Nordic Journal of Linguistics (2026), **49**, 33–67
doi:10.1017/S0332586524000039

CAMBRIDGE UNIVERSITY PRESS

RESEARCH ARTICLE

Argument structure constructions in competition: The Dat-Nom/Nom-Dat alternation in Icelandic

Joren Somers[1], Gard B. Jenset[2] and Jóhanna Barðdal[1]

[1]Linguistics Department, Ghent University, Blandijnberg 2, 9000 Ghent, Belgium and [2]independent scholar
Corresponding author: Joren Somers; Email: joren.somers@ugent.be

(Received 31 August 2023; revised 12 January 2024; accepted 19 January 2024; first published online 06 March 2024)

Abstract
Alternating Dat-Nom/Nom-Dat verbs in Icelandic are notorious for instantiating two diametrically opposed argument structures: the Dat-Nom and the Nom-Dat construction. We conduct a systematic study of the relevant verbs to uncover the factors steering the alternation. This involves a comparison of 15 verbs, five alternating ones, and as a control, five Nom-Dat verbs and five non-alternating Dat-Nom verbs. Our findings show that alternating verbs instantiate the Nom-Dat construction 54% of the time and the Dat-Nom construction 46% of the time on average for four of five verbs when both arguments are full NPs. However, in configurations with a nominative pronoun, the Nom-Dat construction takes precedence over the Dat-Nom construction. Also, for the double-NP configuration, a logistic regression analysis identifies indefiniteness and length as two key predictors, apart from nominative case marking. We demonstrate that the latter systematically correlates with discourse-prominence, which we show, upon closer inspection, correlates with topicality.

Keywords: alternating Dat-Nom/Nom-Dat verbs; dative subjects; logistic regression; neutral word order; pronouns vs. full NPs; subject tests; topicalisation; topicality; usage-based approach

1. Introduction

Modern Icelandic is legendary in the syntactic literature for having non-nominative subject verbs of different types. This includes verbs which select for dative subjects and nominative objects, so-called Dat-Nom verbs. What is less well known is that Dat-Nom verbs in Icelandic divide into two classes with respect to argument structure and the syntactic behaviour of the arguments. One class of Dat-Nom verbs consistently occurs in the Dat-Nom argument structure construction, while another class of verbs alternates between the Dat-Nom and the Nom-Dat argument structure construction (cf. Bernódusson 1982, Jónsson 1997–1998, Barðdal 1999, 2001, 2023: Ch. 3, Platzack 1999, Sigurðsson 2006a, Rott 2013, 2016, Wood & Sigurðsson 2014,

© The Author(s), 2024. Published by Cambridge University Press on behalf of The Nordic Association of Linguists. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/S0332586524000039 Published online by Cambridge University Press

Barðdal, Eythórsson & Dewey 2014, 2019). The difference in behaviour between alternating and non-alternating verbs is illustrated by means of the verbs *nægja* 'find/be sufficient' in (1) and *líka* 'like' in (2) below. The verb *nægja*, being an alternating verb, allows both verbal arguments to take clause-initial position, thus confirming their status as syntactic subjects (cf. 1a–b). At the same time, the other argument is realised in the postverbal slot, which is reserved for objects (for a list of the accepted subject tests in Icelandic, see Section 2 below).[1]

Dat-Nom

(1) a. Gunnari hafði nægt þessi skýring.
*Gunnar.*DAT *had* *sufficed* *this.*NOM *explanation.*NOM
'Gunnar found this explanation sufficient.'

Nom-Dat

(1) b. Þessi skýring hafði nægt Gunnari.
*this.*NOM *explanation.*NOM *had* *sufficed* *Gunnar.*DAT
'This explanation was sufficient for Gunnar.'

Dat-Nom

(2) a. Barninu hafði líkað bragðið illa.
*child.the.*DAT *had* *liked* *taste.the.*NOM *badly*
'The child had not liked the taste.'

*Nom-Dat

(2) b. *Bragðið hafði líkað barninu illa.
*taste.the.*NOM *had* *liked* *child.the.*DAT *badly*
Intended meaning: 'The taste had not been to the child's liking.'

In contrast, in the examples in (2), only the dative of *líka* 'like' may occupy the preverbal position and the nominative the postverbal position (2a), and not vice versa (2b), as is the case with *nægja* 'find/be sufficient' in (1).

By applying a host of accepted subject tests in Icelandic, Barðdal (1999, 2001) was the first to show that either argument of alternating verbs may indeed function as the syntactic subject or the syntactic object. Since then, further work has been carried out on the nature of alternating Dat-Nom/Nom-Dat verbs in Icelandic, including a systematic comparison between the syntactic behaviour of the arguments of classical Dat-Nom verbs and the alternating Dat-Nom/Nom-Dat verbs in Icelandic, also compared to German (cf. Barðdal, Eythórsson & Dewey 2014, 2019, Barðdal 2023, Somers & Barðdal 2023). This work further corroborates the dichotomy between classical Dat-Nom verbs and alternating Dat-Nom/Nom-Dat verbs in Icelandic.

However, what is missing from the literature is a systematic study of how frequently alternating verbs instantiate the Nom-Dat construction and the Dat-Nom construction, respectively, in Icelandic texts. In other words, do all alternating Dat-Nom/Nom-Dat verbs instantiate the two argument structure constructions to the same degree or are the frequencies skewed in favour of one of the argument structure constructions over the other? Further, which factors determine the speakers' choice of one of the two argument structure constructions,

https://doi.org/10.1017/S0332586524000039 Published online by Cambridge University Press

Dat-Nom or Nom-Dat, over the other? One hypothesis is that, other things being equal, the Dat-Nom construction is selected when the dative is topical and that the reverse Nom-Dat construction is selected when the nominative is topical (Barðdal 2001:65; Barðdal, Eythórsson & Dewey 2014, 2019). We return to this point in Sections 4.3 and 5.3 below.

A first attempt at an investigation of this type was carried out by Rott (2013), who has extracted data for eight verbs, i.e. four classical Dat-Nom verbs and four alternating Dat-Nom/Nom-Dat verbs. Rott's study is certainly meritable in that it is the first to lend corpus-based support to the 'alternating predicate puzzle', but it nevertheless suffers from several drawbacks. First, Rott only harvests 50 tokens per verb, and his full dataset only comprises 372 observations. In the present study, however, the number of tokens per verb is increased to 200. Second, Rott also includes clausal nominatives, which are *de facto* considerably longer than nominal arguments, thus being more prone to occurring later in the clause than nominal arguments. In fact, this is exactly what Rott's results show, as 82 out of 87 clausal nominatives occur in postverbal position. This skewness, in turn, greatly inflates the number of Dat-Nom attestations in his sample.

Third, Rott's (2013) study does not specify word order distributions per verb lemma. He thus posits a verb class effect without actually demonstrating that such an effect should exist in the first place. Finally, Rott also does not elaborate on any basic interactions between the argument slots. At least for alternating predicates, he specifies per word order pattern (i.e. Dat-Nom or Nom-Dat) how often each argument is realised as either a full NP, a pronoun, or a clause. However, he fails to disclose how often each of these co-occur with one another, which also makes it difficult to properly assess the scope of his results.

One study that has found homogeneous results for Icelandic alternating verbs, thus corroborating their status as an actual verb class with uniform properties, is that of Bornkessel-Schlesewsky et al. (2011). They have been able to show that alternating verbs consistently trigger a different brain response compared to non-alternating Dat-Nom verbs. However, as was the case for Rott (2013), it is unclear which exact verb types this study is based on, so it is difficult to gauge the scope of these findings. Nevertheless, the uniform electrophysiological response Bornkessel-Schlesewsky et al. have been able to elicit clearly confirms the status of alternating verbs as a syntactically uniform verb class.

The goal of this article is to provide a systematic study of the degree to which the two argument structure constructions are instantiated by alternating verbs in Icelandic. This entails a study which compares nouns with nouns, pronouns with pronouns, and nouns with pronouns. It is also important that both arguments be (pro)nominally realised as opposed to one of the arguments being realised as a clause. Such a study is better designed to control for different factors that may determine the speakers' choice of one argument structure construction over the other.

In the remainder of this article we present a corpus-based study of alternating Dat-Nom/Nom-Dat verbs in Icelandic texts, extracted from the Icelandic Web 2020 corpus (isTenTen20, Jakubíček et al. 2013), which consists of 520 million words. In order to establish a baseline with which our findings for alternating verbs may be compared, we first present results for both ordinary Nom-Dat verbs and

https://doi.org/10.1017/S0332586524000039 Published online by Cambridge University Press

non-alternating Dat-Nom verbs in Icelandic. Our research is based on 15 different verb types, five for each verb class under study. For these, 200 eligible instances are extracted for each lemma, resulting in a total of 3,000 observations. We then proceed to model the data statistically for four out of five alternating verbs, leaving out one outlier. Such an in-depth analysis of the data is crucial in understanding the factors steering the alternation.

This article is organised as follows. In Section 2 we present our object of study, including an overview of the three syntactic verb classes, each selecting for a different argument structure, i.e. the Nom-Dat construction, the Dat-Nom construction, and the alternating Dat-Nom/Nom-Dat constructions. Section 3 gives an overview of the methodology applied, whereas Section 4 presents the results from our study: a baseline for ordinary Nom-Dat verbs and classical Dat-Nom verbs, and the statistics for alternating Dat-Nom/Nom-Dat verbs in relation to these baselines. In Section 5 we single out a set of four alternating verbs, whose behaviour we describe using a logistic regression model. Section 6 summarises the main content and conclusions of the article.

2. Object of study

It is a well-established fact of Icelandic that the subject status of a verbal argument is not necessarily associated with nominative case marking (Andrews 1976, Thráinsson 1979, Zaenen, Maling & Thráinsson 1985, Sigurðsson 1989, Jónsson 1996, Barðdal 2001, inter alia). For these so-called oblique subjects, at least the following nine subjecthood diagnostics have been identified (Andrews 1976, Thráinsson 1979, Zaenen, Maling & Thráinsson 1985, Sigurðsson 1989, Jónsson 1996, Barðdal 2001, 2006, 2023, Barðdal, Eythórsson & Dewey 2019, inter alia):

- first position in declarative clauses
- subject–verb inversion
- first position in subordinate clauses
- subject-to-object raising
- subject-to-subject raising
- long-distance reflexivisation
- clause-bound reflexivisation
- conjunction reduction
- control infinitives

It has been demonstrated that Icelandic oblique subjects pass all of the aforementioned tests, usually referred to in the literature as behavioural tests, as opposed to coding tests (cf. Keenan 1976). Note that the coding test involving subject–verb agreement is not applicable to oblique subjects, as is well known in the literature (Sigurðsson 1990–1991, 2004, inter alia), since agreement is only found with nominative arguments in Icelandic, Germanic, and the Indo-European languages in general (cf. Barðdal 2023:97–98). Moreover, the tests in the bulleted list above confirm the status of oblique subjects as *behavioural* subjects in Icelandic. In this article we intend to lend corpus-based support to the first and the third

https://doi.org/10.1017/S0332586524000039 Published online by Cambridge University Press

behavioural test, i.e. word order distribution in main and subordinate clauses, applying them to Dat-Nom and Dat-Nom/Nom-Dat verbs in Icelandic.

It has already been mentioned above that Dat-Nom verbs come in two different guises: non-alternating Dat-Nom verbs, and alternating Dat-Nom/Nom-Dat verbs. The latter class, which allows for two diametrically opposed case frames, was first discovered by Bernódusson (1982), and it has since been the subject of several studies (Jónsson 1997–1998, Barðdal 1999, 2001, Platzack 1999, Sigurðsson 2006a, Rott 2013, 2016, Barðdal, Eythórsson & Dewey 2014, 2019, Wood & Sigurðsson 2014, Somers & Barðdal 2022, 2023, inter alia). In this article, we refer to them either as 'alternating Dat-Nom/Nom-Dat verbs', or as *nægja*-verbs.

Verbs of the *nægja* type allow both the dative as well as the nominative to take on the role of subject, yet not at the same time. This is manifested in the fact that each of the aforementioned arguments independently passes the subject tests mentioned above, so that, when the dative behaves as the subject, the nominative takes on the role of object, and vice versa (cf. Barðdal 1999, 2001, 2023:Ch. 3, Barðdal, Dewey & Eythórsson 2014, 2019, where it is shown that either argument passes all the subject tests in Icelandic). Examples (1a–b), here repeated as (3a–b), illustrate this phenomenon, in that they show that both arguments may take initial position in declarative clauses without there being a change in meaning or focus.

Dat-Nom

(3) a. Gunnari hafði nægt þessi skýring.
*Gunnar.*DAT *had* *sufficed* *this.*NOM *explanation.*NOM
'Gunnar found this explanation sufficient.'

Nom-Dat

(3) b. Þessi skýring hafði nægt Gunnari.
*this.*NOM *explanation.*NOM *had* *sufficed* *Gunnar.*DAT
'This explanation was sufficient for Gunnar.'

It is clear, however, that the two word orders reflect two different construals of the same event, namely that an experiencer directs his or her attention to a stimulus in (3a) and that a stimulus affects an experiencer in (3b) (cf. Barðdal 2001, 2023:Ch. 3, Barðdal, Eythórsson & Dewey 2019). We take this to be a consequence of the fact that the relevant verbs are force-dynamically neutral in the sense of Talmy (1976), meaning that there is no causal chain found in the event structure of these verbs, as neither of the participants acts upon the other. In other words, there is no causation involved (see also Croft 2012:233, Barðdal 2023:Ch. 3).

Because of their dyadic nature, Barðdal (2001, 2023) and Barðdal, Eythórsson & Dewey (2019) have suggested that alternating verbs of this type in fact instantiate two different argument structure constructions: a Nom-Dat construction that licenses a nominative subject and a dative object, i.e. the stimulus affects the experiencer construal, and a Dat-Nom construction that licenses a dative subject and a nominative object, i.e. the experiencer directs his/her attention towards a stimulus construal. Our approach is fully in line with this analysis, as we subscribe to the view that the subject is the first argument of the argument structure. This is a theory-neutral definition of subject, as all theoretical frameworks employ argument structure or subcategorisation frames in their machinery.

https://doi.org/10.1017/S0332586524000039 Published online by Cambridge University Press

Returning to the examples in (3a–b) above, what speaks against a simple topicalisation analysis is the positioning of the verbal arguments relative to the conjugated verb *hafði* 'had'. In Icelandic the subject must be adjacent to the conjugated verb (unless it is either indefinite or heavy): that is, it must either precede or follow the verb. This is because of the so-called verb-second constraint, which also operates on other Germanic languages (cf. Eythórsson 1995, Axel 2007:27–67, Harbert 2007:398–415, Thráinsson 2007:40–45, inter alia). Had either (3a) or (3b) been a topicalisation of the other, then the nominative in (3a) and the dative in (3b) had been realised in between the conjugated verb *hafði* 'had' and the past participle *nægt* 'sufficed'. This is not the case, though, since both the nominative in (3a) and the dative in (3b) are realised after the non-finite verb, which is an object position.

Consider now the examples in (4a–b) below, which show that attempts at topicalising the object with alternating verbs result in ungrammatical structures in Icelandic (for one exception to this, see (17a–c) below). As already stated above, this involves an inversion of the subject and the verb. Hence, the intended subject argument immediately follows the verb in these examples and the intended object argument occurs in first position.

Intended topicalisation

(4) a. *Gunnari hafði þessi skýring nægt.
Gunnar.DAT had this.NOM explanation.NOM sufficed
Intended meaning: 'To Gunnar, this explanation was sufficient.'

Intended topicalisation

(4) b. *Þessi skýring hafði Gunnari nægt.
this.NOM explanation.NOM had Gunnar.DAT sufficed
Intended meaning: 'This explanation, Gunnar found sufficient.'

Interestingly, not all Dat-Nom verbs allow for the type of alternation shown in (3a–b), as is already mentioned above. Some, such as *líka* 'like', only license dative subjects; their nominative argument invariably behaves as an object with regard to word order distribution. The fact that, for these verbs, subject status is unequivocally associated with the dative case is illustrated by examples (5a–b).

Dat-Nom

(5) a. Barninu hafði líkað bragðið illa.
child.the.DAT had liked taste.the.NOM badly
'The child had not liked the taste.'

*Nom-Dat

(5) b. *Bragðið hafði líkað barninu illa.
taste.the.NOM had liked child.the.DAT badly
Intended meaning: 'The taste had not been to the child's liking.'

Recall that (5b) is ungrammatical because the subject *barninu* 'the child' and the conjugated verb *hafði* 'had' have been separated from one another by the past participle *líkað* 'liked'. If the nominative is realised preverbally for information-structural

https://doi.org/10.1017/S0332586524000039 Published online by Cambridge University Press

reasons, the dative, being the syntactic subject, breaks open the verbal group and is once again reunited with the conjugated verb, as shown in (6).

Topicalisation

(6) Bragðið hafði barninu líkað illa.
taste.the.NOM had child.the.DAT liked badly
'The taste the child had not liked.'

Hence, the example in (6) represents topicalisation and not neutral word order; that is, it is an example of topicalisation that fronts a non-subject constituent to initial position for emphasis (Thráinsson 2007:342). Since the dative subject and the conjugated verb have now been reunited, the example is grammatical. Verbs, like *líka*, that only allow their dative argument to pass the aforementioned subject tests are henceforth called 'non-alternating Dat-Nom verbs', but we will also refer to them as *líka*-verbs in the remainder of this article. Thus, the default argument structure construction that *líka*-verbs instantiate is the Dat-Nom construction. The linear nominative-first order with *líka*-verbs is only used for information-structural purposes (Barðdal, Eythórsson & Dewey 2019). In essence, this means that *líka*-verbs have lexicalised only one of the two construals mentioned above, namely the construal where the experiencer directs his/her attention to a stimulus.[2]

Both alternating Dat-Nom/Nom-Dat verbs, as well as non-alternating Dat-Nom verbs, should be distinguished from ordinary Nom-Dat verbs, or – as we will also be calling them – *hjálpa*-verbs. These are also two-place predicates requiring a nominative and a dative argument, but, crucially, it is the nominative argument that behaves as the syntactic subject and the dative as the object (Barðdal, Eythórsson & Dewey 2019:158), as is evident by the grammaticality of (7a) and the ungrammaticality of (7b).

Nom-Dat

(7) a. Samfélagið verður að hjálpa börnum.
community.the.NOM has to help children.DAT
'The community must help children.'

Intended Dat-Nom

(7) b. *Börnum verður að hjálpa samfélagið.
children.DAT has to help community.the.NOM
Intended meaning: 'Children must get help through the community.'

Topicalisation

(7) c. Börnum verður samfélagið að hjálpa.
children.DAT has community.the.NOM to help
'Children, the community must help.'

Thus, *hjálpa*-verbs constitute the mirror counterpart of the aforementioned *líka*-verbs, in that they exclusively occur in the Nom-Dat argument structure construction, which is the opposite of the Dat-Nom argument structure construction. Also, *hjálpa*-verbs only allow for preposed datives in cases where the dative is topicalised, as is shown in (7c).

https://doi.org/10.1017/S0332586524000039 Published online by Cambridge University Press

In this study we lend corpus-based statistical support to the analysis that the dative and the nominative arguments of *nægja*-verbs are indeed syntactic subjects. This we do by comparing the frequency of topicalised arguments in first position to the frequency of subjects in first position. In other words, if an oblique argument behaves as a subject, it can be expected to be strongly associated with first position in declarative clauses (diagnostic test 1) and first position in subordinate clauses (diagnostic test 3), while topicalised objects would not show the same association. Moreover, since word order in Icelandic is understood to be quite rigid (Thráinsson 2007:342), topicalisation can be expected to be relatively rare, and even less common in subordinate clauses than in main clauses. This is confirmed by Angantýsson's (2020:261) study, although his study is based on acceptability judgements and not corpus frequencies. Nevertheless, empirical studies on how frequent topicalisation actually is are quite scarce for Icelandic.

One study that does include frequency counts is that of Callegari & Ingason (2021). In their diachronic investigation of matrix-clause ditransitive constructions, they explore object topicalisation in Icelandic texts from the twelfth to twenty-first centuries, drawing their data from the IcePaHC corpus (Wallenberg et al. 2011). Callegari & Ingason include both pronominal and nominal objects in their study, i.e. objects realised as both pronouns and full NPs. Out of a total of 1,100 hits, they find 128 instances of object topicalisation, of which 89 have the direct object topicalised (8%), and 39 the indirect object (3.5%). Thus, topicalisation affects approximately 11.5% of the tokens under study, and direct object topicalisation turns out to be more than twice as common as indirect object topicalisation. Callegari & Ingason do not include an unambiguous overview of object topicalisation per century, but a summary graph seems to reveal that, for the twenty-first-century data, both direct objects as well as indirect objects are each topicalised approximately 6% of the time.

Another study worth mentioning in this respect is that of Barðdal & Eythórsson (2012), who map out the word order patterns for monotransitive verbs licensing nominative subjects. Barðdal & Eythórsson's data have also been drawn from the IcePaHC corpus, although an earlier version than that of Callegari & Ingason. The main difference between the two versions is the size of the corpus, with the more recent version containing nearly 60% more texts (1,002,390 vs. 632,000, Einar Freyr Sigurðsson p.c.). With all else being equal, and on the assumption that the smaller version of the corpus is large enough, there is no reason to assume that a comparison involving the frequency of topicalisations between these two studies is not justified. Thus, zooming in on Barðdal & Eythórsson's (2012) results for verb-second clauses (i.e. SVO vs. OVS structures), it turns out that nominative subjects occur 2,327 times (or 80%) in initial position, and 578 times (or 20%) in postverbal position.[3] Therefore, object topicalisation is clearly much more frequent with monotransitives than with ditransitives, at least diachronically. See also Booth & Beck's (2021) study where it is statistically documented that the clausal initial position in Modern Icelandic is a topic position.

It is unclear if the predicates in our study are equally permissive of topicalisation as Barðdal & Eythórsson's (2012) monotransitives and Callegari & Ingason's (2021) ditransitives. For that reason, we map out word order preferences for both the *hjálpa* and *líka* classes and use these counts as a baseline against which word order

https://doi.org/10.1017/S0332586524000039 Published online by Cambridge University Press

preferences for the *nægja* class will be measured. We now turn to a description of our methodology, before we present our findings in Sections 4–5 below.

3. Methodology

This study is based on 15 simple verbs that fall into one of three categories: (i) ordinary Nom-Dat verbs (the *hjálpa* type), (ii) non-alternating Dat-Nom verbs (the *líka* type), and (iii) alternating Dat-Nom/Nom-Dat verbs (the *nægja* type). Our aim was to follow Rott (2013) in our selection of verbs, but some of the verbs he used were too infrequent in the corpus to yield enough eligible tokens. Thus, we complemented the dataset with additional known non-alternating Dat-Nom and alternating Dat-Nom/Nom-Dat verbs (cf. Jónsson 1997–1998, Barðdal 1999:89, 2001:53–58, 2023:81–83). Each category contains five verbs:

(i) Ordinary Nom-Dat verbs: *hjálpa* 'help', *líkjast* 'resemble', *mótmæla* 'contradict', *treysta* 'trust', and *þakka* 'thank'.
(ii) Non-alternating Dat-Nom verbs: *áskotnast* 'receive', *blöskra* 'be shocked, be horrified', *leiðast* 'be bored', *líka* 'like', and *þykja gott/slæmt/* . . . 'think, find, seem good/bad/ . . . '.
(iii) Alternating Dat-Nom/Nom-Dat verbs: *duga* 'suffice, be enough', *dyljast* 'be hidden to somebody, be aware', *endast* 'last', *henta* 'suit, befit', and *nægja* 'be enough, be sufficient'.

We follow Rott (2013:103) in using *blöskra* 'be shocked, be horrified', *leiðast* 'be bored' and *líka* 'like' in the class of non-alternating Dat-Nom verbs, and *henta* 'suit, befit' and *dyljast* 'be hidden to somebody, be aware' in the alternating Dat-Nom/Nom-Dat class.

The analysis is based on a data collection from the Icelandic Web 2020 corpus (isTenTen20, Jakubíček et al. 2013), which consists of approximately 520 million words. The corpus itself has been accessed through the Sketch Engine interface. For each of the aforementioned verbs, a lemmatised search query has been carried out targeting the verb's bare infinitival form. That is also true for the etymologically reflexive *-st*-verbs, as the search engine considers *-st*-forms to be instantiations of the non-suffigated base form. Thus, *líkjast*, *áskotnast*, *leiðast*, *dyljast*, and *endast* were run as *líkja*, *áskotna*, *leiða*, *dylja*, and *enda*, respectively.

The material has subsequently been extracted in one or more files containing 10,000 randomised tokens per verb type, depending on how abundant the data were. In contrast to Rott, who also includes middle field tokens, we only focus on tokens in which the main verb is flanked by either a nominal or a pronominal element. Thus, only instances of the type [Nom-V-Dat] or [Dat-V-Nom] have been taken into account, regardless of clause type. As a consequence, there are no tokens in our dataset of any other kinds of topicalised elements, which in turn excludes, for instance, adverbials.

Contrary to the Mainland Scandinavian languages, Icelandic is a so-called symmetric V2-language, which means that the conjugated verb takes second position both in main clauses as well as in subordinate clauses (Thráinsson 2007:41,

https://doi.org/10.1017/S0332586524000039 Published online by Cambridge University Press

Angantýsson 2020:243). Eligible tokens are therefore not restricted to main clauses only but also include subordinate structures. Per verb type, the first 200 tokens have been withheld for study. Hence, the total number of collected tokens equals 3,000, and the number of collected tokens per verb class equals 1,000.

Per token, all arguments, dative and nominative, have been manually annotated for the following variables: case, (pro)nominality, pronoun type (if applicable), referentiality, person, number, definiteness, animacy, and length. The choice of variables is motivated by two considerations, namely that each of these (a) are well known in the field for affecting word order, (b) may serve as a proxy for discourse prominence or topicality (cf. the discussion in Sections 4.3 and 5.3 below). For each variable in boldface, the relevant values are rendered in small caps, followed by examples in brackets. The nine variables are discussed below.

Case: NOMINATIVE (*þessi sími* 'this phone' NOM.SG, *mín eigin föt* 'my own clothes' NOM.PL) or DATIVE (*hundinum* 'the dog' DAT.SG, *unglingunum* 'the youngsters' DAT.PL).

(Pro)nominality: PRONOUN (*þú* 'you' SG, *ykkur* 'you' PL, *einhverjum* 'some') or full NP (*Ísland* 'Iceland', *ýmsir þingmenn* 'some congressmen' NOM.PL, *bókin* 'the book' NOM.SG).

Pronoun type: PERSONAL (*ég* 'I', *hann* 'he', *þeir* 'they' 3.M), DEMONSTRATIVE (*þessi* 'this', *hinum* 'the other', *slíkur* 'such'), INDEFINITE (*öllum* 'all', *engum* 'no-one', *báðum* 'both'), or RECIPROCAL (*hvert öðru* 'each other' SG.N, *hver annarri* 'each other' SG.F). Reflexives are excluded from study, as they are hypothesised to prefer the postverbal slot. In line with Heylen (2005:103), conjoined pronouns are also excluded, as they arguably lose their pronominal status.

Referentiality: REFERENTIAL or CORRELATIVE. Icelandic allows for the third person personal pronoun *það* 'it' to be used as an expletive or a correlate. Expletives are wholly absent from our dataset, but correlates, which have a clause-anticipating function, show up approximately 300 times. It is hypothesised that such placeholders, given their impoverished semantic status, are inclined to follow the verb, rather than precede it. All remaining arguments, either nominal or pronominal, have been annotated as referential.

Person: FIRST person (*mér* 'me' DAT.SG, *við* 'we', *okkur feðgum* 'us, father and son'), SECOND person (*þú* 'you' SG, *ykkur* 'you' PL, *yður* 'you' PL.HON), or THIRD person (*þeim* 'them' DAT.PL, *henni* 'her' DAT.SG.F, *augum manna* 'people's eyes' DAT.PL).

Number: SINGULAR (*rigningarvatn* 'rainwater' NOM.SG, *stúlkunni* 'the girl' DAT.SG) or PLURAL (*stúlkum* 'girls' DAT.PL, *Hauknum og frú* 'Haukur and his wife' DAT.PL).

Definiteness: DEFINITE or INDEFINITE. Icelandic pronouns are always definite, except for indefinite pronouns (*báðar* 'both' PL.F, *manni* 'one' DAT.SG) and indefinite demonstratives (*slíkt* 'such a thing' SG.N). NPs are considered to be definite if they are preceded by a definite demonstrative pronoun (*þessi hey* 'this hay' NOM.PL) or a possessive pronoun (*þitt fyrirtæki* 'your company' NOM.SG). Constituents that are followed by a cliticised definite article (*blaðinu* 'the paper' DAT.SG), a possessive pronoun (*hlátur hans* 'his laughter' NOM.SG), or a postposed genitive[4] (*stjórn félagsins* 'the company board' NOM.SG) also receive a definite reading. Names of people (*Þorsteini* DAT.SG), institutions (*Fjölmiðlanefnd* 'the Media Committee' NOM.SG),

https://doi.org/10.1017/S0332586524000039 Published online by Cambridge University Press

places (*Keflavík* NOM.SG), and population groups (*Reykvíkingum* 'the people of Reykjavík' DAT.PL) are inherently definite. If a conjoined constituent exhibits conflicting definiteness, in that one conjunct is definite but the other is indefinite, the string is coded for the first conjunct. Thus, the string *4x4 klúbburinn og allir sem elska hálendið í sinni hráustu mynd* 'the 4x4 club and all who love the highlands in their rawest form' is coded as definite, because of the definite status of the first conjunct (*4x4 klúbburinn* 'the 4x4 club').

Animacy: INDIVIDUAL, COLLECTIVE, INANIMATE, NON-INFERABLE or NA. The label 'individual' is used to index constituents referring to humans (*bræðurnir* 'the brothers' NOM.PL), animals (*fuglum* 'birds' DAT.PL), and what Bresnan & Ford (2010:175) call 'humanoid beings' (*Guð* 'God' NOM.SG, *heilagur andi* 'the Holy Spirit' NOM.SG). This includes cells (*krabbameinsfrumur* 'cancer cells' NOM.PL). Groups of individuals are annotated as collective (*fólkið* 'the people' NOM.SG, *Landsbankanum* 'National Bank' DAT.SG). All other constituents, including plants (*þessi jurt* 'this plant' NOM.SG), fungi[5] (*myglusveppum* 'mould' DAT.PL), and dead animals, are labelled 'inanimate'. When animacy cannot unequivocally be determined, we have resorted to the label 'non-inferable'. This, for instance, applies to *lántakanda* 'borrower' DAT.SG, which, in the given context, can both refer to an individual as well as to a corporate borrower. Pronouns that serve as placeholders for a subclause are annotated as 'NA', since their referent is linguistic, and not extra-linguistic. If a conjoined constituent exhibits conflicting animacy, in that one conjunct complies with one label but the other complies with another label, the string is coded for the first conjunct. Thus, the string *bæði einstaklingum og fyrirtækjum* 'both individuals as well as companies' is coded as 'individual', because that is the label that captures the animacy status of the first conjunct (*bæði einstaklingum* 'both individuals').

Length: constituent weight measured in WORDS.

We now turn to our findings and a discussion thereof.

4. Results and discussion

The current section details the results of our study. Section 4.1 establishes a baseline by mapping out word order preferences for both ordinary Nom-Dat verbs, i.e. verbs of the *hjálpa* type, and non-alternating Dat-Nom verbs, or *líka*-verbs. In Section 4.2 we compare the statistics for alternating *nægja*-verbs with the baseline established for *hjálpa*- and *líka*-verbs in Icelandic. Section 4.3 discusses the main implications and conclusions.

4.1 Establishing a baseline: hjálpa- *and* líka-*verbs*

In this section we discuss our findings for both *hjálpa*- and *líka*-verbs. We single out two configurations, i.e. contexts in which both arguments are full NPs (Section 4.1.1) and contexts in which both arguments are pronouns (Section 4.1.2). We leave out a comparison of contexts involving only one pronoun, since Somers & Barðdal (2022:91–92, 95–97) have shown that those frequencies exhibit the same tendencies as documented below. We summarise our conclusions in Section 4.1.3.

https://doi.org/10.1017/S0332586524000039 Published online by Cambridge University Press

Table 1. Nom-Dat verbs and Dat-Nom verbs in the [NP-V-NP] configuration

	Nom-Dat		Dat-Nom			Nom-Dat		Dat-Nom	
Verb	*N*	*f*	*N*	*f*	Verb	*N*	*f*	*N*	*f*
hjálpa	25	100%	0	0%	*áskotnast*	0	0%	48	100%
líkjast	125	100%	0	0%	*blöskra*	0	0%	68	100%
mótmæla	98	98%	2	2%	*leiðast*	0	0%	26	100%
treysta	31	100%	0	0%	*líka*	0	0%	28	100%
þakka	55	100%	0	0%	*þykja*	1	4%	23	96%
Total	334	99%	2	1%	Total	1	1%	193	99%

4.1.1 Word order variation in the [NP-V-NP] configuration

Table 1 presents an overview of the word order distributions for both *hjálpa-* and *líka*-verbs in clauses where both arguments are full NPs. Starting with *hjálpa*-verbs, the general rule is that the dative is realised postverbally: as many as 334 tokens (or 99%) across verbs instantiate the Nom-Dat word order, as opposed to a mere two (or 1%) instantiating the reverse Dat-Nom order. Two examples of the unmarked nominative-before-dative pattern, one with *hjálpa* 'help' and one with *líkjast* 'resemble', are given in (8a–b).

Nom-Dat

(8) a. Listamaður hjálpaði börnum að ...
*artist.*NOM *helped* *children.*DAT *to*
'An artist helped children to ...'

Nom-Dat

(8) b. Einkenni líkjast helst inflúensusýkingu ...
*symptoms.*NOM *resemble* *most* *influenza.infection.*DAT
'Symptoms mostly resemble an influenza infection ...'

The sole *hjálpa*-verb which (marginally) allows datives in initial position is *mótmæla* 'contradict' with the two tokens shown in (9a–b).

Topicalisation

(9) a. Þessari frásögn mótmælti annar sjónarvottur ...
*this.*DAT *narration.*DAT *objected* *another.*NOM *eye.witness.*NOM
'To this narration, another eyewitness objected ...'

Topicalisation

(9) b. Þeirri fyrirhuguðu málsmeðferð mótmæltu ýmsir þingmenn ...
*this.*DAT *intended.*DAT *procedure.*DAT *opposed* *some.*NOM *parliamentarians.*NOM
'This intended procedure, some parliamentarians objected to ...'

Both of these are topicalisations, with the dative occurring in initial position for information-structural purposes. Both tokens also display a discrepancy in definiteness, in that the fronted dative is definite, whereas the postposed nominative

https://doi.org/10.1017/S0332586524000039 Published online by Cambridge University Press

is indefinite. Since definites tend to precede indefinites, this asymmetry is undoubtedly conducive to an inversion of the canonical order of constituents (cf. Siewierska 1993, Lambrecht 1994, 2000, Gregory & Michaelis 2001, inter alia).

Moving on to our findings for *líka*-verbs, it is striking that the acquired figures constitute the mirror image of those obtained for *hjálpa*-verbs, as 193 clauses (or 99%) assign the preverbal slot to the dative. This corroborates the existing analysis of these as being non-alternating Dat-Nom verbs. Two examples of non-alternating Dat-Nom verbs occurring in their neutral Dat-Nom order are presented in (10a–b).

Dat-Nom

(10) a. Fólki líkar laktósafría mjólkin vel.
*people.*DAT *likes* *lactose.free.*NOM *milk.the.*NOM *well*
'People like the lactose-free milk.'

Dat-Nom

(10) b. Mömmunni leiddist enski boltinn.
*mother.the.*DAT *was.bored.by* *English.*NOM *ball.the.*NOM
'The mother disliked English football.'

Only *þykja* returns one token in which the canonical order of constituents is inverted. This example is shown in (11), where the nominative is topicalised to first position, while the dative subject inverts with the finite verb.

Topicalisation

(11) Þetta hey þótti kúnum gott ...
*this.*NOM *hay.*NOM *found* *cows.the.*DAT *good*
'This hay, the cows liked ...'

Observe that, across all ten verbs presented in Table 1, nominal frequencies are generally very high: there are never fewer than 24 attestations per verb, and their total number across all ten verbs amounts to 530. Thus, our findings for both *hjálpa-* and *líka*-verbs in the double-NP configuration can be considered to be very robust.

Finally, both *hjálpa-* and *líka*-verbs show not only a strong verb effect in the [NP-V-NP] configuration but also a robust verb *class* effect, since all verbs prefer either the Nom-Dat or the Dat-Nom order in equal manner.

4.1.2 Word order variation in the [Pro-V-Pro] configuration

Table 2 summarises the results for *hjálpa-* and *líka*-verbs in the [Pro-V-Pro] configuration. For *hjálpa*-verbs, word order preferences in the [Pro-V-Pro] configuration constitute a near-perfect copy of the results presented in Table 1 above. With the exception of *mótmæla*, all *hjálpa*-verbs tend entirely towards the Nom-Dat linear order. Interestingly, the only two attestations of the topicalised Dat-Nom linear order contain a dative demonstrative pronoun in combination with the nominative personal pronoun *ég* 'I'. Both of these examples are given in (12a–b).

https://doi.org/10.1017/S0332586524000039 Published online by Cambridge University Press

Table 2. Nom-Dat verbs and Dat-Nom verbs in the [Pro-V-Pro] configuration

	Nom-Dat		Dat-Nom			Nom-Dat		Dat-Nom	
Verb	*N*	*f*	*N*	*f*	Verb	*N*	*f*	*N*	*f*
hjálpa	68	100%	0	0%	*áskotnast*	1	33%	2	67%
líkjast	6	100%	0	0%	*blöskra*	0	0%	11	100%
mótmæla	25	93%	2	7%	*leiðast*	5	11%	39	89%
treysta	83	100%	0	0%	*líka*	6	8%	71	92%
þakka	56	100%	0	0%	*þykja*	32	35%	60	65%
Total	238	99%	2	1%	Total	44	19%	183	81%

Topicalisation

(12) a. Þessu verð ég að mótmæla og ...
this.DAT must I.NOM to object and
'To this I must oppose and ...'

Topicalisation

(12) b. En hitt, að þetta hafi verið gjört í fullkomnu óþakklæti skólastjóra,
but the.other that this had been done in perfect ingratitude headmaster's
því mótmæli ég algjörlega.
that.DAT oppose I.NOM entirely
'But the other [option], that this was done in a total ingratitude of the headmaster, to that I strongly object.'

In (12a) it is the dative demonstrative *þessu* 'this' which occurs in clause-initial position while the nominative *ég* 'I' inverts with the verb. In (12b) a similar pattern surfaces, this time with the topicalised dative demonstrative *því* 'that' in first position.

Verbs of the *líka* type show considerably more word order variation in the double-pronoun configuration than *hjálpa*-verbs: the Dat-Nom linear order is attested 183 times (or 81%) and the Nom-Dat order 44 times (or 19%). An example of each pattern is provided in (13a) and (13b) respectively.

Dat-Nom

(13) a. Mér þótti hún svo fyndin.
I.DAT thought she.NOM so funny
'I thought she was so funny.'

Topicalisation

(13) b. Það þótti henni ógeðslegt ...
that.NOM thought she.DAT disgusting
'That, she found disgusting ...'

Remarkably, the Nom-Dat pattern for *líka*-verbs is almost uniquely associated with nominative demonstratives: 40 out of 44 tokens occurring with the Nom-Dat linear order are headed by the pronouns *það* 'that' or *þetta* 'this'. This finding is reminiscent of the tendency discussed above for the verb *mótmæla* 'contradict',

https://doi.org/10.1017/S0332586524000039 Published online by Cambridge University Press

which is marginally found in the Dat-Nom linear order, yet only when the dative object is a demonstrative pronoun.

Given the fact that demonstratives convey highly topical information, it is clear that topicality, especially in combination with effects of definiteness and pronominality, may cause changes in the linear order from the neutral Dat-Nom to the topicalised Nom-Dat order. However, the extent to which the word order of different argument structures can be inverted also seems to be dependent on the verb itself. For a more detailed discussion of the effect of nominative demonstratives, see Somers & Barðdal (2022:98–99).

4.1.3 Interim conclusions

The evidence presented in this section is fully in line with the prediction that Icelandic possesses both a class of Nom-Dat verbs as well as a class of Dat-Nom verbs. The former, also referred to as *hjálpa*-verbs, are associated with a nominative-before-dative order, whereas the latter, or *líka*-verbs, instantiate the reverse dative-before-nominative order. Our findings essentially confirm that subjects in Icelandic, regardless of case marking, are very strongly inclined to occupy the preverbal slot (cf. Andrews 1982:428, Sigurðsson 1989:205–206, Jónsson 1996:115, Thráinsson 2007:21, Schätzle 2018, inter alia).

What is especially informative about our results for the [NP-V-NP] configuration, is that Dat-Nom verbs occur with the Dat-Nom linear order to the same degree as ordinary Nom-Dat verbs of the *hjálpa* 'help' type occur with the Nom-Dat linear order. That is, both verb classes realise their syntactic subjects in clause-initial position 99.5% of the time, the nominative for Nom-Dat verbs and the dative for Dat-Nom verbs.

The overwhelming preference of *líka*-verbs for dative-first structures refutes the claim made by Roehm et al. (2007) that non-alternating Dat-Nom verbs in Icelandic are a category in flux, in that they have started adopting the behaviour of alternating Dat-Nom/Nom-Dat verbs. Roehm et al.'s conclusion is based both on an acceptability judgement task as well as on ERP data, but it is unclear exactly which verbs they included in their study. In all likelihood, the situation was exactly the opposite, with Dat-Nom verbs being derived from alternating Dat-Nom/Nom-Dat verbs, through the loss of the Nom-Dat alternant (cf. Barðdal 2023:133–137).

Finally, our data show that topicalisation of this type is very rare in Icelandic. The only verbs found with object topicalisation in the double-NP configuration are *mótmæla* (two tokens) and *þykja* (one token). As for clauses with double pronouns, topicalisation is markedly more frequent with Dat-Nom verbs (44 out of 227 tokens, or 19%) than with Nom-Dat verbs (two out of 240 tokens, or 1%). However, it turns out that almost all fronted nominatives with *líka*-verbs are nominative demonstratives.

4.2 Alternating Dat-Nom/Nom-Dat verbs

In this section we present our findings for the class of alternating Dat-Nom/Nom-Dat verbs, also referred to here as *nægja*-verbs. The organisation of this subsection is as follows: we first discuss the general findings, i.e. the results across all

https://doi.org/10.1017/S0332586524000039 Published online by Cambridge University Press

Table 3. Alternating verbs across configurations

	Nom-Dat		Dat-Nom	
Verb	*N*	*f*	*N*	*f*
duga	180	90%	20	10%
dyljast	150	75%	50	25%
endast	78	39%	122	61%
henta	200	100%	0	0%
nægja	139	69.5%	61	30.5%
Total	747	75%	253	25%

four configurations (Section 4.2.1), after which we turn to word order variation in the [NP-V-NP] configuration (Section 4.2.2), the [Pro-V-Pro] configuration (Section 4.2.3), and finally we discuss configurations where one of the arguments is a pronoun (Section 4.2.4). The results are compared to the baseline set by Nom-Dat *hjálpa*-verbs and Dat-Nom *líka*-verbs.

4.2.1 General findings

The results for the class of *nægja*-verbs, which are presented in Table 3, generally confirm the alternating nature of these predicates: in total, the Nom-Dat linear order is attested 747 times, i.e. ca. 75%, and the Dat-Nom linear order 253 times, i.e. approximately 25% of the time on average across all five predicates. The alternating nature of *nægja*-verbs is also supported by either argument passing the subject tests, see Section 2.

Upon closer inspection, the data in Table 3 reveal three remarkable tendencies. First, the Nom-Dat linear order is generally more common than the Dat-Nom linear order. Secondly, there are notable differences between verbs, in that some seem to allow for word order alternation more readily than others. And, thirdly, it is also remarkable that *henta*, a verb discussed by Barðdal (1999, 2001) as a prime member of the class of alternating verbs, does not yield a single Dat-Nom token.

Our results are generally also less evenly distributed than the ones Rott (2013) documents. He gathered corpus frequencies for the alternating predicates *dyljast* 'be hidden', *henta* 'suit, befit', *veitast* 'find (hard/easy)', and *þóknast* 'satisfy, please', and found that these verbs instantiate the Nom-Dat linear order 76 times, i.e. 51%, and the Dat-Nom linear order 72 times, i.e. 49%. Interestingly, the verb *henta* is included in Rott's dataset, but it is unclear what its frequency distribution is, as he does not display any frequency counts for individual verbs. And, as is already stated in Section 1 above, Rott also includes clausal arguments in his investigation, which makes it even more difficult to compare his findings with ours.

The results most similar to the ones we have obtained here are probably the ones attained by Roehm et al. (2007). Their acceptability judgement task reveals that alternating verbs can be used equally felicitously in both case frames, but participants seemed to prefer the nominative-first structure. In their subsequent

https://doi.org/10.1017/S0332586524000039 Published online by Cambridge University Press

Table 4. Alternating verbs in the [NP-V-NP] configuration

	Nom-Dat		Dat-Nom	
Verb	*N*	*f*	*N*	*f*
duga	33	79%	9	21%
dyljast	2	25%	6	75%
endast	9	29%	22	71%
henta	86	100%	0	0%
nægja	27	54%	23	46%
Total	157	72%	60	28%

ERP-study, alternating verbs even elicited a violation response in the dative-before-nominative configuration, but since it is not made explicit which verbs Roehm et al. actually studied, that claim cannot be verified. In any case, it seems rather unexpected that all alternating verbs should elicit the same response, as the within-class variation is quite substantial, as we document here.

4.2.2 Word order variation in the [NP-V-NP] configuration

In total, alternating verbs are attested 217 times in the [NP-V-NP] configuration; 157 tokens (72%) instantiate the Nom-Dat linear order, and 60 tokens (28%) the Dat-Nom linear order. A more detailed overview of the frequencies per verb can be found in Table 4.

The frequencies in Table 4 are indicative of several different tendencies. First, frequencies in the [NP-V-NP] configuration are much less skewed than for ordinary Nom-Dat verbs or non-alternating Dat-Nom verbs, thereby confirming the generally alternating nature of Dat-Nom/Nom-Dat verbs. A chi-square goodness-of-fit test comparing the two word orders attested with *nægja*-verbs across all five verbs yields a highly significant result with a large effect size ($\chi^2 = 80.14$; df $= 4$; $p_{\text{two-tailed}} < .001$; Cramér's V $= .61$), which should be interpreted as a statistical indication that the distribution of the two word orders cannot be attributed to chance. A more in-depth analysis of the factors driving the alternation is presented in Section 5.

One of these factors, it seems, is verb type: with the exception of *henta*, all verbs are attested at least 21% of the time in either the Dat-Nom or the Nom-Dat linear order, but the degree to which they do so is verb-dependent. The verb *duga*, for instance, is clearly more permissive of clause-initial nominatives, whereas the opposite is true of *dyljast* and *endast*. The verb *nægja* is the most evenly balanced type, favouring a dative-first structure about as often as a nominative-first structure. One example of each word order is given in (14a–b).

Dat-Nom

(14) a. … að Víkingum myndi nægja jafntefli til að …
that Vikings.DAT would suffice tie.NOM in.order to
'… that the Viking team would make do with a tie in order to …'

https://doi.org/10.1017/S0332586524000039 Published online by Cambridge University Press

Nom-Dat

(14) b. En skotfærasafnið hans hefði nægt hverri meðal herdeild.
but munition.collection.the.NOM his had sufficed every.DAT average division
'And his munition collection had been sufficient for every average division.'

Turning to *henta*, the generally skewed frequencies for that verb presented in Table 3 are evidently replicated in the [NP-V-NP] configuration in Table 4, and since nominal frequencies for this verb are very high (86 tokens), its tendency towards the Nom-Dat linear order can be taken to be very robust, which makes this result all the more enticing. Recall that previous research has confirmed *henta*'s status as an alternating verb, as both the nominative as well as the dative independently pass the subjecthood tests presented in Section 2, as is documented by Barðdal (1999, 2001). Clearly, further research is needed to better understand *henta*'s behaviour as an outlier with respect to the word order test.

Also, it is striking how frequencies in the [NP-V-NP] configuration differ from the general frequencies presented in Table 3. For some verbs, like *duga* and *nægja*, the alternation is less skewed in the [NP-V-NP] configuration than it is in general, since the proportional frequencies move closer towards a 50–50 distribution. Other verbs, like *dyljast* and *endast*, tend more towards the Dat-Nom linear order in the [NP-V-NP] configuration.

Finally, our findings for alternating verbs in the [NP-V-NP] configuration tie in nicely with Allen's (1995:108) study of Old English Dat-Nom verbs. Allen (1995) shows that the [NP-V-NP] configuration displays a symmetric distribution between the Nom-Dat linear order and the Dat-Nom linear order (21 vs. 19 attestations). This certainly confirms Allen's (1995:116) claim that her Dat-Nom verbs are indeed alternating verbs in Old English, precisely like *nægja*-verbs in the present study. Unfortunately, exactly like Rott (2013), Allen does not specify how each individual verb weighs in on the alleged verb class effect, so (i) it is unclear whether all verbs in her sample can actually be regarded as alternating, and (ii) if they do, whether they are all equally attracted to both argument structure constructions.

4.2.3 Word order variation in the [Pro-V-Pro] configuration

Table 5 shows that in the [Pro-V-Pro] configuration alternating predicates almost invariably occur in the Nom-Dat linear order: out of 337 attestations, only 19, i.e. 6%, contain a dative in clause-initial position.

Some examples of Dat-Nom word orders involving pronouns are given in (15a–c), while examples of the more abundant Nom-Dat word order are given in (16a–c).

Dat-Nom

(15) a. . . . að honum hafi dulist neitt af þessu.
that he.DAT has been.aware none.NOM of this
'. . . that he was not aware of any of this.'

Dat-Nom

(15) b. Henni duldist það ekki að . . .
she.DAT been.aware it.NOM not that
'She was not aware that . . .'

https://doi.org/10.1017/S0332586524000039 Published online by Cambridge University Press

Dat-Nom

(15) c. Og honum entist hún yfir daginn.
and he.DAT lasted she.NOM over day.the
'And he got her to last throughout the whole day.'

Nom-Dat

(16) a. . . . og það hafi nægt honum.
and it.NOM would.have sufficed he.DAT
'. . . and it would have been enough for him.'

Nom-Dat

(16) b. það myndi ekki duga okkur samt.
it.NOM would not be.enough we.DAT anyway
'Yet, it would still not be enough for us.'

Nom-Dat

(16) c. . . . að þeir mundu endast okkur.
that they.NOM would last we.DAT
'. . . that they would last us.'

Table 5 also shows that the Nom-Dat linear order is not disproportionately associated with any one verb in particular, as frequencies are consistently higher than, or equal to, 92% per verb. In other words, these numbers clearly point towards an overarching verb class effect and not towards individual verb effects.

The findings for the [Pro-V-Pro] configuration also explain at least part of the skewness for alternating predicates in general, as the [Pro-V-Pro] configuration is not only heavily biased towards the Nom-Dat construction but is also very frequent in general, since it accounts for about one-third of all the data collected for *nægja*-verbs (318 tokens out of 1,000).

Given the skewed frequencies in the [Pro-V-Pro] configuration, it should not come as a surprise that tokens containing two personal pronouns show an equal bias: 81 out of 88, or 92%, instantiate the Nom-Dat order (not singled out in Table 5). These findings again mirror Allen's (1995:109) results for 12 Old English alternating verbs, which, in the double personal pronoun configuration, also show a clear tendency towards the Nom-Dat order. This sets them apart from non-alternating *líka*-verbs in configurations with two personal pronouns, as these overwhelmingly tend towards the Dat-Nom order (63 tokens, or 95%) and not to the reverse Nom-Dat order (three tokens, or 5%).

This pronominal skewness with alternating verbs raises the question of whether occurrences with pronouns are perhaps unevenly distributed across the three verb classes in terms of frequency and whether that may possibly explain the high proportion of the Nom-Dat construction here. However, out of 3,000 observations in total for all 15 verbs (1,000 for each verb class) there are 664 Nom-Dat observations, 806 Dat-Nom observations, and 783 alternating Dat-Nom/Nom-Dat observations including at least one pronoun. This shows that alternating verbs are not particularly more frequent with pronouns in general, even though they yield most tokens in the [Pro-V-Pro] configuration (337 for alternating verbs, 227 for classical Dat-Nom verbs, and 240 for ordinary Nom-Dat verbs).

https://doi.org/10.1017/S0332586524000039 Published online by Cambridge University Press

Table 5. Alternating verbs in the [Pro-V-Pro] configuration

	Nom-Dat		Dat-Nom	
Verb	*N*	*f*	*N*	*f*
duga	72	97%	2	3%
dyljast	117	92%	10	8%
endast	30	94%	2	6%
henta	39	100%	0	0%
nægja	60	92%	5	8%
Total	318	94%	19	6%

4.2.4 Configurations with one pronoun and one full NP

The current section zooms in on the two remaining configurations, i.e. contexts containing a nominative pronoun and a dative full NP and contexts with a dative pronoun and a nominative full NP. The results for the former are laid out in Table 6, which shows that *nægja*-verbs are strongly skewed towards the Nom-Dat order when the nominative is pronominal and the dative is a full NP: as many as 73 tokens across verbs (or 92%) allocate the preverbal slot to the nominative. The remaining six observations (or 8%) prefer the Dat-Nom order. This nominative-before-dative skewness is not an idiosyncrasy of individual verbs: it is a commonality of all *nægja*-verbs, thus pointing towards a verb class effect. The results presented in Table 6 are highly reminiscent of our findings for alternating verbs in the double-pronoun configuration (cf. Section 4.2.3 above). Recall that we found alternating verbs to instantiate the Nom-Dat order 94% of the time when both arguments were pronouns.

Our findings for the current configuration also raise the question of how non-alternating *líka*-verbs fare in contexts with a nominative pronoun and a dative full NP, as one might perhaps assume that the skewness found in Table 6 is a general effect of pronouns, not specific to *nægja*-verbs. It turns out that, across verbs, *líka*-verbs would much rather have the dative NP precede the nominative pronoun (47 tokens, or 82%) than the other way around (ten tokens, or 18%) (not shown in any table here). Exactly like in the [Pro-V-Pro] configuration, the bulk of Nom-Dat attestations is due to the effect of nominative demonstratives (eight out of ten, or 80%). Once again, this comparison shows that *líka*-verbs are quite distinct in behaviour from *nægja*-verbs: the former strongly adhere to the dative-before-nominative order irrespective of lexical specifications, while the latter are much more susceptible to pronominal influence. Thus, the behaviour of nominative pronouns, to be in first position with *nægja*-verbs, is not due to a general property of pronouns but represents a fact, specific to *nægja*-verbs.

Let us now explore the results of the most widely attested configuration for alternating verbs, i.e. one in which a dative pronoun enters into competition with a nominative full NP. These numbers are presented in Table 7. In configurations with dative pronouns and nominative NPs, the results show a relatively even distribution across the Dat-Nom and Nom-Dat word order patterns: the former occurs 46% of

https://doi.org/10.1017/S0332586524000039 Published online by Cambridge University Press

Table 6. Alternating verbs in configurations with a nominative pronoun and a dative full NP

	Nom_{PRO}-Dat_{NP}		Dat_{NP}-Nom_{PRO}	
Verb	*N*	*f*	*N*	*f*
duga	20	95%	1	5%
dyljast	16	84%	3	16%
endast	2	100%	0	0%
henta	21	100%	0	0%
nægja	14	87.5%	2	12.5%
Total	73	92%	6	8%

Table 7. Alternating verbs in configurations with a dative pronoun and a nominative full NP

	Nom_{NP}-Dat_{PRO}		Dat_{PRO}-Nom_{NP}	
Verb	*N*	*f*	*N*	*f*
duga	55	87%	8	13%
dyljast	15	33%	31	67%
endast	37	27%	98	73%
henta	54	100%	0	0%
nægja	38	55%	31	45%
Total	199	54%	168	46%

the time and the latter 54%. As soon as *henta* is removed from the dataset, the Dat-Nom order becomes even slightly more common than the Nom-Dat order, reaching a prevalence of 54% (vs. 46% Nom-Dat).

It is striking how well the inter-verb differences uncovered for the current configuration map onto the differences found in the double-NP configuration. That is, the extent to which individual verbs tend to alternate in both of these configurations is nearly identical, at least in relative terms.

Finally, the statistics obtained for alternating Dat-Nom/Nom-Dat verbs, when the dative is a pronoun and the nominative a full NP, deviate considerably from those of their non-alternating Dat-Nom counterparts: in configurations with a dative pronoun and a nominative full NP, *líka*-verbs opt for the Dat-Nom order 508 times (or 97%), yet only 14 times (or 3%) for the Nom-Dat order. Once more, this underscores the split of the overarching class of Dat-Nom verbs into alternating *nægja*-verbs and non-alternating *líka*-verbs. It also confirms that the behaviour of nominative pronouns in the [Pro-V-Pro] configuration is not a general effect of pronouns.

https://doi.org/10.1017/S0332586524000039 Published online by Cambridge University Press

Table 8. Proportional prevalence of Nom-Dat vs. Dat-Nom linear order in the [NP-V-NP] configuration for *hjálpa*-, *líka*-, and *nægja*-verbs, and for *nægja*-verbs excluding *henta*

	Nom-Dat	Dat-Nom
hjálpa-verbs	99.5%	0.5%
líka-verbs	0.5%	99.5%
nægja-verbs	72%	28%
nægja-verbs (excluding *henta*)	54%	46%

4.3 Interim conclusions

The findings presented in this section confirm that Icelandic indeed possesses a class of alternating Dat-Nom/Nom-Dat verbs, as we have here, for the first time in the literature, established with statistics that both arguments of *nægja*-verbs pass the word order test (excluding *henta*). This is evident from the fact that, in the [NP-V-NP] configuration, the Nom-Dat linear order is attested 72% of the time, and the Dat-Nom linear order 28% of the time. This is very different from both *hjálpa*- and *líka*-verbs, as 99.5% of all instances involving full NPs show up with the Dat-Nom vs. the Nom-Dat linear order, respectively, for the two verb classes, as is reiterated in Table 8.

We base our conclusions of neutral word order on attestations where both arguments are lexically realised as full NPs, as pronouns clearly impose an information-structural bias on word order, for instance, inducing topicalisation. Furthermore, Table 8 also shows that the results are all the more powerful once *henta*, the outlier, is excluded from the statistics, yielding 54% Nom-Dat and 46% Dat-Nom linear order.

The fact that *henta* consistently occurs in the nominative-before-dative linear order is a compelling result in itself. Its word order bias can be explained in two ways: (i) our sample is off, or (ii) *henta* is not an alternating verb. The former would be indicative of a discrepancy between what is theoretically possible and what is actually attested, the latter of a potential linguistic change, but both hypotheses warrant further investigation.

We now summarise the results obtained for the remaining three configurations, which all involve at least one pronoun. These essentially show that alternating *nægja*-verbs swing towards the Nom-Dat order whenever nominative pronouns are involved. Both the double-pronoun configuration as well as the configuration involving a nominative pronoun and a dative full NP instantiate the Nom-Dat order in more than 92% of all cases. One might perhaps believe that this pronominal effect with *nægja*-verbs is a derivative of animacy, as pronouns tend to refer to animate entities, which in turn tend to be subjects, thus occurring clause-initially (cf. Du Bois 1987). This, however, is not the case for our dataset, as 99% of pronominal nominatives, even excluding correlates, are inanimate. Thus, the pronominal skewness towards the Nom-Dat order cannot be attributed to animacy.

Non-alternating *líka*-verbs are not altogether immune to the influence of nominative pronouns (and especially nominative demonstratives), but they only allow for 19% topicalisation in the [Pro-V-Pro] configuration and 18% in

https://doi.org/10.1017/S0332586524000039 Published online by Cambridge University Press

configurations with a nominative pronoun and a dative full NP. Alternating *nægja*-verbs in configurations with a dative pronoun and a nominative NP virtually mimic the frequencies obtained for the [NP-V-NP] configuration. Thus, our results confirm the status of *nægja*-verbs as a class in their own right, different from non-alternating *líka*-verbs.

In the next section, we investigate the factors underlying each word order pattern of alternating verbs by appealing to a set of variables that are known to influence linearisation, including animacy, definiteness, referentiality, pronoun type, person, number, and length. Conveniently, several of these may also serve as a proxy for topicality, as topics tend to be animate, definite, referential, pronominal, and short (Givón 1976:152, Arnold et al. 2000:34, Croft 2003:178–179, Rosenbach 2008:156, Arnold et al. 2013:406, Cristofaro 2013:74, 2019:28, Booth & Beck 2021:11, inter alia). It has indeed been argued in the literature that, other things being equal, alternating predicates allocate the preverbal slot to the argument that is most topical in the discourse (Barðdal 2001, Barðdal, Eythórsson & Dewey 2014, 2019). By modelling the word order variation statistically, we aim to uncover which factors have a direct bearing on linearisation. Additionally, we hypothesise the results to converge towards those values that have been shown to correlate with topicality (cf. supra). That way, we bring into the equation a variable that has not been explicitly factored in, but that may still have an influence on the alternation under study.

5. Statistical modelling

In what follows, we investigate the factors guiding the word order variation in the set of alternating Dat-Nom/Nom-Dat verbs by means of two logistic regression models. This involves a comparison across all configurations, on the one hand, and across double NPs, on the other. The reason why we single out double NPs is because pronouns are well known to skew word order preferences (Du Bois 1987, Croft 2012, Cristofaro 2013, 2019, Booth & Beck 2021, inter alia). Importantly, since *henta* behaves as a clear outlier, it is excluded from the remainder of the analysis. As such, the results presented in the current section are only based on our findings for the verbs *duga*, *dyljast*, *endast*, and *nægja*.

The purpose of the logistic regression analyses is to identify, and quantify, nuanced empirical interactions between the factors that we hypothesise are involved in the alternation. These include the variables discussed in Section 3, for which the dataset is annotated, namely case marking, pronominality, pronoun type (if applicable), referentiality, person, number, definiteness, animacy, and length. The dependent variable is argument position, either first or second. Binary logistic regression is a probabilistic algorithm that models the outcome as a probability, conditional on the value of the predictor variables (Harrell 2015:219). Although binary logistic regression makes relatively few assumptions, as with any regression model, collinearity (i.e. correlation) among the predictors can be a concern (Harrell 2015:255).

We have chosen to use ordinary logistic regression rather than mixed-effect/multilevel models. The reason for this is simple: the natural group-level (or random

https://doi.org/10.1017/S0332586524000039 Published online by Cambridge University Press

effect) in such a model would be verbs, but a much larger number of verbs would need to be included to defend the added complexity of mixed-effect models (Gelman & Hill 2007:247). When the random effect variable has few levels, mixed-effects logistic models reduce to ordinary logistic models with only fixed effects, in the absence of meaningful group-level variation.

For evaluating the logistic regression models, we rely on a combination of inspecting model residuals (Gelman & Hill 2007:97–101) and measures of predictive capability, in particular the c-index, since formal tests of fit are often inappropriate for logistic regression (Harrell 2015:236). The c-index measures the proportion of correctly classified responses when comparing the predictions of the model with the observed values in the dataset (Harrell 2015:257). While several other measures exist (Harrell 2015:256–257), we have chosen to report the c-index since it has a reasonably intuitive interpretation. A c-index value of .5 indicates random choice, 1.0 indicates perfect prediction, and .8 and above is often taken as an indication of good predictive capability (Baayen 2008:284, Harrell 2015:257). However, it is worth noting that there is some arbitrariness in these thresholds and in medicine, for example, the threshold for an acceptable model is usually taken to be .7 (Hartman et al. 2023, White et al. 2023). Importantly, over-reliance on a single measure can be detrimental, as noted by White et al. (2023), which is why we use the c-index alongside inspection of model residuals, bearing in mind that although a higher c-index is better than a lower one, it does not tell the whole story.

The logistic regression models have all been fitted in R using the *rms* package. A close inspection of the binned model residuals shows no signs of structural problems. Due to the skewed (i.e. non-symmetrical) distribution in the underlying data set, the length variable has been transformed by taking the natural logarithm of the observed data, which reshapes the data by adjusting the scale, resulting in a more symmetrical distribution. Although other logarithm bases would work equally well for the data transformation, the natural logarithm benefits from being directly interpretable in the model as proportional differences (Gelman & Hill 2007:60–61). A positive regression coefficient indicates that a variable is associated with the first argument position, while a negative value signals association with the second argument position.

Section 5.1 presents the results of the logistic regression analysis modelling all the data obtained across configurations, whereas Section 5.2 singles out the tokens instantiating the double-NP configuration. As such, the first model is based on all 800 observations and the second on 131 observations. In Section 5.3 we discuss how our findings tie in with the concept of topicality.

5.1 Across configurations

The output of the first logistic regression model, which builds on all 200 observations per verb type, is presented in Table 9. Recall that *henta*, the outlier, has been excluded. The c-index of the model is .794, a value that is only decimals away from what is commonly taken as a good predictive capability (Baayen 2008:204, Harrell 2015:257).

Table 9 shows the logistic regression coefficient (β), standard error (SE), z-score (Z) and p-value (p) for eight variables, seven of which exert a significant influence on

https://doi.org/10.1017/S0332586524000039 Published online by Cambridge University Press

Table 9. Results of the logistic regression model for alternating verbs across configurations excluding *henta* ($N = 800$). Significant *p*-values are in boldface

	β	SE	*Z*	*p*
Intercept	−2.63	.46	−5.8	**<.0001**
Animacy = individual	.77	.34	2.3	**<.05**
Animacy = inanimate	−.11	.53	−.2	>.8
Animacy = NA	1.33	.64	2.1	**<.05**
Case = nominative	2.26	.46	5.0	**<.0001**
Definiteness = indefinite	−1.08	.14	−7.9	**<.0001**
Length	−.83	.10	−8.1	**<.0001**
Number = singular	−.30	.14	−2.1	**<.05**
Person	.74	.10	7.2	**<.0001**

the alternation. Only animacy (value: inanimate) does not have any predictive power, presumably because it strongly correlates with nominative case. Positive regression coefficients indicate an association with the first argument slot, while negative regression coefficients indicate an association with the second slot. As such, animacy (value: individual), animacy (value: NA), case (value: nominative), and person (modelled numerically, see below) are associated with the first argument position, whereas definiteness (value: indefinite), length, and number (value: singular) are tied to the second argument position. Importantly, factors generating a significant effect do not necessarily correlate with one other. As an example, the second argument is very often either indefinite or long, but it is not necessarily associated with both properties at the same time. All the logistic regression coefficients represent the log-odds ratio of switching from second to first position.

Starting with the variables tied to the first argument position, the model attributes a major effect to nominative case marking and animacy (value: NA), which generate a coefficient of 2.26 and 1.33, respectively (corresponding to a 56.5% and 33.2% increase in the likelihood of switching from second to first argument slot). The fact that the former, nominative case, accounts for such a large portion of the variation is hardly surprising, as approximately two thirds of all tokens (i.e. 547 out of 800) with *duga*, *dyljast*, *endast*, and *nægja* instantiate the Nom-Dat order (cf. Section 4.2.1 above). Nevertheless, the question remains what exactly this means. That is, are alternating verbs, in fact, more strongly drawn to the Nom-Dat order than they are to the reverse Dat-Nom order? Or is the Nom-Dat bias in our sample merely the result of chance? We return to this point in Section 5.2 below.

The second variable closely tied to the first argument position, i.e. animacy (value: NA), captures all instantiations of correlative *það* 'it', which is a third person personal pronoun anticipating a subclause. Correlative *það*, exactly like expletive and existential *það*, is indeed well known to occur clause-initially in Icelandic (Rögnvaldsson 2002, Thráinsson 2007:366–367, inter alia). Still, the fact that this correlate so willingly takes initial position is remarkable, as it goes against the expectation that semantically impoverished units should take a less prominent

https://doi.org/10.1017/S0332586524000039 Published online by Cambridge University Press

position than arguments referring to extralinguistic entities (cf. Siewierska 1993:831). Somers & Barðdal (2023:19–21), for instance, have shown that the strong inclination of Icelandic correlates towards the nominative-before-dative order is something that sets these apart from their German counterparts, since the latter are much more permissive of alternation.

The third variable whose connection with the first argument position is significant is person. We decided to encode this variable as numeric, not categorical, and every unit increase in person (i.e. from first to second, and from second to third) yields an increase in association with the first slot. The effect is quantified by the coefficient as .74. This means that the likelihood of an argument being realised preverbally increases by approximately 18.5% for every one-unit increase along the scale. Thus, constituents are overall more likely to be allotted the first slot if their referent is non-local, third person, as opposed to local, first and second person. Interestingly, this constitutes a violation of the person hierarchy, which stipulates that local pronouns should take precedence over both non-local pronouns and full NPs (Silverstein 1976, Siewierska 1993:831, Croft 2003:130, Haude & Witzlack-Makarevich 2016, inter alia).

The reason, we believe, that non-local referents, i.e. third person, are more likely to occur in the first slot, despite the person hierarchy, is due to the effect of correlates, which nearly always occur preverbally (140 out of 148 cases, or 95%), as opposed to postverbally (eight out of 148 cases, or 5%). The number of local pronouns in our dataset is quite limited (a mere 250 across argument positions). One reason why local pronouns are so rare is because these are never nominative in our dataset (however, see below). Local pronouns are also more strongly tied to the second slot (197 out of 250 cases, or 79%). Interestingly, 118/197 postverbal (dative) local pronouns compete with a nominative pronoun, and nominative pronouns show a strong tendency to take first position in any case.

Moreover, in accordance with the person hierarchy, if a nominative is a first or second person pronoun, i.e. a local pronoun, only the Nom-Dat word order is acceptable with alternating predicates in Icelandic. One such attested example, cited from Barðdal & Eythórsson (2003), is given in (17a) with the nominative *við* 'we' in first position. As is shown in (17b), if the nominative is a local pronoun, the Dat-Nom construction is ungrammatical, since the nominative *við* 'we' cannot occur in the object position immediately following the non-finite verb. This analysis is further confirmed by the example in (17c), which shows that if the dative occurs in first position, the local pronoun must invert with the finite verb, which is a clear-cut subject property. For obvious pragmatic reasons, the constructed examples in (17b–c) render the dative *fólki* 'people' as a definite NP instead of an indefinite one.

Neutral Nom-Dat word order

(17) a. Við föllum víst alls ekki fólki í geð.
we.NOM fall.1PL surely all not people.DAT in liking
'We are supposedly not popular among people at all.'
(www.hi.is/~hoski/Haskolasaga.pdf, here cited from Barðdal & Eythórsson 2003:764)

https://doi.org/10.1017/S0332586524000039 Published online by Cambridge University Press

Intended neutral Dat-Nom word order

(17) b. *Fólkinu höfum/hefur víst alls ekki fallið
people.the.DAT have.1PL/has.3SG surely all not fallen.PTCP
við í geð.
we.NOM in liking
Intended meaning: 'The people are supposedly not popular among us at all.'

Topicalisation

(17) c. Fólkinu höfum við víst alls ekki fallið í geð.
people.the.DAT have.1PL we.NOM surely not all fallen.PTCP in liking
'Among people, we are supposedly not popular at all.'

Recall that at the beginning of this article (see examples 3–4), we provided evidence for the analysis that alternating verbs may instantiate two diametrically opposite argument structure constructions, Dat-Nom and Nom-Dat, and that neither structure is a topicalisation of the other. This is invariably true in all cases except when the nominative is a local pronoun, as in (17) above.

Returning now to the logistic regression analysis, another variable that shows a mild preference for the preverbal slot is animacy (value: individual), with a coefficient of .77. Evidently, this is fully in line with the expectation that animate beings should take precedence over both collectives and inanimates (Allan 1987, Siewierska 1993, Dahl & Fraurud 1996, inter alia). Rott (2013) also found an effect for animacy in his study of four Icelandic alternating verbs. More specifically, he has shown that the nominative is hardly ever animate, but that it invariably precedes the dative when it is. Our data are indicative of a similar trend: out of 800 nominative constituents, a mere 14 are animate. Of these, 11 (or 79%) are attested in the Nom-Dat order, and three (or 21%) in the reverse Dat-Nom order. The same holds, *mutatis mutandis*, for the dative: it is hardly ever inanimate (19 out of 800 tokens), but when it is, it shows a very strong preference for the Nom-Dat order (15 tokens, or 79%) and not the Dat-Nom order (four tokens, or 21%).[6] Observe that we hereby dispel the myth that the dative of Dat-Nom verbs is animate by definition (see Kutscher 2009:24, Verhoeven 2009, 2015, Rott 2013:93). The tendency for the dative of such verbs to be animate is indeed very strong, but it is by no means an absolute.

Two variables that associate with the postverbal slot are length and definiteness (value: indefinite) (coefficients −.83 and −1.08). Again, these facts rhyme well with what is known in the literature, namely that indefinite and heavy constituents tend to occur later in the clause (see Behaghel 1909/10, Allan 1987, Siewierska 1993, Arnold et al. 2000, inter alia). Specifically with regard to the Dat-Nom/Nom-Dat alternation, Rott (2013) has also found length to exert an effect, but his evidence for this factor only relates to the distribution of clausal nominatives, which greatly prefer the Dat-Nom order (82 out of 87 tokens, or 94%) to the Nom-Dat order (five out of 87 tokens, or 6%). Our own study conclusively confirms that length is a factor even when both arguments are (pro)nominal.

Rott also finds an effect for definiteness, but the results are again clouded by the high number of clausal nominatives in his dataset. Starting with the Nom-Dat order, he reports that 70 out of 73 nominatives (or 96%) are definite, compared to 55 out of 78 datives (or 71%). Thus, the Nom-Dat order clearly correlates with definite nominatives. As for the Dat-Nom order, 48 out of 94 datives (or 51%) are definite,

https://doi.org/10.1017/S0332586524000039 Published online by Cambridge University Press

Table 10. Results of the logistic regression model for alternating verbs in the double-NP configuration excluding *henta* ($N = 131$). Significant *p*-values are in boldface

	β	SE	*Z*	*p*
Intercept	.10	.22	.4	**>.7**
Case = nominative	1.16	.22	5.4	**<.0001**
Definiteness = indefinite	−1.15	.18	−6.4	**<.0001**
Length	−.54	.12	−4.4	**<.0001**
Number = singular	−.36	.21	−1.7	>.1

compared to ten out of 12 nominatives (or 83%). However, since the Dat-Nom order so strongly correlates with clausal nominatives already (cf. supra), the speaker's choice for this alternant has already been accounted for.

Additionally, our model also connects singular constituents with the second argument position, but their effect is considerably weaker (coefficient −.30) than for the remaining variables.

5.2 Double-NP configuration

The results of the second logistic regression model, which is solely based on the 131 tokens containing double NPs, are shown in Table 10. The c-index is .7, indicating a lower predictive power than the previous model (whose c-index is .794) although still in a range that would be considered acceptable in some fields. A lower c-index value is not unexpected given that the dataset on which the model is trained is smaller, and as such we consider it worth reporting on, especially since the model residuals do not indicate any particular issues.

A comparison with the first model yields fascinating results. First, the current model singles out three variables that were also identified as strong predictors by the first model, i.e. nominative case marking, indefiniteness, and length (coefficients 1.16, −1.15, and −.54, respectively, corresponding to a 29% increase, a 28.8% decrease, and a 13.5% decrease in the likelihood of switching from second to first argument slot). The effects of case marking and length in particular seem to be somewhat mitigated compared to the first model, but both still generate highly significant *p*-values. Second, the weaker predictors pertaining to both animacy and number no longer appear to have any predictive power, and third, the person variable has been levelled out, as full NPs are self-evidently always third person.

The second logistic regression analysis essentially reveals two tendencies. As with the first analysis, it demonstrates the importance of nominative case marking for the first argument slot. Recall that the four alternating verbs under study yield a total of 71 Nom-Dat tokens and 60 Dat-Nom tokens. The difference between both these subsets is taken to be statistically significant, but, again, it remains to be investigated (i) whether a different sample would equally single out nominative case marking as a significant predictor, and (ii) whether other alternating verbs are equally sensitive to the effect of nominative case marking.

https://doi.org/10.1017/S0332586524000039 Published online by Cambridge University Press

Second, the analysis shows that both indefinite and lengthy constituents have a proclivity for the second argument slot. The model neither reveals whether these variables are interrelated, nor whether they correlate with case marking. We discuss these issues further in the following subsection.

5.3 Interim conclusions

The current section has provided an in-depth statistical analysis of the alternating verbs *duga*, *dyljast*, *endast*, and *nægja* by means of two logistic regression models, one scrutinising the results across configurations ($n = 800$), another exploring the results for the double-NP configuration ($n = 131$). We have identified several well-known predictors from studies of word order as playing a role here, most notably animacy, indefiniteness, and length. Both models also single out nominative case marking as a predictor for the first argument slot.

Starting with the first three predictors, we have found animate (dative) constituents to prefer the first argument slot and indefinite and long constituents to prefer the second argument slot. These findings are evidently interesting in themselves, but the key question is whether there is a greater generalisation to be made. In fact, as is already mentioned in Sections 3 and 4.3 above, these three factors, animacy, indefiniteness, and length, are all proxies for topicality. That is, topical constituents tend to be animate rather than inanimate, and non-topical constituents tend to be indefinite and long rather than definite and short (Givón 1976:152, Croft 2003:178–179, Arnold et al. 2013:406, Cristofaro 2013:74, Booth & Beck 2021:11, inter alia). Thus, apart from their relevance as individual predictors, the factors in question also suggest that word order with alternating verbs is a derivative of discourse prominence (cf. Barðdal 2001, Barðdal, Eythórsson & Dewey 2014, 2019).

Turning now to the last predictor correlating with the first argument slot, nominative case marking, the question arises whether the nominative is a factor in itself or whether the case marking is an epiphenomenon of other factors, such as, for instance, pronominality. Out of 331 nominative pronouns in first position, 42%, or 140 instances, are correlates, which show a clear preference for first position in Icelandic in any case (Rögnvaldsson 2002, Thráinsson 2007:366–367, inter alia). A closer inspection of the [Pro-V-Pro] configuration, including correlates, reveals that nominative pronouns in first position show an average length of 1.1 words, while dative pronouns in second position have an average word length of 1.9. This suggests that the nominative-first effect with two pronouns is a consequence of length.

However, such a length effect with nominatives in first position is not found for the double-NP configuration. Instead, nominatives in first position turn out to be definite in 43 out of 71 instances (61%), which clearly makes them topical. What is more, a gauge at the 28 indefinite examples of nominatives in first position reveals that they are either specific or simply more topical than the dative, as is evident from example (18) below, despite being indefinite.

https://doi.org/10.1017/S0332586524000039 Published online by Cambridge University Press

Nom-Dat

(18) Eyjan er mjög grösug, og eyjarbúar nota ekki **rigningarvatn**, heldur eru þar uppsprettur og brunnar, því að þar **rignir** ekki það mikið, að
rigningarvatn mundi duga Grímseyingum
*rain.water.*NOM *would* *suffice.*INF *Grímsey.islanders.*DAT
til daglegrar neyzlu.
'The island is very grassy and the islanders do not use rainwater, instead there are wells and springs; since it does not rain that much there, *rainwater would not suffice the inhabitants of Grímsey* for daily use.'

Observe that the indefinite nominative, *rigningarvatn* 'rainwater', in the Nom-Dat example above reiterates information mentioned earlier in the discourse, with both the former (accusative) *rigningarvatn* 'rainwater' and *rignir* 'rains' rendering the latter (nominative) *rigningarvatn* 'rainwater' highly topical, despite it being indefinite. This not only shows that topicality may not simply be reduced to one (or more) of its proxies but also that it has considerable explanatory power of its own. At least for the [NP-V-NP] configuration, it seems promising to explicitly factor in topicality as a predictor variable, because the effect of nominative case uncovered in the present study appears to be an epiphenomenon of a topic-first effect rather than a veritable nominative-first effect in itself.

6. Summary and conclusions

In this article we have succeeded in lending empirical support to the claim that behavioural subjects in Modern Icelandic are strongly tied to clause-initial position, regardless of whether these are marked in the nominative or the dative case. For this purpose, we have extracted 200 examples of 15 verbs each from the Icelandic Web 2020 corpus, thus amounting to a total of 3,000 tokens, all occurring with a dative and a nominative. The first class consists of five ordinary Nom-Dat verbs like *hjálpa* 'help', the second consists of five classical Dat-Nom verbs like *líka* 'like', and the third one of five alternating Dat-Nom/Nom-Dat verbs like *nægja* 'find/be sufficient'.

The dataset has been annotated for nine variables: case marking, (pro) nominality, type of pronoun, referentiality, person, number, definiteness, animacy, and length. Our goal has been to provide statistical evidence of our alternating analysis for *nægja*-verbs, namely that these verbs alternate between two word orders, dative-before-nominative and nominative-before-dative, due to the fact that they instantiate two diametrically opposite argument structures, i.e. Dat-Nom and Nom-Dat.

We first establish a baseline with the help of ordinary Nom-Dat verbs, or *hjálpa*-verbs, and non-alternating Dat-Nom verbs, or *líka*-verbs, in configurations with two full NPs. It turns out that both these verb classes, *hjálpa*- and *líka*-verbs, realise the syntactic subject clause-initially 99.5% of the time. In contrast, for alternating Dat-Nom/Nom-Dat verbs, i.e. *nægja*-verbs, our findings generally confirm that the subject is the first argument of the argument structure, be that the dative or the nominative.

When *nægja*-verbs occur with two full NPs, their distribution is considerably less skewed towards one of the two argument structure constructions than with either

https://doi.org/10.1017/S0332586524000039 Published online by Cambridge University Press

hjálpa- or *líka*-verbs. There are, however, substantial differences found across verbs, with the Nom-Dat case frame attested more frequently than the Dat-Nom case frame, or in 72% vs. 28% of the cases. This number of 28% Dat-Nom is considerably higher than the 0.5% baseline for topicalisation documented with *hjálpa-* and *líka*-verbs above, and it is also noticeably higher than the 8% topicalisation documented by Callegari & Ingason (2021). This, in turn, rules out a topicalisation analysis of dative-before-nominative word orders with *nægja*-verbs. As a matter of fact, there is one particular verb, *henta*, that behaves unexpectedly in that it occurs consistently with the Nom-Dat linear order, irrespective of whether the two arguments are realised as full NPs or as pronouns. Thus, when recalculating the numbers for full NPs without the outlier, *henta*, the distribution amounts to 54% Nom-Dat vs. 46% Dat-Nom. Again, this rules out a topicalisation analysis of the dative-before-nominative order altogether.

Our analysis of *nægja*-verbs has also shown that their word order distributions are considerably more prone to pronominal influence than the ones attested for either *hjálpa-* or *líka*-verbs. More specifically, in contexts where the nominative is pronominal, *nægja*-verbs strongly prefer the nominative to precede the dative. However, contexts in which a dative pronoun enters into competition with a nominative full NP show the same word order distributions as the [NP-V-NP] configuration. These findings confirm the status of alternating Dat-Nom/Nom-Dat verbs as a syntactic class in their own right, distinct from non-alternating Dat-Nom verbs.

Finally, we have modelled the word order variation of *nægja*-verbs statistically. Recall that we removed *henta* from our dataset, as its frequencies were unexpectedly skewed. Across configurations, word order patterns are prone to a host of factors, including nominative case marking, indefiniteness, length, animacy, and person. For the double-NP configuration, the logistic regression analysis has identified nominative case marking, indefiniteness, and length as the most important predictors.

As it turns out, the factors underlying the variation in word order, both across configurations as well as in the double-NP configuration, converge nicely in that all these appear to reflect topicality in one way or another. After all, topicality is highly interwoven with animate, pronominal, definite, and short constituents. The only two exceptions to the topical-first trend that we have uncovered involve nominative case marking and person. Starting with person, third person arguments are generally relatively equally divided across the two positions, except for correlates, which occur in first position 95% of the time. Thus, we believe that the third person effect, detected in the logistic regression analysis for first position, is an epiphenomenon of this.

Turning to nominative case marking, we have also shown that in the double-pronoun configuration, which favours the Nom-Dat word order, the preverbal nominative pronoun is considerably shorter than the postverbal dative pronoun, indeed suggesting that the real issue here is length rather than case marking. Regarding the configuration with double NPs, 61% of the nominatives in first position are definite, again confirming the role of topicality. The remaining 39% of the preverbal nominatives are indefinite, yet an initial inspection of these instances shows that the majority are topical, although some are specific. This again validates

https://doi.org/10.1017/S0332586524000039 Published online by Cambridge University Press

the role of topicality, also for double NPs, confirming that the strongly observed nominative-first effect is an artefact of topicality.

To conclude, comparing *nægja*- and *líka*-verbs, we have shown that the former, but not the latter, have a choice between two alternating constructions, Dat-Nom and Nom-Dat. It turns out that well-worn pragmatic factors such as topicality govern the choice between the two diametrically opposite constructions with *nægja*-verbs. In contrast, with *líka*-verbs, the grammar does not provide this option to begin with, meaning that this verb class is confined to the Dat-Nom argument structure construction.

Regarding future research, the most pressing issue at this point is a comparison of the behaviour of alternating Dat-Nom/Nom-Dat verbs across the languages where such a class has been shown to exist, for instance Russian, Lithuanian, Romanian, Latin, and Ancient Greek (cf. Barðdal 2023:Ch. 3 and the references therein). A particularly promising comparison is one between Modern Icelandic and Modern German, due to their close kinship. For a first attempt at such a venture, see Somers & Barðdal (2023), although a more fine-grained analysis of the relevant data is needed to improve our understanding of the factors at play.

Acknowledgements. This is a heavily revised version of Somers & Barðdal (2022) in *Working Papers in Scandinavian Syntax* (WPSS 107). For comments and/or discussions, we thank Johan Brandtler, Ludovic De Cuypere, Torsten Leuschner, the editor, Marit Julien, three anonymous reviewers of *NJL*, and the audiences at *Constructions in the Nordics 3* in Kiel in September 2022, at the Belgian Taaldag in Liège in October 2022, at the *North by Northwest* seminar at Lyon University in November 2022, at the VII CONECT Internacional in Brazil in November 2022, at the Amazonicas IX in Bogotá, Colombia, in June 2023, and the audience at the Forschungskaleidoskop seminar at Hamburg University in June 2023. This research is a part of a larger project on Language Productivity at Work (Co-PI Jóhanna Barðdal), generously funded by Ghent University's Special Research Fund's Concerted Research Action Scheme (BOF-GOA grant no. 01G01319).

Authors' contributions. All three authors designed the research and planned and wrote the manuscript; JB and JS gathered the data; JS coded the data; GBJ performed the computational analysis; all authors contributed equally to the discussion and the interpretation of the results.

Notes

1 Glossing abbreviations follow the Leipzig Glossing Rules. The abbreviations used in this article are DAT = dative, F = feminine, HON = honorific, INF = infinitive, M = masculine, N = neuter, NOM = nominative, PL = plural, SG = singular.

2 Regarding the differences in argument structure patterns between *nægja*- and *líka*-verbs, see Wood & Sigurðsson (2014) for a putative analysis in terms of semantic differences of the two verb classes and Barðdal, Eythórsson & Dewey (2019:153–161) for refuting their argumentation.

3 These numbers are taken from Le Mair et al. (2017:131), which contains slightly updated numbers compared to the ones in Barðdal & Eythórsson (2012).

4 Sigurðsson (2006b:217) stresses that not every [Noun]-[Genitive] combination naturally receives a definite reading. This is particularly true of [Noun]-[Noun] combinations whose possessee is unmarked for definiteness, e.g. *bók kennara* 'a teacher's book'.

5 Fungi are strictly speaking neither plant nor human, but, for our purposes, labelling them as inanimate seems to be most fitting.

6 Note that these numbers do not include *henta*, which counts 32 tokens with inanimate datives. This brings the total number of inanimate datives with *nægja*-verbs to 51, or 5.1% of all datives in this class.

https://doi.org/10.1017/S0332586524000039 Published online by Cambridge University Press

References

Allan, Keith. 1987. Hierarchies and the choice of left conjuncts. *Journal of Linguistics* 23. 51–77.

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 *Papers from the 6th Annual Meeting of the North Eastern Linguistic Society* (Montreal Working Papers in Linguistics 6), 1–21.

Andrews, Avery. 1982. The representation of case in Modern Icelandic. In Joan Bresnan (ed.), *The mental representation of grammatical relations*, 427–503. Cambridge, MA: MIT Press.

Angantýsson, Ásgrímur. 2020. The distribution of embedded verb second and verb third in Modern Icelandic. In Rebecca Woods & Sam Wolfe (eds.), *Rethinking verb second*, 240–264. Oxford: Oxford University Press.

Arnold, E. Jennifer, Anthony Losongco, Thomas Wasow & Ryan Ginstrom. 2000. Heaviness vs. newness: The effects of structural complexity and discourse status on constituent ordering. *Language* 76(1). 28–55.

Arnold, Jennifer E., Elsi Kaiser, Jason M. Kahn & Lucy K. Kim. 2013. Information structure: Linguistic, cognitive, and processing approaches. *Wiley Interdisciplinary Reviews: Cognitive Science* 4(4). 403–413.

Axel, Katrin. 2007. *Studies on Old High German syntax: Left sentence periphery, verb placement and verb-second*. Amsterdam: John Benjamins.

Baayen, R. Harald. 2008. *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.

Barðdal, Jóhanna. 1999. The dual nature of Icelandic psych-verbs. *Working Papers in Scandinavian Syntax* 64. 79–101.

Barðdal, Jóhanna. 2001. The perplexity of Dat-Nom verbs in Icelandic. *Nordic Journal of Linguistics* 24. 47–70.

Barðdal, Jóhanna. 2006. Construction-specific properties of syntactic subjects in Icelandic and German. *Cognitive Linguistics* 17(1). 39–106.

Barðdal, Jóhanna. 2023. *Oblique subjects in Germanic: Their status, history and reconstruction*. Berlin: Mouton de Gruyter.

Barðdal, Jóhanna & Thórhallur Eythórsson. 2003. Icelandic vs. German: Oblique subjects, agreement and expletives. *Chicago Linguistic Society* 39(1). 755–773.

Barðdal, Jóhanna & Thórhallur Eythórsson. 2012. 'Hungering and lusting for women and fleshly delicacies': Reconstructing grammatical relations for Proto-Germanic. *Transactions of the Philological Society* 110(3). 363–393.

Barðdal, Jóhanna, Thórhallur Eythórsson & Tonya Kim Dewey. 2014. Alternating predicates in Icelandic and German: A sign-based construction grammar account. *Working Papers in Scandinavian Syntax* 93. 50–101.

Barðdal, Jóhanna, Thórhallur Eythórsson & Tonya Kim Dewey. 2019. The alternating predicate puzzle: Dat-Nom vs. Nom-Dat in Icelandic and German. *Constructions and Frames* 11(1). 107–170.

Behaghel, Otto. 1909/10. Beziehungen zwischen Umfang und Reihenfolge von Satzgliedern. *Indogermanische Forschungen* 25. 110–142.

Bernódusson, Helgi. 1982. Ópersónulegar setningar [Impersonal sentences]. University of Iceland MA thesis.

Booth, Hannah & Christin Beck. 2021. Verb-second and verb-first in the history of Icelandic. *Journal of Historical Syntax* 5(28). 1–53.

Bornkessel-Schlesewsky, Ina, Franziska Kretzschmar, Sarah Tune, Luming Wang, Safiye Genç, Markus Philipp, Dietmar Roehm & Matthias Schlesewsky. 2011. Think globally: Cross-linguistic variation in electrophysiological activity during sentence comprehension. *Brain and Language* 117(3). 133–152.

Bresnan, Joan & Marilyn Ford. 2010. Predicting syntax: Processing dative constructions in American and Australian varieties of English. *Language* 86(1). 168–213.

Callegari, Elena & Anton Karl Ingason. 2021. Topicalization: The IO/DO asymmetry in Icelandic. *Working Papers in Scandinavian Syntax* 105. 1–17.

Cristofaro, Sonia. 2013. The referential hierarchy: Reviewing the evidence in diachronic perspective. In Dik Bakker & Martin Haspelmath (eds.), *Languages across boundaries: Studies in memory of Anna Siewierska*, 69–94. Berlin: De Gruyter Mouton.

https://doi.org/10.1017/S0332586524000039 Published online by Cambridge University Press

Cristofaro, Sonia. 2019. Taking diachronic evidence seriously: Result-oriented vs. source-oriented explanations of typological universals. In Karsten Schmidtke-Bode, Natalia Levshina, Susanne Maria Michaelis & Ilja A. Seržant (eds.), *Explanation in typology: Diachronic sources, functional motivations and the nature of the evidence*, 25–46. Berlin: Language Science Press.

Croft, William. 2003. *Typology and universals*, 2nd edn. Cambridge: Cambridge University Press.

Croft, William. 2012. *Verbs: Aspect and clausal structure*. Oxford: Oxford University Press.

Dahl, Östen & Kari Fraurud. 1996. Animacy in grammar and discourse. In Thorstein Fretheim & Jeanette K. Gundel (eds.), *Reference and referent accessibility*, 47–64. Amsterdam: John Benjamins.

Du Bois, John W. 1987. The discourse basis of ergativity. *Language* 63(4), 805–855.

Eythórsson, Thórhallur. 1995. *Verbal syntax in the early Germanic languages*. Cornell University PhD dissertation.

Gelman, Andrew & Jennifer Hill. 2007. *Data analysis using regression and multilevel/hierarchical models*. Cambridge: Cambridge University Press.

Givón, Talmy. 1976. Topic, pronoun, and grammatical agreement. In Charles N. Li (ed.), *Subject and topic*, 149–188. New York: Academic Press.

Gregory, Michelle L. & Laura A. Michaelis. 2001. Topicalization and left-dislocation: A functional opposition revisited. *Journal of Pragmatics* 33(11). 1665–1706.

Harbert, Wayne. 2007. *The Germanic languages*. Cambridge: Cambridge University Press.

Harrell, Frank E. 2015. *Regression modeling strategies: With applications to linear models, logistic regression, and survival analysis*, 2nd edn. Cham: Springer.

Hartman, Nicolas, Sehee Kim, Kevin He & John D. Kalbfleisch. 2023. Pitfalls of the concordance index for survival outcomes. *Statistics in Medicine* 42. 2179–2190.

Haude, Katharina & Alena Witzlack-Makarevich. 2016. Referential hierarchies and alignment: An overview. *Linguistics* 54(3). 433–441.

Heylen, Kris. 2005. *Zur Abfolge (pro)nominaler Satzglieder im Deutschen: Eine korpusbasierte Analyse der relativen Abfolge von nominalem Subjekt und pronominalem Objekt im Mittelfeld*. KU Leuven PhD dissertation.

Jakubíček, Miloš, Adam Kilgarriff, Vojtěch Kovář, Pavel Rychlý & Vít Suchomel. 2013. The TenTen corpus family. In Andrew Hardie & Robbie Love (eds.), *7th International Corpus Linguistics Conference*, 125–127. Lancaster: University of Lancaster.

Jónsson, Jóhannes Gísli. 1996. *Clausal architecture and case in Icelandic*. University of Massachusetts, Amherst, PhD dissertation.

Jónsson, Jóhannes Gísli. 1997–1998. Sagnir með aukafallsfrumlagi [Verbs selecting for oblique subjects]. *Íslenskt mál og almenn málfræði* 19–20. 11–43.

Keenan, Edward L. 1976. Towards a universal definition of subject. In Charles N. Li (ed.), *Subject and topic*, 303–333. New York: Academic Press.

Kutscher, Silvia. 2009. *Kausalität und Argumentrealisierung: Zur Konstruktionsvarianz bei Psychverben am Beispiel europäischer Sprachen*. Tübingen: Niemeyer.

Lambrecht, Knud. 1994. *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents*. Cambridge: Cambridge University Press.

Lambrecht, Knud. 2000. When subjects behave like objects: An analysis of the merging of S and O in sentence-focus constructions across languages. *Studies in Language* 24(3). 611–682.

Le Mair, Esther, Cynthia A. Johnson, Michael Frotscher, Thórhallur Eythórsson & Jóhanna Barðdal. 2017. Position as a behavioral property of subjects: The case of Old Irish. *Indogermanische Forschungen* 122. 111–142.

Platzack, Christer. 1999. The subject of Icelandic psych-verbs: A minimalistic account. *Working Papers in Scandinavian Syntax* 64. 103–115.

Roehm, Dietmar, Matthias Schlesewsky & Ina Bornkessel-Schlesewsky. 2007. *Position or morphology? An electrophysiological examination of incremental argument interpretation in Icelandic*. Poster presented at the 20th CUNY Conference on Human Sentence Processing, UC San Diego, California, 29–31 March 2007.

Rögnvaldsson, Eiríkur. 2002. *ÞAÐ* í fornu máli – og síðar [*IT* in Old Icelandic – and later]. *Íslenskt mál og almenn málfræði* 24. 7–30.

Rosenbach, Anette. 2008. Animacy and grammatical variation: Findings from English genitive variation. *Lingua* 118(2). 151–171.

https://doi.org/10.1017/S0332586524000039 Published online by Cambridge University Press

Rott, Julian A. 2013. Syntactic prominence in Icelandic experiencer arguments: Quirky subjects vs. dative objects. *STUF – Language Typology and Universals* 66(2). 91–111.
Rott, Julian A. 2016. Germanic psych processing: Evidence for the status of dative experiencers in Icelandic and German. In Christel Stolz & Thomas Stolz (eds.), *From Africa via the Americas to Iceland*, 215–320. Bochum: Dr N. Brockmeyer.
Schätzle, Christin. 2018. *Dative subjects: Historical change visualized*. University of Konstanz PhD dissertation.
Siewierska, Anna. 1993. Syntactic weight vs. information structure and word order variation in Polish. *Journal of Linguistics* 29. 233–265.
Sigurðsson, Halldór Ármann. 1989. *Verbal syntax and case in Icelandic*. Lund University PhD dissertation.
Sigurðsson, Halldór Ármann. 1990–1991. Beygingarsamræmi [Agreement]. *Íslenskt mál og almenn málfræði* 12–13. 31–77.
Sigurðsson, Halldór Ármann. 2004. Icelandic non-nominative subjects: Facts and implications. In Peri Bhaskararao & Karumuri V. Subbarao (eds.), *Non-nominative Subjects* 2, 137–159. Amsterdam: John Benjamins.
Sigurðsson, Halldór Ármann. 2006a. The nominative puzzle and the low nominative hypothesis. *Linguistic Inquiry* 37. 289–308.
Sigurðsson, Halldór Ármann. 2006b. The Icelandic noun phrase: Central traits. *Arkiv för Nordisk Filologi* 121. 193–236.
Silverstein, Michael. 1976. Hierarchy of features and ergativity. In Robert M. W. Dixon (ed.), *Grammatical categories in Australian languages*, 112–171. Canberra: Australian Institute of Aboriginal Studies.
Somers, Joren & Jóhanna Barðdal. 2022. Alternating Dat-Nom/Nom-Dat verbs in Icelandic: An exploratory corpus-based analysis. *Working Papers in Scandinavian Syntax* 107. 83–110.
Somers, Joren & Jóhanna Barðdal. 2023. Comparing the argument structure of alternating Dat-Nom/Nom-Dat predicates in German and Icelandic. *Working Papers in Scandinavian Syntax* 108. 1–25.
Talmy, Leonard. 1976. Semantic causative types. In Masayoshi Shibatani (ed.), *The grammar of causative constructions*, 43–116. New York: Academic Press.
Thráinsson, Höskuldur. 1979. *On complementation in Icelandic*. New York: Garland.
Thráinsson, Höskuldur. 2007. *The syntax of Icelandic*. Cambridge: Cambridge University Press.
Verhoeven, Elisabeth. 2009. Subjects, agents, experiencers, and animates in competition: Modern Greek argument order. *Linguistische Berichte* 219. 355–376.
Verhoeven, Elisabeth. 2015. Thematic asymmetries do matter! A corpus study of German word order. *Journal of Germanic Linguistics* 27(1). 45–104.
Wallenberg, Joel C., Anton Karl Ingason, Einar Freyr Sigurðsson & Eiríkur Rögnvaldsson. 2011. Icelandic Parsed Historical Corpus (IcePaHC). Version 0.9.
White, Nicole, Rex Parsons, Gary Collins & Adrian Barnett. 2023. Evidence of questionable research practices in clinical prediction models. *BMC Medicine* 21. 339.
Wood, Jim & Halldór Ármann Sigurðsson. 2014. *Let*-causatives and (a)symmetric DAT-NOM constructions. *Syntax* 17(3). 269–298.
Zaenen, Annie, Joan Maling & Höskuldur Thráinsson. 1985. Case and grammatical functions: The Icelandic passive. *Natural Language and Linguistic Theory* 3. 441–483.

Cite this article: Somers J, Jenset GB, and Barðdal J (2026). Argument structure constructions in competition: The Dat-Nom/Nom-Dat alternation in Icelandic. *Nordic Journal of Linguistics* **49**, 33–67. https://doi.org/10.1017/S0332586524000039

https://doi.org/10.1017/S0332586524000039 Published online by Cambridge University Press