Alternating Dat-Nom/Nom-Dat Verbs in Icelandic: An Exploratory Corpus-Based Analysis*

Joren Somers & Jóhanna Barðdal
Ghent University

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. Since the discovery of this verb class in Icelandic, considerable work has been carried out on different aspects of the nature of these verbs in Icelandic and related languages. Yet, what is missing from the literature is a systematic study of the distribution of the relevant verbs across the two argument structure constructions in language use and whether all alternating verbs instantiate both argument structure constructions to the same degree. For this purpose, we have carried out a study of 15 verbs, five alternating ones, and as a control, five ordinary Nom-Dat verbs and five non-alternating Dat-Nom verbs. Our findings show that alternating verbs instantiate the Nom-Dat construction in 54% of the cases, and the Dat-Nom construction in 46% of the cases on average for four of the five verbs when both arguments are full NPs, although considerable statistical differences are found between the five verbs. Another remarkable finding is that when the two arguments are pronouns, the Nom-Dat construction takes precedence over the Dat-Nom construction.

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 1997–98, Barðdal 1999, 2001, 2022: Ch. 3, Platzack 1999, Sigurðsson 2006, Rott 2013, 2016, Wood & Sigurðsson 2014, 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' and *líka* 'like'. The verb *nægja*, being an alternating verb, allows both verbal arguments to take clause-

* For comments and/or discussions, we thank Johan Brandtler, Ludovic De Cuypere, Gard B. Jenset, Torsten Leuschner, 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, and at the VII CONECT Internacional in Brazil in November 2022. 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 nr. 01G01319).

initial position, thus confirming their status as syntactic subjects. At the same time, the other argument is realised in the postverbal slot, which is reserved for objects. The verb *líka* does not allow for this kind of alternation, as the dative must never occupy the postverbal object slot:

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

b. *Þessi skýring hafði nægt Gunnari.*
this.NOM explanation.NOM had sufficed Gunnar.DAT
'These explanations were not sufficient for Gunnar'

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

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.'

The fact that either argument of alternating verbs may function as the syntactic subject or the syntactic object was first documented by Barðdal (1999, 2001) with respect to a host of accepted subject tests for Icelandic. 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). 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, what determines the speakers' choice of one of the two argument structure constructions, Dat-Nom or Nom-Dat, over the other?

A first attempt at an investigation of this type was carried out by Rott (2013). He extracted his data from a corpus of 70 million words and collected tokens for eight verbs in total, 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 harvested 50 tokens per verb, and his full dataset only comprised 372 observations. Another disadvantage of Rott's study is that it also includes clausal arguments, i.e. instances where the nominative is realised as a clause, as opposed to when it is realised as a nominal argument. Since clausal arguments are *de facto* considerably longer than nominal arguments, clausal arguments are

more prone to occurring later in the clause than nominal arguments. As a consequence, clausal arguments should show a greater tendency to be realised as objects, as objects generally occur later in the clause than subjects in Icelandic. 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, since 82 out of 94 Dat-Nom attestations can probably be attributed to a length effect.

Another limitation of Rott's (2013) study is that it does not specify word order distributions per verb lemma, thus positing a verb class effect without actually demonstrating that such an effect should exist in the first place. Finally, Rott also does not investigate 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. For (pro)nominal constituents, he also specifies the type of constituent (proper noun, definite NP, indefinite NP; personal pronoun, demonstrative pronoun, indefinite pronoun). 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.

The goal of this paper 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 and pronouns with pronouns, instead of mixing the two types of argument realisations with each other. 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 speakers' choice of one argument structure construction over the other.

In the remainder of this paper, 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. However, in order to establish a baseline with which our findings for alternating verbs may be compared, we first present an identical study involving both classical Dat-Nom verbs and ordinary Nom-Dat verbs in Icelandic. Thus, the study makes use of five verbs for each of the three argument structure classes, 15 verbs in total. For these, 200 eligible instances are extracted for each lemma, resulting in a total of 3,000 observations. We show that the baseline established for ordinary Nom-Dat verbs is also upheld for classical Dat-Nom verbs in Icelandic, while alternating Dat-Nom/Nom-Dat verbs deviate substantially from this baseline.

This paper is organised as follows: in Section 2 we present our object of study, including an overview of the three verb classes selecting for the Nom-Dat construction, the Dat-Nom construction and the alternating Dat-Nom/Nom-Dat constructions. We then present our hypothesis in Section 3, that the three types of verbs show variation in word order distribution depending on which argument structure construction they select for. Section 4 gives an overview of the methodology applied, whereas Section 5 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. Section 6 summarises the main content and conclusions of the paper.

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 et al. 1985, Sigurðsson 1989, Jónsson 1996, *inter alia*). For these so-called quirky or oblique subjects, at least the following nine subjecthood diagnostics have been identified (Andrews 1976, Thráinsson 1979, Zaenen et al. 1985, Sigurðsson 1989, Jónsson 1996, Barðdal 2001, Barðdal 2006, 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 reflexivization
- clause-bound reflexivization
- 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). As such, these tests confirm the status of oblique subjects as *behavioural* subjects in Icelandic (see the references listed above for a more detailed discussion). In this paper we wish to lend corpus-based support to the first and the third test in the bulleted list above, 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 1997–98, Barðdal 1999, 2001, Platzack 1999, Sigurðsson 2006, Rott 2013, 2016, Wood & Sigurðsson 2014, Barðdal, Eythórsson & Dewey 2014, 2019, *inter alia*). In this paper, we either refer to them as **alternating Dat-Nom/Nom-Dat verbs**, or as ***nægja*-verbs** (due to the even distribution of frequencies below).

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, 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.

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

b. *Þessi skýring hafði nægt Gunnari.*
this.NOM explanation.NOM had sufficed Gunnar.DAT
'These explanations were not sufficient for Gunnar'

What speaks against a simple topicalisation analysis of the examples above 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, 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.

Because of their dyadic nature, Barðdal (2001) 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 licences a nominative subject and a dative object, and a Dat-Nom construction that licences a dative subject and a nominative object. Our approach is fully in line with this analysis, and we subscribe to the view that the subject of alternating predicates is constructionally determined.

Interestingly, not all Dat-Nom verbs allow for alternation, as is already mentioned above. Some, such as *líka* 'like' only licence 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 the following examples:

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

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 (4b) is infelicitous because the subject *barninu* 'the child' and the conjugated verb *hafði* 'had' have been separated from one another by the past participle *líkað* 'liked'. In case the nominative is realised preverbally for information-structural reasons, the dative, being the syntactic subject, breaks open the verbal group and is once again reunited with the conjugated verb:

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

The example in (5) is topicalisation and not neutral word order; that is, it is a topicalisation construction 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 that, like *líka*, 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 paper. In construction grammar terms, it can thus be stated that the default argument structure construction *líka*-verbs occur in is the Dat-Nom construction, and that the linear nominative-first order is only used for information-structural purposes (Barðdal, Eythórsson & Dewey 2019).

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 (6a) and the ungrammaticality of (6b) below:

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

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

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 (6b–c).

3 Hypotheses

In this study we endeavour to lend corpus-based statistical support to the analysis that the dative 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. Thus, our aim is to corroborate Thráinsson's (2007: 21) claim that Icelandic is a subject-first language, and that this inclination is not sensitive to case marking.

As is already pointed out above, Icelandic, like several other languages, allows for a constituent other than the subject to be fronted to initial position for information-structural purposes, a phenomenon also known as topicalisation, However, 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 it is based on acceptability judgements and not corpus frequencies. Nevertheless, empirical studies on how frequent topicalisation actually is, are quite scarce.

One study that does include frequency counts, is Callegari & Ingason (2021). In their diachronic investigation of matrix-clause ditransitive constructions, they explore object topicalisation in 12th to 21st century Icelandic texts, 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 21st-century data, both direct objects as well as indirect objects are each topicalised approximately 6% of the time.

It is unclear if the predicates in our study are equally permissive of topicalisation as Callegari & Ingasson's (2021) ditransitive verbs are. For that reason, we map out word order preferences for the *hjálpa*-class and use these counts as a first baseline against which word order preferences for the *líka*- and the *nægja*-classes will be measured. Our expectations regarding word order preferences for *hjálpa*-verbs are captured in Hypothesis 1:

H1 Verbs of the *hjálpa-type* are hypothesised to show a strong preference for the Nom-Dat linear order, as they select for the Nom-Dat argument structure construction. This means that they generally realise the behavioural subject, which is encoded in the nominative, in clause-initial position.

Mutatis mutandis, the same prediction is expected to hold for non-alternating Dat-Nom verbs, which is captured in Hypothesis 2:

H2 Verbs of the *líka*-type are hypothesised to show a strong preference for the Dat-Nom linear order, as they select for the Dat-Nom argument structure construction. This means that they generally realise the behavioural subject, which is encoded in the dative, in clause-initial position.

It has already been pointed out that *nægja*-verbs constitute somewhat of an intermediate category between *hjálpa*- and *líka*-verbs, as both of their core arguments pass the subject tests. Therefore, this class of verbs is expected to deviate significantly from the baseline set by either the *hjálpa*- or the *líka*-class. This expectation is captured in Hypothesis 3:

H3 Verbs of the *nægja*-type are hypothesised to show a significantly less skewed preference for either the Nom-Dat linear order or the Dat-Nom linear order, as they are hypothesised to be able to instantiate both the Dat-Nom and the Nom-Dat argument structure constructions. As subjecthood is constructionally determined, both the nominative as well as the dative are expected to occur in clause-initial position with notable frequency.

We now turn to a description of our methodology, before we present our findings in Section 5 below.

4 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 1997–98, Barðdal 1999: 89, 2001: 53–58). Each category contains five verb types:

(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* ‘think, find, seem’
(iii) Alternating Dat-Nom/Nom-Dat verbs: *duga* ‘suffice, be enough’, *dyljast* ‘be hidden to sby, be aware’, *endast* ‘last’, *henta* ‘suit, befit’, *nægja* ‘be enough, be sufficient’

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

One or more files have subsequently been downloaded of 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. 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, 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.

All tokens have been annotated for the following variables: case, (pro)nominality, pronoun type (if applicable), referentiality, person, number, definiteness, animacy, and length, although only the first three are investigated in this study. Each of these three is discussed in turn below:

(i) Case: **nominative** or **dative**
(ii) (Pro)nominality: **pronoun** (*þú* 'you', *ykkur* 'you' 2p.acc/dat.pl, *einhverjum* 'some') or full **NP** (*Ísland* 'Iceland', *ýmsir þingmenn* 'some congressmen', *bókin* 'the book')
(iii) Pronoun type: **personal** (*ég* 'I', *hann* 'hann', *þeir* 'they' 3p.m), **demonstrative** (*þessi* 'this', *hinum* 'the other', *slíkur* 'such'), **indefinite** (*öllum* 'all', *engum* 'no-one', *báðum* 'both'), or **reciprocal** (*hvert öðru* 'each other' neut., *hver annarri* 'each other' fem.). Reflexives are excluded from study, as they are hypothesised to prefer the post-verbal slot. In line with Heylen (2005: 103), conjoined pronouns are also excluded, as they arguably lose their pronominal status

We now turn to our findings and a discussion thereof.

5 Results and discussion

In the following three subsections, 5.1, 5.2 and 5.3, we present our findings for each of the three verb classes. We start with *hjálpa*-verbs, to establish a baseline for ordinary Nom-Dat verbs, from there proceeding towards *líka*-verbs, also to establish a baseline but this time for classical Dat-Nom verbs in Icelandic. In 5.3 we then compare the statistics for alternating *nægja*-verbs with the baselines established for *hjálpa*- and *líka*-verbs in Icelandic.

5.1 Non-alternating Nom-Dat verbs

In the first section below, we give an outline of our findings with *hjálpa*-verbs in general. We discuss our findings for two different configurations, namely when both arguments are full NPs as opposed to when both arguments are pronouns. Finally, we summarise our conclusions for *hjálpa*-verbs in Section 5.1.4.

5.1.1 General findings

As is evident from Table 1, *hjálpa*-verbs show a very robust preference for the Nom-Dat linear order across configurations: no less than 989 out of 1,000 tokens prefer the nominative argument to precede the dative, rather than the other way around. The verbs *líkjast* 'resemble', *þakka* 'thank', and *treysta* 'trust' are absolute in this respect, as they do not yield a single Dat-Nom token.

Table 1. Ordinary Nom-Dat verbs across word order configuration

	Nom-Dat		Dat-Nom	
Verb	**N**	**f**	**N**	**f**
hjálpa	199	99.5%	1	0.5%
líkjast	200	100%	0	0%
mótmæla	190	95%	10	5%
treysta	200	100%	0	0%
þakka	200	100%	0	0%
Total	989	99%	11	1%

The only two verbs with which the Dat-Nom linear order is found are *hjálpa* 'help' (one attestation) and *mótmæla* 'contradict' (ten attestations). Interestingly, the one Dat-Nom attestation for *hjálpa*, shown in (7) below, comes from a biblical text (most likely a translation), which, due to its inherently archaic style, underlines its particular status.

(7) *Hann aumkast yfir bágstadda og snauða, og fátækum hjálpar hann.*
he takes.pity over the.needy and the.impoverished and the.poor.DAT helps he.NOM
'He takes pity on the deprived and the impoverished, and the poor he helps.'

Also, for *mótmæla*, the Dat-Nom linear order seems to represent topicalisations, i.e. a word order pattern that allows the canonical order of constituents to be inverted to signal a constituent's pragmatic salience. This can also be deduced from the fact that all datives in clause-initial position are either demonstrative pronouns (e.g. *því* 'that'; six tokens), or definite NPs (e.g. *þessu fólki* 'these people'; four tokens), as is shown in (8a–b), respectively:

(8) a. *... en því mótmælti Endurvinnslan.*
but that.DAT objected recycling.company.NOM
'... but to that the recycling company objected.'

b. *Þessu fólki mótmælti ég hvar sem ég gat.*
these.DAT people.DAT opposed I.NOM where which I could
'To these people, I objected wherever I could.'

We, thus, conclude that the overwhelming number of attestations of the Nom-Dat linear order corroborates the assumption that this word order represents neutral word order for *hjálpa*-verbs. As such, these findings confirm the already established fact in Icelandic that *hjálpa*-verbs indeed select for the Nom-Dat argument structure construction.

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

Table 2 presents an overview of the word order variation (or rather the lack thereof) in the [NP-V-NP] configuration with *hjálpa*-verbs. The general rule in the [NP-V-NP] configuration is to realise the dative postverbally, as the examples with *hjálpa* 'help' and *líkjast* 'resemble' in (9) below show:

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

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

Table 2 reveals an overwhelming tendency towards the Nom-Dat linear order, which follows naturally from the heavily skewed frequencies for *hjálpa*-verbs in general, as discussed above, but the data still reveal two noticeable trends.

Table 2. Ordinary Nom-Dat verbs in the [NP-V-NP] configuration

	Nom-Dat		**Dat-Nom**	
Verb	**N**	**f**	**N**	**f**
hjálpa	25	100%	0	0%
líkjast	125	100%	0	0%
mótmæla	98	98%	2	2%
treysta	31	100%	0	0%
þakka	55	100%	0	0%
Total	334	99%	2	1%

First, nominal frequencies in the [NP-V-NP] configuration are generally very high; there are never fewer than 25 attestations per verb, and their total number across all five verbs amounts to 336, which is equal to approximately one third of all the tokens collected for this verb class. Thus, our findings for *hjálpa*-verbs in the double-NP configuration can be considered to be very robust.

Secondly, it is worth noting that the [NP-V-NP] configuration seems to further amplify the inclination of these verbs towards the Nom-Dat linear order; again the sole verb that (marginally) allows datives in initial position is *mótmæla* 'contradict' with the following two tokens:

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

b. *Þeirri fyrirhuguðu málsmeðferð mótmæltu ýmsir þingmenn ...*
the.DAT intended.DAT procedure opposed some.NOM parliamentarians
'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 is indefinite. Such an asymmetry is

undoubtedly conducive to an inversion of the canonical order of constituents (cf. Siewierska 1993, Lambrecht 1994, 2000, Gregory & Michaelis 2001, *inter alia*).

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

Word order preferences in the [Pro-V-Pro] configuration, as they are presented in Table 3, constitute a near-perfect copy of the results presented in Tables 1–2 above.

Table 3. Ordinary Nom-Dat verbs in the [Pro-V-Pro] configuration

	Nom-Dat		**Dat-Nom**	
Verb	**N**	**f**	**N**	**f**
hjálpa	68	100%	0	0%
líkjast	6	100%	0	0%
mótmæla	25	93%	2	7%
treysta	83	100%	0	0%
þakka	56	100%	0	0%
Total	238	99%	2	1%

With the exception of *mótmæla*, all *hjálpa*-verbs tend entirely towards the Nom-Dat linear order. Interestingly, the only two attestations of the Dat-Nom linear order contain a dative demonstrative pronoun in combination with the nominative personal pronoun *ég* 'I', which again clearly points towards an effect of topicality. We demonstrate this with one example of each of the two configurations below, the Nom-Dat order in (11a) and the Dat-Nom order in (11b):

(11) a. *Ég mótmælti þessu og benti á að ...*
I.NOM objected this.DAT and pointed on that
'I objected to this and pointed out that ...'

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

5.1.4 Interim conclusion

The evidence presented in this section is fully in line with the prediction that Icelandic possesses a class of Nom-Dat verbs, as the Nom-Dat linear order is attested in 989 out of 1,000 times across configurations (i.e. 99%), and 334 out of 336 times in the [NP-V-NP] configuration in particular (i.e. 99.5%). Hypothesis 1 is thus borne out.

Furthermore, our data show that object topicalisation in Icelandic is very rare; it is mostly associated with pronominality (nine out of 11 cases), but the verb *mótmæla* also, albeit

marginally, allows for topicalisation in the [NP-V-NP] configuration (two out of 11 cases). Interestingly, topicalisation is much less frequent in our dataset than in the aforementioned Callegari & Ingason (2021) study, as their data for the 21st century seem to reveal that both direct objects as well as indirect objects allow for topicalisation approximately 6% of the time. In other words, their ditransitive verbs seem to allow for topicalisation more readily than the *hjálpa*-verbs in our study.

We do not know on how many verbs Callegari & Ingason base their topicalisation study, except that it involves all the ditransitive verbs found in the IcePaHC corpus, spanning from the 12th to 20th century Icelandic. Yet, our goal here is not to draw definite conclusions about the prevalence of topicalisation in Icelandic in general, but only to set a baseline for Nom-Dat verbs, for a comparison with alternating Dat-Nom/Nom-Dat verbs below. This is the reason why we compare our numbers with Callegari & Ingason's numbers. We also believe that five Nom-Dat verbs occurring in a set of 1,000 tokens in total is adequate to establish such a baseline for this narrowly defined verb class.

Thus, we conclude that the evidence presented here firmly shows: (i) that Icelandic has an overwhelming tendency for nominative subjects to precede dative objects in the linear order, (ii) that this effect not only plays out at the level of individual verbs, but also that there is an overarching verb class effect, and (iii) that pronouns only mildly swing a verb's preference for a given linear order of constituents (see, however, discussion below).

5.2 Non-alternating Dat-Nom verbs

In this subsection we turn to word order preferences for *líka*-verbs, i.e. verbs selecting for the Dat-Nom argument structure construction. The obvious question is whether dative subjects show the same tendency as nominative subjects with *hjálpa*-verbs to occur in initial position. Thus, these findings constitute the second baseline against which we compare our findings for alternating Dat-Nom/Nom-Dat verbs. We start with an overview of our general results, before we present the two configuration-specific findings involving full NPs vs. pronouns. A special discussion of the effect of demonstratives in the nominative case is also included.

5.2.1 General findings

The frequencies obtained for the non-alternating Dat-Nom verbs in our sample virtually mirror the ones found for the Nom-Dat verbs: out of 1,000 attestations, 931 realise the dative argument in initial position. We take this overwhelming tendency for these verbs to show up with the Dat-Nom linear order to suggest that the Dat-Nom order is, indeed, the neutral word order for this verb class. Table 4 presents an overview of the individual frequencies per verb type.

The fact that *líka*-verbs generally show a very strong inclination towards the Dat-Nom linear order indeed corroborates the assumption that these verbs select for the Dat-Nom argument structure construction, as is already established in the literature on Icelandic syntax. However, the total number of tokens showing the inverted order of constituents is remarkably higher than for Nom-Dat verbs, as the Nom-Dat linear order is attested 69 times (i.e. 7%), as opposed to only 11 attestations of the Dat-Nom linear order for *hjálpa*-verbs (i.e. 1%).

Table 4. Non-alternating Dat-Nom verbs across word order configurations

	Nom-Dat		Dat-Nom	
Verb	**N**	**f**	**N**	**f**
áskotnast	3	1.5%	197	98.5%
blöskra	1	0.5%	199	99.5%
leiðast	7	3.5%	193	96.5%
líka	7	3.5%	193	96.5%
þykja	51	25.5%	149	74.5%
Total	69	7%	931	93%

What is remarkable about these results is the way in which they relate to the numbers presented in the aforementioned study by Callegari & Ingason (2021). Recall that their 21st-century data show topicalisation to occur approximately 6% of the time, both for direct objects and indirect objects. These findings tie in nicely with what we find for Dat-Nom verbs in general, which topicalise the nominative argument approximately 7% of the time (see Table 4), yet they differ starkly from what we attest for Nom-Dat verbs, which topicalise the dative only 1% of the time (see Table 1).

Regardless of the differences between how often the dative of Nom-Dat verbs is topicalised as opposed to the nominative of Dat-Nom verbs, the 7% mean for *líka*-verbs mentioned above is inflated considerably by the high number of Nom-Dat attestations for *þykja* 'think, find, seem' (25.5%). Recalculating the frequencies, without the outlier *þykja*, the number of Nom-Dat attestations for *líka*-verbs drops to 2.25%, which is markedly less than the 6% of object topicalisation Callegari & Ingason documented for their dataset. Also, when zooming in on the high share of Nom-Dat word order attestations with *þykja*, it is striking that 49 out of 51 Nom-Dat tokens found with this verb have their nominative slot filled by either a definite pronoun (41 tokens), or a definite full NP (eight tokens), configurations which are shown in (12a–b), respectively. An array of studies have shown definiteness and pronominality to be key factors in word order variation (cf. Siewierska 1993, Lambrecht 1994, 2000, Gregory & Michaelis 2001, *inter alia*). The verb *þykja* is clearly particularly sensitive to this tendency.

(12) a. *Það þótti honum óskaplega mikið varið í ...*
that.NOM thought he.DAT incredibly much worthy in
'That (which) he felt was extremely worthy ...'

b. *Þetta hey þótti kúnum gott ...*
this.NOM hay.NOM found cows.the.DAT good
'This hay, the cows like ...'

Zooming in further on the Nom-Dat tokens with pronominal nominatives, another remarkable tendency surfaces, again with *þykja*: 37 out of 41 tokens are demonstrative pronouns. This finding is reminiscent of the tendency discussed in Section 5.1.1 above for the verb *mótmæla* 'contradict', which is marginally found in the Dat-Nom linear order, mostly when the dative is a demonstrative pronoun. This is shown in example (13) below:

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

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.

It is also worth pointing out that the results presented in Table 4 are very much in line with Rott's (2013) empirical analysis of four Icelandic *líka*-verbs, viz. *blöskra* 'be shocked, be horrified', *gremjast* 'resent, be annoyed', *líka* 'like', and *leiðast* 'be bored', for which he found that the dative argument was realised preverbally 162 times (96%), but postverbally only seven times (4%). Rott does not include any frequencies for individual verbs, but since his results are equally skewed as ours, it is reasonable to assume that the verb class effect he uncovers may also be dependent on individual verb effects. Recall that it is case marking and argument structure that motivates our verb class categorisation, not the behaviour of individual verbs.

Finally, 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.

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

In more than one way, the figures presented in Table 5 constitute the mirror image of those presented in Table 2. First, all *líka*-verbs show a very robust preference for the Dat-Nom linear order, which corroborates the existing analysis of these as being non-alternating Dat-Nom verbs; only *þykja* returns one token in which the canonical order of constituents is inverted. This example, which has already been discussed below as (12b), is here repeated as (14):

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

Thus, the variance observed in Table 4 is almost non-existent in Table 5, which, again, confirms the status of *líka*-verbs as unequivocal non-alternating dative-subject predicates.

Secondly, both the number of attestations per verb as well as the total number of tokens in the [NP-V-NP] configuration in general is quite high, which means that the proportional frequencies for this verb class in this configuration (one vs. 193 tokens) are as such both trustworthy and reliable.

Table 5. Non-alternating Dat-Nom verbs in the [NP-V-NP] configuration

	Nom-Dat		Dat-Nom	
Verb	**N**	**f**	**N**	**f**
áskotnast	0	0%	48	100%
blöskra	0	0%	68	100%
leiðast	0	0%	26	100%
líka	0	0%	28	100%
þykja	1	4%	23	96%
Total	1	99%	193	1%

Finally, *líka*-verbs, exactly like *hjálpa*-verbs, not only show a strong verb effect in the [NP-V-NP] configuration, but also a robust verb *class* effect, since all verbs prefer the Dat-Nom linear order in equal manner. This shows, once again, that these verbs not only instantiate the Dat-Nom argument structure construction but also that they instantiate only that argument structure and not the Nom-Dat one.

5.2.3 The effect of nominative demonstratives

It has already been pointed out in Section 5.2.1 above that the skewed general frequencies for *þykja* are largely due to the influence of nominative demonstratives. Therefore, it seems worth investigating to what extent the Nom-Dat linear order for *líka*-verbs in general is associated with nominative demonstratives. In order to do so, let us briefly revisit the nominal frequencies capturing the prevalence of the Nom-Dat linear order with these verbs, presented in Table 4, in order to compare them with the number of Nom-Dat attestations containing nominative demonstratives in particular. These numbers are presented in Table 6.

Table 6. Non-alternating Dat-Nom verbs occurring with nominative demonstratives compared to the total number of Nom-Dat attestations

	Nom_{dem}-Dat	Nom-Dat	
Verb	**N**	**N**	**f**
áskotnast	0	3	0%
blöskra	0	1	0%
leiðast	5	7	71%
líka	6	7	86%
þykja	37	51	73%
Total	48	69	70%

Table 6 shows that, for non-alternating Dat-Nom verbs of the *líka* type, the Nom-Dat linear order is indeed strongly associated with nominative demonstratives: out of 69 attestations, 48 contain the demonstratives *það* 'it' or *þetta* 'that'. With the exception of *áskotnast* and *blöskra*, which are generally not found with the Nom-Dat linear order anyway, proportional frequencies are relatively evenly distributed across types, ranging from 71% for *leiðast* to 86% for *líka*. Even though the total numbers for *leiðast* and *líka* are low, it seems clear that nominative

demonstratives trigger the use of the topicalisation construction, as opposed to the neutral word order found with Dat-Nom verbs.[1]

Now that we have established how permissive the initial slot is of nominative demonstratives in the Nom-Dat linear order, let us compare these numbers to the prevalence of nominative demonstratives in the second slot, given in Table 7. Thus, we repeat the numbers from the first column in Table 6, also occurring in the first column in Table 7. The results are, as a matter of fact, quite remarkable. First, Table 7 shows that nominative demonstratives are not uniquely bound to clause-initial position. In fact, nominative demonstratives are far more common in the Dat-Nom linear order than in the Nom-Dat linear order, as the former is attested with nominative demonstratives 136 times, but the latter only 48 times. Thus, the post-verbal position is still more strongly associated with nominative demonstratives than the preverbal position.

Table 7. Non-alternating Dat-Nom verbs in instances involving nominative demonstratives

	Nom$_{dem}$-Dat		**Dat-Nom$_{dem}$**	
Verb	**N**	**f**	**N**	**f**
áskotnast	-	-	-	-
blöskra	0	0%	12	100%
leiðast	5	17%	25	83%
líka	6	9%	61	91%
þykja	37	49%	38	51%
Total	48	26%	136	74%

Secondly, and perhaps more interestingly, the degree to which nominative demonstratives tend to occupy initial position seems to be verb-dependent, with some verbs allowing nominative demonstratives in postverbal position only (*blöskra*), some allowing them to take initial position only marginally (*leiðast*, *líka*), and some allowing them to occupy either slot more or less equally often (*þykja*). These frequencies may be the result of interaction effects between the argument slots that only a more in-depth statistical analysis can reveal, which again means that these results are clearly in want of further investigation.

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

Turning to the [Pro-V-Pro] configuration in general, Table 8 summarises the results obtained for this configuration with *líka*-verbs. In total, the Nom-Dat linear order is attested 44 times (20%), and the Dat-Nom linear order 183 times (80%). As was mentioned in Section 5.2.1, the former is almost uniquely associated with nominative demonstratives: 43 out of 44 tokens occurring with the Nom-Dat linear order are headed by the pronouns *það* ‘it’ or *þetta* ‘that’.

[1] Johan Brandtler points out to us (p.c.) that Swedish *tycka* ‘think, believe’ shows a similar pattern, in that an object pronoun is more natural in first position than the subject, especially if the object pronoun refers to a clause, e.g. *Det tycker jag också* vs. *Jag tycker det också* ‘I also believe that’. Note that Swedish *tycka* and Icelandic *þykja* are cognates, so the question arises whether this may be a very old word order pattern with this verb. Even so, it remains to be investigated whether topicalisation of object pronouns is somehow enabled by specific verb semantics.

What this essentially means is that only nominative demonstratives are able to significantly swing a verb's inherent word order preference, and that even non-alternating verbs are not immune to their influence.

Table 8. Non-alternating Dat-Nom verbs in the [Pro-V-Pro] configuration

	Nom-Dat		Dat-Nom	
Verb	**N**	**f**	**N**	**f**
áskotnast	1	33%	2	67%
blöskra	0	0%	11	100%
leiðast	5	11%	39	89%
líka	6	8%	71	92%
þykja	32	35%	60	65%
Total	44	19%	183	81%

Zooming in on tokens with two *personal* pronouns (not singled out in Table 8), another interesting tendency surfaces: two personal pronouns are attested 64 times, and only once (2%) do they prefer the Nom-Dat linear order over the Dat-Nom word order (98%). This result is telling as it stands in stark contrast with the results for alternating predicates in contexts with two personal pronouns, as they almost invariably realise the Nom-Dat word order pattern (see Section 5.3.3 below).

5.2.5 Interim conclusion

The results presented in this section are indicative of several tendencies. First, we have corroborated with corpus frequencies the established analysis that Icelandic indeed possesses a class of Dat-Nom verbs whose non-canonically case-marked subject is very strongly associated with the preverbal slot: dative subjects take initial position in 931 out of 1,000 tokens across all configurations. In the [NP-V-NP] configuration, the Dat-Nom linear order is attested even more frequently, showing up in 193 out of 194 tokens (99.5%). Hypothesis 2 is thus confirmed.

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.

Finally, the inverted order of constituents, involving topicalisation, is mostly brought about by nominative demonstratives (48 out of 69 tokens), but this tendency is essentially unidirectional, as nominative demonstratives also (and, in fact, more frequently) occur in postverbal position. Why some verbs are more permissive of clause-initial nominative demonstratives than others is a question that remains unanswered at present.

5.3 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 follows that of sections

5.1 and 5.2 above: we first discuss the general findings, after which we turn to word order variation in the [NP-V-NP] configuration, and finally, the [Pro-V-Pro] configuration. The results are compared to the baseline set by Nom-Dat *hjálpa*-verbs and Dat-Nom *líka*-verbs. The main implications and conclusions are discussed in 5.3.4.

5.3.1 General findings

The results for the class of *nægja*-verbs, as they are presented in Table 9, 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.

Table 9. Alternating verbs across word order 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%

Upon closer inspection, the data 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 (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 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.

5.3.2 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 10.

Table 10. Alternating verbs in the [NP-V-NP] configurations

	Nom-Dat		**Dat-Nom**	
Verb	**N**	**f**	**N**	**f**
duga	33	79%	9	21%
dyljast	2	25%	6	75%
endast	9	30%	22	70%
henta	86	100%	0	0%
nægja	27	54%	23	46%
Total	157	72%	60	28%

The frequencies in Table 10 are indicative of several different tendencies. First, frequencies in the [NP-V-NP] configuration are much less skewed than for Nom-Dat verbs or non-alternating Dat-Nom verbs, thereby confirming the generally alternating nature of Dat-Nom/Nom-Dat verbs. A chi-squared goodness-of-fit test comparing both word orders across verbs yields a highly significant result ($X^2 = 43.36$; df = 1; $p_{\text{two-tailed}} < 0.001$), which should be interpreted as a statistical indication that the distribution of both word orders cannot be attributed to chance. This does not mean that the verbs in question do not alternate, but rather that there are factors guiding the alternation that have yet to be uncovered.

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 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 (15a–b) below:

(15) 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 ...'

b. *En skotfærasafnið hans hefði nægt hverri meðal herdeild.*
but munition.collection.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 9 are evidently replicated in the [NP-V-NP] configuration, 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, future 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 9. 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. It is evident that a more in-depth analysis of this class of verbs is needed in order to lay out a more detailed picture of the alternation and the degree to which every factor impacts the competition between the two diametrically opposed argument structure constructions.

Finally, our findings for alternating verbs in the [NP-V-NP] configuration tie in nicely with Allen's (1995: 108) study on 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 attestations vs. 19 attestations). This certainly suggests that Allen's Dat-Nom verbs are indeed alternating verbs, as Allen (1995: 116) herself assumes. 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. This is evidently not a trivial matter, as if we were to remove *henta* from our sample on the assumption that it is not an alternating verb, the four remaining verbs would together instantiate the Nom-Dat linear order 71 times, and the Dat-Nom linear order 60 times. If one consequently fails to break these numbers down and present verb type-specific counts, as we have done, one obscures any verb-specific tendencies, thereby creating the impression that all verbs occur in either construction approximately equally often.

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

Table 11 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 (16) below, while examples of the more abundant Nom-Dat word order are given in (17):

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

b. *Henni duldist það ekki að ...*
she.DAT be.aware it.NOM not that
'She was aware that ...'

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

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

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

c. *Og vonuðumst til að þeir mundu endast okkur.*
and hoped to that they.NOM would last us.DAT
'And hoped that they would last us.'

Table 11 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).

Table 11. Alternating predicates 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%

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 construction (not singled out in Table 11). 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. Thus, with tokens containing two personal pronouns, alternating verbs clearly behave as Nom-Dat verbs, and not as Dat-Nom verbs, as the latter tend almost uniquely towards the Dat-Nom linear order across configurations.

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. 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, see next section for a further discussion).

5.3.4 Interim conclusions

The findings presented in this section show that Icelandic indeed possesses a class of alternating Dat-Nom/Nom-Dat verbs. This is evident from the different behaviour of *nægja*- vs. *líka*-verbs documented above. For instance, 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, which is very different from both *líka*- and *hjálpa* verbs. For these two verb classes, it is clear that the Dat-Nom and the Nom-Dat linear orders represent neutral word order, as 99,5% of all instances involving full NPs show up with the Dat-Nom vs. the Nom-Dat linear order respectively, as is shown in Table 12. 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.

Furthermore, Table 12 also shows that the numbers for alternating Dat-Nom/Nom-Dat verbs deviate significantly from the baseline set by *hjálpa*- and *líka*-verbs. 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. Hypothesis 3 is therefore largely borne out that there are two neutral word orders for *nægja*-verbs, and thus that these verbs may instantiate both the Dat-Nom and the Nom-Dat argument structure constructions. In contrast, *líka*-verbs only instantiate the Dat-Nom argument structure construction, with the Nom-Dat linear order representing topicalisation, and vice versa for *hjálpa*-verbs.

Table 12. 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%

We now turn to the question asked in Section 1, namely which factors determine the speakers' choice of one of the two argument structure constructions, Dat-Nom or Nom-Dat, over the other with alternating verbs. We have shown here that alternating Dat-Nom/Nom-Dat verbs are much more sensitive to the distinction between nominal and pronominal influence than *hjálpa*- and *líka*-verbs are. There is thus no doubt that for the [Pro-V-Pro] configuration, alternating verbs instantiate the Nom-Dat construction to a much greater degree than *líka*-verbs, which in turn make extensive use of the topicalised Nom-Dat linear order, as is evident from Table 13.

Taking a closer look at the proportions between the three verb classes, as represented in Table 13, *nægja*-verbs instantiate the Nom-Dat argument structure construction in 95% of the

cases in which the two arguments are lexically realised as pronouns. Corresponding numbers for *hjálpa*- and *líka*- verbs are 99.2% vs. 20%, respectively. That *hjálpa*-verbs show a 99.2% prevalence for the Nom-Dat linear order is, of course, expected since the Nom-Dat linear order represents neutral word order for *hjálpa*-verbs. Hence, these instances simply represent the ordinary Nom-Dat argument structure construction for these verbs with subject status and topicality coinciding in one argument, the nominative. Thus, the really interesting comparison to be carried out here is between classical Dat-Nom verbs of the *líka*-type and alternating Dat-Nom/Nom-Dat verbs of the *nægja*-type.

Table 13. Nom-Dat vs Dat-Nom linear order in the [Pro-V-Pro] configuration for *hjálpa*-, *líka*-, and *nægja*-verbs, and for *nægja*-verbs excluding *henta*

	Nom-Dat	**Dat-Nom**
hjálpa-verbs	99.2%	0.8%
líka-verbs	20.0%	80.0%
nægja-verbs	94.4%	5.6%
nægja-verbs (excluding *henta*)	95.0%	5.0%

For *líka*-verbs, as much as 20% of the instances in the [Pro-V-Pro] configuration are topicalised Nom-Dat structures, while 95% of the instances with *nægja*-verbs in the same configuration instantiate the Nom-Dat argument structure construction. These numbers simply show that when both arguments of *nægja*-verbs are pronouns, the Dat-Nom argument structure construction is more or less excluded. The same cannot be said about *líka*-verbs with which the Dat-Nom argument structure construction is employed in 80% of the cases where two pronouns are involved. Moreover, as is discussed in the preceding section, when the two pronouns are both lexically realised as **personal** pronouns, in 81 out of 88 cases, or in 92%, the Nom-Dat argument structure construction is chosen over the Dat-Nom one.

These facts tie in with Barðdal's (2001: 65) claim that discourse factors, or more closely topicality, really is the issue when Icelandic speakers choose between the two argument structure constructions. That is, they choose the Dat-Nom construction when the dative is topical and the Nom-Dat construction when the nominative is topical, except for when both arguments are realised as pronouns, including personal pronouns. In such cases, the nominative clearly takes precedence over the dative, irrespective of whether the nominative is in the 1st, 2nd or 3rd person (cf. discussion in Barðdal & Eythórsson 2003).

Another compelling result yielded by this study of alternating verbs concerns *henta* and its categorical behaviour as a Nom-Dat verb. This 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.

One study that has found homogeneous results for Icelandic alternating verbs, thus corroborating their status as an actual verb class with uniform properties, is Bornkessel-Schlesewsky et al. (2011). They were 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 Roehm et al. (2007) and Rott (2013), it is unclear which exact verb types this study is based on, so that it is difficult to gauge the scope of these findings. Nevertheless, the uniform

electrophysiological response Bornkessel-Schlesewsky et al. were able to elicit seems to mitigate the conclusion that alternation might be a gradient property.

6 Summary and conclusions

In this paper we have succeeded in lending empirical support to the claim that behavioural subjects in Icelandic are strongly tied to clause-initial position, and that this tendency is not sensitive to case marking. For this purpose, we have extracted 200 examples of 15 verbs from the Icelandic Web 2020 corpus, 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'.

Our dataset is annotated for three variables, (i) case marking, (ii) (pro)nominality, i.e. whether the arguments are full NPs or pronouns, and (iii) the type of pronoun. This, in addition, of course, to verb class. We also put forward three hypotheses:

(i) That ordinary Nom-Dat verbs like *hjálpa* show a strong preference for the Nom-Dat linear order, since these verbs instantiate the Nom-Dat argument structure construction
(ii) That classical Dat-Nom verbs like *líka* show a strong preference for the Dat-Nom linear order, since these verbs instantiate the Dat-Nom argument structure construction
(iii) That alternating Dat-Nom/Nom-Dat verbs like *nægja* show a much less skewed preference for either of the two linear orders, since these verbs may instantiate either the Dat-Nom or the Nom-Dat argument structure constructions

We first establish a baseline for ordinary Nom-Dat verbs, i.e. the *hjálpa*-verbs, with two full NPs. For this configuration, the nominative subject is realised clause-initially 99.5% of the time. For classical Dat-Nom verbs, i.e. *líka*-verbs, the dative subject is also realised clause-initially in 99.5% of the time in the same configuration, i.e. when both arguments are full NPs. However, *líka*-verbs' propensity for the Dat-Nom construction may occasionally be swung by nominative demonstratives and definite NPs.

In contrast, for alternating Dat-Nom/Nom-Dat verbs, i.e. *nægja*-verbs, our findings generally confirm that subjecthood is constructionally determined. 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 *hjálpa*- or *líka*-verbs. There are, however, considerable 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. There is one particular verb, *henta* itself, as a matter of fact, 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. However, when recalculating the numbers for full NPs without the outlier, *henta*, the distribution amounts to 54% Nom-Dat vs. 46% Dat-Nom. These facts are in line with Barðdal's (2001) claims that the choice of construction is determined by the topicality of the two arguments, that the Nom-Dat construction is used when the nominative is topical and the Dat-Nom construction when the dative is topical.

There is one major exception to the above-mentioned distribution of alternating verbs across the two argument structure constructions, and this involves cases where the two arguments are pronouns, including personal pronouns. In such cases, the Nom-Dat construction takes a clear precedence over the Dat-Nom construction, as is also discussed by Barðdal & Eythórsson (2003) for Icelandic and Allen (1995) for Old English. This certainly is a topic in need of further investigation.

Finally, the results we have obtained for *henta* prompt the use of experimental methods, as this verb interestingly passes all subjecthood tests for both of its nominal arguments, but in our dataset it has nevertheless been found to occur solely in the Nom-Dat construction. This raises questions about the correlation between corpus frequencies and neutral word order, which in turn motivates the use of experimental methods, as these may help establish if there is indeed a major mismatch between corpus frequencies and acceptability with regard to *henta*.

References

Allen, Cynthia L. 1995. *Case marking and reanalysis: Grammatical relations from Old to Early Modern English*. Oxford: Oxford University Press.

Andrews, Avery D. 1976. The VP complement analysis in Modern Icelandic. *North Eastern Linguistic Society* 6: 1–21.

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

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

Barðdal, Jóhanna. 1999. The dual nature of Icelandic psych-verbs. *Working Papers in Scandinavian Syntax* 64: 78–101.

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. 2022. *Oblique subjects in Germanic: Their status, history and Reconstruction.* To appear with 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 & 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.

Bernódusson, Helgi. 1982. Ópersónulegar setningar [Impersonal sentences]. University of Iceland Master's Thesis.

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.

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

Eythórsson, Thórhallur. 1995. *Verbal Syntax in the Early Germanic Languages*. Cornell University Ph.D. Dissertation.

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.

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 Ph.D. Dissertation.

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

Jónsson, Jóhannes Gísli. 1996. *Clausal architecture and case in Icelandic*. University of Massachusetts, Amherst, Ph.D. Dissertation.

Jónsson, Jóhannes Gísli. 1997–98. Sagnir með aukafallsfrumlagi [Verbs selecting for oblique subjects]. *Íslenskt mál* 19–20: 11–43.

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

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.

Platzack, Christer. 1999. The subject of Icelandic psych-verbs: A Minimalistic account. *Working Papers in Scandinavian Syntax* 64: 103–116.

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, March 29th–31st, 2007.

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 *From Africa via the Americas to Iceland*, ed. by Christel Stolz & Thomas Stolz, 215–320. Bochum: Dr. N. Brockmeyer.

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 Doctoral Dissertation.

Sigurðsson, Halldór Ármann. 2006. The nominative puzzle and the low nominative hypothesis. *Linguistic Inquiry* 37: 289–308.

Thráinsson, Höskuldur. 1979. *On complementation in Icelandic*. New York: Garland.

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

Wallenberg, Joel C., Anton Karl Ingason, Einar Freyr Sigurðsson & Eiríkur Rögnvaldsson. 2011. Icelandic Parsed Historical Corpus (IcePaHC). Version 0.9.

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.

Joren Somers
Ghent University
joren.somers@ugent.be

Jóhanna Barðdal
Ghent University
johanna.barddal@ugent.be