constructions with other types of predicate. I show some examples in (21)-(23). This generalisation to all predicate types accounts for the significant increase in expletive *það* in the data for 1901-2008 (see Table 3).

- (21) **Það** þurfti ekki að vitja um Hans...
 EXPL need.PST NEG to check.INF about Hans
 ‘It is not necessary to check up on Hans...’ (1883, Voggur.81)
- (22) ... **það** var barið; Geimundur og Snjólaug
 EXPL be.PST knock.PASS.PTCP Geimundur and Snjólaug
 komu jafnsnemma til dyranna...
 come.pst same-time to doors.def
 ‘...someone knocked; Geimundur and Snjólaug came to the door at the same time...’ (1902, Fossar.1623)
- (23) **Það** má reyna að telja einhverjum öðrum en
 EXPL may try.INF to tell.INF someone other than
 mér trú um það.
 I.dat truth about DEM
 ‘One may try to tell someone other than me the truth about that.’ (1908, Ofurefli.1638)

3.4 Expletive *það* in presentational constructions

Rögnavaldsson (2002) found that expletive *það* is absent in presentational constructions in early Icelandic, first appearing in c.1500, with the verbs *vera* ‘be’ or *verða* ‘become’. This claim is supported by the IcePaHC data, where *það* is absent in the period 1150-1350, see Table 4. Later in the diachrony,

²³ See Booth (forthcoming) for discussion of inherently subjectless predicates in Icelandic.

increases in expletive *það* are visible in the data for 1751-1900 and 1901-2008. The frequency of *það* in presentationals for 1901-2008 is higher (94.5%) than the frequency observed for impersonal constructions (79.1%), cf. Table 3.

Period	Prefinite EXPL (V2)	No EXPL (V1)	Total	%EXPL	χ^2
1150-1350	0	39	39	0.0%	***
1351-1550	5	33	38	13.2%	***
1551-1750	4	21	25	16.0%	***
1751-1900	35	19	54	64.8%	ns
1901-2008	86	5	91	94.5%	***
All periods	130	117	247	52.6%	

Table 4: Proportion of prefinite expletive *það* in presentational constructions in IcePaHC, 1150-2008

The first examples with the expletive occur in the period 1351-1550, e.g. (24)-(26). For those examples which do not start a new chapter, I provide the preceding context to show that *það* cannot be analysed as anaphoric.

- (24) **Það** var einn ríkur maður, að
 EXPL be.PST.SG one.NOM.SG rich.NOM.SG man.NOM.SG COMP
 hann trúði ei rétt eftir því sem honum
 he believe.PST NEG rightly after DEM REL him
 bar...
 deserve.PST

‘There was one rich man, that he did not rightly believe in that which he deserved...’ (1475, *Ævintyri*.175)²⁴

- (25) **Það** var einn kvinna er fastaði
 EXPL be.PST.SG one.NOM.SG woman.NOM.SG REL fast.PST
 við brauð og vatn fyrir Marju messu
 with bread and water for Mary.GEN mass
 Magdalena].
 Magdalene.GEN
 ‘There was one woman who fasted with bread and water for
 Mary Magdalene’s mass.’ (1475, *Ævintyri*.17)²⁵

- (26) Þar voru og rómverskir og júdear og þeir sem
 there be.PST also romans and jews and DEM REL
 kallaðir voru proseliti, en það voru þeir
 call.PASS.PTCP be.PST proselytes but DEM be.PST DEM
 sem snúist höfðu frá heiðnum siðum og
 REL turn.PST.PTCP have.PST from heathen ways and
 til júdeverskra hluta. **Það** voru og þeir
 to jewish things EXPL be.PST.PL also DEM.NOM.PL
 er sum ritning kalla Gethinn...
 REL some writings call.PRS Gethinn
 ‘There were also Romans and Jews and those who were called
 Proselytes, but that was those who had turned from heathen
 customs to Jewish ways. There were also those whom certain
 writings call Gethinn...’ (1525, *Georgius*.694)

The five examples with the expletive in the period 1351-1500 are attested in texts from 1475-1525, thus at the very end of the second period. All five occur with the verb *vera* ‘be’, in line with Rögnvaldsson’s (2002) finding. In all five, the verb agrees in number with the postfinite subject noun phrase, as in present-day Icelandic. The two texts in which the examples are attested are

²⁴ This sentence begins a new chapter, see footnote 21.

²⁵ This sentence also begins a new chapter, see footnote 21. *Einn* is masculine in form and so unexpected as a modifier of *kvinna* (FEM); this is however the text as in *IcePaHC*.

translations rather than native Icelandic compositions: *Ævintyri* is a translation from English and *Georgius* from Low German. I return to this observation below.

Of the four examples with expletive *það* in 1551-1750, two occur in texts which – though not direct translations – are acknowledged to have influence from German and/or Danish. This applies to (27) from *Indiafari*, recognised to have some German influence, and (28) from *Vidalin*, which has possible German and Danish influence.²⁶

(27) Ei minnist ég múrshæðina, en víst
 NEG remember.PRS I wall-height.DEF but certainly
 meina ég það fimmtán faðma hæð vera.
 think.PRS I DEM fifteen fathoms.GEN height be.INF
Það eru margir kimar niður við
 EXPL be.PRS.PL many.NOM.PL bilges.NOM.PL down by
 sjómálið...
 high-waterline.DEF
 ‘I do not remember the wall height, but I certainly think it to
 be fifteen fathoms in height. There are many bilges down by
 the high waterline...’ (1661, *Indiafari*.73.1287)

(28) Hvað er nú framar? Þú ert forsmáður, líður
 what be.PRS now further you be.PRS manager pass.PRS
 órétt, enginn vill kannast við þig, það
 injustice no-one wish.PRS reckon.INF with you EXPL
 er ekki nýtt í heiminum.
 be.PRS.SG nothing.NOM.SG new.NOM.SG in world.DEF

²⁶ The *Indiafari* extract reports speech of German-speaking soldiers, and there is also some German influence in the narrative, see:
<https://github.com/antonkarl/icecorpus/blob/master/info/1661.indiafari.bio-tra.info> [last accessed 26.09.2018].

For details on German and Danish influence in *Vidalin*, see:
<https://github.com/antonkarl/icecorpus/blob/master/info/1720.vidalin.rel-ser.info> [last accessed 26.09.2018].

‘What more is there? You are the manager, you pass injustice, no-one wishes to reckon with you, there is nothing new in the world.’ (1720, Vidalin.603)

The remaining two examples in 1551-1750 represent the first presentational constructions with *það* in native Icelandic compositions, shown here in (29) and (30).

(29) Þar með gjörði hún so að inna: “**Það**
 there with do.PST she so to tell.INF EXPL
 vekjast upp harmar í hjarta mínu...
 awake.PRS.MID.PL up sorrows.NOM.PL in heart my
 ‘Therewith she went to tell the following: there awakens
 sorrow in my heart...’ (1675, Armann.113.775)

(30) En síðan tóku þeir að smá-dirfast, báðu
 but then begin.PST they to small-dare.INF ask.PST
 þeir kónginn um Viðey, **það** væri
 they king.DEF about Viðey EXPL be.PST.SBJV.SG
 einn klén hestahólmi, svo
 one.SG.NOM small.SG.NOM horse-holm.SG.NOM so
 sem vera mætti fyrir eina tvo hesta eða þrjá...
 as be.INF might for one two horses or three
 ‘But then they started to be slightly daring, they asked the king
 about Viðey, there would be one small horseholm, such as
 could be for one, two or three horses...’ (1680, Skalholt.127)

There are, however, earlier examples of presentational constructions with a clause-initial *þar*, formally identical to the locative adverb *þar* ‘there’, which is plausibly expletive, e.g. (31)-(32). Both *Græn* and *Gunnar* are considered native Icelandic compositions.

(31) **Þar** var suðurmaður einn í
 there be.PST.SG south-man.NOM.SG one.NOM.SG in
ferð er Tyrkir hét.
 company REL Tyrkir be-called.PST

‘There was one southern man in the company, who was called Tyrkir.’ (1250, *Græn*, Chapter 2)²⁷

- (32) **Þar** var ein hella stór
there be.PST.SG one.NOM.SG rock-slab.NOM.SG big.NOM.SG
á vellinum.
on field.DEF
‘There was a certain big slab of rock in the field.’
(1400, Gunnar 679)

In (31)-(32), there is no specified location in the preceding context to which *þar* could have anaphoric reference. Furthermore, each example has a location specified later in the sentence (shown in italics). These two observations I take as indication that *þar* is an expletive, rather than a locative adverb. This contrasts with present-day Icelandic, where only *það* ‘it’ is available as an expletive in presentational constructions. If one excepts such examples as involving an expletive *þar*, this challenges standard accounts in two ways. Firstly, it is widely assumed that Old Icelandic had no expletive in presentational constructions (see section 2.3). Secondly, the possibility that *þar* was available as an expletive in Icelandic presentational constructions historically is not considered in the literature.

Later examples with a plausibly expletive clause-initial *þar* also exist, e.g. (33)-(34). Again, each example features a locative later in the sentence.

- (33) **Þar** var gnægð grjóts og
there be.PST.SG abundance.NOM.SG stone.GEN and

²⁷ *Græn* refers to *Grænlendinga saga* (c.1250) which is one of the texts which was manually examined for expletive contexts in Booth (forthcoming). The edition used was that available on the online text repository *Icelandic Netútgáfan*: <https://www.snerpa.is/net/index.html> [last accessed 31.07.2018].

viður hér og þar við sjávarströndina með
 wood.NOM.SG here and there by sea-beach.DEF with
björgunum.
 cliffs.DEF
 ‘There was an abundance of stone and some wood here and
 there by the beach with the cliffs.’ (1650, Illugi.563)

- (34) **Þar** bjó í Kaupinhafn einn
 there live.PST.SG in Copenhagen one.NOM.SG
 slátrari sá er Kristján hét...
 butcher.NOM.SG DEM REL Kristján be-called.PST
 ‘There lived in Copenhagen one butcher who was called
 Kristján...’ (1661, Indiafari.35.258)

If one accepts the analysis of *þar* as an expletive in (33)-(34), then these examples, together with examples like (31)-(32) above, indicate that there was a stage in Icelandic at which both *þar* ‘there’ and *það* ‘it’ were available as expletives in presentational constructions. Cases of competition between two expletive forms have been observed elsewhere in historical Germanic; see e.g. Faarlund (1990, pp. 70–72) and Kinn (2011) on Middle Norwegian; Brevik (1983, pp. 257–259, 324) and Pfenninger (2009, pp. 54–56) on English. I leave further examination of the Icelandic case of competition for future research.

4. Discussion

4.1 A contact-induced change?

Returning to the status of *það* as an expletive, the earliest examples of presentational constructions with *það* in IcePaHC were found in translation texts, see (24)-(26), or in texts which are assumed to have some influence from other languages, see (27)-(28). This is in line with the study by Rögnvaldsson (2002), who found the earliest examples of presentational constructions with *það* in the English translation *Ævintýri*, see (24)-(25). This observation leads Eythórssón & Sigurðardóttir (2016, p. 102) to claim that it is “likely that the filler *það* found its way into Icelandic due to foreign influence”. This seems a rather bold leap and is unlikely for a number of reasons.

Firstly, expletive *það* was not wholly absent before this point. As I showed in section 3.3, there are examples of impersonal constructions in Old Icelandic (1150-1350) where expletive *það* is present, e.g. (15)-(17). The later appearance of *það* in other construction types does not, therefore, represent a complete innovation in the language.

Secondly, as I argued in section 3.4, an expletive *þar* is attested in presentational constructions before the first appearances of expletive *það* in this context. So the availability of an expletive in Icelandic presentational constructions was already a feature of the grammar, before examples of expletive *það* appear in translation texts.

Thirdly, the example in (18) with expletive *það* casts further doubt on the contact theory. This example comes from the *Ævintýri* text, the English

translation which has been at the centre of the contact claim. (18) is an impersonal passive, a construction which is not grammatical in English; this instance of expletive *það* thus cannot be accounted for in terms of a wholesale borrowing from English.²⁸

Finally, typological evidence shows that structural borrowings occur relatively rarely and typically require intense contact between speaker populations (e.g. Thomason & Kaufman, 1988), circumstances not documented for this period. Translation could be considered relatively intense contact, and it is plausible that the examples in (24)-(25) represent the first instances of expletive *það* in Icelandic presentational constructions, on the model of English. However, the claim that the expletive became a fully productive feature of Icelandic grammar from these sparse contact-related examples is implausible, based on our current understanding of language change, where change is standardly acknowledged to work from the bottom-up, rather than in a top-down fashion. Moreover, as the results presented in Table 4 show, there is a significant time gap between the expletives in this text (*Ævintýri*, c.1475), and the development whereby expletive *það* becomes frequent in presentational constructions (in the data for 1751-2008).

For these reasons, an analysis in terms of spread within the language seems more plausible than an exclusively contact-induced explanation for the

²⁸ I thank an anonymous reviewer for pointing this out.

emergence of expletive *það* in Icelandic. In Booth (forthcoming), I present such an account which argues that expletive *það* spread to new construction types – including presentational constructions where it out-competed the earlier expletive *þar* – on the model of earlier topicless constructions with cataphoric *það* and a clausal argument, e.g. (10)-(12).

4.2 *A closer look at distribution*

In section 2.3, I introduced the Prefinite First Hypothesis which states that prefinite expletives emerge before subject expletives in historical Germanic. Though widely accepted, this hypothesis has not been tested for Icelandic. In present-day Icelandic, expletive *það* is restricted to the clause-initial prefinite position in impersonal and presentational constructions, thus behaving like a prefinite expletive (see section 2.2). Given this distribution in the present-day language, if the Prefinite First Hypothesis holds for Icelandic, then one expects no instances of the expletive in postfinite position (i.e. as an unambiguous subject expletive) in the earlier periods.

I test this hypothesis against the IcePaHC data. Of all instances of expletive *það*, I examine to what extent these are restricted to the prefinite position.²⁹ I examine impersonal and presentational constructions together. Contra the

²⁹ The positional distribution of cataphoric *það* in the history of Icelandic is investigated in Booth (forthcoming).

hypothesis, postfinite expletives do occur in all IcePaHC periods, although prefinite expletives are much more frequent than postfinite expletives, see Table 5.

Period	Prefinite EXPL	Postfinite EXPL	Total	%Prefinite	χ^2
1150-1350	6	1	7	85.7%	ns
1351-1550	9	3	12	75.0%	**
1551-1750	6	3	9	66.7%	***
1751-1900	52	7	59	88.1%	ns
1901-2008	154	1	155	99.4%	**
All periods	227	15	242	93.8%	

Table 5: Position of expletive *það* in impersonal and presentational constructions in IcePaHC, 1150-2008

I provide select examples of postfinite expletives in (35)-(38).

- (35) Var **það** síð dags.
 be.PST EXPL late day.GEN
 ‘It was late in the day.’ (1350, Finnbogi.635.548)
- (36) En frá keisaranum er **það** að segja.
 but from emperor.DEF be.PRS EXPL to say.INF
 ‘But one speaks of the emperor.’ (1525, Erasmus.402)
- (37) Margt er **það** líka, sem enn er
 much.NOM.SG be.PRS.SG EXPL also REL still be.PRS
 óljóst í ævisögu jarðarinnar...
 unclear in biography earth.GEN.DEF
 ‘There is also much which is still unclear in the biography of
 the Earth...’ (1835, Jonasedli.141)

- (38) Sömuleiðis var það hneisa fyrir landið,
 same-time be.PST.SG EXPL shame.NOM.SG for land.DEF
 ef höfuðsmaður bæri íslenzkum
 if chieftain bear.PST.SBJV Icelandic.DAT
 veizlum illa söguna meðal hærri
 celebrations.DAT bad.ACC story.ACC.DEF among higher
 stéttar manna erlendis...
 classes men.GEN abroad
 ‘At the same time there was shame across the land, if a
 chieftain would bring to Icelandic celebrations the bad story
 among higher classes of men abroad...’
 (1882, Torfhildur.132.1860)

Moreover, there are also examples of expletive *þar* in postfinite position in presentational constructions in the earlier data, e.g. (39)-(40).

- (39) Talaði Víglundur það einn tíma að hann
 tell.PST Víglundur CATPH one time COMP he.NOM
 vildi að þau myndu sína
 wish.PSTCOMP they.NOM bind.PRS.SBJV their-own
 ást með fastmælum en Ketilríður gaf sér
 love with fast-talk but Ketilríður give.PST REFL
 fátt að: “Eru þar,” segir hún
 few PTCL be.PST.PL EXPL say.PRS she.NOM
 “margir hlutir í móti”.
 many.NOM.PL thing.NOM.PL against
 ‘Víglundur said one time that he wished that they would seal
 their love with an engagement but Ketilríður took coldly to
 this: "There are", she says, "many things against that." ’
 (1400, Viglundur.539)
- (40) Gekk hann í móti sínum fóstbróður með
 go.PST he.NOM towards his-own foster-brother with
 allmikilli gleði, og var þar allmikill
 all-great joy and be.PST.SG EXPL all-great.NOM.SG
 fagnafundur...
 joyful-meeting.NOM.SG
 ‘He went towards his foster brother with great joy and there
 was a greatly joyful meeting....’ (1480, Jarlmann.569)

These examples are evidence that *það* cannot be exclusively accounted for as a structurally motivated filler for V2. Nevertheless, given that the expletive

overwhelmingly prefers the prefinite position in all periods, it does not seem reasonable to claim that *það* qualifies as a fully-fledged subject expletive in the historical data.

A final observation of the findings in Table 5 is that the positional distribution of expletive *það* appears to undergo a change as of 1901; in the data for 1901-2008, virtually all instances of *það* occur in prefinite position, in contrast to the earlier periods which show a more mixed picture.³⁰ The IcePaHC data therefore indicates that the restricted positional distribution standardly claimed in the literature for *það* holds for modern Icelandic (post-1901), but not for older stages (pre-1901).

5. Conclusion

The corpus study presented in this paper sheds new light on the status of expletives in historical Icelandic. Two claims in the literature were confirmed by the IcePaHC data: (1) cataphoric *það* is indeed frequently present in constructions with a clausal argument in the earliest texts; (2) expletive *það* was confirmed to be absent in presentational constructions prior to c.1475, when it begins to appear in translation texts. However, contrary to the

³⁰ The postfinite expletive for 1901-2008 is in a text from 1902, i.e. at the very beginning of the period.

literature it was shown that expletive *það* is already attested in impersonal constructions in Old Icelandic (1150-1350).

I argued that an earlier form *þar* ‘there’ was available as an expletive in presentational constructions and persists in later periods, thus overlapping with expletive *það*. The availability of these two options challenges previous accounts of historical Icelandic, which neglect the availability of *þar* as an expletive. Furthermore, in light of the overall corpus findings, a recent claim by Eythórsson & Sigurðardóttir (2016) that the emergence of expletive *það* is a contact phenomenon was argued to be implausible.

More broadly, the IcePaHC findings pose a challenge for the Prefinite First Hypothesis assumed by many for the diachrony of expletives in Germanic. The corpus findings indicate that the Icelandic expletive did not start out exclusively as a prefinite expletive, contra the predictions of the hypothesis. The positional prefinite restriction standardly claimed for the Icelandic expletive is thus in fact a relatively recent phenomenon (post-1901). In sum, the Icelandic data – together with other recent studies which pose similar problems (Axel, 2007 on German; Kinn, 2016 on Norwegian) – indicates that the Prefinite First Hypothesis should be revisited, particularly in light of the empirical opportunities offered by the current availability of historical syntactic corpora for many Germanic languages.

References

- Axel, K. (2007). *Studies on Old High German Syntax: Left Sentence Periphery, Verb Placement and Verb-Second*. Amsterdam: John Benjamins.
- Booth, H. (forthcoming). *Expletives and Clause Structure: Syntactic Change in Icelandic*. Doctoral dissertation, University of Manchester.
- Breivik, L. E. (1983). *Existential there: A Synchronic and Diachronic Study*. Bergen: Department of English, University of Bergen.
- Chomsky, N. (1981). *Lectures on Government and Binding*. Dordrecht: Foris publications.
- Eythórsson, T. & Sigurðardóttir, S. S. (2016). A brief history of Icelandic weather verbs: syntax, semantics and argument structure. *Working Papers in Scandinavian Syntax*, 96, 91–125.
- Faarlund, J. T. (1990). *Syntactic Change: Toward a Theory of Historical Syntax*. Berlin; New York: de Gruyter.
- Falk, C. (1993). *Non-referential Subjects in the History of Swedish*. Lund: Department of Scandinavian Languages, University of Lund.
- Haiman, J. (1974). *Targets and Syntactic Change*. The Hague: Mouton.
- Hróarsdóttir, Þ. (1998). *Setningafræðilegar Breytingar á 19. öld: Þróun Þriggja Málbreytinga*. Reykjavík: Institute of Linguistics, University of Iceland.
- Ingham, R. (2001). The structure and function of expletive *there* in pre-modern English. *Reading Working Papers in Linguistics*, 5, 231–249.
- Kinn, K. (2011). Overt non-referential subjects and subject-verb agreement in Middle Norwegian. *Working Papers in Scandinavian Syntax*, 88, 21–50.
- Kinn, K. (2016). Referential vs. non-referential null subjects in Middle Norwegian. *Nordic Journal of Linguistics*, 39(3), 277–310.
- Kroch, A. & Taylor, A. (2000). The Penn-Helsinki Parsed Corpus of Middle English (PPCME2). Philadelphia: Department of Linguistics, University of Pennsylvania, second edition. <http://www.ling.upenn.edu/hist-corpora/>.

- Lenerz, J. (1985). Zur Theorie syntaktischen Wandels: das expletive *es* in der Geschichte des Deutschen. In W. Abraham (Ed.), *Erklärende Syntax des Deutschen*, (pp. 99–136). Tübingen: Gunter Narr.
- Maling, J. (1988). Variations on a theme: Existential sentences in Swedish and Icelandic. *McGill Working Papers in Linguistics. Special Issue on Comparative Germanic Syntax*, 168–191.
- Pfenninger, S. E. (2009). *Grammaticalization Paths of English and High German Existential Constructions: A Corpus-based Study*. Bern: Peter Lang.
- Platzack, C. (1983). Existential sentences in English, German, Icelandic and Swedish. In F. Karlsson (Ed.), *Papers from the 7th Scandinavian Conference of Linguistics*, (pp. 80–100). Helsinki: Department of General Linguistics, University of Helsinki.
- Randall, B. (2005). *CorpusSearch2 User's Guide*. Philadelphia: Department of Linguistics, University of Pennsylvania. <http://corpus-search.sourceforge.net/CS-manual/Contents.html>.
- Richards, M. & Biberauer, T. (2005). Explaining Expl. In M. den Dikken & C. Tortora (Eds.), *The Function of Function Words and Functional Categories*, (pp. 115–153). Amsterdam: John Benjamins.
- Rögnvaldsson, E. (2002). ÞAÐ í fornu máli - og síðar. *Íslenskt mál*, 24, 7–30.
- Santorini, B. (2010). *Annotation manual for the Penn Historical Corpora and the PCEEC*. Philadelphia: Department of Linguistics, University of Pennsylvania. <http://www.ling.upenn.edu/hist-corpora/annotation/index.html>.
- Sells, P. (2005). The peripherality of the Icelandic expletive. In M. Butt & T. Holloway King (Eds.), *Proceedings of the LFG '05 Conference* (pp. 408–428). Stanford, CA: CSLI Publications.
- Sigurðsson, H. Á. (1990). V1 declaratives and verb raising in Icelandic. In J. Maling & A. Zaenen (Eds.), *Syntax and Semantics: Modern Icelandic Syntax* (pp. 41–69). San Diego, CA: Academic Press.
- Sigurðsson, H. Á. (2007). Argument features, clausal structure and the computation. In T. Bhattachayra, E. Reuland & G. Spathas (Eds.), *Argument Structure* (pp. 121–158). Amsterdam: John Benjamins.
- Thomason, S. G. & Kaufman, T. (1988). *Language Contact, Creolization, and Genetic Linguistics*. Berkeley, CA: University of California Press.
- Thráinsson, H. (1979). *On Complementation in Icelandic*. New York: Garland.

- Thráinsson, H. (2007). *The Syntax of Icelandic*. Cambridge: Cambridge University Press.
- Vikner, S. (1995). *Verb Movement and Expletive Subjects in the Germanic Languages*. Oxford: Oxford University Press.
- Wallenberg, J. C., Ingason, A. K., Sigurðsson, E. F. & Rögnvaldsson, E. (2011). Icelandic Parsed Historical Corpus (IcePaHC), version 0.9. http://linguist.is/icelandic_treebank.
- Weenen, A. de Leeuw van (Ed.). (1993). *The Icelandic Homily Book: Perg. 15 4^o in the Royal Library, Stockholm*. Reykjavík: Stofnun Árna Magnússonar á Íslandi.
- Zaenen, A., Engdahl, E. & Maling, J. (2017). Subject properties in presentational sentences in Icelandic and Swedish. *Bergen Language and Linguistics Studies*, 8(1), 260–281.