Dative Subjects and the Rise of Positional Licensing in Icelandic

Hannah Booth
University of Manchester

Christin Schätzle
University of Konstanz

Kersti Börjars
University of Manchester

Miriam Butt
University of Konstanz

Abstract

We present the results of research on two areas of Icelandic historical syntax: dative subjects and V1 word order. These strands of syntax had previously been examined independently, but were found to be intimately connected as part of a broader collaboration between theoretical and computational linguistics involving the Icelandic Parsed Historical Corpus (IcePaHC). The interaction we found between V1 declaratives and dative subjects provides evidence for: a) changes over time with respect to the association of dative arguments with the subject role (contra Barðdal and Eythórsson 2009); b) the gradual development of left peripheral structure and the rise of positional licensing (in line with Kiparsky 1995, 1997). We provide an analysis of positional licensing in LFG terms and account for the newly observed complex interaction between datives, subjects and word order presented in this paper.

1   Introduction

Icelandic is standardly acknowledged to be the most conservative of the Germanic languages; Modern Icelandic has retained a complex morphological case system, three genders, as well as rich verbal morphology with distinctions for person, number, tense and mood. The language is also conservative in the sense that it has not undergone certain diachronic developments exhibited in other Germanic languages, such as the innovation of an indefinite article. Nevertheless, various morphosyntactic changes have been observed over the course of the ten centuries of attested written Icelandic. Among the most salient of these are a reduced freedom in word order (Rögnvaldsson, 1995), which we interpret as increased configurationality, a decrease in verb-first declaratives (Sigurðsson, 1990), an increase in dative subjects and the rise of overt expletive elements (Rögnvaldsson, 2002).

Overall, however, morphosyntactic change in Icelandic remains relatively understudied. In particular, the interaction between the various different changes has scarcely been explored. Moreover, most pre-existing studies typically compare data from the modern language with that from the earliest attested stage ('Old Icelandic', c. 1150-1350), thus leaving the diachronic detail of the intervening periods unclear. We present a corpus linguistic study using the Icelandic Parsed Historical Corpus ('IcePaHC', Wallenberg et al., 2011). In this study, we set out to investigate the interaction between word order, subjects and datives. The IcePaHC data allows us to conduct our investigation across all attested stages of the language (1150-2008), thereby capturing an unusually comprehensive level of diachronic detail. Our study reveals an interaction between the increasingly close association of certain types of

dative arguments with subjects and a developing preference for subjects to occur clause-initially. Overall, we see our data as providing evidence for the gradual development of left peripheral structure and the rise of positional licensing in the history of Icelandic. Building on Kiparsky's (1995; 1997) proposals for positional licensing, we propose an LFG analysis that lays the groundwork for capturing the complex interrelation between case, word order, lexical semantics and information structure across the whole Icelandic diachrony.

2 Background

Modern and Old Icelandic both exhibit the verb-second (V2) constraint, which maximally allows one constituent in the clause-initial prefinite position (Rögnvaldsson and Thráinsson, 1990; Rögnvaldsson, 1995). Compare the subject-initial clause in (1-a) with (1-b) and (1-c), which each have a fronted non-subject constituent in initial position, the verb in second position and the subject in postfinite position:

(1) a. Ég **gleymdi** þeim fljótt.   **S-V-O**
         I.NOM forget.PST.1SG they.DAT quickly
         'I quickly forgot them.'
     b. Þeim **gleymdi** ég fljótt.   **O-V-S**
         they.DAT forget.PST.1SG I.NOM quickly
         'Them, I quickly forgot.'
     c. Fljótt **gleymdi** ég þeim.   **X(P)-V-S**
         quickly forget.PST.1SG I.NOM they.DAT
         'Quickly I forgot them.'

Despite being a V2 language, modern Icelandic still allows for verb-first (V1) order in declarative sentences (Sigurðsson, 1990), as exhibited in (2):

(2) a. **Sá** ég þá á svipstundu villu míns
         see.PST.1SG I.NOM then on moment.DAT error.ACC my.GEN
         vegar.
         way.GEN
         'I then saw at once the error of my ways.' (IcePaHC: Margasaga, 1985)
     b. **Hugði** ég mér gott til hans liðveislu.
         think.PST.1SG I.NOM I.DAT good to he.GEN assistance.GEN
         'I expect to benefit from his assistance.' (IcePaHC: Ofsi, 2008)

V1 declaratives are typical for older stages of Germanic (see e.g. Hinterhölzl and Petrova, 2010), but not generally for modern Germanic languages, where this order has either been lost or reduced to certain contexts, such as the context of the lead-in to jokes in German ('Came a woman through the door...'). Icelandic is exceptional in that such structures are still a general part of the modern language. Nevertheless, a decrease in V1 over time has been noted (Sigurðsson, 1990). This has previously been explained in connection with the rise of the expletive element

það occurring in clause-initial position (Franco, 2008; Sigurðsson, 1990), although the exact interaction between V1 and *það* has not been closely examined.

Another striking feature of Icelandic morphosyntax are non-nominative subjects, including dative subjects. The synchronic existence of dative subjects has been well established for modern Icelandic by showing that dative arguments as in the examples in (3) qualify as subjects on the grounds of various classic tests for subjecthood in Icelandic (e.g., Andrews, 1976; Zaenen et al., 1985):

(3)   a.   **Mér** fannst   eins og **þeim**   stæði
           I.DAT seem.PST.3.SG as   if   they.DAT stand.PST.SBJV.3SG
           stuggur   af mér.
           aversion.NOM of I.DAT
           'It seemed to me as if they were frightened of me.'
           (IcePaHC: Mamma, 2008)
       b.   ... **mér** leið   illa.
            I.DAT feel.PST.3SG badly
           'I was feeling bad.' (IcePaHC: Ofsi, 2008)

The historical origin of dative subjects has attracted a good deal of research in recent years. A major point of debate concerns whether dative subjects are a common Proto Indo-European feature or are in fact a more recent innovation. The so-called 'Oblique Subject Hypothesis' (e.g., Barðdal and Eythórsson, 2003, 2009; Barðdal et al., 2012) argues for the Proto Indo-European inheritance of dative subjects, mainly drawing on the continuous existence of a monolithic dative subject construction throughout the Icelandic diachrony. This hypothesis challenges the more traditional 'Object-to-Subject Hypothesis' (cf. Cole et al., 1980; Haspelmath, 2001), which generally takes dative subjects to be the result of the gradual reanalysis of former objects. Evidence for the Object-to-Subject Hypothesis has been adduced from a related branch of Indo-European: the diachrony of Indo-Aryan case. While Old Indo-Aryan shows no evidence for dative subjects (Hock, 1990), dative objects were gradually reanalyzed as dative subjects during the New Indo-Aryan period, a process connected to lexical semantic shifts of individual verbs (cf. Deo, 2003; Butt and Deo, 2013), even after the original case system was lost in the course of Middle Indo-Aryan.

Interestingly, the Icelandic attestation only goes back to the 12th century, which is around the time that dative subjects begin to be possible in Indo-Aryan. Moreover, the Icelandic situation with respect to dative subjects is not as stable as it should be, assuming them to be an old part of the language. There is a change which has been dubbed 'Dative Substitution' or 'Dative Sickness', which can be traced back to the 19th century, but is currently still in progress (see e.g. Svavarsdóttir, 1982; Smith, 1996; Jónsson, 2003; Barðdal, 2011). This change describes a process whereby accusative experiencer subjects are systematically replaced by datives as given in (4). Dative substitution has been assumed to be connected to the increasing systematic association of dative case with experiencer semantics (Smith, 1996; Jónsson, 2003)

and presumably began in the latter part of the 19th century (Barðdal, 2011).

(4)   a.   **Mig** langar   að fara.
            I.ACC long.PRES to go.INF
            'I long to go.'
   b.   **Mér** langar   að fara.
            I.DAT long.PRES to go.INF
            'I long to go.'

<div align="right">(Smith, 1996, 22)</div>

In addition, Schätzle et al. (2015) found that the distribution of dative subjects changes over the centuries. As with dative substitution, datives are increasingly associated with experiencers, and experiencer subjects in turn increase over time due to the lexicalization of former middle formations as psych predicates.

The examination of the distribution of dative subjects over time in conjunction with the Object-to-Subject hypothesis also necessitates a study of word order patterns. In conducting these studies, we discovered a previously unnoticed connection between the two exceptional features of Icelandic described above. We detail these empirical observations in the next section.

3   Corpus study: Findings and Conclusions

We investigated the interaction between subject position, verb placement, expletive *það* and subject case in matrix declarative clauses in several detailed corpus studies. This paper brings together two strands of research that were originally conducted independently: 1) the interaction between V1 and expletives (the Manchester team); 2) factors conditioning the distribution of dative subjects (the Konstanz team). In looking at these phenomena as part of a larger collaborative enterprise exploring Visual Analytics for Linguistics (LingVis), we discovered that the phenomena were more closely interlinked than previously noted.[1]

Our historical data is based on IcePaHC. IcePaHC is a valuable resource as it covers all attested stages of Icelandic, with 61 text extracts dating from the 12th to the 21st century CE. At the same time, however, we recognize some limitations of the corpus: the texts included represent a very small sample of the wealth of attested historical Icelandic, and there is a genre issue in that certain genres are over-represented and others under-represented in individual periods. The corpus is syntactically annotated in the Penn Treebank-format (Marcus et al., 1993) and provides information about clause types, grammatical relations, constituent order, noun types (proper nouns, empty/overt pronouns and expletives, etc.) and case, see e.g. the sample annotation in Figure 1, which shows a matrix IP with a clause-initial pronominal dative subject (*mér* 'me') followed by an inflected form (*finnst*), an old middle, of the verb *finna* 'seem, think, feel':

---

[1]We concentrate here on providing the relevant data and figures; for our most recent work on LingVis for historical linguistics with reference to Icelandic, see Schätzle et al. (2017).

```
(IP-MAT-SPE (NP-SBJ (PRO-D Mér-mér))
    (VBPI finnst-finna)
    (CP-ADV-SPE (WADVP-1 0)
      (C sem-sem)
      (IP-SUB-SPE (ADVP *T*-1)
              (NP-SBJ (PRO-N ég-ég))
              (BEPS sé-vera) (VBN sloppinn-sleppa)
              (PP (P úr-úr) (NP (NP-POS (ONE+Q-G einhvers-einhver)
              (N-G konar-konar)) (N-D fangelsi-fangelsi)))))
      (. .-.))
  (ID 1882.TORFHILDUR.NAR-FIC,.603))
```

Figure 1: Sample annotation from IcePaHC

We extracted the constructions relevant to our study and on which we base our diachronic claims using the CorpusSearch query language specific to Penn-style treebanks (Randall, 2000) and our own Perl scripts.[2] In total, we identified 65,394 declarative matrix clauses in the corpus in which the subject was distinctly annotated for case. These clauses form the foundation for the frequency calculations in our study. The obtained frequencies were divided into time periods as suggested by the existing literature on historical Scandinavian (Haugen, 1984), see e.g., the first column in Table 1.[3] We also conducted χ^2-tests to examine whether the observed distributions differed from what could be expected given the overall distributions of the constructions in the whole corpus (p<0.05*, p<0.01**, p<0.001***).

3.1 Subject positions, verb placement and expletives

At first, we looked at the distribution of subjects, i.e. whether they occur before or after the finite verb. The diachronic distributions of prefinite and postfinite subjects are displayed in Table 1, with the percentage of prefinite subjects given in the penultimate column. Texts in the first time period (1150–1349) have an almost equal share of pre- and postfinite subjects. Over time however, subjects are increasingly realized in the prefinite position, with a highly significant increase of prefinite subjects from 57.6% to 73.0% in the period post-1900. Given the striking increase in prefinite subjects as of 1900, we wondered if what we see in the data indicates the development of a designated subject position. In order to test this hypothesis, we investigated the options for verb placement in IcePaHC and their conditioning factors in more detail.

Butt et al. (2014) show in a corpus study of V1 matrix declaratives in IcePaHC that although V1 is attested throughout the history of Icelandic, V1 constructions undergo a marked decrease as of 1900. We confirmed their findings with our

---

[2]It is worth noting that the corpus query language does not always yield all the relevant constructions; we therefore rely mainly on our own Perl scripts.

[3]Note that the third time period in this classification, 1550–1749, is affected by a genre effect inherent in IcePaHC which causes deviating percentages from the overall developments throughout our study (cf. Butt et al. 2014; Schätzle et al. 2015). This time period is mainly represented by religious and legal texts, while narrative texts (i.e. sagas/modern fiction) dominate the other stages.

Period	prefinite	postfinite	Total	% prefinite	χ^2
1150-1349	7045	6672	13717	51.4%	***
1350-1549	10091	8258	18349	55.0%	***
1550-1749	6076	5134	11210	54.2%	***
1750-1899	6490	4767	11257	57.6%	
1900-2008	7924	2937	10861	73.0%	***

Table 1: Positions for all subjects across IcePaHC.

dataset and provide the respective frequencies in Table 2. The initial share of 20.6% V1 clauses decreases significantly over time, with a particularly striking drop in frequency between the last two time stages, from 18.4% to merely 2.7%.

Period	V1	non V1	Total	% V1	χ^2
1150-1349	2829	10888	13718	20.6%	***
1350-1549	3656	14693	18349	19.9%	***
1550-1749	1654	9556	11210	14.8%	***
1750-1899	2072	9185	11257	18.4%	***
1900-2008	292	10569	10861	2.7%	***

Table 2: Distribution of V1 matrix declaratives in IcePaHC.

As mentioned in section 2, the decrease in V1 has been previously connected to a simultaneous increase in the frequency of the expletive element *það* (Franco, 2008; Sigurðsson, 1990). In present-day Icelandic, expletive *það* appears in clause-initial position in a range of construction types, including presentationals, e.g. (5-a), but in older stages of the language it is typically absent, e.g (5-b):[4]

(5) a. **Það** var   töluverður   snjór   yfir öllu.
         EXPL be.PST.3SG considerable.NOM snow.NOM over everything.DAT
         'There was a considerable amount of snow over everything.'
         (IcePaHC: Ofsi, 2008)
     b. _____ Var   þá gleði   mikil   í kóngs   höll.
         ØEXPL be.PST.3SG then joy.NOM great.NOM in king.GEN hall.DAT
         'There was then great joy in the king's hall.'
         (IePaHC: Jarlmann, 1480)

A connection between a decrease in V1 and the rise of clause-initial expletives has previously been observed by Axel (2007) for historical German, though the connection has not been closely investigated for Icelandic. Our second corpus study therefore examined the proportion of instances in which expletive *það* is present in the clause-initial prefinite position rendering the sentence V2 (5-a), compared

---

[4]ØEXPL is used descriptively to mark the absence of *það* in a context where it would be expected in Modern Icelandic. This is not a theoretical statement on the possibility of 'null expletives'.

to those instances in which *það* is absent, resulting in a V1 structure (5-b). The proportions are compared across the five time periods in order to assess whether a rise in expletive *það* could explain the decrease in V1. The findings from IcePaHC are shown in Table 3:

Period	prefinite expl (V2)	no expl (V1)	Total	% expl	χ^2
1150-1349	16	153	169	9.5%	***
1350-1549	26	205	231	11.3%	***
1550-1749	13	87	100	13.0%	***
1750-1899	59	92	151	39.1%	
1900-2008	160	28	188	85.1%	***

Table 3: Distribution of prefinite expletives in IcePaHC.

The data in Table 3 indicates that there is a striking increase in expletive *það* as of 1900. This coincides with the decrease in V1 observed in Table 2 and thus it seems reasonable to conclude that an increase in expletive *það* in the clause-initial prefinite position rendering older V1 structures as V2 is a factor behind the decrease in V1.

There are, however, two issues which indicate that the expletive story cannot fully account for the decrease in V1. The first issue is that not all V1 clauses in IcePaHC are constructions which lack an expletive *það*. An example is (6), which has a prototypical referential subject in postfinite position (*drottinn*), and in which an overt expletive is therefore not expected to occur. Such examples indicate that the rise of expletive *það* can only be part of the explanation for the decrease in V1.

(6) Sýndi   **drottinn**   mikla   miskunn   vin
     show.PST.3SG lord.NOM.DEF great.ACC mercy.ACC friend.DAT
     sínum   sankti   Georgíum
     his-own.DAT saint.DAT George.DAT
     'The Lord showed great mercy to his friend St. George.'
     (IcePaHC: Georgius, 1525)

Recall that our diachronic findings for subject positions outlined in Table 1 led us to ask whether we are observing the emergence of a new designated subject position (the prefinite position) as of 1900. This raises a second issue with the V1-expletive connection. It is well known that expletive *það* in Modern Icelandic is restricted to the clause-initial prefinite position in almost all construction types (for recent accounts see Sells, 2005; Sigurðsson, 2007; Thráinsson, 2007). Compare the impersonal passive construction in (7-a) which has a clause-initial *það*, with the identical construction with fronting of NEG in (7-b), in which *það* is absent:

(7) a. **Það** var   ekki minnst   á önnur   dýr.
         EXPL be.PST.3SG NEG mention.PTCP on other.ACC animals.ACC
         'There was no mention of other animals.'

b. Ekki var _____ minnst á önnur
   NEG be.PST.3SG ØEXPL mention.PASS.PTCP on other.ACC
   dýr.
   animals.ACC
   'There was no mention of other animals.' (IcePaHC: Sagan, 1985)

In this respect, expletive *það* does not behave like a subject, as in Icelandic subjects typically invert with the verb and occur in postfinite position in non-subject fronting contexts like (6) or (7-b). Compare the examples in (7) with the Swedish impersonal passive in (8), in which expletive *det* behaves more like a subject, occurring in both pre- and postfinite position:

(8) a. **Det** dansades   i går.   (Swedish)
         EXPL dance.PST.PASS yesterday
         'There was dancing yesterday.'
     b. I går   dansades   **det**.   (Swedish)
         yesterday dance.PST.PASS EXPL
         'Yesterday there was dancing.'

Data like (7) and (8) are standardly cited to support the claim that Swedish expletive *det* qualifies as a syntactic subject, whereas Icelandic expletive *það* does not, e.g. Platzack (1983) and Maling (1988); cf. Faarlund (1990) 'expletive topic'.

   The positional distribution contrast between Icelandic *það* and Swedish *det*, together with the grammaticality of so-called "Transitive Expletive Constructions" in Icelandic (9), which are not permitted in Swedish, lead us to conclude that expletive *það* is not fulfilling the role of a subject in Icelandic. In Transitive Expletive Constructions such as (9), the expletive co-occurs with an overt thematic subject (*margir jólasveinar* 'many Christmas trolls'):

(9) **Það** hafa   [**margir jólasveinar**]   borðað
     EXPL have.PRES.3PL many.NOM Christmas-trolls.NOM eat.PST.PTCP
     búðing.
     pudding.ACC
     'Many Christmas trolls have eaten pudding.'
     (Bobaljik and Jonas, 1996, 209)

   In contrast, Sells (2005), working within LFG, assumes that *það* is always a SUBJ in Icelandic. He treats Transitive Expletive Constructions as in (9) as cases where the thematic subject of the clause (*margir jólasveinar*) is unified with the information coming from the expletive *það*. Sells does not provide direct evidence for the subjecthood of *það*, but assumes it. In contrast, we see examples as in (9) as another indication that *það* is in fact not acting as a canonical SUBJ.

   We conclude that the fact that expletive *það* – which does not straightforwardly qualify as a subject – is increasingly appearing in the clause-initial prefinite position in the historical data does not provide evidence for a newly designated prefinite

subject position. We must therefore revise our initial hypothesis and search for an alternative account.

3.2 V1 and Information Structure

The alternative account which we think likely has already been articulated in the existing literature (e.g. Hinterhölzl and Petrova, 2010) and has been picked up on and formalized within LFG by Sells (2005). We rely heavily on this part of his formal analysis in the account we develop in the next section. The core observation as to the function of the expletive that lies at the heart of this account was made by Rögnvaldsson and Thráinsson (1990). They observe that the function of the expletive is to license a clause in which there is no topic, not even the subject (which tends to be the default topic).[5]

We therefore propose an information structural motivation for the rise of clause-initial *það*, which in turn contributes to the decrease of V1. Our proposal follows Hinterhölzl & Petrova's (2010) information structural account for V1 in historical German, which assumes that the verb served as an information structural boundary, separating topic and comment. V1 clauses thus mark topicless sentences (e.g. presentationals) which place the whole clause into the scope of the assertion, i.e. the whole clause is focused (cf. Hinterhölzl and Petrova 2010). Contrary to Hinterhölzl and Petrova, who assume an articulated structure for Old High German, we assume a flat sentence structure for older Icelandic, but one in which verb placement was used for information structural purposes. In the course of the history of Icelandic, the prefinite position was eventually identified as the preferred topic position and became fixed as part of a more general growth of syntactic structure (cf. Kiparsky (1997) and references therein). Expletive *það*, meanwhile, emerges as a filler for this topic position in topicless sentences, rendering the original V1 structures as V2.[6] Subjects moreover are usually topics (Givón, 1990) and become more firmly associated with the clause-inital position accordingly.

3.3 Dative subjects and word order

Cross-linguistically, there is a well-known trade-off between word order, case and/or agreement to mark grammatical relations (e.g., see Kiparsky 1987, 1988, 1997). Having concluded that structural changes led to the emergence of a clause-initial topic position which frequently houses subjects (subjects often being topical), we examined the diachrony of subject case in IcePaHC, in order to see whether any other changes coincide with the observed structural changes. We based ourselves on the initial findings of Schätzle et al. (2015) on the diachrony of dative subjects, but took those investigations further with respect to overall subject case distributions

---

[5]See also Zaenen (1983).

[6]Maling (1980) argues that Stylistic Fronting is another way to fill subject/topicless clauses, converting a V1 structure into a V2 clause.

and the interaction of dative subjects and voice. Our results are shown in Tables 4 and 5, respectively:[7]

Period	Nom	Dat	Acc	Gen	Total	% Dat	χ^2
1150-1349	13028	535	128	26	13718	3.9%	
1350-1549	17596	591	142	20	18349	3.2%	***
1550-1749	10676	417	106	11	11210	3.7%	
1750-1899	10733	428	80	16	11257	3.8%	
1900-2008	10113	626	115	7	10861	5.8%	***

Table 4: Diachronic distribution of subject case in IcePaHC.

Period	active	middle	passive	Total	% middle	χ^2
1150-1349	397	66	72	535	12.3%	***
1350-1549	417	80	94	591	13.5%	***
1550-1749	239	69	109	417	16.6%	***
1750-1899	273	88	67	428	20.6%	
1900-2008	315	239	72	626	38.2%	***

Table 5: Diachronic distribution of voice across dative subjects in IcePaHC.

The share of dative subjects is stable until the very last time period, in which dative subjects undergo a marked and significant increase from 3.8 to 5.8% as of 1900. Schätzle et al. (2015) have shown that the observed increase of dative subjects is driven by lexical semantic factors, as dative subjects occur more often with so-called psych or experiencer predicates over time. Further evidence for the increasing correlation between lexical semantic factors and dative subjects comes from the growing use of verbs carrying middle morphology (marked via the suffix -*st*) together with dative subjects, which furthermore increases significantly as of the last time period, see Table 5. These formerly middle verbs are mainly lexicalized experiencer predicates (cf. Schätzle et al. 2015). These findings tally well with the establishment of a systematic association of dative case with lexical semantic conditions, i.e. experiencer/goal semantics, over the history of Icelandic, as middle constructions are generally characterized by the structural absence of an agent argument (Kaufmann, 2007; Wood, 2015; Schätzle et al., 2015).

Subsequently, we analyzed the interaction of dative case with subjects and structural position (pre- versus postfinite position) in IcePaHC. Table 6 shows the positions in which dative subjects occur throughout IcePaHC. The last column in

---

[7]The original dative subject sentences as annotated in IcePaHC have undergone an additional process of manual correction, as we encountered several annotation mistakes during our studies. Mainly dative objects had been erroneously annotated as subjects and some nominative proper nouns had been wrongly tagged as datives. The annotation of the pronoun *þér* is furthermore problematic in IcePaHC, because it can be either the 3SG.DAT of the personal pronoun *þú* 'you' or an honorific form of address in nominative case in the older Icelandic texts.

Table 6 moreover displays the original relative frequencies of all prefinite subjects as given in Table 1 for a better comparison.

In contrast to the equal distribution between pre- and postfinite position that we have shown for all subjects, dative subjects are preferably realized in the postfinite position in older stages of Icelandic, with a share of only 24.5% of prefinite subjects in the first time stage. Nevertheless, dative subjects also conform to the overall development of the prefinite position as a fixed topic/subject position, becoming increasingly realized in the prefinite position over time. Again, the most striking increase takes place as of 1900. However, dative subjects lag behind all subjects overall in the increasing preference for the prefinite position, with the prefinite position only becoming dominant for dative subjects as of 1900 with a share of 56.4% (compare 55.0% in Table 1 for all subjects for 1350–1549 already).

Period	prefin (Dat)	postfin (Dat)	Total	% prefin (Dat)	χ^2	% prefin (all)
1150-1349	131	404	535	24.5%	***	51.4%
1350-1549	126	465	591	21.3%	***	55.0%
1550-1749	119	298	417	28.5%	*	54.2%
1750-1899	151	277	428	35.3%		57.6%
1900-2008	353	273	626	56.4%	***	73.0%

Table 6: Subject positions for dative subjects across IcePaHC.

A closer examination of V1 sentences containing dative subjects offered further insights. We found that dative subjects have a stronger overall tendency to occur in V1 constructions than all subjects, see Table 7. Despite the fluctuations that we see in the V1 distribution for dative subjects over time, V1 sentences with dative subjects are decreasing overall. Once more, we observe a very striking drop in V1 sentences in our data set as of 1900 (from 29.4% to only 3.2%).

Period	V1 (Dat)	non V1 (Dat)	Total	% V1 (Dat)	χ^2	% V1 (all)
1150-1349	173	362	535	32.3%	**	20.6%
1350-1549	254	337	591	43.0%	***	19.9%
1550-1749	106	311	417	25.4%		14.8%
1750-1899	126	302	428	29.4%		18.4%
1900-2008	20	606	626	3.2%	***	2.7%

Table 7: Dative subjects in V1 declaratives across IcePaHC.

In sum, the positioning of dative subjects deviates from the overall development of structure, in that their preference to occur clause-initially is considerably weaker than for all subjects. We explain this deviation in terms of the dual linking possibilities associated with experiencer/goal arguments, which are compatible

with both subject and object realization. We claim that the dative experiencers must first become more firmly linked to subjects than objects, before they are able to conform to the overall structural licensing configuration. Our hypotheses and analyses regarding the interaction of positional licensing and subject case in the history of Icelandic on the basis of our corpus evidence are detailed in section 4.

4   Formal Analysis

In this section, we build on the existing formal proposals by Sells (2005) on the syntax of Icelandic with respect to the expletive *það*. Unlike Sells (2005), we do not assume that *það* fulfils the role of a SUBJ (see section 3.1). However, in line with Sells (and previous literature), we see *það* as playing an information structural role in that it licenses topicless clauses. We integrate the patterns found for dative subjects into the analysis by also drawing on Kiparsky's (1997) ideas with respect to the rise of positional licensing and the concomitant "growth" of syntactic structure. Kiparsky argues that the rise of positional licensing correlates with the loss of morphology in the history of English. These factors led to the firm establishment of the functional category I, which is seen as having been only optional in Old English. Icelandic differs sharply from English in that it has not seen a radical loss of morphological marking. However, Modern Icelandic exhibits a fairly fixed word order and it is clear that its word order has become more fixed over time. Our findings confirm this overall take on Icelandic, showing that subjects become more firmly associated with the prefinite position over time (cf. Table 1). Kiparsky argues that while morphological marking has not been lost in Icelandic, it is recessive in determing grammatical relations; *positional licensing* of grammatical relations on the other hand is dominant, thus accounting for the increase in word order rigidity. Kiparsky also provides an analysis of experiencer subjects in Old English, which show variable linking to SUBJ vs. OBJ. Again, we build on these insights in the development of our analysis.

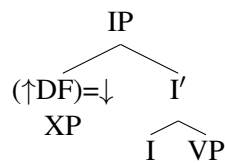## 4.1   Rise of Positional Licensing — LFG

Kiparsky (1995) argues that the Germanic languages developed structure and associated functional categories which were not present in its Indo-European ancestor. In particular, he adduces data primarily from Hulk and van Kemenade (1995) and van Gelderen (1993) to show that Old English lacked a clause-level functional projection C and that this developed at a later stage. He takes parallel data from Old Icelandic and Old High German to indicate that C was present in these languages. The idea that a flat structure lacking functional categories develops into a more articulated structure making use of functional categories has been applied to Icelandic noun phrases by Börjars et al. (2016). We develop an analysis that brings together the factor of positioning for information structural reasons with the increasing use of position for the licensing of grammatical relations, as proposed by Kiparsky. We

assume the restrictive approach to functional categories within LFG, where a functional category is assumed when a functional feature is associated with a structural position (see Kroeger (1993, 6–7) and Börjars et al. (1999)).

For the earliest data, we assume a flat tree in which grammatical relations are indicated by case, along the lines of the analysis developed for Urdu by Butt and King (2004). As is still currently the case in Urdu/Hindi (Gambhir, 1981), word order in Icelandic was used to signal information structural content. In particular, we follow Hinterhölzl & Petrova's (2010) information structural account for V1 in historical German which assumes that the finite verb served as an information structural boundary, separating topic and comment. V1 clauses thus mark topicless sentences (e.g. presentationals).
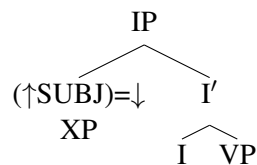
A number of factors contributed to the growth of structure from this point. The information-structural role of the finite verb leads to the initial position becoming increasingly associated with a specific discourse function capturing given or topical information. We follow Kiparsky (1997), who provides evidence for a scenario in which I was optional in Old English, but became a necessary part of the clause over the history of English. This analysis is extended to Icelandic, yielding the clausal structure in (10) (where DF=Discourse Function).

(10)

$$
\begin{array}{c}
\text{IP} \\
\diagdown \\
\underset{\text{XP}}{(\uparrow\text{DF})=\downarrow} \quad \underset{\overset{\frown}{\text{I} \quad \text{VP}}}{\text{I}'}
\end{array}
$$

The subject is often also the topical constituent of a clause. This is because subjects tend to encode the more agentive, sentient clausal participant. Agentive and sentient entities tend to make better topics than inanimate patients. Given this crosslinguistic tendency, it has been argued that subjects are in fact grammaticalized topics (Givón, 1990). This would then lead to the initial position becoming associated with the subject function over time as shown in (11).

(11)

$$
\begin{array}{c}
\text{IP} \\
\diagdown \\
\underset{\text{XP}}{(\uparrow\text{SUBJ})=\downarrow} \quad \underset{\overset{\frown}{\text{I} \quad \text{VP}}}{\text{I}'}
\end{array}
$$

However, this is not quite the case in Icelandic, as we have clauses like the Transitive Expletive Construction, repeated in (12) where the thematic subject *margir jólasveinar* is not in prefinite position. Or V1 constructions as in (13), where the subject is again in the postfinite position.

117

(12) **Það** hafa   [**margir**   **jólasveinar**]   borðað
      EXPL have.PRES.3PL many.NOM Christmas-trolls.NOM eat.PST.PTCP
      búðing.
      pudding.ACC
      'Many Christmas trolls have eaten pudding.'
      (Bobaljik and Jonas, 1996, 209)

(13) **Hafði**   hann   ekki sextíu menn   með sér.
      have.PST.3SG he.NOM NEG sixty   men.ACC with REFL.DAT
      'He didn't have sixty men with him.' (IcePaHC: Fossar, 1902)

Sells (2001, 2005) proposes (14) as the blueprint for clausal structure in Icelandic. The idea is that the prefinite position in SpecIP is associated with a discourse function. Subjects can appear here, but are not restricted to this position. In particular, a subject can appear in the immediately postfinite position when the prefinite position is occupied by the expletive *það* or in V1 sentences. He argues for a set of linear constraints governing the clause structure of Icelandic.

(14)
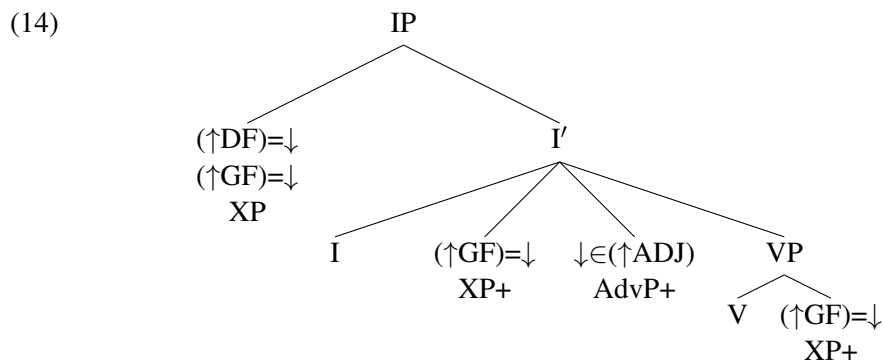
Unlike Sells, we do not see expletive *það* as a subject. Instead, we posit the possibilities shown in (15) for the SpecIP position in Icelandic. For one, it can be a topical position that can host subject topics, but also other topical grammatical functions, though subjects are preferred for the reasons stated above.

Alternatively, this position can be filled by the expletive *það*. In this case, we explicitly state that the sentence has no topic. The position may also remain unfilled, leading to V1 structures, which also do not contain a topic.

(15)
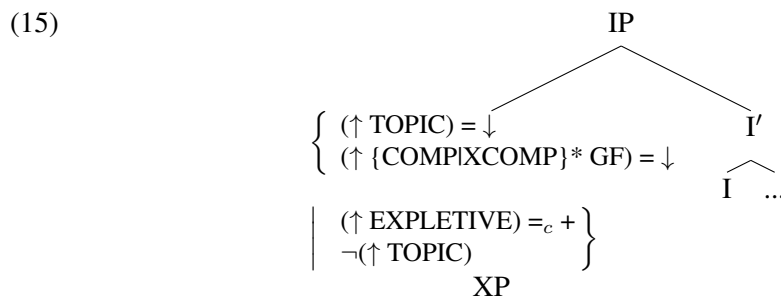
However, an overall reorganization of the information structural packaging in Icelandic is leading to a decrease in V1 constructions in favor of an obligatory prefinite position that hosts topics and, with increasing preference, subjects (cf. Table 6). When clauses do not contain a topic, this position is filled with the expletive.

In sum, our corpus data supports Kiparsky's analysis of the rise of positional licensing in Icelandic.[8] In the next section, we turn to understanding the role of dative subjects with respect to positional licensing.

4.2   Dative Subjects and Positional Licensing

Recall that dative subjects have undergone some change over time. For one, dative case has become more firmly associated with experiencers over the history of Icelandic. For another, dative experiencer arguments become increasingly associated with the prefinite position, which in turn is a topical position that is increasingly becoming associated with subjects.

Experiencer predicates involve an experiencer and a stimulus. Neither experiencers nor stimuli are prototypical agents or patients in the sense of Dowty (1991) or Van Valin and La Polla (1997), and are in principle open to variable linking. There are several sources for experiencer predicates, among them former middles in Icelandic, see Schätzle et al. (2015). The relationship between an experiencer and a stimulus is essentially a stative one and is often historically derived from locative predication where the experiencer is encoded as a location at which the stimulus is placed (cf. Butt and Deo (2013) for Sanskrit), as shown in (16).

(16)   ...og þó er   mér grunur
        and still be.PRES.3.SG I.DAT suspicion.NOM
        '...and I am still suspicious' (Lit. suspicion is to me) (IcePaHC: Morkin,
        1275)

The variable linking possibilities are easily demonstrated with respect to the former locative predications. (17) shows the linking for a locative predication, where a stimulus (theme) is associated with a location. Notice, however, that as per the classic analysis of locative inversion (Bresnan and Kanerva, 1989), this configuration also allows for an alternative linking by which the location is linked to a subject and the theme to an object.

---

[8]A related topic is the issue of null pronouns/arguments. Icelandic has allowed for various types of argument drop throughout its history, see e.g. Sigurðsson (1993). In an IcePaHC study, Kinn et al. (2016) show that the frequency of referential null subjects has decreased over the history of Icelandic, again with the period as of 1900 showing the most significant changes. This finding also supports the overall rise of positional licensing.

(17)   Experiencer Verb   <   theme   location >
                                [−r]   [−o]
                                 | |
                                SUBJ   OBL
                    OR   OBJ   SUBJ
                                NOM   DAT/GEN

A hallmark of experiencers is that they tend to be sentient, which is in turn a property more commonly associated with Proto-Agents rather than Proto-Patients (see Dowty, 1991). Assuming the argument hierarchy in (18), the location can therefore over time be reinterpreted as a sentient goal/experiencer, resulting in the alternative linking in (17) being preferred for experiencer predicates, as in (19).

(18) $agent > goal/experiencer > instrument > theme/patient > location$

(19)   Experiencer Verb   <   experiencer   theme/stimulus   >
                                   [−o]   [−r]
                                    | |
                                   SUBJ   OBJ
                                  DAT/GEN   NOM

With respect to Icelandic, we postulate that topical dative arguments were placed prefinitely in what was established as a topical SpecIP position. In turn, this position becomes more firmly associated with subjects as of 1900, but recall that our data showed that dative subjects consistently lag behind other subjects in appearing in the prefinite position. We explain this as follows: dative experiencers are not prototypical subjects and may retain some object properties. Over time, dative experiencers become more firmly established as subjects, with the linking possibility in (19) becoming established as dominant linking for experiencer predicates. This is in line with what Allen (1995) has shown for oblique experiencers in English. Allen demonstrates that oblique experiencers already have some subject properties in Old English, and that when nominative subjects become obligatory in the 14th century, the experiencer subjects also conform and are realized as nominatives.

Once the prefinite position becomes more firmly established as a subject position in Icelandic (with over 70% of subjects appearing there in the data as of 1900), dative subjects also increasingly occur there, in order to conform to the overall structural change in the language. However, as they are not prototypical subjects, they follow suit, rather than lead the change; as non-canonical subjects, they eventually follow the language's overall positional licensing.

5   Summary and Conclusions

Kiparsky (1997) explicitly discusses experiencer subjects. These present an interesting difficulty for his linking system, which is based on binary features, like LFG's standard Mapping Theory. These features are associated with position, thematic role,

case and agreement morphology (for a summary of Kiparsky's theory of linking, see Butt (2006)). In Kiparsky's analysis, the Icelandic experiencer argument is assigned features that are in principle not compatible with a subject analysis, because the features associated with the dative preclude a linking to subject. However, he sees the linking to subject as being "rescued" by the features associated with SpecIP. Positional nominative case can be assigned in this position and the features associated with positional nominative case "override" the inherent features of the dative argument. In Kiparsky's analysis, morphology is thus "recessive", as it does not determine the grammatical relation status of an argument. Rather, position is dominant due to the rise of positional licensing in Germanic.

Our corpus study provides further evidence for the rise of positional licensing in Icelandic, but we also suggest that the prefinite SpecIP position is not yet exclusively a subject position, as V1 has not been lost (as yet) and the expletive *það* co-occurs with postfinite thematic subjects. We have further shown evidence for changes in the status of dative arguments. These become more firmly associated with experiencer semantics, which in turn become more firmly associated with a linking to subjects. As the prefinite position is becoming more established as a dedicated subject position, dative subjects are also increasingly found in this position.

Our analysis posits a complex interaction between case, word order, lexical semantics and information structure, but separates out lexical semantics from positional licensing. That is, unlike in Kiparsky's system, where position and case interact directly, LFG's linking theory does not factor in position. Rather, position and grammatical functions interact, so that grammatical functions act as a mediator between position and lexical semantics. Information structure in turn interacts with word order. With respect to this interaction, much more needs to be done to understand how information structure was expressed in Old Icelandic vs. the newer developments in the language.

References

Allen, Cynthia L. 1995. *Case Marking and Reanalysis: Grammatical Relations from Old to Early Modern English*. Oxford: Oxford University Press.

Andrews, Avery D. 1976. The VP Complement Analysis in Modern Icelandic. In *Proceedings of the North East Linguistic Society (NELS)*, volume 6, pages 1–21, reprinted 1990 with minor revisions in *Modern Icelandic Syntax*, 165–185.

Axel, Katrin. 2007. *Studies on Old High German Syntax: Left Sentence Periphery, Verb Placement and Verb-Second*. Amsterdam: John Benjamins.

Barðdal, Jóhanna. 2011. The Rise of Dative Substitution in the History of Icelandic: A Diachronic Construction Grammar Account. *Lingua* 121(1), 60–79.

Barðdal, Jóhanna and Eythórsson, Thórhallur. 2003. The change that never happened: the story of oblique subjects. *Journal of Linguistics* 39, 439–472.

Barðdal, Jóhanna and Eythórsson, Thórhallur. 2009. The Origin of the Oblique Subject Construction: An Indo-European Comparison. In V. Bubenik, J. Hewson

and S. Rose (eds.), *Grammatical Change in Indo-European Languages*, pages 179–193, Amsterdam: John Benjamins.

Barðdal, Jóhanna, Smitherman, Thomas, Bjarnadóttir, Valgerður, Danesi, Serena, Jenset, Gard B. and McGillivray, Barbara. 2012. Reconstructing Constructional Semantics: The Dative Subject Construction in Old Norse-Icelandic, Latin, Ancient Greek, Old Russian and Old Lithuanian. *Studies in Language* 36(3), 511–547.

Bobaljik, Jonathan and Jonas, Dianne. 1996. Subject Positions and the Roles of TP. *Linguistic Inquiry* 27(2), 195–236.

Börjars, Kersti, Chisarik, Erika and Payne, John. 1999. On the justification for functional categories in LFG. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG99 Conference*, Stanford, CA: CSLI Publications.

Börjars, Kersti, Harries, Pauline and Vincent, Nigel. 2016. Growing syntax: the development of a DP in Northern Germanic. *Language* 91, e1–37.

Bresnan, Joan and Kanerva, Jonni. 1989. Locative Inversion in Chicheŵa: A Case Study of Factorization in Grammar. *Linguistic Inquiry* 20, 1–50.

Butt, Miriam. 2006. *Theories of Case*. Cambridge: Cambridge University Press.

Butt, Miriam, Bögel, Tina, Kotcheva, Kristina, Schätzle, Christin, Rohrdantz, Christian, Sacha, Dominik, Dehe, Nicole and Keim, Daniel. 2014. V1 in Icelandic: A multifactorical visualization of historical data. In *Proceedings of VisLR: Visualization as added value in the development, use and evaluation of Language Resources*, Workshop at the 9th edition of the Language Resources and Evaluation Conference (LREC 2014), Reykjavik, Iceland.

Butt, Miriam and Deo, Ashwini. 2013. A Historical Perspective on Dative Subjects in Indo-Aryan, Paper presented at the LFG13 Conference.

Butt, Miriam and King, Tracy Holloway. 2004. The Status of Case. In Veneeta Dayal and Anoop Mahajan (eds.), *Clause Structure in South Asian Languages*, pages 153–198, Berlin: Kluwer Academic Publishers.

Cole, Peter, Harbert, Wayne, Hermon, Gabriella and Sridhar, S. N. 1980. The acquisition of subjecthood. *Language* 56(4), 719–743.

Deo, Ashwini. 2003. Valency Change and Case Marking: Marathi Dative Experiencers, handout from the Pioneer Workshop on Case, Valency and Transitivity.

Dowty, David. 1991. Thematic proto-roles and argument selection. *Language* 67(3), 547–619.

Faarlund, Jan Terje. 1990. *Syntactic Change : Toward a theory of historical syntax*. Berlin ; New York: de Gruyter.

Franco, Irene. 2008. V1, V2 and criterial movement in Icelandic. *Studies in Linguistics* 2, 141 – 164.

Gambhir, Vijay. 1981. *Syntactic Restrictions and Discourse Functions of Word Order in Standard Hindi*. Ph. D.thesis, University of Pennsylvania, Philadelphia.

Givón, Talmy. 1990. *Syntax: a functional typological introduction*. Amsterdam: John Benjamins.

Haspelmath, Martin. 2001. Non-Canonical Marking of Core Arguments in European

Languages. In A. Y. Aikhenvald, R.M.W. Dixon and M. Onishi (eds.), *Non-Canonical Marking of Subjects and Objects*, pages 53–83, Amsterdam: John Benjamins.

Haugen, Einar. 1984. *Die skandinavischen Sprachen: Eine Einführung in ihre Geschichte*. Hamburg: Buske.

Hinterhölzl, Roland and Petrova, Svetlana. 2010. From V1 to V2 in West Germanic. *Lingua* 120.2, 315–328.

Hock, Hans Henrich. 1990. Oblique Subjects in Sanskrit. In M. Verma and KP Mohanan (eds.), *Experiencer Subjects in South Asian Languages*, pages 119–139, Stanford: CSLI Publications.

Hulk, Aafke and van Kemenade, Ans. 1995. Verb-second, *Pro*-drop, Functional Projections and Language Change. In A. Battye and Ian Roberts (eds.), *Clause Structure and Language Change*, pages 227–256, Oxford: Oxford University Press.

Jónsson, Jóhannes Gísli. 2003. Not so quirky: on subject case in Icelandic. In E. Brandner and H. Zinsmeister (eds.), *New Perspectives on Case and Case Theory*, pages 129–164, Stanford: CSLI Publications.

Kaufmann, Ingrid. 2007. Middle Voice. *Lingua* 117, 1677–1714.

Kinn, Kari, Rusten, Kristian A. and Walkden, George. 2016. Null subjects in early Icelandic. *Journal of Germanic Linguistics* 28(1), 31–80.

Kiparsky, Paul. 1987. Morphology and grammatical relations, Stanford University.

Kiparsky, Paul. 1988. Agreement and Linking Theory, Stanford University.

Kiparsky, Paul. 1995. Indo-European origins of Germanic syntax. In Ian G Roberts and Adrian Battye (eds.), *Clause structure and language change*, pages 140–167, Oxford: Oxford University Press.

Kiparsky, Paul. 1997. The Rise of Positional Licensing. In Ans van Kemenade and Nigel Vincent (eds.), *Parameters of Morphosyntactic Change*, pages 460–494, Cambridge: Cambridge University Press.

Kroeger, Paul R. 1993. *Phrase structure and grammatical relations in Tagalog*. Stanford, CA: CSLI Publications.

Maling, Joan. 1980. Inversion in embedded clauses in Modern Icelandic. *Íslenskt mál og almenn málfræði* 2, 174–193.

Maling, Joan. 1988. Variations on a theme: Existential sentences in Swedish and Icelandic. *McGill Working Papers in Linguistics* pages 168–191.

Marcus, Mitchell P., Marcinkiewicz, Mary Ann and Santorini, Beatrice. 1993. Building a large annotated corpus of English: the Penn Treebank. *Computational Linguistics* 19(2), 313–330.

Platzack, Christer. 1983. Existential sentences in english, german, icelandic and swedish. In *Papers from the 7th Scandinavian Conference of Linguistics. Department of General Linguistics, University of Helsinki*, pages 80–100.

Randall, Beth. 2000. CorpusSearch: a Java program for searching syntactically annotated corpora, philadelphia: Dept. of Linguistics, University of Pennsylvania.

Rögnvaldsson, Eiríkur. 1995. Old Icelandic: A Non-Configurational Language?

NOWELE 26(95), 3–29.

Rögnvaldsson, Eiríkur. 2002. ÞAÐ í fornu máli – og síðar. *Íslenskt mál* 24, 7–30.

Rögnvaldsson, Eiríkur and Thráinsson, Höskuldur. 1990. On Icelandic Word Order once more. *Syntax and Semantics* 24, 3–40.

Schätzle, Christin, Butt, Miriam and Kotcheva, Kristina. 2015. The Diachrony of Dative Subjects and the Middle in Icelandic: A Corpus Study. In M. Butt and T. H. King (eds.), *Proceedings of the LFG15 Conference*, CSLI Publications.

Schätzle, Christin, Hund, Michael, Dennig, Frederik L., Butt, Miriam and Keim, Daniel A. 2017. HistoBankVis: Detecting Language Change via Data Visualization. In Gerlof Bouma and Yvonne Adesam (eds.), *Proceedings of the NoDaLiDa 2017 Workshop Processing Historical Language*, NEALT Proceedings Series 32, pages 32–39.

Sells, Peter. 2001. *Structure, Alignment and Optimality in Swedish*. Stanford: CSLI Publications.

Sells, Peter. 2005. The peripherality of the Icelandic expletive. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG '05 Conference*, pages 408–428, Standford: CSLI Publications.

Sigurðsson, Halldór Ármann. 1990. V1 declaratives and verb raising in Icelandic. In Joan Maling and Annie Zaenen (eds.), *Modern Icelandic Syntax (Syntax and Semantics 24)*, pages 41–69, San Diego: Academic Press.

Sigurðsson, Halldór Ármann. 1993. Argument-drop in old Icelandic. *Lingua* 89, 247–280.

Sigurðsson, Halldór Ármann. 2007. Argument features, clausal structure and the computation. In Tammoy Bhattachayra, Eric Reuland and Giorgos Spathas (eds.), *Argument structure*, pages 121–158, Amsterdam: John Benjamins.

Smith, Henry. 1996. *Restrictiveness in Case Theory*. Cambridge: Cambridge University Press.

Svavarsdóttir, Ásta. 1982. Þágufallssýki 4, 159–189.

Thráinsson, Höskuldur. 2007. *The Syntax of Icelandic*. Cambridge: Cambridge University Press.

van Gelderen, Elly. 1993. *The Rise of Functional Categories*. Amsterdam: John Benjamins.

Van Valin, Robert D. and La Polla, Randy J. 1997. *Syntax: Structure, Meaning and Form*. Cambridge: Cambridge University Press.

Wallenberg, Joel C., Ingason, Anton Karl, Sigurðsson, Einar Freyr and Rögnvaldsson, Eiríkur. 2011. Icelandic Parced Historical Corpus (IcePaHC).

Wood, Jim. 2015. *Icelandic Morphosyntax and Argument Structure*. New York: Springer.

Zaenen, Annie. 1983. On syntactic binding. *Linguistic Inquiry* 14, 469–504.

Zaenen, Annie, Maling, Joan and Thráinsson, Höskuldur. 1985. Case and grammatical functions: the Icelandic passive. *Natural Language and Linguistic Theory* 3(4), 441–483.