

**THE DIACHRONY OF DATIVE SUBJECTS AND THE
MIDDLE IN ICELANDIC: A CORPUS STUDY**

Christin Schätzle Miriam Butt
University of Konstanz University of Konstanz

Kristina Kotcheva
Universität Freiburg

Proceedings of the LFG15 Conference

Miriam Butt and Tracy Holloway King (Editors)

2015

CSLI Publications

<http://csli-publications.stanford.edu/>

Abstract

Icelandic is very well known for non-nominative subjects. In recent years, it has been proposed that dative subjects are a Proto-Indo European feature, and that a Dative Subject Construction has been inherited through the ages in the daughter language families (Barðdal and Eythórsson, 2009; Barðdal et al., 2012). We conduct a corpus study and show that while dative subjects can indeed already be found in the earliest attested Icelandic texts, their distribution has been changing over the last millenium. In particular, their use in middles has increased significantly. We explain our findings via an increased use of experiencer subjects combined with a more regular association of experiencer arguments with dative case. We provide a formal analysis within LFG's Mapping or Linking Theory that draws on Barron's (2001) analysis of the diachronic development of raising verbs in Latin. Overall, we see our work as providing evidence against dative subjects in Icelandic as being due to an inherited monolithic Dative Subject Construction.

1 Introduction

Dative subjects are common in a variety of modern Indo-European languages. In recent years, their origin has been discussed controversially. The so-called Oblique Subject or Semantic Alignment Hypothesis (see, e.g. Barðdal and Eythórsson 2009; Barðdal et al. 2012) takes dative subjects to be a common Proto Indo-European inheritance. In contrast, the so-called Object-to-Subject Hypothesis assumes that dative subjects were innovated at a later stage in that former dative objects were reanalyzed as subjects (see, e.g. Haspelmath 2001). The latter is the more standard hypothesis and is supported by the empirical facts in Indo-Aryan, where no evidence for dative subjects can be adduced for Old Indo-Aryan (Hock, 1990), but where there is evidence for objects being reanalyzed as dative subjects in several stages (Deo, 2003; Butt and Deo, 2013). The conditioning factors involve verbal lexical semantics and the increasing systematic association of dative case with experiencers arguments.

The oldest attested texts for Icelandic stem from the 12th century. This is about the time that new case markers entered the system in Indo-Aryan (the old system having eroded over the course of Middle Indo-Aryan). It is from the 11th to 12th centuries onwards that dative subjects begin to be possible in Indo-Aryan and evidence for continuing change in progress can be found in Indo-Aryan.

Given the Indo-Aryan situation, we asked ourselves whether a similar time line might not be at work in Icelandic and whether evidence for a change in progress with respect to dative subjects could be found in Icelandic. We investigated this

[†]We would like to thank the audiences of LFG15 for valuable feedback and comments on this work. We would also like to thank the two reviewers and Jóhannes Gísli Jónsson, Anna Kibort, Joan Maling and Annie Zaenen, who took valuable time to engage with our paper and provide us with invaluable feedback. We would also like to acknowledge the DFG-funded project on *Visual Analysis of Language Change and Use Patterns* which supported Christin Schätzle's part of the research.

by conducting a corpus study of the distribution of dative subjects in the Icelandic Parsed Historical Corpus (IcePaHC; Wallenberg et al. 2011). Our aim was to identify factors conditioning the distribution of dative subjects in Icelandic.

Our results show that the use of dative subjects overall is on the rise and is strongly associated with the middle morpheme *-st*.¹ These dative subject middles are mainly found with psych predicates, raising predicates and some specialized lexicalized usages like ‘seem’, which originally meant ‘like’. Nominative subjects are correspondingly decreasing.

The overall picture that emerges from the corpus study suggests that dative case is becoming more systematically associated with lexical semantic factors. This does not support the idea of dative subjects as a stable, common Proto-Indo-European inheritance. Rather, the dative is part of a complex case system that encodes (lexical) semantic distinctions. Our data can be understood via Kibort’s Mapping Theory (cf. Kibort 2013, 2014) in combination with Kaufmann’s (2007) analysis of the middle and Barron’s (2001) LFG analysis of the diachronic change of verbs of perception to raising verbs and epistemic modals. Our findings are also in line with Jónsson (2003), who argues that lexical semantics are the major conditioning factor for case in Icelandic and that Icelandic is seeing a change by which dative subjects are becoming systematically associated with experiencers and goals.

2 Icelandic: Basic Data

Icelandic is an accusative language, with nominative being the default subject case, and accusative the default for direct objects. Although nominative is the most common morphological case for subjects, the synchronic existence of non-nominative subjects is well established (cf. Andrews 1976, 2001; Zaenen et al. 1985). The word order of Icelandic is fairly fixed.

2.1 Non-nominative subjects in Icelandic

Non-nominative subjects, including dative subjects are attested in the earliest Icelandic texts, which date from the 12th century. Old Scandinavian, the ancestor language of Icelandic, is said to display a high degree of correlation between morphological case and semantic role (Faarlund, 1994, 59). Nominative case was characteristic of the agent role, dative was typically used for marking the semantic roles of instrumental, recipient and ablative. Accusative case was associated with themes, and genitive with partitives (in a wider sense). There was no one-to-one relation of semantic role and morphological case. Both subjects and (direct) objects could be in any of the four cases (Thráinsson 1994, 175; Barðdal 2001).

In Modern Icelandic, all non-nominative subjects are said to be non-agentive (Thráinsson 1994, 176, Jónsson 2003). Dative subjects have been associated broadly

¹Wood (2015) analyzes this as a clitic, albeit as one with reduced mobility. The precise morphological status of *-st* is not relevant for the discussion in this paper.

conjunction with a past participle.²

- (3) a. einhver barði strákana í skólanum
somebody.NOM hit.PAST.3.SG the.boy.ACC.PL in the.school.DAT.SG
'Somebody hit the boys in school.'
- b. strákarnir voru barðir í
the.boy.NOM.PL be.PAST.3.PL hit.PPART.M.NOM.PL in
skólanum
the.school.DAT.SG
'The boys were hit in school.' (Thráinsson, 1994, 177)

In contrast, dative and genitive objects generally preserve their case marking under passivization, as shown in (4) for a dative.³

- (4) a. Skipstjórinn sökkti skipinu.
the.captain.NOM.SG sink.PAST.3.SG the.ship.DAT.SG
'The captain sank the ship.'
- b. Skipinu var sökkt af
the.ship.DAT.SG be.PAST.3.SG sink.PPART.N.NOM.SG by
skipstjóranum.
the.captain.DAT.SG
'The ship was sunk by the captain.' (Zaenen and Maling, 1984)

Besides active and passive, Icelandic verbs can be marked for middle voice via the suffix *-st* (Sigurðsson, 1989; Anderson, 1990; Wood, 2015), as shown in (5).⁴ As with passives, the accusative object of a transitive verb in an active clause is realised as a nominative subject in its middle counterpart.

- (5) a. Ég fann hestinn.
I.NOM find.PAST.1.SG the.horse.ACC.SG
'I found the horse.'
- b. Hesturinn fannst.
the.horse.NOM.SG find.PAST.MID.3.SG
'The horse got found.' (Sigurðsson, 1989, 243)

Dative objects vary with respect to middle formation. When the object is a dative theme/patient as in (4-a), the dative case is not preserved, as shown in (6). When the dative marks benefactives or goals, as in (7), the dative is preserved.

²The past participle agrees in number and gender with the nominative subject of the passive clause, but not with non-nominative subjects as per the general rules of Icelandic verb agreement, where they display invariant third person singular agreement (Thráinsson, 1994, 177).

³Prepositions select for case and the 'by' in passive by-phrases selects for a dative.

⁴This distribution of the Icelandic *-st* morphology is complex and not all of the verbs carrying *-st* are middles. However, the great majority are compatible with middle meanings as identified crosslinguistically (Kemmer, 1993; Kaufmann, 2007). Also see Wood (2015).

- (6) a. Peir sökktu bátnum.
 they.NOM sink.PAST.3.PL the.ship.DAT.SG
 ‘They sank the ship.’
 b. Baturinn sökk.
 the.ship.NOM.SG sink.PAST.3.SG
 ‘The ship sank.’ (Sigurðsson, 1989, 277)
- (7) a. Pétur bauð mér vinnu.
 Peter.NOM offer.PAST.3.SG I.DAT job.ACC.SG
 ‘Peter offered me a job.’
 b. Mér bauðst vinna.
 I.DAT offer.PAST.MID job.NOM.SG
 ‘I got the opportunity to get a job.’ (Sigurðsson, 1989, 260)

Wood (2015) argues that datives as in (7) share characteristics with dative goal objects in ditransitives and should be analyzed as applied objects added to a predication via a specialized functional head (Appl). While building on Wood’s basic insights, we provide an alternative analysis of (6) and (7) in section 4.

3 Corpus Study

The goal of our corpus study was to investigate whether dative subjects as a construction have been stable over time in Icelandic or whether evidence for structural change could be found. We based our study on the IcePaHC, which includes texts from the earliest attested stages of Icelandic up to modern times, i.e. from the 12th to the 21st century. There are 60 texts from different genres (mainly Sagas) containing approximately one million words. The corpus is annotated syntactically in the Penn Treebank style (Marcus et al., 1993), which includes the annotation of case and grammatical relations. As a sample, we provide the annotation of (8) in Figure 1. The sentence contains a dative subject with the predicate *finna* ‘seem’.

- (8) Mér finnst sem ég sé
 I.DAT seem.PRES.MID.SG like I.NOM be.SBJNC.PRES.1.SG
 sloppinn úr einhvers konar
 escape.PPART.M.NOM.SG from some.GEN.SG kind.GEN.SG
 fangelsi.
 prison.DAT.SG
 ‘It seems to me as if I have escaped from some kind of prison.’
Brynjólfur Sveinsson biskup, 1882

As can be seen, there is a matrix IP with a pronominal dative subject (*mér* ‘I.DAT’). The verb is an inflected middle form: *finnst*. The annotation also provides a lemma, in this case the infinitive form *finna*. The verb embeds a CP that contains an IP with a pronominal nominative subject, and so on, building a hierarchical constituency and dependency structure.

```
(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

3.1 Corpus Study Part I — Conditioning Factors

The goal of the first part of our corpus study was to identify factors which condition the appearance of dative subjects in Icelandic. We examined the interaction between dative subject case and factors which have previously been implicated in the literature (e.g. Barðdal et al. 2012; Barðdal and Eythórsson 2009; Barðdal 2011; Svenonius 2002; Maling 2002; Svenonius 2006; Jónsson 2003; Jónsson 2013), namely semantic verb class and voice in all sentences of the corpus.⁵

In order to investigate whether verb class is a significant conditioning factor, we manually annotated the IcePaHC with additional information about semantic verb classes. We based our annotations on Levin’s classification for English (Levin 1993) in combination with previous literature on verb classes in Icelandic (Barðdal et al. 2012). An example of this augmented annotation is provided in Figure 2, in which the verb *finna* is classified as a psych verb (PSY).⁶

```
(IP-MAT-SPE (NP-SBJ (PRO-D Mér-mér))
  (VBPI-PSY finnst-finna)
  . . . .
```

Figure 2: Sample annotation of (8) with additional verb class information.

In accordance with the literature on Icelandic (e.g. Haugen 1984), we divided our data and results according to the time stages displayed in Table 1, which shows

⁵In our study, we took the syntactic annotation in the corpus at face value. However, we did find instances of erroneous annotation. We were able to identify some of these as part of the corpus study and duly disregarded them. Others were identified in retrospect by two critical readers: Joan Maling and Jóhannes Gísli Jónsson. Erroneous annotations include, for example, dative subjects on ditransitives and accusative subjects in conjunction with middle morphology, both of which did not and do not exist in Icelandic. We have been in touch with the corpus developers and the errors brought out by our study have now been corrected.

⁶The following categories for semantic verb classes were used in the corpus study: psych, sending and carrying, exerting force, appearance, communication, change of state, combining and attaching, involving the body, judgement, change of possession, removing, verbs with predicative complements, existence, perception, motion, aspect, desire, putting, measure, destruction, social interaction, throwing, ingestion, lingering and rushing, creation and transformation, emission, killing, concealment, search, cutting, posture, contact by impact, learning, poking, image creation, lodging, and cognition. These verb classes are too fine grained and will be revised as part of future work.

our results with respect to voice. We used χ^2 to test whether the observed distributions differ from what could be expected (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$).

Time	active	passive	middle	χ^2
1150-1350	64.4%	18.2%	17.4%	***
1350-1550	66.8%	17.5%	15.7%	***
1550-1750	46.1%	28.8%	25.1%	***
1750-1900	53.1%	20.8%	26.2%	
1900-present	43.2%	14.3%	42.5%	***
all	55.3%	19.5%	25.5%	

Table 1: Diachronic distribution of dative subject predicates by voice

The data confirm that voice is relevant. There are changes in the frequency of actives, passives and middles occurring with dative subjects over time. Dative subjects appear most often with actives, but also fairly often with middle morphology. Dative subjects in active constructions are mostly experiencers and are mainly found with psych predicates (mean=53.4%). An illustrative example is given in (9).

- (9) Það líkaði biskupi illa.
 this.NOM.SG please.PAST.3.SG bishop.DAT.SG badly
 ‘The bishop disapproved of this.’ *Brynjólfur Sveinsson biskup, 1882*

Over the whole time span, dative subjects are found less often with passives. These passive dative subjects are mainly associated with verbs of communication (mean=14.8%), e.g. *tilkynna* ‘announce’ in (10), and change of possession (mean=14.6%) as *kaupa* ‘buy’ in (11), but are also decreasing within these verb classes over time in keeping with the overall decrease of passive dative subjects.

- (10) en drottningu var tilkynnt að
 and queen.DAT.SG be.PAST.3.SG announce.PPART.N.NOM.SG that
 Þorsteinn væri afturkominn.
 Þ.NOM.SG be.PAST.SUBJ.3.SG return.PPART.M.NOM.SG
 ‘And it was announced to the queen that Þorsteinn had returned.’
Sögubáttur af Ármanni og Þorsteini gála, 1675
- (11) Var þessu nú keypt.
 be.PAST.3.SG this.N.DAT.SG now buy.PPART.N.NOM.SG
 ‘This was now bought.’ *Gunnars saga Keldugnúpsfífls, 1400*

While dative subjects with actives and passives show an overall decrease, middle morphology with dative subjects is on the increase. Dative subjects with middles also occur most often with psych predicates (56.2%), as in (12).

- (12) Jú, mér leiðist ósköp, segir Sigríður.
 yes I.DAT.SG be.bored.PAST.MID.SG much say.PRES.3.SG S.NOM.SG
 ‘Yes, I am very bored, Sigríður said.’ *Piltur og stúlka, 1850*

Note that the data in the table would seem to indicate that significant changes happened in the third time stage (1550–1750). However, as discussed in other work (Butt et al., 2014), these deviating percentages are due to a genre effect. While Sagas predominate in the other time stages, the third time stage is mainly represented via religious and legal texts in IcePaHC. The genre effect is clearly visible in the glyph visualization shown in Figure 3. We also used this visualization to explore the interaction between semantic verb class and dative subject marking.

3.2 Semantic Verb Classes

Since lexical semantics plays a role in the appearance of dative subjects, we are investigating the interaction between semantic verb class and dative subjects. Our data has so far yielded some suggestive evidence, but no significant results.

Figure 3 shows a visualization of the entire IcePaHC. The texts are ordered chronologically on the vertical axis. The texts are arranged horizontally according to genre. Narrative sagas are to the left, the religious and legal texts more towards the center and the right. As can be seen, there is a time stage in which the religious and legal texts predominate. This is where the genre effect skews the data.

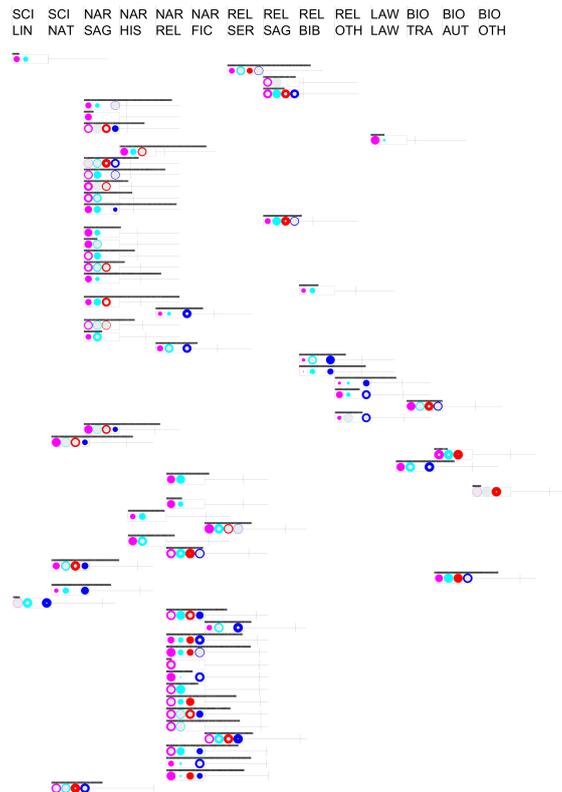


Figure 3: Multifactorial visualization of dative subjects in IcePaHC

Figure 4 displays one glyph from Figure 3. Each glyph represents one of the texts within IcePaHC. On top of the glyph is a horizontal bar which visualizes the text length. The horizontal bar consists of light and dark gray stripes. Each stripe represents a sentence. Light gray stripes indicate those sentences which contain a dative subject. Mousing over the gray stripes allows one to see the sentence involved. The circles represent the appearance of dative subjects with different verb classes. The semantic verb classes are as per the umbrella classifications in Barðdal et al. (2012) and the higher categorization which she assumes in Barðdal (2011).⁷ They can be expanded into subclasses per mouse click, as shown in Figure 4. The circular glyphs are redundantly coded by color and indicate whether a given verb class or category appears more or less often than expected based on the text length and with respect to the entire corpus. If the circular glyphs are filled in from the outside, the given class or category occurs more often than expected. If the glyph is filled from the inside out, the given class or category occurs less often than expected (Schätzle et al., 2014).

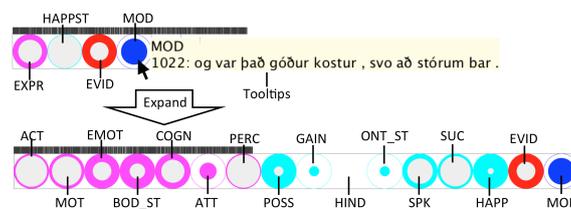


Figure 4: Text glyph visualization of dative subjects in a text from IcePaHC according to lexical semantic verb class and category

The visualization follows an ‘overview first — details on demand’ approach (Keim et al., 2008) allowing for interactive exploratory access to the data which is particularly useful for the analysis of a multifactorial diachronic data set. Interactive visualizations such as the one in Figure 3 are also good in situations where a purely quantitative approach to data is not fine-grained enough because there are too few instances per category to allow for firm statistical conclusions.

The visualization shows a preponderance of dative subjects in experiencer verbs (pink) as well as “happenstance” verbs (light blue), which include verbs of possession and gain and communication (SPK). The other major categories in evidence are verbs of evidentiality (dubbed thus by Barðdal; red) and modals (dark blue), which appear more frequently in the latter stages of the language.

⁷They also include verbs of motion and activity verbs. In future work, we will experiment with a different classification of verbs, in particular, we will experiment with broader umbrella categories, making it more likely that significant patterns can be found. Also note that Jónsson (2003) shows that further factors play a role. For example, verbs of strong positive emotion never take a dative subject, but verbs of negative emotion do.

3.3 Corpus Study Part II — Voice

The first part of our study shows that dative subjects are on the increase in a subset of structural and lexical semantic contexts. Our results suggest that voice plays a major role in the diachronic development of dative subjects. In a second study, we therefore studied potential factors affecting the appearance of dative subjects with different voices. We quantitatively analyzed case, word order, and transitivity according to voice in all matrix declarative sentences in IcePaHC. Moreover, the following verb types in IcePaHC were analyzed separately: main verbs, modals, ‘have’, ‘do’, ‘become’, and ‘be’. Again, we divided the data into the time stages suggested for Icelandic and conducted a χ^2 -test for significance as before.

We first established the overall occurrence of dative subjects in comparison to other types of subjects. We extracted all subject cases in declarative matrix sentences from the corpus. We identified 65568 declarative matrix sentences in which the subjects were distinctly case marked by the annotation format of the corpus excluding ambiguous or non-marked subjects. Their diachronic distribution with respect to subject case is displayed in Table 2.

Time	NOM	ACC	DAT	GEN	χ^2
1150-1350	94.7%	0.9%	4.2%	0.2%	
1350-1550	96.1%	0.7%	3.1%	0.1%	***
1550-1750	95.0%	0.9%	4.1%	0.1%	
1750-1900	95.3%	0.6%	4.0%	0.1%	*
1900-present	93.1%	1.0%	6.0%	0.1%	***
all	95.1%	0.8%	4.2%	0.1%	

Table 2: Diachronic distribution of subject case

As expected, subjects are most often nominative. A small percentage of subjects bear dative case. Accusative and genitive subjects are rare. Diachronically, nominatives only decrease marginally, but significantly in the last time stage, while accusative and dative subjects increase highly significantly. The accusative subject numbers are small, so it is not clear whether this result should be weighed strongly.

This contrasts with Barðdal’s (2008) corpus-based analysis of subject case in Old and Modern Icelandic based on texts coming from similar genres as the texts in IcePaHC. She shows that nominative subjects increase from 76.3% in Old to 85.0% in Modern Icelandic and that dative subjects decrease from 18.4% to 10.3% which clearly contradicts our findings. Accusative and genitive subjects are slightly reduced over time, but rarely appear overall. Additionally, the ratio of non-nominative subjects as opposed to nominative subjects is considerably higher in her study, i.e. 23.7%/15.0% to 76.3%/85.0% versus 4.9% to 95.1% in our study. However, Barðdal only considered subjects of transitive predicates while we did not distinguish between different valencies.

We then further identified 17523 matrix declarative sentences with distinctly case marked subjects and objects of transitive predicates. Our motivation was to

find out whether the case marking system found in the corpus corresponds to what is generally assumed about subject and object marking in Icelandic and to find out how dative case marked NPs in general develop over time. The distribution of object case in the history of Icelandic is shown in Table 3.

Time	NOM	ACC	DAT	GEN	χ^2
1150-1350	4.2%	72.0%	19.5%	4.3%	***
1350-1550	3.6%	69.7%	24.1%	2.6%	*
1550-1750	4.8%	71.3%	21.4%	2.5%	*
1750-1900	3.8%	67.3%	25.7%	3.2%	*
1900-present	4.9%	67.1%	25.6%	2.4%	***
all	4.2%	69.8%	22.9%	3.1%	

Table 3: Diachronic distribution of object case

As expected, objects are most often accusative, but dative objects are also common. Nominative and genitive objects appear much more infrequently. Genitive and accusative objects are slightly reduced over time, while dative objects increase significantly. Nominative objects stay fairly stable over the whole time span.

A data set examined by Barðdal (2001) yielded 93–94% nominative subjects, 67–69% accusative objects and about 25% dative objects, the numbers being stable in both historical and modern stages of Icelandic. These findings tally with our study. However, we do find evidence for change over time.

We examined subject and object cases individually according to different verb types. We found that verb type has no effect on the investigated factors case, voice, word order, and transitivity. Hence, we present our findings for main verbs in the following as they represent the majority of all verbs. We found 51209 matrix declarative sentences with case marked subjects and main verbs in the corpus. The distribution of case marking on subjects and objects for main verbs only is similar to the findings for all verbs (Tables 2 and 3). This is true for the distribution of subject case in active sentences as well (Table 2). However, passive sentences deviate from this overall pattern, see Table 4. Over time, dative subjects increase significantly in passives and nominative subjects decrease slightly.^{8,9}

With respect to middles, as shown in Table 5, the increase in dative subjects is more marked and more significant.¹⁰ Strikingly, nominative subjects are reduced

⁸Accusative subjects with passives are actually ungrammatical, see e.g. Svenonius (2006). The examples we found appear to be falsely annotated nominative and dative subjects. Genitive subjects are indeed possible in passives, but are rarely used.

⁹The first part of our corpus study found that passives were decreasing, in particular with respect to two verb classes. The difference arises from the material under consideration. In this second part of the study, we looked only at matrix declaratives. In the first part, we also included participial constructions and small clauses.

¹⁰Further annotation mistakes were uncovered here. The genitive subjects in Table 5 in the third time stage should have been annotated as attributes, not subjects. Genitive subjects do not occur with middles, and neither do accusative subjects, these also represent annotation errors.

Time	NOM	ACC	DAT	GEN	χ^2
1150-1350	86.4%	0.3%	10.9%	2.4%	*
1350-1550	83.7%	0.0%	14.5%	1.8%	
1550-1750	83.3%	0.2%	16.0%	0.5%	
1750-1900	80.9%	0.0%	16.8%	2.3%	
1900-present	82.2%	0.8%	16.4%	0.6%	*
all	83.8%	0.3%	14.4%	1.6%	

Table 4: Diachronic distribution of subject case in passive sentences

by 17%, while middle morphology in the corpus is strongly associated with dative subjects, showing an increase of 17%. The middle verbs found in the corpus are mainly experiencer predicates, see e.g. (13), and raising verbs. As shown in (14), these are often based diachronically on former psych predicates or verbs of perception. In (14) the raising predicate is *finna* ‘seem’, which used to mean ‘like’ and still has a modern psych verb interpretation as ‘feel/sense’ in certain contexts.

Time	NOM	ACC	DAT	GEN	χ^2
1150-1350	92.8%	0.2%	7.0%	0.0%	***
1350-1550	94.4%	0.2%	5.4%	0.0%	***
1550-1750	92.3%	0.1%	7.5%	0.1%	*
1750-1900	89.5%	0.0%	10.5%	0.0%	
1900-present	75.5%	0.1%	24.4%	0.0%	***
all	89.3%	0.1%	10.6%	0.0%	

Table 5: Diachronic distribution of subject case with middles

- (13) Mest af öllu gramdist honum það thó að sjá
 most of all.DAT resent.PAST.3.SG he.DAT that still to see.INF
 menn, sem ... bregðast sér
 man.NOM.PL REL.PRO fail.PRES.MID self.3.SG.DAT
 ‘But most of all it annoyed him to see men who ... fail him.’
Leysing, 1907
- (14) Ekki finnast mér Sóróasters trúarbrögð eins
 not seem.PRES.MID.SG I.DAT S.GEN.SG religion.NOM.SG one
 herleg og þau sem Óðinn
 glorious.M.NOM.SG and that.N.NOM.SG like Ó.NOM.SG
 hafði og kenndi.
 have.PAST.3.SG and teach.PAST.3.SG
 ‘Zoroaster’s teachings don’t seem to me as glorious as the ones that Odin
 had and taught.’ *Fimmbæðra saga, 1790*

As the increase in dative subjects is strongly correlated with a rise in middle

forms, we took a closer look at the interaction of dative subjects and middles in comparison with the interaction of dative subjects with other voices of main verbs in matrix declaratives. As shown in Table 6 and as expected, the use of dative subjects with middles increases over time. Additionally, active constructions appear less often with dative subjects.

Time	active	passive	middle	χ^2
1150-1350	70.4%	16.6%	13.0%	***
1350-1550	69.3%	16.4%	14.3%	***
1550-1750	49.2%	21.8%	29.0%	***
1750-1900	58.5%	24.3%	17.2%	
1900-present	41.0%	47.1%	11.8%	***
all	58.0%	25.9%	16.1%	

Table 6: Diachronic distribution of dative subject predicates by voice

Most of the dative subjects are found with intransitive verbs, see Table 7. Transitive predicates with dative subjects have the following mean object distribution: nominative 89.6%, accusative 6.7%, dative 2.2%, and genitive 1.5%.¹¹ With respect to word order, no significant deviances with respect to dative subjects from what was found for the whole corpus by Butt et al. (2014) were identified.

Time	intransitive	transitive	ditransitive	χ^2
1150-1350	72.3%	27.4%	0.3%	
1350-1550	69.0%	31.0%	0.0%	
1550-1750	68.5%	31.5%	0.0%	
1750-1900	78.3%	21.7%	0.0%	
1900-present	79.2%	20.8%	0.0%	*
all	73.8%	26.1%	0.1%	

Table 7: Diachronic distribution of dative subject according to transitivity

Our corpus study shows that dative subjects are on the increase overall. Additionally, there is a strong association of dative subjects with middle morphology. These middle forms are mainly found on psych and raising predicates. Nominative subjects with middles are correspondingly decreasing. Other possible factors, e.g. word order, modality, or verbal semantic class, are not significant.

¹¹In fact, the number of genitives should be 0 as this combination is not possible, pointing to further annotation errors. The dative subjects found with ditransitives listed in Table 7 have also all been identified as annotation errors.

4 Analysis and Discussion

The corpus studies have yielded a rich set of observations that need to be understood further. However, a central result of our study is that dative subjects are on the increase overall in Icelandic and that the main factor governing this increase is middle morphology on verbs. In this section we concentrate on understanding this result and put forward a proposal that involves three major factors. One is an increasingly systematic association of dative case with experiencers in Icelandic (Smith, 1994; Jónsson, 2003). The second is the interaction between middle formation and the systematic association of experiencers and goals with the dative. The third is the development of new lexicalizations that instantiate dative experiencer verbs via middle formation and diachronic reanalysis.

4.1 Case and Thematic Role

In analyzing Dative Sickness in Icelandic, Smith (1996) describes it as a change in progress in which datives are becoming systematically associated with either goals or experiencers in Modern Icelandic. This Dative Sickness accounts for some of the increase in dative subjects in Icelandic. Our studies did not find this to be a significant factor with respect to its application to individual verbs. However, we hypothesize that in combination with middle formation, the systematic association of experiencers/goals with datives becomes potent.

4.2 Dative Objects

Objects in Icelandic exhibit a variety of case marking (Maling, 2002). As with other parts of the case marking system, the association between case marking and thematic roles or other event semantic factors is mostly regular (cf. discussions in Zaenen et al. (1985); Faarlund (1994); Jónsson (2003)). Additionally, Svenonius (2002) shows that dative and accusative objects alternate systematically in certain verbs whereby the dative is used when the external argument (the subject) is only part of the initiating or causing event, but does “accompany” the object to the result state. For example, verbs of ballistic motion in which a ball is thrown or an arrow loosed would take a dative object. On the other hand, when a ball is rolled alongside so that the roller and the ball end up in the same place, an accusative would be used. Svenonius’ generalization can motivate the dative on objects as in (6) (the sinkers do not accompany the boat under the water).

Wood (2015) argues that dative objects can be either direct or applied objects. Examples such as (6) or (14) are instances of direct objects, datives as in (7) are applied objects. We adopt Svenonius’ and Wood’s basic insights and show how these can be used to understand the increase of datives with middle morphology.

4.3 Middle Formation and Datives

In providing an analysis of the middle, we draw on Kaufmann’s (2007) analysis in conjunction with Kibort’s (2013; 2014) revised Linking Theory. Kaufmann shows that in middles the agent argument is in principle present on a pragmatic or inferential level; however, it is not accessible for structural purposes. This analysis tallies well with the known facts for Icelandic by which middles are incompatible with agentive modifiers or by-phrases (Wood 2015, 139 and references therein). This stands in stark contrast to the passive, where the agent argument is still structurally available, it is just not realized as a subject, but as an OBL or an adjunct.

We posit the linking schema in Table 8 for middles. It shows that an agent is involved in the event in principle. However, this agent is not available for linking in the middle. The effect is that the agent is not structurally present.^{12,13}

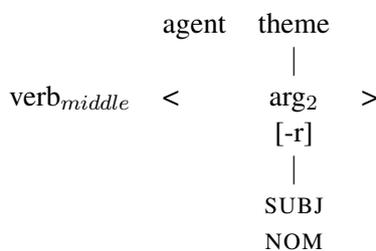


Table 8: Linking Schema for a Middle

When this linking schema applies to (6), the only argument for linking is the boat and it is realized as a nominative subject. A different situation obtains when the dative marks benefactive or goal arguments, as in (15).

- (15) a. Pétur bauð mér vinnu.
 Peter.NOM offer.PAST.3.SG I.DAT job.ACC.SG
 ‘Peter offered me a job.’
- b. Mér bauðst vinna.
 I.DAT offer.PAST.MID job.NOM.SG
 ‘I got the opportunity to get a job.’ (Sigurðsson, 1989, 260)

If we adopt Wood’s analysis, in examples as in (15), the dative is an applied argument. As per standard LFG (Dalrymple, 2001), applied arguments are treated as unrestricted [-r] arguments. This also fits in with the general observation that the goal argument in Icelandic is linked to the direct object in actives (Zaenen et al., 1985; Bresnan, 2001) and the accusative object is the secondary object (OBJ_θ).

As shown in Table 9, when middle formation applies, the agent is not available for linking. Instead, the applied [-r] object is linked to the subject. The remaining argument is linked to the direct object. Case marking here is sensitive to various

¹²Wood (2015) essentially analyzes the middle *-st* as instantiating the agent argument.

¹³For a different LFG analysis of middles, see Arka (2015).

	agent	goal _{appl}	theme
offer _{middle}	<	arg ₂	arg ₃ >
		[-r]	[+o]
		SUBJ	OBJ
		DAT	NOM

Table 9: Linking Schema for the middle of ‘offer’

factors. For one, the subject is a goal and as such attracts dative case. For another, there is a constraint in Icelandic such that if the subject is non-nominative, then the object must be nominative (Jónsson, 2003).¹⁴ The active secondary accusative object is thus realized as a nominative direct object in the middle.

Examples as in (15) add to the numbers of dative subjects found with middles. However, we suggest that the major factor governing the increase of dative subjects with middles is the formation of new experiencer predicates.

4.4 Verbs of Perception and Experiencer Predicates

A majority of the verbs with middle morphology and dative subjects in the corpus are experiencer predicates as in (16) (cf. also (8)). These experiencer predicates are generally cognate with a verb of perception, as shown in (17). A synchronic and diachronic relationship between experiencer predicates, including verbs like ‘seem’, and verbs of perception or feeling as in (17) can be observed crosslinguistically.

- (16) Eiríki finnst þetta ómögulegt
Eric.DAT.SG find.PRES.MID.SG this.NOM.SG impossible
‘Eric finds this impossible.’ (Jónsson, 2003, 131)

- (17) Haraldur fann það.
Harald.NOM.SG feel.PAST.3.SG it.ACC
‘Harald felt/sensed it.’ *Morkinskinna, 1275*

Barron (2001) proposes a diachronic LFG-based analysis for understanding the diachronic development from verbs of perception to experiencer and raising predicates. She identifies three constraints that must obtain for the diachronic development: 1) presence of secondary predication; 2) suppression of the perceiver argument; 3) shift from a physical to a mental process.

The shift from a physical to a mental process is given by pairs like (18) vs. (17). We propose that the second constraint is fulfilled by middle formation. (19) shows that *finna* can be used in contexts with secondary predication.

¹⁴There are some exceptions to this for restricted set of verbs involving either archaic usages or verbs with Acc-Acc patterns where the accusative subject is replaced with a dative via Dative Sickness (Wood, 2015, 45–46).

- (18) Hann fann hestbein.
 he.NOM.SG find.PAST.3.SG horse.bones.ACC.PL
 ‘He found horse bones.’ *Grettis saga Ásmundarsonar, 1310*
- (19) og fann hertoginn að hann mundi skýr
 and find.PAST.3.SG the.duke.NOM that he.NOM would intelligent.NOM
 og málsnjallur vera.
 and speech.excellent.NOM be.
 ‘And the duke found that he would be intelligent and eloquent.’
Fimmbraeðra saga, 1790

The examples in (17)–(19) all involve nominative subjects, but (16) has a dative subject. Working on Latin, Barron suggests that the dative perceiver/experiencer is first introduced as an optional argument in contexts where the nominative perceiver had been suppressed at argument structure (via passivization in her data). Although not realized syntactically, the perceiver is nevertheless part of the event semantics and can hence be introduced back into the syntax optionally. In Latin, this is achieved via a dative NP.

We suggest that essentially the same process is at work in Icelandic and that experiencers are (re)introduced in a clause via an applied dative. Consider (17). In middle formation, the agent and perceiver *Harald* is not expressed as per the linking schema in Table 4.3. However, the experiencer verb ‘feel/sense’ implies an experiencer/perceiver. This experiencer argument can be (re)introduced overtly via a dative applied object, as shown in Table (19).¹⁵ The linking then proceeds exactly as in Table 9.

	agent	experiencer _{appl}	theme	
finna _{middle}	<	arg ₂	arg ₃	>
		[-r]	[+o]	
		SUBJ	OBJ	
		DAT	NOM	

Table 10: Linking Schema for experiencer ‘find’

This analysis accounts for the fact that experiencer verbs in Icelandic are often found with the *-st* middle morphology and makes sense of our corpus studies, which show that an increase in dative subjects was mainly associated with middles.

¹⁵See (Wood, 2015, 231) for argumentation along similar lines. Also see Maling’s (1990) example (61), which shows that the dative argument in experiencer middles is not derivable from the corresponding active clause.

5 Conclusion

The overall picture that emerges from our diachronic corpus study suggests that dative case is becoming more systematically associated with lexical semantic meanings and that the overall system is quite complex. On the one hand, datives are becoming more systematically associated with experiencers and goals. On the other hand, the dative is used when the external initiator of an event only initiates an event, but does accompany the object to a result state. Both types of datives feed into valency alternations conditioned by voice. Our corpus study implicates middles as a major factor in the increase of dative subjects over time. We suggested that this is due to experiencers being introduced via applied datives into experiencer/perception predications. In analogy to developments in Latin, we show how such experiencer predicates can arise out of middles for certain verbs and can thus account for the preponderance of middle morphology with experiencer predicates.

While it could be possible that a single Dative Subject Construction has been inherited from Proto-Indo European and now instantiates the various dative subject examples, the data instead point to a complex system in which lexical and event semantic considerations interact with systematic but variable linking possibilities between arguments and grammatical relations, giving rise to several different realizational possibilities and the concomitant introduction of new predicational meanings such as those associated with dative subject experiencer predicates based on verbs of perception.

References

- Anderson, Stephen R. 1990. The grammar of Icelandic verbs in *-st*. In J. Maling and A. Zaenen (eds.), *Modern Icelandic Syntax*, pages 235–273, San Diego: Academic 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.
- Andrews, Avery D. 2001. Non-canonical A/S Marking in Icelandic. In M. Noonan (ed.), *Non-Canonical Marking of Subjects and Objects*, pages 85–111, Amsterdam: John Benjamins.
- Arka, I Wayan. 2015. Constructed Middles in Marori: an LFG Analysis. This Volume.
- Barðdal, Jóhanna. 2001. *Case in Icelandic — A Synchronic, Diachronic and Comparative Approach*. Ph. D. thesis, Lund University.
- Barðdal, Jóhanna. 2008. *Productivity: Evidence from Case and Argument Structure in Icelandic*. 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. 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, Jensen, 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.
- Barron, Julia. 2001. Perception and Raising Verbs: Synchronic and Diachronic Relationships. In M. Butt and T. H. King (eds.), *Time Over Matter*, pages 73–103, Stanford: CSLI Publications.
- Bresnan, Joan. 2001. *Lexical-Functional Syntax*. Oxford: Blackwell.
- Butt, Miriam, Bögel, Tina, Kotcheva, Kristina, Schätzle, Christin, Rohrdantz, Christian, Sacha, Dominik, Dehe, Nicole and Keim, Daniel. 2014. V1 in Icelandic: A multifactorial 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.
- Dalrymple, Mary. 2001. *Lexical Functional Grammar*. New York: Academic Press.
- Deo, Ashwini. 2003. Valency Change and Case Marking: Marathi Dative Experiencers, handout from the Pioneer Workshop on Case, Valency and Transitivity.
- Faarlund, Jan Terje. 1994. Old and Middle Scandinavian. In E. König and J. van der Auwera (eds.), *The Germanic Languages*, Chapter 3, pages 38–71, New York: Routledge.
- 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.
- 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.
- 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.
- Jónsson, Jóhannes Gísli. 2013. Dative vs. accusative and the nature of inherent case. In B. Fernández and R. Etxepare (eds.), *Variation in Datives. A Micro-Comparative Perspective*, pages 144–180, Oxford: Oxford University Press.
- Kaufmann, Ingrid. 2007. Middle Voice. *Lingua* 117, 1677–1714.
- Keim, Daniel, Andrienko, Gennady, Fekete, Jean-Daniel, Górg, Carsten, Kohlham-

- mer, Jörn and Melançon, Guy. 2008. Visual Analytics: Definition, Process, and Challenges. In A. Kerren, J. T. Stasko, J.-D. Fekete and C. North (eds.), *Information Visualization*, pages 154–175, Berlin: Springer.
- Kemmer, Suzanne. 1993. *The Middle Voice*. Amsterdam: John Benjamins.
- Kibort, Anna. 2013. Objects and Lexical Mapping Theory. In M. Butt and T. H. King (eds.), *Proceedings of the LFG13 Conference*, CSLI Publications.
- Kibort, Anna. 2014. Mapping out a construction inventory with (Lexical) Mapping Theory. In M. Butt and T. H. King (eds.), *Proceedings of the LFG14 Conference*, pages 262–282, CSLI Publications.
- Levin, Beth. 1993. *English Verb Classes and Alternations: A Preliminary Investigation*. Chicago: Chicago University Press.
- Maling, Joan. 1990. The hierarchical assignment of grammatical cases, handout from talk at Stofnun Sigurðar Nordals.
- Maling, Joan. 2002. Það rignir þágufalli á Íslandi. Verbs with Dative Objects in Icelandic. *Íslenskt mál* 24, 31–105.
- 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.
- Schätzle, Christin, Sacha, Domnik and Butt, Miriam. 2014. Diachronic Visualization of Dative Subjects in Icelandic, poster presentation at the Workshop on Big Data Visual Computing, *44th Annual Meeting of the Gesellschaft für Informatik*.
- Sigurðsson, Halldór Ármann. 1989. *Verbal Syntax and Case in Icelandic: A Comparative GB Approach*. Ph. D. thesis, Lund University.
- Smith, Henry. 1994. "Dative Sickness" in Germanic. *Natural Language and Linguistic Theory* 12(4), 675–736.
- Smith, Henry. 1996. *Restrictiveness in Case Theory*. Cambridge: Cambridge University Press.
- Svenonius, Peter. 2002. Icelandic case and the structure of events. *Journal of Comparative Germanic Linguistics* 5, 197–225.
- Svenonius, Peter. 2006. Case Alternations and the Icelandic Passive and Middle. In S. Manninen, D. Nelson, K. Hieta, E. Kaiser and V. Vihman (eds.), *Passives and Impersonals in European Languages*, Amsterdam: John Benjamins.
- Thráinsson, Höskuldur. 1994. Icelandic. In E. König and J. van der Auwera (eds.), *The Germanic Languages*, Chapter 6, pages 142–188, London: Routledge.
- 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 and Maling, Joan. 1984. Unaccusative, passive, and quirky case. In M. Cobler, S. MacKaye and M. T. Wescoat (eds.), *Proceedings of the Third West Coast Conference on Formal Linguistics (WCCFL)*, pages 317–329.
- Zaenen, Annie, Maling, Joan and Thráinsson, Höskuldur. 1985. Case and grammatical functions: the Icelandic passive. *Natural Language and Linguistic Theory* 3(4), 441–483.