Promotor	Prof. dr. Jóhanna Barðdal Vakgroep Taalkunde
Copromotor	Prof. dr. Torsten Leuschner Vakgroep Taalkunde
Decaan	Prof. dr. Gita Deneckere
Rector	Prof. dr. Rik Van de Walle

voor mijn geliefde ouders, Ingrid Boone & Robert Somers

The best of all is to be loved by your father and your mother, and your family. You can't expect much more than that.

– Roald Dahl

Front cover: *Blue and Grey* (Mark Rothko, 1962)
Back cover: *No. 14* (Mark Rothko, 1960)

FACULTEIT LETTEREN
EN WIJSBEGEERTE

Alternating Dat-Nom/ Nom-Dat verbs in German and Icelandic

A corpus study

Joren Somers

Proefschrift voorgelegd tot het behalen van de graad van Doctor in de Taalkunde
2024

UNIVERSITEIT
GENT

Preface

Graag neem ik de gelegenheid te baat om mijn dankbaarheid te uiten ten aanzien van iedereen die me bij de totstandkoming van deze thesis heeft bijgestaan. Dat is in de eerste plaats mijn eerste promotor, Jóhanna Barðdal, die ik bewonder om haar positiviteit, haar gedrevenheid, haar diepgewortelde passie voor taalkunde, haar oprechte interesse in anderen en haar durf. De vanzelfsprekendheid waarmee zij de afgelopen jaren tijd voor mij vrijmaakte wanneer dat voor mij belangrijk was, vind ik bijna ontroerend. Ze heeft me er ook steeds van overtuigd dat mijn stem in het wetenschappelijke debat waarde had, waardoor ik amper te maken kreeg met *writer's block*, existentiële twijfel of eender welke andere kwaal waaraan men als schrijver wel eens gaat lijden. Daarnaast hebben we ook ontzettend veel plezier gehad samen en naar mijn gevoel zijn we echt maatjes geworden nadat we samen op congres waren geweest in Kiel, ondertussen goed anderhalf jaar geleden. Na al die jaren van intense samenwerking ben ik er ook van overtuigd geraakt dat we in feite heel goed op elkaar lijken (zoals volgens mij wel meer doctorandi op hun promotoren lijken): we zijn beide waardeloze ochtendmensen, we houden niet van betutteling, we zijn oprecht verwonderd door taalkundige eigenaardigheden, vooral als ze etymologie of syntaxis betreffen, en we werken allebei graag en veel. Jóhanna, mijn ontmoeting met jou, die ondertussen dateert van iets meer dan vijf jaar geleden, reken ik tot de belangrijkste uit mijn leven. Voor je nabijheid, je steun en je vertrouwen dank ik je oprecht. Hjartans þakkir!

Een tweede woord van dank richt ik aan mijn tweede promotor, Torsten Leuschner. Torsten is voor mij onnavolgbaar in zijn nauwgezetheid en in zijn kennis van de Duitse taalkunde en ik begreep al snel dat ik bij hem in veilige handen was. Ook heeft hij altijd met oprechte interesse mijn teksten gelezen en wist hij mijn bijwijlen barokke en onstabiele taalbouwsels om te vormen tot stevigere structuren. Als ik met hem van gedachten gewisseld had, zag ik de zaken ook steeds helderder. En ik voelde dat hij oprecht geloofde in de goede afloop van dit doctoraat, ook als ik dat zelf soms moeilijker vond. Herzlichen Dank, Torsten!

Dank ook aan Ludovic De Cuypere, mijn derde DBC-lid, die me introduceerde tot de discipline van de kwantitatieve onderzoeksmethoden en mijn analyses runde alsof het kinderspel was. Zonder zijn deskundigheid en statistische bijstand deed ik waarschijnlijk nog steeds chi-kwadraattoetsen op een kladblaadje.

Dit doctoraat zou er ook nooit gekomen zijn als de voortreffelijke Peter Lauwers al die jaren geleden geen GOA-beurs had gewonnen. Peter is een formidabel taalkundige, die als geen ander het belang kent van mildheid. Ik bedank hem voor zijn flexibiliteit toen dit onderzoek uiteindelijk een andere richting uitging dan we oorspronkelijk voor ogen hadden. En ook voor zijn open deur en voor die drie extra maanden, die een wereld van verschil maakten. In één adem wens ik ook alle andere leden van ons GOA-consortium te bedanken: Renata, Miriam, Timothy, Rob, Anne-Sophie, Mark, Anna, Sven, Margot, Emmaline, Martin, Mariia en mijn bestie, Anouk, met wie ik de afgelopen jaren geweldig veel lol heb gehad. Haar directheid en haar vele woorden van steun ("You got this!") waren zo nu en dan verlossend.

Ook de juryleden die dit doctoraat zullen beoordelen, ben ik dankbaar. In alfabetische volgorde: Hans Boas, Tanja Mortelmans, Marleen Van Peteghem en Klaas Willems. Van ieder van hen heb ik de afgelopen jaren iets geleerd, hetzij omdat ik les bij hen volgde, hetzij omdat ze me op belangrijke momenten feedback gaven, of allebei. Ik heb het grootste respect voor hen.

Daarnaast waren er ook zo vele andere collega's (van wie ik zelfs niet altijd de naam ken), maar wier vriendelijkheid en loutere aanwezigheid de Blandijn tot een gezellige werkplek maakten. Ik dank in het bijzonder mijn collega's Scandinavistiek. Zij hebben me koffie leren drinken en elke fika ondergedompeld in een warm bad Zweeds. Ik dank ook Flor voor de gezellige pauzes, die steevast een grotere hap uit onze namiddag namen dan we voor ogen hadden. Zijn taalkundig doorzicht vind ik benijdenswaardig. Ook Gil en Antoine waren er, en alle collega's Spaans. Het was altijd prettig even bij hen te verpozen, daar op de zuidflank van de Blandijn. Een speciaal woord is er voor Linde, mijn meest gelijkgezinde en best geklede collega. We hadden nooit veel woorden nodig om elkaar te begrijpen. En Geertrui, onze sfeer beherende en immer pragmatische vakgroepsecretaris. Ik mis onze gezellige babbels en persoonlijke gesprekken nu al. Natuurlijk waren er ook alle andere collega's van het ATP: Annemie, Chris, Els, Gitte, Linde, Peter, Raf, Tanja, en zovele anderen.

Het is onmogelijk een boeiend en bloeiend professioneel leven te leiden zonder dat je je daarin gesteund weet door geliefden, vrienden en naasten. Ik neem het risico zo veel mogelijk mensen bij naam te noemen en er zo misschien enkele over het hoofd te zien. De eerste tot wie ik me richt is Tom. De afgelopen jaren waren beslist niet de

gemakkelijkste, en we hebben moeten aanvaarden dat we elkaar in soms fundamentele opzichten niet kunnen geven waar we elk zo naar verlangen. Toch is onze relatie heerlijk in haar onvolmaaktheid omdat we elke dag om het hardst lachen en omdat we elkaars trouwste supporter zijn, ook al voelen we soms weerstand bij de keuzes van de ander. Tom, jij bent het die mijn wereld heeft opengebroken en die mij heeft gevormd. Je hebt me geleerd wat humor is en getoond hoe louterend theater kan zijn. Ik bewonder je en heb je lief.

Sylvester noem ik ook, en wel voor zijn aanstekelijke verwondering en zijn ontroerende vragen, waar ik vaak geen antwoord op heb. Ik bewonder hem meer dan eender welke andere persoon die ik ken.

Lieve PDL-vrienden, niets is er beter dan in jullie gezelschap in het glooiende landschap van Zuid-Engeland te vertoeven, een plek waaraan we allemaal ons hart verloren hebben. Ik hoop er nog vaak met jullie heen te trekken, want ik denk dat ik daar het gelukkigst ben. Een speciaal woord richt ik aan Lise, mijn gefelligste en meest attente vriendin. Steeds was ze er met bemoedigende woorden en nuance. En ze herinnerde me eraan dat de aanhouder wint. *Slow and steady wins the race!* Wannes en Ruben horen ook in dit rijtje thuis. Zij zijn de enige twee mensen die dit onderzoek waarschijnlijk nog beter begrijpen dan ikzelf. Als ik bij jullie ben, voel ik mij gezien en gehoord. En Laurens, die ik niet vaak genoeg zie, maar die er altijd is geweest sinds we samen kampen bouwden in de tuinen van onze ouders. Fien, jij bent er al het langst en ik hoop dat je nog minstens even lang blijft. Gaétan, jij bent er pas sinds kort en ik voel dat we nog lang niet zijn uitgepraat.

Anouk, jij zorgde ervoor dat ik de afgelopen jaren steeds uitkeek naar maandagochtend, wanneer we samen al Valeirend de week inzetten. Er zijn weinig mensen met wie ik meer tijd spendeerde dan met jou, en nog krijg ik er geen genoeg van. Heerlijk is ook mijn vriendschap met Dominique, het grootste verteltalent en met voorsprong de meest vermakelijke mens die ik ken. In jouw gezelschap smaakt de koffie het best. En Onur, Maarten, Hugo en Raf, wanneer is het volgende feestje?

Tine was er ook, en vooral op momenten waarop ik haar het meeste nodig had (al gold dat omgekeerd misschien ook wel). Je scherpe analyses en snedige reacties maken van jou een van mijn boeiendste vrienden. En Liesbet, die de perfecte combinatie van bedachtzaamheid en ondernemingsgezindheid belichaamt: of in Knossos of in Drongen, steeds wordt er hartelijk gelaën. Tot slot ook Alexia, mijn favoriete badmintonbuddy. Zij geeft de beste knuffels.

Ik vermeld ook graag mijn meter en peter, twee van de meest optimistische en veerkrachtige mensen die ik ken, en mijn schoonouders, Veerle en Marcel, die me al vlug

in hun hart sloten, net zoals ik hen in het mijne. Ook Pieter, Hans en Sara; ik denk dat we meer op elkaar lijken dan we durven toe te geven.

Het allerlaatste woord richt ik aan mijn ouders, Ingrid Boone en Robert Somers, die ik zo veel liever zie dan ik met woorden kan zeggen. Dat ik me een vol mens voel, tevreden en vrolijk, heb ik aan hen te danken. En hoe ouder ik word, hoe beter ik besef dat hun ouderlijke toewijding en onbaatzuchtigheid grenzen aan het bovenmenselijke. Ik voel dat ik hen bén, in de bedachtzaamheid waarmee ik formuleer, in de felheid waarin ik af en toe verval, in de manier waarop ik mijn handen vouw als ik in de zetel zit, in ingehouden ontroering, in soberheid. Jullie zijn mijn ankerpunt en aan jullie draag ik dit werk op.

Joren Somers, 12 mei 2024

Table of contents

Preface ..... *ix*

Table of contents ..... *xiii*

Chapter 1 Introduction *1*

1.1 The issue **1**

1.2 Argument structure, event structure and subjecthood **3**

1.3 Syntactic subjecthood **6**

1.3.1 Establishing syntactic subjecthood 6
1.3.2 Benefits of a syntactic approach to subjecthood 8

1.4 Subjecthood through a Construction Grammar lens **10**

1.4.1 Construction Grammar: what and why? 10
1.4.2 Grammatical relations as (generalised language-specific) relations or (construction-specific) roles 13
1.4.3 Formalising the alternating Dat-Nom/Nom-Dat constructions 15

1.5 Alternating verbs and verb classes **17**

1.6 Where this study fits in **18**

1.6.1 Spotlighting the word order test 18
1.6.2 Previous studies on German 19
1.6.3 Previous studies on Icelandic 21
1.6.4 Research questions 21
1.6.5 A brief word on the data collection process 23

1.7 Genesis of the thesis **25**

1.8 Overview **27**

References **28**

Chapter 2 Onpersoonlijke datief-nominatiefwerkwoorden in het Hedendaagse Duits: Definitie, inventaris en semantische classificatie 33

2.1 Het concept 33

2.2 De reikwijdte van de onpersoonlijke datief-nominatiefconstructie in het Hedendaagse Duits 36

2.3 Naar een inventaris van alle onpersoonlijke datief-nominatiefwerkwoorden 40

2.3.1 Het uitgangspunt 40
2.3.2 Lexicografische zoekopdracht 42
2.3.3 Aanvullende zoekstrategieën 43
2.3.4 Besluit 45

2.4 Semantische classificatie 46

2.4.1 Clustering in zestien klassen 46
2.4.2 Algemene tendensen 50
2.4.3 Bespreking per klasse 52
2.4.4 Besluit 53

2.5 Slotbemerking 54

Referenties 55

Appendix 1 57

Appendix 2 62

Chapter 3 Linearization and unaccusativity: The relative order of dative and nominative arguments with five German verbs of success and failure 69

3.1 Introduction 69

3.2 Word order patterns with verbs of success and failure 71

3.2.1 Dative-before-nominative 71
3.2.2 Dative-before-nominative and nominative-before-dative 72

3.3 Hypotheses 75

3.4 Methodology 75

3.4.1 Verb selection 75
3.4.2 Data cleaning 76
3.4.3 Data annotation 77

3.5 Results and discussion 77

3.5.1 General frequencies 78

3.5.2 Word order variation in the [NP-V-NP] configuration 78
3.5.3 Word order variation in the [Pro-V-Pro] configuration 80
3.5.4 Competition between nominative NPs and dative pronouns 81
3.5.5 Competition between dative NPs and nominative pronouns 82

3.6 Conclusions and outlook 83

References 84

Chapter 4 Argument structure constructions in competition: The Dat-Nom/Nom-Dat alternation in Icelandic 87

4.1 Introduction 87

4.2 Object of study 90

4.3 Methodology 96

4.4 Results and discussion 99

4.4.1 Establishing a baseline: *hjálþa*- and *líka*-verbs 99
4.4.2 Alternating Dat-Nom/Nom-Dat verbs 104
4.4.3 Interim conclusions 110

4.5 Statistical modelling 112

4.5.1 Across configurations 114
4.5.2 Double-NP configuration 117
4.5.3 Interim conclusions 118

4.6 Summary and conclusions 120

References 122

Chapter 5 Subjecthood and argument structure of synonymous Dat-Nom/Nom-Dat verbs across German and Icelandic 127

5.1 Introduction 128

5.2 Methodology 131

5.2.1 Dataset 131
5.2.2 Data cleaning 134
5.2.3 Data annotation 135

5.3 Findings 137

5.3.1 Verbs of the *hjálþa*- and *helfen*-type 137
5.3.2 Verbs of the *líka*- and *gefallen* type 141

5.3.3 Verbs of the *nægja*- and *genügen*-type 146
5.3.4 Personal pronouns 151

5.4 Conditional inference tree 154

5.5 Summary and conclusions 158

References 162

Chapter 6 A corpus-based analysis of the Dat-Nom/ Nom-Dat alternation in German 165

6.1 Introduction 166

6.2 State-of-the-art 168

6.2.1 The scope of Dat-Nom/Nom-Dat verbs 168
6.2.2 Dat-Nom/Nom-Dat verbs: definition 173

6.3 Methodology 177

6.3.1 Data inspection 178
6.3.2 Data cleaning 179
6.3.3 Data annotation 182

6.4 Results and discussion 184

6.4.1 Case marking 185
6.4.2 Internal order of the two arguments 194

6.5 Summary and conclusions 204

References 205

Appendix 3 211

Appendix 4 215

Chapter 7 Discussion and general conclusion 217

7.1 Introduction 217

7.2 Main findings 218

7.3 Comparing the targets to the baseline 222

7.3.1 Icelandic 222
7.3.2 German 225

7.4 Theoretical implications 233

7.4.1 General implications....233
7.4.2 Implications for the constructional representation of polysemous verbs....235
7.4.3 Against an allostructional account of the Dat-Nom/Nom-Dat alternation....239

7.5 Outlook....240

References....242

Chapter 1
Introduction

1.1 The issue

The present thesis investigates word order with alternating Dat-Nom/Nom-Dat verbs in Modern Icelandic and Modern German, two closely related members of the Germanic language family. It has been argued that such alternating Dat-Nom/Nom-Dat verbs exist in both of these languages (Bernódusson 1982; Jónsson 1997–1998; Barðdal 1999, 2001, 2023: Ch. 3; Eythórsson & Barðdal 2005; Platzack 1999; Primus 1999; Sigurðsson 2006; Rott 2013, 2016; Wood & Sigurðsson 2014; Barðdal, Eythórsson & Dewey 2014, 2019; Somers & Barðdal 2022, 2023). Such verbs, which require both a dative and a nominative argument, may instantiate two parallel linearisation patterns: dative-before-nominative and nominative-before-dative. Crucially, neither of these can be regarded as a topicalisation of the other, as informal enquiries with native speakers confirm: they conceive of both alternants as equally neutral in the sense that there is no information-structural interference. As a result, alternating verbs boast two argument structures, the Dat-Nom and the Nom-Dat argument structure constructions. Two examples of such alternating verbs, one featuring Icelandic *nægja* 'be enough, be sufficient' and another featuring German *gefallen* 'like, please', are presented under (1–2) below:

Dat-Nom/Nom-Dat

(1) a. *Elsu* *hefur* *nægt* *ferðin á hálendin.*
Elsa.DAT has sufficed trip.the.NOM in highlands.the
'Elsa found the trip into the highlands sufficient.'

(1) b. *Ferðin á hálendinu* *hefur* *nægt* *Elsu.*
trip.the.NOM in highlands.the has sufficed Elsa.DAT
'The trip into the highlands was sufficient for Elsa.'

(2) a. *Der Frau* *hat* *das Buch* *gefallen.*
the.DAT woman has the.NOM book liked
'The woman liked the book.'

(2) b. *Das Buch* *hat* *der Frau* *gefallen.*
the.NOM book has the.DAT woman pleased
'The book pleased the woman.'

In addition to verbs of this alternating type, Icelandic, but not German, possesses a class of non-alternating Dat-Nom verbs. These instantiate but a single Dat-Nom argument structure, which is illustrated by examples (3a–c) below:

Dat-Nom

(3) a. *Manninum* *hefur* *líkað* *bókin* *vel.*
man.the.DAT has liked book.the.NOM well
'The man has liked the book.'

(3) b. **Bókin* *hefur* *líkað* *manninum* *vel.*
book.the.NOM has liked man.the.DAT well
Intended meaning: 'The book has pleased the man.'

(3) c. *BÓKIN* *hefur* *manninum* *líkað* *vel.*
book.the.NOM has man.the.DAT liked well
'It is the book that the man has liked.'

The grammaticality of (3a) and the ungrammaticality of (3b) confirm the status of Icelandic *líka* 'like' as a non-alternating Dat-Nom verb. Note that in (3b) it is not the sentence-initial position of the nominative that renders the example problematic, but rather the fact that the dative follows the non-finite verb *líkað* 'liked'. Sentence-initial nominatives with Dat-Nom verbs are surely grammatical, as example (3c) illustrates, but only when the dative is in immediate adjacency to the finite verb, thus confirming its role as an inverted subject. Furthermore, an inversion of the canonical order of constituents is only permissible under very specific pragmatic conditions, as with topicalisation. For a more in-depth discussion of argument structure and grammatical relations, see Sections 1.2–1.3 below.

Third, German and Icelandic also have at their disposal a class of ordinary **Nom-Dat** verbs. These equally license a nominative and a dative argument, but, crucially, they correlate with a single Nom-Dat argument structure, as is illustrated by examples (4–5) below:

Nom-Dat

(4) a. *Der Dirigent* *dankt* *dem Orchester.*
the.NOM conductor thanks the.DAT orchestra
'The conductor thanks the orchestra.

(4) b. *DEM ORCHESTER* *dankt* *der Dirigent.*
the.DAT orchestra thanks the.NOM conductor
'It is the orchestra that the conductor thanks.'

(5) a. *Stjórnandinn* *þakkar* *hljómsveitinni.*
conductor.the.NOM thanks orchestra.the.DAT
'The conductor thanks the orchestra.'

(5) b. *HLJÓMSVEITINNI* *þakkar* *stjórnandinn.*
orchestra.the.DAT thanks conductor.the.NOM
'It is the orchestra that the conductor thanks.'

Ceteris paribus, Icelandic *þakka* and German *danken* both realise the nominative before the dative. This is also shown under (4a) and (5a). The reverse dative-before-nominative order, albeit grammatical, is only used for topicalisation (cf. 4b and 5b). As such, Nom-Dat verbs build the mirror image of non-alternating Dat-Nom verbs in that they possess but a single argument structure, i.e. Nom-Dat.

1.2 Argument structure, event structure and subjecthood

In the previous section, it has been shown that verbs licensing both a nominative and a dative pertain to one of three classes: Dat-Nom verbs, Nom-Dat verbs, and Dat-Nom/Nom-Dat verbs. As set out in Section 1.1, this thesis adopts the view that Icelandic possesses each of these classes, while German only possesses the latter two. Table 1.1 provides an overview of the different verb classes and their respective presence or absence in the languages under study:

Table 1.1 Dat-Nom, Nom-Dat and Dat-Nom/Nom-Dat verbs across Icelandic and German

	Icelandic	**German**
Dat-Nom verbs	✓	
Nom-Dat verbs	✓	✓
Dat-Nom/Nom-Dat verbs	✓	✓

For Icelandic, the existence of both types of Dat-Nom verbs, alternating and non-alternating ones, has been acknowledged by several scholars of that language (Bernódusson 1982; Jónsson 1997–1998; Barðdal 1999, 2001, 2023: Ch. 3; Platzack 1999; Sigurðsson

2006; Rott 2013, 2016; Wood & Sigurðsson 2014; Barðdal, Eythórsson & Dewey 2014, 2019; Somers & Barðdal 2022, 2023). For German, however, the existence of a class of alternating Dat-Nom/Nom-Dat verbs has not even been considered in the literature. Instead, the consensual view is that all verbs licensing a nominative and a dative in that language are Nom-Dat verbs, of which some license an inversion of the canonical word order pattern through a 'mandatory topicalisation' of the dative argument, thus yielding a dative-before-nominative surface order (Bader and Häussler 2010; Haspelmath 2001; Kempen and Harbusch 2005; Verhoeven 2015, *inter alia*). The current thesis diverges from the standard syntactic analysis of German in two ways: (1) it assumes the existence of two syntactic verb classes, Nom-Dat (with verbs like *danken* 'thank') and Dat-Nom/Nom-Dat (with verbs like *gefallen* 'like, please'), instead of a single Nom-Dat class and (2) it assumes that verbs like *gefallen* do not require mandatory topicalisation of the dative argument, but alternate instead between two argument structures, Dat-Nom and Nom-Dat, and, consequently, license two word order patterns corresponding to each of these: dative-before-nominative and nominative-before-dative.

In line with Croft (1998, 2001: 164; 2012: 233), Barðdal (2001: 60–66; 2023: 40–43) and Barðdal, Eythórsson & Dewey (2014: 56–60; 2019: 117–122), I argue that a given verb's argument structure is a direct reflection of its event structure. For highly agentive verbs, like *kick*, this involves a causal chain, meaning that an instigator acts upon an endpoint. Thus, the instigator maps onto the subject and the endpoint onto the object of the clause. Similarly, for verbs like *bother*, the stimulus acts upon the endpoint, the animate being, yielding the same argument structure as with *kick*. For stative verbs, and typically verbs expressing emotions, two construals are available: (1) a stimulus affects an experiencer (*frighten*-verbs) and (2) an experiencer directs their attention to a stimulus (*fear*-verbs).

For the two German classes distinguished above, this means that Nom-Dat verbs like *danken* adhere to the causative construal exemplified above with *kick* where an instigator may act upon an endpoint. Hence, the nominative is mapped onto the subject role and the dative onto the object role. Contrariwise, alternating Dat-Nom/Nom-Dat verbs like *gefallen* are compatible with the two construals given above for verbs expressing emotions, namely the nominative stimulus affects a dative experiencer or a dative experiencer directs their attention to a nominative stimulus. This explains why alternating verbs correspond to two opposite argument structure structures, and not just one.

Thus, argument structure has a bearing on grammatical relations, as grammatical relations are assigned in accordance with it. The interrelation posited between event structure, argument structure and grammatical relations is represented in (6) below:

(6) event structure → argument structure → grammatical relations

As argued by Barðdal, Eythórsson & Dewey (2019: 120–122) and Barðdal (2023: Ch. 3), Dat-Nom/Nom-Dat predicates are force-dynamically neutral in the sense of Talmy (1976, 1985, 1988), which means that they are not readily associated with one unambiguous event

structure. The fact that alternating verbs lack such a causal chain of events is a direct consequence of their being low in agentivity, as such verbs generally pertain to mental states, cognitive processes, contexts of gain and benefit, or happenstance events.

In contrast, ordinary Nom-Dat verbs and non-alternating Dat-Nom verbs each correspond to a single underlying event structure. For Nom-Dat verbs, the construal directly reflects a chain of events in which a nominative referent targets a dative referent. Examples of such verbs include German *vertrauen* 'trust' and *gehorchen* 'obey' or Icelandic *gleyma* 'forget' and *þakka* 'thank'. As for non-alternating Dat-Nom verbs, the relevant construal of the event is one in which a dative referent directs their attention to a nominative referent, giving rise to a Dat-Nom argument structure. Semantically, non-alternating Dat-Nom verbs show a clear overlap with alternating Dat-Nom/Nom-Dat verbs: both these classes are equally low in agentivity and non-alternating Dat-Nom verbs also express meanings very similar to the ones discussed in the context of Dat-Nom/Nom-Dat verbs above, comprising verbs like *líka* 'like', *áskotnast* 'receive' or *þykja* 'think, find, seem'. As argued by both Croft (1998) and Barðdal (2001: 57, 2023: 72), it is a lexical idiosyncrasy of any given verb, Dat-Nom or Dat-Nom/Nom-Dat, whether or not it alternates and, thus, whether it corresponds to one type of construal or two.

The fact that different verbs may be associated with different argument structures has important theoretical implications. Following Eythórsson & Barðdal (2005: 831), Barðdal & Eythórsson (2006: 180; 2012: 368–369; 2020: 205), Barðdal, Eythórsson & Dewey (2019: 120), Johnson et al. (2019: 487), Barðdal et al. (2023: 329–333) and Barðdal (2023: Ch. 2), I subscribe to the view that the first argument of the argument structure corresponds to the syntactic subject, while any remaining arguments function as objects. The original definition of subjecthood rooted in argument structure, as proposed by Eythórsson & Barðdal (2005), is presented in (7) below:

(7) The subject of a predicate is the leftmost argument of its subcategorization frame, here labeled ARG1. The internal order of the arguments is in turn determined by the causal conceptual structure of the predicate and the force-dynamic relations between the participants of the event denoted by each predicate.

Returning to the three verb classes shown in Table 1.1, this means that Dat-Nom verbs license a dative subject and a nominative object, whereas Nom-Dat verbs license a nominative subject and a dative object. For alternating Dat-Nom/Nom-Dat verbs, subjecthood is variable in that the relevant verbs may either take a dative subject and a concomitant nominative object (Dat-Nom) or a nominative subject and a concomitant dative object (Nom-Dat).

1.3 Syntactic subjecthood

By defining the subject as the first argument of the argument structure, I adopt a syntactic concept of subjecthood. Syntactic subjects do not necessarily display the coding properties associated with subjects, i.e. nominative case and subject-verb agreement, but they do behave like subjects syntactically. What exactly this means is discussed in Section 1.3.1. I then go on to review the benefits associated with a syntactic approach to subjecthood in Section 1.3.2.

1.3.1 Establishing syntactic subjecthood

In order to establish syntactic subjecthood, linguists resort to a set of tests, also known as subjecthood diagnostics. The following subject tests have been identified for Icelandic and German alike, with long-distance reflexivisation solely applying to Icelandic (Barðdal 2006; Barðdal, Eythórsson & Dewey 2019; Barðdal 2023: Ch. 6):

Table 1.2 Subjecthood diagnostics applying to syntactic subjects in Icelandic and German

	Icelandic	**German**
first position	✓	
subject-verb inversion	✓	✓
subject-to-object raising	✓	✓
subject-to-subject raising	✓	✓
conjunction reduction	✓	(✓)
control infinitives	(✓)	(✓)
clause-bound reflexivisation	✓	✓
long-distance reflexivisation	✓	N/A

Two of these diagnostics, i.e. conjunction reduction and control infinitives, hold a special status with regard to German and/or Icelandic. Hence, they are bracketed in Table 1.2 above. Starting with control infinitives, it is generally assumed that oblique subjects may be left unexpressed in control infinitives in Icelandic, but not in German (Zaenen, Maling & Thráinsson 1985: 106, 126; Sigurðsson 1989: 207; 2002: 694; Fanselow 2002: 242–245; Bayer 2004: 54–57; Haider 2005: 5; Wunderlich 2008: 588, *inter alia*). Such an analysis is consistent with the consensual view that only the former language possesses oblique subjects (see references immediately above). However, as shown by Barðdal & Eythórsson (2006), the differences between these languages are in fact gradient, not categorical.

An acceptability judgement task carried out by Eythórsson & Barðdal (2005) involving twelve control constructions for eight different German verbs reveals that participants

differ widely in their approval of such structures, with the least acceptable infinitive yielding an acceptability of 12.5% and the most acceptable infinitive an acceptability of 86.5%. The average approval rate across verbs amounts to 39.3%. A corresponding study of Icelandic control infinitives carried out by Barðdal & Eythórsson (2006) generated highly similar results: out of three constructions for two different verbs, approval rates vary from 6.5% to 57.1%, with an average of 26.4% across test items. Clearly, some individuals (and some constructions) are more permissive of control than others, although such users and such constructions may be found in Icelandic and German alike. Hence, the great Icelandic–German divide does not exist (Barðdal & Eythórsson 2005; Barðdal 2023).

As for conjunction reduction, the established view is that it serves as a subject test in Icelandic and German, but that only nominative subjects pass the test in the latter language (Sigurðsson 2002: 694–695; Fanselow 2002: 239–240; Bayer 2004: 54–55, 57–58, *inter alia*). Subjects in the former language, by contrast, are easily reducible on identity with other subjects, regardless of any potential case distinctions. This process is exemplified in (8) below, using an example sentence from Barðdal, Eythórsson & Dewey (2019: 123):

(8)	*Ég*	*kynntist*	*fólkinu,*	*og*	___	*hefur*	*líkað*
	I.NOM	got.to.know	people.the.DAT	and	Ø.DAT	have	liked
	það	*vel.*					
	it.NOM	well					
	'I got to know the people and liked them a lot.'						

Example (8) shows that the nominative subject *ég* in the first conjunct and the dative subject *mér* in the second conjunct share a syntactic pivot, as one argument may freely be reduced on identity with the other, despite any potential conflict in case marking.

In German, by contrast, it is indeed questionable whether corresponding examples like the one under (9), where a nominative is deleted on identity with a dative, are grammatical:

(9)	*??Mir*	*gefällt*	*das Buch*	*und*	___	*kaufe*	*es.*
	me.DAT	likes	the.NOM book	and	Ø.NOM	buy	it.ACC
	'I like the book and I buy it.'						

However, this may be because ellipsis in German appears to be more constrained than it is in Icelandic. More specifically, conjunction reduction in German is restricted to arguments sharing both a syntactic and a morphological pivot (Barðdal 2006; Barðdal & Eythórsson 2006; Barðdal, Eythórsson & Dewey 2019: 114).

Indeed, as shown by Barðdal (2006: 60–62, 2023: 287–288), examples in which a subject-like oblique is deleted on identity with an unequivocal dative object do not render an indisputably grammatical result:

(10)	*??Mir*	*verzieh*	*er*	*und*	___	*hat*	*er*	*gefallen.*
	me.DAT	pardoned	he.NOM	and	Ø.DAT	has	he.NOM	pleased
	'He pardoned me and pleased me.'							

Conjunction reduction therefore seems to be liable to the same inter-speaker variation as control infinitives. Furthermore, as argued by Barðdal (2006: 60–62), conjunction reduction in German is sensitive to case marking and grammatical relations alike. If conjunction reduction were a mere case test, (10) would have to be entirely grammatical, quod non.

1.3.2 Benefits of a syntactic approach to subjecthood

Adopting a syntactic approach to subjecthood is in many ways beneficial. As argued by Eythórsson & Barðdal (2005) and Barðdal (2023: 39), all of the subject tests automatically target the first argument of the argument structure. This means that a definition of subject rooted in argument structure is naturally parsimonious, because it assumes that the diagnostics shown in Table 1.2 above indeed apply to the leftmost argument of the argument structure. Thus, there is no need to explicitly incorporate the tests into one's definition of subject.

For German in particular, a syntactic approach to subjecthood also resolves a long-standing incoherence whereby the subject is defined as being in the nominative case even though that argument does not automatically fulfil the subject requirement of occurring in clause-initial position in word orders which are information-structurally neutral. This is particularly relevant for alternating Dat-Nom/Nom-Dat verbs instantiating the Dat-Nom argument structure, the very existence of which shows that German cannot at the same time be a subject-first language and a nominative-first language.[1]

Defining the subject as the first argument of the argument structure is also advantageous from a semantic point of view, because it allows one to view the subject as the entity from whose perspective the state-of-affairs is presented: the subject is the default theme of a clause, whereas the object is the default rheme.

Another reason why a definition of the subject in terms of argument structure is preferable is because it is sufficiently broad, licensing both canonically as well as non-canonically case-marked subjects. A definition appealing to morphology alone, viz. nominative case marking, loses out on a behavioural generalisation, apart from (mistakenly) assuming a one-to-one correspondence between (morphological) case and (syntactic) behaviour (Barðdal 2000: 109–110).

[1] For references on German as a language featuring SVO in main clauses, see Greenberg (1966: 109), Comrie (1981: 89), Dryer (2013: Ch. 81) and Barðdal (2023: 291), *inter alia*.

In addition, as argued by Barðdal (2000: 102), scholars equating subjects with nominatives might erroneously believe that nominative subjects form a homogeneous class, which is clearly not the case. As an example, Barðdal shows that some nominative-subject verbs in Icelandic allow passivisation, whereas others do not. The same holds for German. Thus, even if one appeals to a morphological criterion to delineate subjects, one still has to account for the fact that two nominative subjects might exhibit blatantly different behaviours. Also, it is unclear why a morphological criterion should eclipse one or more syntactic criteria in determining subjecthood, as pointed out by Barðdal (2000). The axiom that the subject is by definition marked for nominative has permeated much of the German work on subjecthood, including Zaenen, Maling & Thráinsson (1985), Bayer (2004), Haider (2005: 5), Eisenberg (2013: 278), and many others. Constituents may or may not bear nominative case, and they may or may not behave as subjects, but one does not necessarily entail the other.

Sixth, a definition of subject rooted in argument structure is perfectly falsifiable. The most direct reflection of argument structure is neutral word order. Both usage-based linguistics and psycholinguistics can aid in gaining insight into word order patterns, either by scrutinising corpus data or by appealing to grammaticality judgement tasks. The former approach builds on the assumption that neutral word order is considerably more frequent than topicalisation structures, the latter on the fact that native speakers will consistently rate the acceptability of structures involving neutral word order higher. Specifically within the context of subjects, linguists can also appeal to any of the remaining subject tests. As argued at the beginning of this section, the subject, being the first argument of the argument structure, is expected to satisfy all of the remaining subject tests as well. However, due to a restriction on ellipsis in German, the scholarship of that language has argued against a subject analysis of oblique subjects (cf. my discussion of control infinitives and conjunction reduction in Section 1.3.1 above). Nevertheless, the subject behaviour of a given argument should also be testable using additional data in support of any of the remaining subjecthood diagnostics.

Finally, as argued by Eythórsson & Barðdal (2005: 831) and Barðdal (2023: 39), a definition of subject grounded in argument structure is entirely theory-neutral, as all linguistic frameworks acknowledge the role of argument structure or subcategorisation frames for grammatical relations. As a consequence, such a definition should be easily translatable into other frameworks.

1.4 Subjecthood through a Construction Grammar lens

The current study adopts a constructional approach to language. What exactly this means and why I have decided in favour of Construction Grammar is outlined in Section 1.4.1. The next section, Section 1.4.2, motivates my choice of Radical Construction Grammar in particular. Section 1.4.3, the concluding section, formalises the alternating Dat-Nom/Nom-Dat constructions.

1.4.1 Construction Grammar: what and why?

Construction Grammar is a linguistic framework revolving around the notion of the 'construction', which is characterised as a form-meaning correspondence (Goldberg 1995, 2003, 2006, 2019 *inter alia*). Constructions are assumed to be the basic units of language, found at all levels of description, from the smallest morpheme all the way up to the most intricate syntactic pattern. The best known definition of the construction is most likely the one offered by Goldberg (1995: 4):

> C is a CONSTRUCTION iff_{def} C is a form-meaning pair $<F_i, S_i>$ such that some aspect of F_i or some aspect of S_i is not strictly predictable from C's component parts or from other previously established constructions.

Goldberg's (1995) definition crucially hinges on the non-compositional nature of the form-meaning pair, which means that only those patterns count as constructions that possess an element of form or meaning that is not strictly predictable from the construction's component parts. As such, constructions constitute Saussurian signs whose form and meaning are conventionally associated with one another. However, as from 2006, Goldberg expanded her definition of construction to compositional form-meaning pairings occurring sufficiently frequently (Goldberg 2006: 5), thus abandoning the original definition hinging on the non-compositional nature of the construction. Goldberg's revised definition ties in well with Croft's (2001: 19) assumption that a construction is any type of conventionalised form-meaning pairing. This is also the view I subscribe to.

The present work is concerned with a specific type of construction known as an *argument structure construction*. According to Goldberg (2019: 28), argument structure constructions "determine basic clause types in each language and constrain the interpretation of 'who did what to whom'." As such, they constitute basic syntactic patterns, which also express basic scenes. Goldberg (1995: 225–226) states that these scenes may refer to acts of transfer, cause and effect, experience, and so on. Among the most widely known argument structure constructions are the ditransitive construction

(e.g. *John gave Jack a kiss*), the caused-motion construction (e.g. *Thomas chased Elisabeth out of the room*) and the resultative construction (e.g. *Bob painted the front door dark green*).

The main claim at the heart of this thesis is that there are certain verbs, both in German and in Icelandic, that are not just associated with one argument structure construction, but two: the Dat-Nom argument structure construction and the Nom-Dat argument structure construction. The reason why alternating verbs possess two argument structures is twofold: (1) syntactically, the Dat-Nom and the Nom-Dat alternants of alternating verbs are both equally neutral and (2) semantically, hypothesised Dat-Nom/Nom-Dat verbs indeed express very basic scenes, which mostly relate to experience, even though they may also describe success, failure, perception, bodily states, cognitive processes and many more. The semantics of the verbs associating with the alternating Dat-Nom/Nom-Dat constructions are further discussed in Section 1.5 below as well as in Chapter 2. Of course, the idea that verbs may be used in more than one argument structure construction is not new: Goldberg (1995: 11) has shown that English *kick* is compatible with no less than eight argument structure constructions. However, what is remarkable about the alternating Dat-Nom/Nom-Dat constructions is that these involve the same number of participants, which furthermore retain their case marking across both constructions, Dat-Nom and Nom-Dat.

Let us now turn to the reasons for turning to a Construction Grammar approach. First, it is well known that Construction Grammar arose out of the need to account for non-core patterns in language, more specifically idioms (Croft 2001: 15). One of the first research articles drawing attention to this issue is Fillmore et al. (1988), who argue that the grammatical analysis of any given language should go beyond a merely componential approach associating lexical items with grammatical rules. The idea behind the constructional approach is that a model of grammar ought not to brush aside non-core cases, as these, too, are part of the language system and, hence, should be properly accounted for. Kay & Fillmore (1999: 1) aptly describe this ambition by stating that Construction Grammar is a framework that seeks to account for the system as a whole. Since verbs licensing oblique subjects also constitute a non-core case in the grammar, it should not come as a surprise that Construction Grammar is particularly relevant for modelling and accounting for the type of data discussed here. After all, non-canonical subjects are part of the language just as much as canonical subjects are, so their existence should not be explained in an ad hoc fashion, but systematically. This point is further elaborated on in Section 1.4.2 below, where the choice for Radical Construction Grammar in particular is justified.

Secondly, the importance of idioms in cognitive theories is evidently related to the importance of figurative language in such frameworks. As argued by Langacker (1987: 38), cognitive grammarians, contrary to proponents of more generative approaches, acknowledge that figurative language may interact with grammatical phenomena. This premise is also crucial to the present study, which assumes certain verbs to interact with

the Dat-Nom/Nom-Dat alternation in some of their metaphorical senses, but never in literal usage (see Section 1.6.5 below for more information on the role of polysemy in the present study). Consider in this respect examples (11–13) below:

Agentive and literal

Nom-Dat

(11) *Die Maus de Soto* *entgeht* *einem Fuchs.*
the.NOM mouse de Soto escapes a.DAT fox
'The mouse de Soto escapes a fox.'

Nom-Dat

(12) *Ramses II.* *entging* *nur knapp* *der Gefangenschaft.*
Ramesses II.NOM escaped only narrowly the.DAT captivity
'Ramesses II only narrowly escaped captivity.'

Non-agentive and figurative

Dat-Nom

(13) a. *Den Staaten entgingen* *dadurch*
the.DAT states missed.out.on as.a.result
Einkommenssteuern in Milliardenhöhe.
income.taxes.NOM in billion.height
'The states lost billions in income tax as a result.'

Nom-Dat

(13) b. *Rund 150 Milliarden Euro* *entgingen* *der EU* *jährlich.*
around 150 billions.NOM euro missed.out.on the.DAT EU annually
'About 150 billion euros would be lost to the EU annually.'

Examples (11–13) involve the verb *entgehen*. Each illustrates a different sense: literal 'escape' in (11), figurative 'escape' in (12) and figurative 'miss out on' in (13). For polysemous verbs of the *entgehen*-type, I hypothesise literal senses to solely associate with the Nom-Dat argument structure construction, as these are invariably agentive. This is what is shown in example (11). Figurative senses, by contrast, may or may not associate with the Dat-Nom/Nom-Dat alternation, depending on whether or not they are agentive. Hence, figurative yet agentive 'escape', as illustrated in (12), is compatible with the non-alternating Nom-Dat construction alone, whereas figurative and non-agentive 'miss out on', as shown in (13), is compatible with either construction, Dat-Nom or Nom-Dat. These examples confirm the relevance of Langacker's claim for the present study: figurative language may indeed interact with grammatical phenomena, hence showing that it is unsound to treat semantics outside of the grammar.

Third, as argued by Goldberg (2006: 10), Construction Grammar offers a "what you see is what you get approach" to language (see also Langacker 1987: 27). This means that Construction Grammar, contrary to Generative Grammar, need not resort to syntactic transformations to account for the existence of certain patterns. One such transformation

is believed to derive double-object constructions from *to*-datives, the latter being more basic than the former (Goldberg 1995: 103–108). However, Goldberg shows that such an analysis is theoretically inadequate, since metaphorical ditransitives only exist in double-object constructions, and not in *to*-dative constructions:

(14) a. *She gave me a headache/a kiss/an idea.* (Goldberg 1995: 106)

(14) b. **She gave a headache/a kiss/an idea to me.* (Goldberg 1995: 106)

By considering double-object ditransitives and prepositional-object ditransitives as constructions in their own right, one does justice to the empirical facts, which show that one's existence is not tied to the other.

Specifically within the context of the current study, a Construction Grammar approach allows one to consider German SVO patterns as constructions in their own right. Such a position clearly differs from generative approaches, which maintain that German is essentially an SOV language and that SVO patterns are the result of a derivational operation (Bach 1962, Bierwisch 1963, Koster 1975, *inter alia*). As argued by Boas & Ziem (2018: 4), it is difficult to advocate for a generative stance in this respect, as it is empirically impossible to determine which constructions should be more basic than others. Boas & Ziem's emphasis on empirical falsifiability is not a trivial matter: it ties in well with Croft's (2001: 9) premise that linguistics is first and foremost an empirical science.

Fourth, I adopt a Construction Grammar approach because I subscribe to the view that, as Goldberg puts it, "knowledge of language is knowledge" (2006: 59; 2019: 52). This means that our knowledge of language is parallelled in other cognitive faculties (Bybee 1998: 249–252). Such a view clearly contrasts with more modular approaches, which maintain that the language faculty is associated with a so-called Language Acquisition Device.

Fifth and finally, Construction Grammar may account uniformly for linguistic data at all levels of granularity, as every unit may be considered a construction. These include morphemes, words, phrases, clauses and even texts. Croft (2001: 17) as well as Croft & Cruse (2004: 255) note that this particular advantage of Construction Grammar is often referred to as the *syntax-lexicon continuum* (cf. also Langacker 1987: 35).

1.4.2 Grammatical relations as (generalised language-specific) relations or (construction-specific) roles

This thesis is specifically couched in a Radical Construction Grammar framework (Croft 2001, Barðdal 2006, 2023), my choice for which is motivated in three ways. First, Radical Construction Grammar constitutes a usage-based version of Construction Grammar. Usage-based approaches assign a major role to language use, which is believed to shape the language system in a bottom-up fashion (Barðdal & Gildea 2015). As a consequence,

Radical Construction Grammarians believe that the study of language should proceed in a data-driven way. This stance is fully in line with Croft's aforementioned premise that linguistics ought to be an empirical science. Secondly, usage-based approaches do not seek to impose onto language a logical system à la generative grammar. Instead, advocates of usage-based approaches believe that the language system is to a certain extent idiosyncratic (cf. Goldberg's 1995 patterns which cannot be explained by rules and Croft's 1998 claims that the assignment of argument structure to lexical verbs is in part idiosyncratic).

Thirdly, Radical Construction Grammar offers a particularly attractive approach to grammatical relations because it assumes grammatical relations to be construction-specific roles rather than generalised language-specific relations (Croft 2001; Barðdal 2006, 2023; Barðdal, Eythórsson & Dewey 2014, 2019; Kyröläinen 2015; Witzlack-Makarevich 2019, *inter alia*). This is essentially also what Langacker (1987: 38) argues by maintaining that linguistic units are only meaningful in context. Radical Construction Grammar clearly applies this creed to syntax. However, vanilla Construction Grammar has assumed grammatical relations to exist independently of constructions. One problem with this view is that one would not expect subjects to behave differently across constructions, for instance with regard to case marking. After all, the independently posited subject role would have to be nominative and oblique at the same time, which is an incongruity (Barðdal & Gildea 2015: 30–31; Barðdal 2023: 34–38). Thus, subjects are not only language-specific, but also construction-specific.

A constructionist view of subjecthood not only accounts for the fact that subjects may display different *coding* properties, but also for the variation in their *behavioural* properties, both across languages as well as language-internally (Barðdal 2006; Barðal & Eythórsson 2006; Barðdal & Gildea 2015: 30; Barðdal 2023: 46–65). Starting with the former, it has already been mentioned in Section 1.3 above that Icelandic subjects pass nine different behavioural tests, whereas their German counterparts only pass eight (Barðdal, Eythórsson & Dewey 2014, 2019; Barðdal 2023: Ch. 6). The reason is that the ninth test, i.e. long-distance reflexivisation, is only applicable to the former language, but not to the latter, as the following two example sentences from Barðdal (2006: 53) illustrate:

Icelandic

(15)	*Hans$_i$*	*vildi*	*að*	*Anna*	*gæfi*	*sér$_i$*	*smá*	*köku.*
	Hans	wanted	that	Anna	would.give	self	a.little	cake
	'Hans wanted Anna to give him a little cake.'							

German

(16)	*Hans$_i$*	*wollte,*	*dass*	*Anna*	**sich$_i$/ihm$_i$*	*ein bisschen Torte*	*gäbe.*
	Hans	wanted	that	Anna	self	a little cake	gave
	'Hans wanted Anna to give him a little cake.'						

Examples (15–16) show that subjects in Icelandic may bind reflexive pronouns in subordinate clauses, whereas subjects in German may only bind anaphoric pronouns

across clause boundaries. The fact that the present test yields an infelicitous result for German does not however exclude a subject analysis of *Hans* in (16); it merely demonstrates that the syntax of German is constrained differently than that of Icelandic (see also Barðdal 2023: 38). Note also that a constructional definition of subject does not rule out comparative research on grammatical relations, neither language-internally nor cross-linguistically or diachronically (Barðdal 2023: ibd.). In fact, as argued by both Rögnvaldsson (1996: 54–55) and Barðdal (2000: 96), one should only assume formally similar structures to behave differently if such an assumption is independently motivated.

The hypothesis that subject status is constructionally determined is readily applicable to the present study. Starting with ordinary Nom-Dat verbs, I argue that these instantiate a [Nominative-Dative] argument structure construction. Non-alternating Dat-Nom verbs, however, select for a [Dative-Nominative] argument structure construction. Alternating Dat-Nom/Nom-Dat verbs may instantiate either argument structure construction, [Dative-Nominative] or [Nominative-Dative], since either argument, dative or nominative, may be assigned the subject role. The next section further develops the concept of constructionally determined subjecthood.

1.4.3 Formalising the alternating Dat-Nom/Nom-Dat constructions

The current subsection models the interaction between alternating verbs and their respective argument structure constructions, Dat-Nom and Nom-Dat. The formalism has been adopted from Barðdal (2023: 121–124), while the semantic frames stem from Framenet (Ruppenhofer et al. 2016) and Johnson et al. (2019). As shown in Figure 1.1, a verbal lexeme compatible with the alternating constructions simultaneously interacts with either of these:

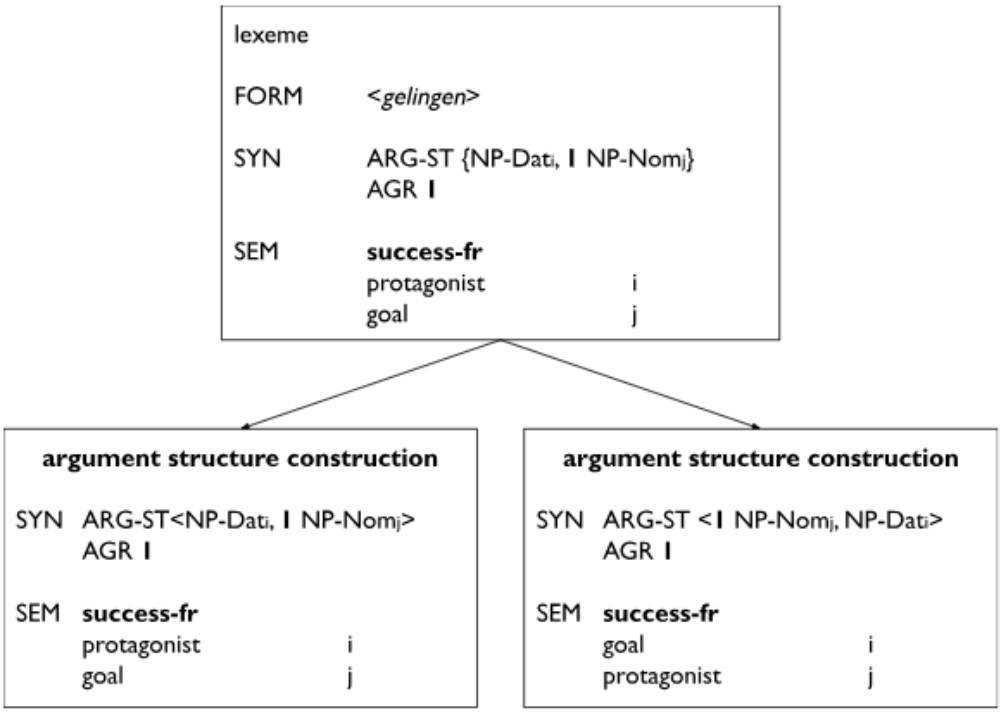


Figure 1.1 The compatibility of *gelingen* with both of its argument structure constructions, Dat-Nom and Nom-Dat

Figure 1.1 represents the lexical entry of the alternating verb *gelingen* 'succeed', in the upper part of the figure, along with both its argument structure constructions: the bottom left Dat-Nom construction and the bottom right Nom-Dat construction. The lexical entry comprises three fields: the FORM field, the SYN field and the SEM field. The FORM field captures the phonology of the verb and is rendered in between angled brackets. The SYN field represents both arguments that make up the argument structure, i.e. the dative and the nominative. Crucially, the list is unordered, as is customary in Construction Grammar (Müller 2012, *inter alia*), hence its being rendered in between curly brackets. In addition, the SYN field stipulates that the argument triggering agreement (AGR) is the nominative. The last field is the SEM field, which is associated with the success frame in this particular case. In line with Johnson et al. (2019), I posit both a protagonist and a goal as being a part of the semantic frame. The former corresponds to the argument marked in the dative and the latter to the argument marked in the nominative. As soon as the lexical entry is mapped onto either argument structure construction, Dat-Nom or Nom-Dat, the unordered list of participants is ordered and the relative order of the dative and the nominative arguments is fixed. Two examples illustrating either argument structure construction are supplied in (17a–b) below:

Dat-Nom

(17)	a.	*Mark*	*gelingt*	*die Flucht.*
		Mark.DAT	succeeds	the.NOM escape
		'Mark manages to escape.'		

Nom-Dat

(17)	b.	*Dieses Vorhaben*	*gelang*	*den Ruswilern*	*nun*
		this.NOM endeavour	succeeded	the.DAT Ruswilers.DAT	now
		vollumfänglich.			
		in.full			
		'This endeavour was of major success for the Ruswil team.'			

The model presented above solely applies to monosemous verbs; as soon as a verb exhibits multiple senses (see Section 1.6.5 below), there are two further scenarios: either the verb licenses multiple senses which are each compatible with either argument structure construction, or it licenses multiple senses of which only some are compatible with either argument structure construction, while others are compatible with the Nom-Dat argument structure construction only. For a formalisation of the different senses and their relation to different argument structure construction, see Section 7.4.2 below.

1.5 Alternating verbs and verb classes

Alternating verbs are verbs licensing two argument structure constructions. Hence, they build a class that is determined by syntactic subcategorisation, in the sense of argument structure. According to Levin (1993), the syntactic behaviour of a verb follows directly from its semantics. It should therefore be possible to detect a semantic criterion to identify alternating verbs. Levin's assumption about the interrelation between verb semantics and verb behaviour is the starting point of the German component of this thesis, since no exhaustive list of alternating verbs in that language is available in the literature. Of course, Levin's study is based on verb classes in English, but the supposed link between verb semantics and verb behaviour is one that is assumed to extend to other languages as well (cf. her discussion of the conative alternation in Warlpiri) (Levin 1993: 10–11).

Analysis of the list of known alternating verbs in German, as it is offered by Barðdal (2004), reveals that each of these display a thematic asymmetry between the dative and the nominative, in that the dative, being the lower-ranked case, is associated with the higher-ranked semantic role, which usually refers to an experiencer (see, however, Chapter 6, for a critical assessment of the application of the term 'experiencer' in the literature). By compiling a list of the German verbs exhibiting such an asymmetry, I endeavour to capture all verbs which are potentially alternating. Additionally, I break down the retrieved inventory into smaller semantic categories based on the work of Barðdal (2004), Barðdal et al. (2012) and Barðdal et al. (2016). However, I do not contend that alternation is more pervasive in certain semantic subclasses as opposed to others. Instead,

the classification presented in Chapter 2 is intended as a comprehensive overview of the data to the reader.

To summarise, I subscribe to Levin's view that semantics may provide useful pointers on the syntactic behaviour of certain verbs at the macrolevel, but I do not hypothesise alternation to be more prevalent in certain semantic pockets, thus refuting her view at the microlevel.

1.6 Where this study fits in

The goal of this thesis is to provide a systematic study of word order variation with German and Icelandic Dat-Nom/Nom-Dat verbs in verb-second clauses, thereby lending support to the positional test presented under (1). In doing so, the current study not only provides linearisation data for a host of verbs in both these languages, it also harnesses these data to make a meaningful contribution to ongoing discussions of subjecthood. The value of this dissertation is therefore both empirical and theoretical. The remainder of this section further motivates the choice of word order as a subject test (Section 1.6.1), followed by a discussion of extant work on word order, first with regard to German (Section 1.6.2), then with regard to Icelandic (Section 1.6.3). Section 1.6.4 provides an overview of the research questions on which this thesis builds. The final section, Section 1.6.5, offers a brief word on the data collection process.

1.6.1 Spotlighting the word order test

The reasons why this dissertation focuses on the word order test are threefold. First, word order provides a window onto argument structure, which, in turn, reflects the default order of constituents, i.e. linearisation under neutral conditions. As illustrated by (4b) and (5b) above, pragmatic factors may trigger a rearrangement of the linear order through object topicalisation, but such an operation is optional. Hence, attestations of neutral word order in corpora are bound to be markedly more frequent than structures involving topicalisation. Thus, the argument is essentially statistical in nature.

At this point, one may wonder what the difference is between an argument structure construction and a word order pattern. As argued in Section 1.4.1 above, the Dat-Nom and the Nom-Dat constructions are abstract grammatical units, involving basic syntactic patterns carrying basic semantics. It is therefore clear that alternating verbs are associated with two abstract argument structure constructions, and not just one. However, a word order pattern, i.e. dative-before-nominative or nominative-before-dative, is an instantiation

of an argument structure construction. Such a view is, without a doubt, fully in line with a constructional approach to language, which maintains that constructions are found at various levels of schematicity. As an example, one may consider Boas & Ziem's (2018: 13) distinction between the (abstract) ditransitive construction and (concrete) instantiations of such a construction. The former is referred to as a 'construction', the latter as a 'construct'.

Second, empirical research on German and Icelandic word order variation with non-canonically case marked verbs is comparatively scarce. For German, there are three notable studies within this area and for Icelandic there are two, only one of which is synchronic in nature. Also, the datasets on which the extant studies build are relatively small, containing approximately 1,000 data points for a set of 10 to 20 verbs at most. Furthermore, none of the available work reckons with notable differences between individual verbs, as it posits a verb-class effect without demonstrating why such an effect should exist in the first place. The current study demonstrates that there are surely greater generalisations to be made across verbs, but that it is nevertheless unsound to disregard intra-class variation.

Third, this study fills an important lacuna in existing research, as it brings to the fore a subject test for which empirical data are still largely lacking (see however Sections 1.6.2 and 1.6.3 below). Previous studies have already supplied data for raising constructions, control infinitives, conjunction reduction and clause-bound reflexivisation (Eythórsson & Barðdal 2005: 846–859, Barðdal 2006, Barðdal, Eythórsson & Dewey 2019, Barðdal 2023: Ch. 6). An in-depth study of the word order test is therefore the next logical step.

Finally, by spotlighting the word order test, I focus on a diagnostic which has traditionally not been considered a valid subjecthood test, as German word order is taken to be free rather than fixed (cf. Reis 1982: 191). Instead, subjecthood in German has always been established by resorting to the control and conjunction reduction tests, even though it is unclear why these should have a higher validity than any of the remaining diagnostics (Barðdal & Eythórsson 2018; Barðdal 2023: 51; Kyröläinen 2015: 253–254). I argue that word order may indeed serve as a valid subjecthood test. I also introduce into my study two sets of control verbs, one set for German and another for Icelandic, against which the findings from the target verbs are compared.

1.6.2 Previous studies on German

Three German corpus studies that report on valuable numbers are Bader & Häussler (2010), Verhoeven (2015) and Bader (2020). Bader & Häussler (2010) carry out a large-scale corpus study across both verb-second and verb-final clauses. Part of their study is also dedicated to the influence of non-canonical case marking on linearisation. In total, they collect 34 verb-second and 142 verb-final tokens displaying a thematic asymmetry between

the nominative and the accusative and the nominative and the dative, respectively. As many as 18 verb-second tokens, or 53%, instantiate the nominative-before-oblique order, while the remaining 16 tokens, or 47%, instantiate the reverse oblique-before-nominative order. Regarding the verb-final tokens, the numbers are as follows: 41, or 29%, realise the nominative before the oblique argument, whereas 101 tokens, or 71%, opt for the oblique-first order.

Bader & Häussler confirm the alternating hypothesis for both verb-second as well as verb-final clauses, but their study is flawed in four respects. First, they only include tokens in which the (dative or accusative) oblique argument, but not the nominative argument, is systematically definite, potentially inflating the share of oblique-before-nominative tokens. Second, their total number of tokens is very limited, as they only manage to collect 34 verb-second and 142 verb-final tokens. Third, they do not distinguish between dative-experiencers and accusative-experiencers, assigning to a single class all verbs displaying a thematic asymmetry. And finally, they do not report on any verb-specific tendencies. Sound though it is to single out verbs pertaining to the same class, it is not a given that all verbs within that class should exhibit the same behaviour.

Verhoeven (2015) is the second corpus study scrutinising the influence of non-canonical case marking on linearisation. She finds that ten verbs taking a dative experiencer and a nominative stimulus instantiate the dative-before-nominative order 38.5% of the time (448 out of 1,164 tokens) and the Nom-Dat order 61.5% of the time (716 out of 1,164 tokens). All of these are verb-second tokens containing two full NPs. Verhoeven's is the first study to validate the effect of non-canonical case marking on word order without obscuring the difference between dative-experiencer and accusative-experiencer verbs. Also, her sample is substantially larger than Bader & Häussler's (2010). On the downside, Verhoeven's sample comprises a mere ten verbs and she does not spell out how each of these potentially alternating verbs weighs in on the numbers she obtains, as her data are not coded for verb lemma.

The last relevant study is Bader (2020), who discusses word order in the German middlefield for clauses containing either an accusative or a dative pronoun and a nominative full NP. For a subset of 571 tokens containing a non-canonically case marked (dative or accusative) verb, Bader finds 92% of clauses to instantiate an oblique-before-nominative order. For another subset of 636 unaccusative verbs, whose oblique argument takes dative case by definition, he finds 91% of tokens that instantiate the dative-before-nominative order. At first blush, Bader's results seem to suggest that non-canonically case marked verbs are non-alternating Dat-Nom and Acc-Nom verbs, instead of alternating Dat-Nom/Nom-Dat verbs and alternating Acc-Nom/Nom-Acc verbs, respectively. However, it is of course well known that oblique pronouns precede nominative full NPs in the middlefield anyway (Shannon 2000; Kempen & Harbusch 2004; Heylen 2005, *inter alia*). Investigating neutral word order patterns, regardless of the topological field under study, requires a comparison that only involves full NPs.

1.6.3 Previous studies on Icelandic

For Icelandic, two studies have focused on word order in oblique-subject constructions so far: Barðdal & Eythórsson (2012) and Rott (2013). Barðdal & Eythórsson's is a diachronic study providing word order statistics for 12th- to 20th-century Icelandic collected from the Icelandic Parsed Historical Corpus. The data show that oblique subjects in the history of Icelandic have always had a proclivity for the preverbal slot, with 96 tokens, or 94%, instantiating the SVO alignment pattern as opposed to 6 tokens, or 4%, occurring in the reverse OVS pattern. A comparison with a baseline set by nominative subject verbs reveals that the latter are in fact more permissive of topicalisation, as these instantiate the SVO pattern 2,327 times, or 80%, and the OVS pattern 578 times, or 20%.

Rott (2013) studies the word order distributions of eight dative-subject verbs in Present-Day Icelandic, four each of the alternating Dat-Nom/Nom-Dat type and the non-alternating Dat-Nom type. His data stem from the Leipzig Corpora Collection and comprise both tokens in which either argument flanks the conjugated verb (i.e. Dat-V-Nom or Nom-V-Dat) as well as contexts of inversion (i.e. V-Dat-Nom or V-Nom-Dat). Rott empirically confirms that the distinction between both classes of Dat-Nom verbs is justified, as verbs of the non-alternating type never license inversion with the nominative and never realise the dative in an object position when the nominative is topicalised (for an example, see 3b above). In addition, non-alternating verbs almost always have the dative in first position, with the nominative occupying the middlefield. The alternating verbs in Rott's study behave radically differently: (1) they license inversion with either argument, dative or nominative, (2) they allow for either argument to be realised in an object position, and (3) either argument takes clause-initial position approximately equally often.

However, Rott's sample is fairly limited in size, as it only comprises 372 observations. Secondly, precisely like all three German studies discussed in Section 1.6.2, he obscures any intra-class variation by not discussing statistics for individual verbs, but only for the class as a whole. And thirdly, Rott also includes clausal nominatives, which greatly inflates the number of Dat-Nom attestations in his dataset, since clausal constituents tend to be longer than nominal constituents, thus tending towards clause-final position.

1.6.4 Research questions

The present thesis is dedicated to seven research questions, divided across both German and Icelandic. The first six questions pertain to either of these languages in isolation, with the first four questions focusing on German and the following two on Icelandic. The final question, question seven, addresses a cross-linguistic comparison of both languages under investigation.

i. German
   - How many, and which, alternating Dat-Nom/Nom-Dat verbs are there in Modern German?
   - To which semantic domains do German Dat-Nom/Nom-Dat verbs belong?
   - Which factors determine the speaker's choice for one linearisation pattern over the other?
   - Are all German Dat-Nom verbs genuine Dat-Nom/Nom-Dat verbs insofar as both arguments, dative and nominative, pass the word order test, or do some behave as either non-alternating Dat-Nom verbs or as ordinary Nom-Dat verbs?

ii. Icelandic
   - How do alternating verbs in Icelandic behave vis-à-vis the word order test?
   - Which factors determine the speaker's choice of one linearisation pattern over the other?

iii. Cross-linguistic comparison
   - How do alternating verbs behave across Icelandic and German? Are there any cross-linguistic generalisations to be made?

First, how many, and which, alternating Dat-Nom/Nom-Dat verbs exist in Modern German? Barðdal (2004) has shown that there are at least 45 such verbs, but their exact number remains unknown. The first goal of this study is to expand upon Barðdal's list in order to compile an exhaustive inventory of all German verbs alternating between the Dat-Nom and the Nom-Dat orders.

Second, what are the semantic domains to which German Dat-Nom/Nom-Dat verbs pertain? Barðdal (2004) and Barðdal et al. (2016) have shown that German non-canonically case-marked verbs, regardless of the case of the oblique argument and the verbal case frame, belong to very specific semantic classes. Nevertheless, the question remains what the specific subspectrum for alternating Dat-Nom/Nom-Dat verbs looks like. In addition, it is unclear how the size of each semantic field relates to another. That is: are all semantic classes represented equally, boasting roughly the same number of candidate verbs, or do some classes have more members than others?

Third, are all German Dat-Nom verbs genuine Dat-Nom/Nom-Dat verbs in that both arguments, dative and nominative, pass the word order test, or do some behave as either non-alternating Dat-Nom verbs or as ordinary Nom-Dat verbs? As mentioned in Sections 1.1 and 1.2 above, I adopt as a working hypothesis the view that all German Dat-Nom verbs are alternating Dat-Nom/Nom-Dat verbs (Eythórsson & Barðdal 2005: 868; Barðdal, Eythórsson & Dewey 2014, 2019: 131–148; Rott 2016: 239–249; Barðdal 2023: Ch. 6). Thus, both nominal arguments are expected to pass the word order test by occupying the preverbal slot with reasonable frequency. In order to determine what counts as

'reasonable frequency', the data are compared to a control set of ordinary Nom-Dat verbs like *danken* 'thank' and others of that type. The extent to which the latter instantiate the topicalised dative-before-nominative order may serve as a baseline against which to compare frequencies for the Dat-Nom/Nom-Dat target verbs.

Fourth, which factors determine the speaker's choice for one linearisation pattern over the other? In order to address this question, I appeal to a set of variables that have been identified in the literature as relevant predictors underlying word order variation, viz. case, (pro)nominality, pronoun type (if applicable), referentiality, person, number, definiteness, animacy and length. By additionally coding the data for verb lemma, I aim to uncover whether and, if so, to which extent individual verbs weigh in on the choice between the Dat-Nom and the Nom-Dat order.

Fifth, how do alternating verbs behave in Icelandic? Even though the study of alternating predicates finds its origin in that language, empirical data are so far lacking entirely. By sampling a set of Dat-Nom/Nom-Dat verbs and then comparing these findings to a double baseline set by both non-alternating Dat-Nom verbs and ordinary Nom-Dat verbs, the aim is to document word order variation for Icelandic in very much the same way as for German (cf. research question three above).

The sixth research question is the Icelandic counterpart of question four above on German, as it seeks to uncover the linguistic motivation for one word order pattern compared to the other, i.e. Dat-Nom vs. Nom-Dat. The predictor variables selected for German are also applied to Icelandic.

The final question this dissertation seeks to address is how alternating verbs behave across both Icelandic and German. Do these predicates license a comparable variation in word order, or are numbers markedly more skewed in one language compared to the other? Additionally, do the predictor variables have a comparable degree of explanatory power in both languages? Or do certain variables have more explanatory power in one language compared to the other?

1.6.5 A brief word on the data collection process

The present thesis studies linearisation patterns with verbs displaying a thematic asymmetry between the dative and the nominative arguments. Such an asymmetry is not only inherent to (monosemous) verbs such as *gelingen*, but also to polysemous verbs. The latter come in two types: polysemous verbs such as *genügen* 'be enough, be sufficient; fulfil' and polysemous verbs such as *passen* 'fit (clothes); be suitable, be convenient'. Starting with *genügen*, only its former sense, 'be enough, be sufficient', displays a thematic asymmetry between the dative and the nominative argument, not its latter sense, 'fulfil', which is also more agentive. The 'fulfil' sense is exemplified in (18a–b) and the 'be enough, be sufficient' sense in (19a–b):

Nom-Dat

Neutral nominative-before-dative order

(18) a. *Die Zimmer* *genügen* *hohen Ansprüchen von Bewohnern und Gästen.*
the.NOM rooms fulfil high.DAT standards.DAT of residents and guests
'The rooms meet the high standards of residents and guests.'

Nom-Dat

Topicalised dative-before-nominative order

(18) b. *Diesen Anforderungen* *genügte*
these.DAT requirements fulfilled
der Sachvortrag des Klägers *nicht.*
the.NOM factual.submission the.GEN plaintiff not
'These requirements, the plaintiff's factual submission did not meet.'

Dat-Nom/Nom-Dat

Neutral nominative-before-dative order

(19) a. *8 GB* *möchten* *dem Nutzer* *genügen.*
eight gigabytes.NOM should the.DAT user suffice
'Eight gigabytes should be enough for the user.'

Dat-Nom/Nom-Dat

Neutral dative-before-nominative order

(19) b. *Danny* *genügte* *das Ergebnis.*
Danny.DAT sufficed the.NOM result
'Danny was satisfied with the result.'

Comparing examples (18–19), it stands out that, regardless of the linear order, (18a–b) do not display an asymmetric mapping of the thematic roles onto their syntactic arguments, whereas (19a–b) do. Hence, (18a) is an instance of neutral nominative-before-dative word order, whereas (18b) is an instance of topicalisation, in which the dative is fronted for pragmatic reasons. Example (19a), in contrast, instantiates the Nom-Dat argument structure construction for the alternating 'be enough, be sufficient' sense, whereas example (19b) instantiates the Dat-Nom argument structure construction for that very same sense.

Thus, it is evident that the data collection cannot take any potential variation in word order with verbs of the *genügen*-type at face value. After all, either order, nominative-before-dative as well as dative-before-nominative, might also feature a sense lacking a thematic asymmetry. Hence, for every token of the *genügen*-type collected, I systematically eliminated all agentive senses, regardless of the word order pattern they occurred in.

Let us now consider the polysemous verb *passen*, whose literal 'fit' sense is represented in examples (20a–b) below, while examples (21a–b) exhibit the metaphorical 'be suitable, be convenient' sense:

Dat-Nom/Nom-Dat

Neutral nominative-before-dative order

(20)	a.	*Nicht jeder Schuh*	*passt*	*einem*	*auch.*
		not every.NOM shoe	fits	one.DAT	also
		'Not every shoe fits you.'			

Dat-Nom/Nom-Dat

Neutral dative-before-nominative order

(20)	b.	*Steven*	*passte*	*die Kleidung*	*wie angegossen.*
		Steven.DAT	fitted	the.NOM clothes	like moulded.on
		'Steven's clothes fitted him like a glove.'			

Dat-Nom/Nom-Dat

Neutral nominative-before-dative order

(21)	a.	*Die großzügigen Ausnahmeregelungen*	*passten*	
		the.NOM generous exemptions	suited	
		den Umweltschützern	*nicht.*	
		the.DAT environmentalists.DAT	not	
		'The generous exemptions did not suit the environmentalists.'		

Dat-Nom/Nom-Dat

Neutral dative-before-nominative order

(21)	b.	*Ihm*	*passte*	*der Zeitpunkt*	*nicht.*
		him.DAT	suited	the.NOM time	not
		'He did not find the timing suitable.'			

In contrast to *genügen*, either sense of polysemous *passen* may be associated with either the Dat-Nom construction or the Nom-Dat construction. As a result, tokens featuring either the dative-before-nominative order as well as the nominative-before-dative order have readily been included into the dataset.

All further details concerning the data collection and annotation processes are discussed in Chapters 2, 4, 5 and 6 below.

1.7 Genesis of the thesis

This thesis is article-based with each of the five content chapters having been submitted to peer-reviewed journals. At the time of writing, three papers have already been published. The remaining two have been accepted for publication. The bibliographic entries for each of the five papers are presented below:

- Chapter 2 = Somers, Joren. 2021. Onpersoonlijke datief-nominatiefwerkwoorden in het hedendaagse Duits: Definitie, inventaris en semantische classificatie. *Handelingen van de Koninklijke Zuid-Nederlandse Maatschappij voor Taal- en Letterkunde en Geschiedenis* 75: 211–238.
- Chapter 3 = Somers, Joren. 2023. Linearization and unaccusativity: The relative order of dative and nominative arguments with five German verbs of success and failure. *Papers of the Linguistic Society of Belgium* 17: 54–75.
- Chapter 4 = Somers, Joren, Gard B. Jenset & Jóhanna Barðdal. 2024. Argument structure constructions in competition: The Dat-Nom/Nom-Dat alternation in Icelandic. *Nordic Journal of Linguistics*: online first article. DOI: https://doi.org/10.1017/S0332586524000039.
- Chapter 5 = Somers, Joren, Gard B. Jenset & Jóhanna Barðdal. Subjecthood and argument structure of synonymous Dat-Nom/Nom-Dat verbs across German and Icelandic. Accepted for publication in *Lingua*.
- Chapter 6 = Somers, Joren, Torsten Leuschner, Ludovic De Cuypere & Jóhanna Barðdal. A corpus-based analysis of the Dat-Nom/Nom-Dat alternation in German. Accepted for publication in *Zeitschrift für Sprachwissenschaft*.

As may be deduced from the bibliographic entries above, the order in which the articles have been incorporated into this thesis respects the order in which they were written, with the first article having been written and published first, and the last article having been finalised last. As for the three co-authored papers, i.e. Somers, Jenset & Barðdal (2024), Somers, Jenset & Barðdal (accepted) and Somers et al. (accepted), the contributions of each of the authors is as follows:

- Somers, Jenset & Barðdal (2024): All three authors designed the research and planned and wrote the manuscript; Jóhanna Barðdal and Joren Somers gathered the data; Joren Somers coded the data; Gard B. Jenset performed the computational analysis; all authors contributed equally to the discussion and the interpretation of the results.
- Somers, Jenset & Barðdal (accepted): All three authors designed the research and planned and wrote the manuscript; Jóhanna Barðdal and Joren Somers gathered the data; Joren Somers coded the data; Gard B. Jenset performed the computational analysis; all authors contributed equally to the discussion and the interpretation of the results.
- Somers, Leuschner, De Cuypere & Barðdal (accepted): All four authors designed the research and planned and wrote the manuscript; Joren Somers gathered and coded the data; Ludovic De Cuypere performed the computational analysis; all authors contributed equally to the discussion and the interpretation of the results.

1.8 Overview

This thesis is organised into five content chapters. Chapter 2 introduces the concept of alternating verbs for in German, which are delineated from similar structures comprising a so-called free dative. The second part of this chapter is dedicated to the compilation of an exhaustive list of alternating Dat-Nom/Nom-Dat verbs. With the help of Barðdal (2004), the online version of the *Duden Universalwörterbuch* and a host of secondary sources, I manage to identify no fewer than 146 verbs showing potentially alternating behaviour in one or more of their senses. Finally, I cluster the obtained verbs semantically using a typology suggested by Barðdal et al. (2016). The results show a remarkable Zipfian distribution, with heavy clustering around the semantic fields of emotion, cognition, and gain and benefit.

Chapter 3 constitutes a case study of a cluster of five closely related German verbs, i.e. *gelingen* 'to succeed', *glücken* 'to succeed', *missglücken* 'to fail', *misslingen* 'to fail', and *missraten* 'to fail'. Not only do these verbs exhibit a high degree of semantic coherence, they are also members of the class of unaccusatives, testifying to their syntactic relatedness. The results of this second corpus study show that all five verbs systematically alternate between the Dat-Nom and the Nom-Dat argument structures in a configuration with double NPs.

Chapter 4 introduces the Icelandic dataset, which comprises a total of 15 verbs: five Dat-Nom verbs, five Nom-Dat verbs and five Dat-Nom/Nom-Dat verbs. It is established that each of these syntactic classes correlates with a different argument structure construction. The verbs in question then serve as input for the first corpus study, which builds on 3,000 random observations extracted from the isTenTen20 corpus, 1,000 for each class. By comparing the findings for alternating Dat-Nom/Nom-Dat verbs to the baseline set by both non-alternating Dat-Nom verbs and ordinary Nom-Dat verbs, it is shown that Dat-Nom/Nom-Dat verbs behave markedly differently vis-à-vis the word order test. A logistic regression analysis interprets the variation in word order for this class in light of nine different variables: case, (pro)nominality, pronoun type (if applicable), referentiality, person, number, definiteness, animacy and length.

Chapter 5 once again brings to the fore the 15 Icelandic verbs first analysed in Chapter 4, now comparing their behaviour to cognate or (near-)synonymous verbs in Modern German. Each Icelandic verb is matched to a German verb acquired in Chapter 2. The cross-Germanic comparison discussed in the present chapter reveals important similarities and remarkable differences between both languages. One example of parallel behaviour across Icelandic and German relates to the effect of nominative pronouns, which exhibit a strong preference for the preverbal argument slot. A key difference between the two languages under investigation is their sensitivity to a pronoun's referential status. More specifically, it is striking that nominative correlates, i.e. third-person pronouns acting as

placeholders for a subclause, are consistently assigned preverbal position in Icelandic, while they are much more sensitive to alternation in German.

The final chapter, Chapter 6, which is again committed to German, breaks down into three parts. The first part is dedicated to a discussion of the terms 'psych verbs', 'experiential verbs', 'unaccusative verbs' and '*piacere* verbs'. It is argued that the categories referred to by these terms share a meronymic relation with one another, and that each can be subsumed under the umbrella of Dat-Nom/Nom-Dat verbs. The second part of Chapter 6 discusses the main semantic and syntactic features of Dat-Nom/Nom-Dat verbs. Part three is entirely devoted to the second corpus study, which takes as its starting point the verbs listed in Chapter 2. All 146 of these have been subjected to a data inspection stage, upon which 76 have been retained. The German corpus study applies the exact same methodology as the Icelandic study discussed in Chapter 4. Furthermore, it is based on the deTenTen13 corpus, which belongs to the same family of corpora as the isTenTen20 employed for Icelandic, thus rendering the results perfectly comparable across languages. In total, the German study yields 13,292 data points, which are first analysed bivariately by scrutinising the correlation between case marking on the one hand and (pro)nominality, pronoun type (if applicable), person, number, definiteness, animacy and number, on the other. A mixed-effects logistic regression analysis using the verb as a random intercept identifies the variables that determine the variation between both word order patterns, dative-before-nominative and nominative-before-dative.

In conclusion, I summarise the main findings of this thesis, while also interpreting them in light of the working hypotheses. I also provide an outlook on future research within the area of non-canonical case marking in general and of alternating Dat-Nom/Nom-Dat verbs in particular.

References

Bach, Emmon. 1962. The order of elements in a transformational grammar of German. *Language* 38(3): 263–269.

Bader, Markus. 2020. The position of object pronouns in the German middlefield. *Linguistics* 58(4): 1059–1115.

Bader, Markus and Jana Häussler. 2010. Word order in German: A corpus study. *Lingua* 120: 717–762.

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

Barðdal, Jóhanna. 2000. The subject is nominative! On obsolete axioms and their deep-rootedness. In Carl-Erik Lindberg & Steffen Nordahl Lund (eds.), *17th Scandinavian Conference of Linguistics*, 93–117. Odense: Institute of Language and Communication.

Barðdal, Jóhanna. 2001. The perplexity of Dat-Nom verbs in Icelandic. *Nordic Journal of Linguistics* 24: 47–70.
Barðdal, Jóhanna. 2004. The semantics of the impersonal construction in Icelandic, German and Faroese: Beyond thematic roles. In Werner Abraham (ed.), *Focus on Germanic typology*, 105–137. Berlin: Akademie Verlag.
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. 2006. Control infinitives and case in Germanic: 'Performance error' or marginally acceptable constructions. In Leonid Kulikov, Andrej Malchukov & Peter de Swart (eds.), *Case, Valency and Transitivity*, 147–177. Amsterdam: John Benjamins.
Barðdal, Jóhanna & Thórhallur Eythórsson. 2009. The origin of the oblique-subject construction: An Indo-European comparison. In Vít Bubeník, John Hewson & Sarah Rose (eds.), *Grammatical Change in Indo-European Languages*, 179–193. Amsterdam: John Benjamins.
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. 2018. What is a subject? The nature and validity of subject tests. In Jóhanna Barðdal, Na'ama Pat-El, & Stephen Mark Carey (eds.), *Non-Canonically Case-Marked Subjects: The Reykjavík – Eyjafjallajökull Papers*, 257–273. Amsterdam: John Benjamins.
Barðdal, Jóhanna & Thórhallur Eythórsson. 2020. How to identify cognates in syntax: Taking Watkins' legacy one step further. In Jóhanna Barðdal, Spike Gildea & Eugenio R. Lujan (eds.), *Reconstructing Syntax*, 197–238. Leiden: Brill.
Barðdal, Jóhanna & Spike Gildea. 2015. Diachronic Construction Grammar: Epistemological context, basic assumptions and historical implications. In Jóhanna Barðdal, Elena Smirnova, Lotte Sommerer & Spike Gildea (eds.), *Diachronic Construction Grammar*, 1–50. Amsterdam: John Benjamins.
Barðdal, Jóhanna, Carlee Arnett, Stephen Mark Carey, Þórhallur Eyþórsson, Gard B. Jenset, Guus Kroonen & Adam Oberlin. 2016. Dative subjects in Germanic: A computational analysis of lexical semantic verb classes across time and space. *Language Typology and Universals* 69(1): 49–84.
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.
Bayer, Josef. 2004. Non-nominative subjects in comparison. In Peri Bhaskararao & Karumuri Venkata Subbarao (eds.), *Non-Nominative Subjects I*, 31–58. Amsterdam: John Benjamins.
Bernódusson, Helgi. 1982. Ópersónulegar setningar [Impersonal sentences]. Reykjavík: University of Iceland Master's Thesis.
Bierwisch, Manfred. 1963. *Grammatik des deutschen Verbs*. Berlin: Akademie Verlag.
Boas, Hans C. & Alexander Ziem. 2018. Approaching German syntax from a constructionist perspective. In Hans C. Boas & Alexander Ziem (eds.), *Constructionist Approaches to Syntactic Structures in German*, 1–46. Berlin/Boston: De Gruyter Mouton.
Bybee, Joan. 1998. A functionalist approach to grammar and its evolution. *Evolution of Communication* 2(2): 249–278.
Croft, William. 1998. Event structure in argument linking. In Miriam Butt & Wilhelm Geuder (eds.), *The projection of Arguments: Lexical and Compositional Factors*, 1–43. Stanford: CSLI.

Croft, William. 2001. *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford: Oxford University Press.
Croft, William. 2012. *Verbs: Aspect and Causal Structure*. Oxford: Oxford University Press.
Croft, William & D. Alan Cruse. 2004. *Cognitive Linguistics*. Cambridge: Cambridge University Press.
Dryer, Matthew S. 2013. Order of Subject, Object and Verb. In: Dryer, Matthew S. & Haspelmath, Martin (eds.) *WALS Online* (v2020.3) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.7385533, (Available online at http://wals.info/chapter/81, Accessed on 2024-02-22.).
Eisenberg, Peter. 2013. *Grundriss der deutschen Grammatik. 2: Der Satz*. Stuttgart: Metzler.
Eythórsson, Thórhallur & Jóhanna Barðdal. 2005. Oblique subjects: A common Germanic inheritance. *Language* 81(4): 824–881.
Fanselow, Gisbert. 2002. Quirky subjects and other specifiers. In Ingrid Kaufmann & Barbara Stiebels (eds.), *More than Words: A Festschrift for Dieter Wunderlich*, 227–250. Berlin: Akademie Verlag.
Fanselow, Gisbert. 2003. Zur Generierung der Abfolge der Satzglieder im Deutschen. *Neue Beiträge zur Germanistik* 112: 9–47.
Fillmore, Charles J., Paul Kay & Mary Kay O'Connor. 1988. Regularity and idiomaticity in grammatical constructions: The case of 'let alone'. *Language* 64: 501–538.
Goldberg, Adele. 1995. *Constructions: A Construction-Grammar Approach to Argument Structure*. Chicago: Chicago University Press.
Goldberg, Adele. 2003. Constructions: A new theoretical approach to language. *Trends in Cognitive Sciences* 7(5): 219–224.
Goldberg, Adele. 2006. *Constructions at Work: The Nature of Generalization in Language*. Oxford: Oxford University Press.
Goldberg, Adele. 2019. *Explain me this: Creativity, competition, and the partial productivity of constructions*. Princeton: Princeton University Press.
Greenberg, Joseph H. 1966. *Universals of Language*. Cambridge, MA: The MIT Press.
Haider, Hubert. 2005. How to turn German into Icelandic – and derive the OV–VO contrasts. *Journal of Comparative Germanic Linguistics* 8: 1–53.
Haider, Hubert. 2010. *The Syntax of German*. Cambridge: Cambridge University Press.
Haspelmath, Martin. 2001. Non-canonical marking of core arguments in European languages. In Alexandra Aikhenvald, Robert M. W. Dixon & Masayuki Onishi (eds.), *Non-canonical subjects and objects*, 53–83. Amsterdam: John Benjamins.
Heylen, Kris. 2005. *Zur Abfolge (pro)nominaler Satzglieder im Deutschen: Eine korpusbasierte Analyse der relativen Abfolge von nominalen Subjekt und pronominalem Objekt im Mittelfeld*. PhD dissertation, Katholieke Universiteit Leuven.
Johnson, Cynthia A., Peter Alexander Kerkhof, Leonid Kulikov, Esther Le Mair and Jóhanna Barðdal. 2019. How to succeed in Indo-European without really trying? Argument structure, conceptual metaphor and semantic change. *Diachronica* 36(4): 463–508.
Jónsson, Jóhannes Gísli. 1997–98. Sagnir með aukafallsfrumlagi [Verbs selecting for oblique subjects]. *Íslenskt mál* 19–20: 11–43.
Kay, Paul & Charles J. Fillmore. 1999. Grammatical constructions and linguistic generalizations: The 'What's X doing Y' construction. *Language* 75: 1–33.
Kempen, Gerard & Karin Harbusch. 2004. A corpus study into word order variation in German subordinate clauses: Animacy effects linearization independently of grammatical function assignment. In Thomas Pechmann & Christopher Habel (eds.), *Multidisciplinary Approaches to Language Production*, 173–181. Berlin: Mouton de Gruyter.
Kempen, Gerard & Karin Harbusch. 2005. The relationship between grammaticality ratings and corpus frequencies: A case study into word order variability in the midfield of German clauses. In Stephan Kepser & Marga Reis (eds.), *Linguistic Evidence*, 329–349. Berlin: Mouton de Gruyter.
Koster, Jan. 1975. Dutch as an SOV language. *Linguistic Analysis* 1: 111–136.

Kyröläinen, Aki-Juhani. 2015. From canon to monolith to clusters: A constructionist model of subjecthood in Russian. In Marja-Liisa Helasvuo & Tuomas Huumo (eds.), Subjects in Constructions – Canonical and Non-Canonical, 253–291. Amsterdam: John Benjamins.

Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar, Vol. I: Theoretical Prerequisites*. Stanford: Stanford University Press.

Levin, Beth. 1993. *English Verb Classes and Alternations: A Preliminary Investigation*. Chicago: University of Chicago Press.

Müller, Stefan. 2012. *On the Copula, Specificational Constructions and Type Shifting*. Berlin: Freie Universität Berlin Master's Thesis.

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

Reis, Marga. 1982. Zum Subjektbegriff im Deutschen. In Werner Abraham (ed.), *Satzglieder im Deutschen: Vorschläge zur syntaktischen, semantischen und pragmatischen Fundierung*, 171–211. Tübingen: Gunter Narr.

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.

Rögnvaldsson, Eiríkur. 1996. Frumlag og fall að fornu [Subject and case in Icelandic middle ages]. *Íslenskt mál og almenn málfræði* 18: 37–69.

Ruppenhofer, Josef, Michael Ellsworth, Miriam R. L. Petruck, Christopher R. Johnson, Collin F. Baker & Jan Scheffczyk. 2016. *Framenet II: Extended Theory and Practice*. Berkeley: International Computer Science Institute.

Shannon, Thomas F. (2000). On the order of (pro)nominal arguments in Dutch and German. In Thomas F. Shannon and Johan P. Snapper (eds.), *The Berkeley Conference on Dutch Linguistics 1997: The Dutch Language at the Millennium*, 145–196. New York: University Press of America.

Sigurðsson, Halldór Ármann. 1989. *Verbal Syntax and Case in Icelandic*. PhD dissertation, Lund University.

Sigurðsson, Halldór Ármann. 2002. To be an oblique subject: Russian vs. Icelandic. *Natural Language and Linguistic Theory* 20(4): 691–724.

Sigurðsson, Halldór Ármann. 2006. The Nominative Puzzle and the Low Nominative Hypothesis. *Linguistic Inquiry* 37: 289–308.

Somers, Joren and 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.

Talmy, Leonard. 1985. Force dynamics in language and thought. *Chicago Linguistic Society* 21: 293–337.

Talmy, Leonard. 1988. Force dynamics in language and cognition. *Cognitive Science* 12(1): 49–100.

Verhoeven, Elisabeth. 2015. Thematic asymmetries do matter! A corpus study of German word order. *Journal of Germanic Linguistics* 27(1): 45–104.

Witzlack-Makarevich, Alena. 2019. Argument selectors: A new perspective on grammatical relations: An introduction. In Alena Witzlack-Makarevich & Balthasar Bickel (eds.), *Argument Selectors: A New Perspective on Grammatical Relations*, 1–38. Amsterdam: John Benjamins.

Wood, Jim & Halldór Ármann Sigurðsson. 2014. *Let*-causatives and (a)symmetric DAT-NOM constructions. *Syntax* 17(3): 269–298.

Wunderlich, Dieter. 2008. The force of lexical case: German and Icelandic compared. In Kristin Hanson & Sharon Inkelas (eds.), *The Nature of the Word: Essays in Honor of Paul Kiparsky*, 587–620. Massachusetts: MIT.
Zaenen, Annie, Joan Maling & Höskuldur Thráinsson. 1985. Case and grammatical Functions: The Icelandic passive. *Natural Language and Linguistic Theory* 3: 441–483.

Chapter 2
Onpersoonlijke datief-nominatiefwerkwoorden in het Hedendaagse Duits: Definitie, inventaris en semantische classificatie

Several studies have drawn attention to the fact that subjecthood is not necessarily associated with nominative case-marking (Andrews 1976; Þráinsson 1979; Zaenen, Maling & Þráinsson 1985; Sigurðsson 1989, *inter alia*). This article seeks to contribute to the study of non-canonically case marked subjects by zooming in on nominative-dative verbs in Present-Day with a subject-like oblique argument (henceforth: impersonal dative-nominative verbs). The aim of this article is twofold. First, I present an exhaustive list of all German impersonal dative-nominative verbs, and secondly, I aim to classify them according to the semantic typology suggested by Barðdal et al. (2016). The results show that impersonal dative-nominative verbs are far more numerous than hitherto assumed, and that their semantic distribution gravitates towards meanings pertaining to emotions, cognitive processes, and contexts of gain and benefit.

2.1 Het concept

Het voorwerp van deze studie zijn werkwoorden met **subjectachtige datieven**. Het betreft met andere woorden structuren met niet-canonieke casusmarkering waarvan het datiefargument een rol vervult die veelal met het onderwerp geassocieerd wordt. In de

Relationele Grammatica worden dergelijke werkwoorden ook wel aangeduid als inversiepredicaten (Blake 2001: 79), maar in navolging van Barðdal (2004) zal ik de term **onpersoonlijke datief-nominatiefwerkwoorden** hanteren.[1]

De volgende voorbeelden mogen de zonet geijkte begrippen verder illustreren:

(1) a. *Die Kinder* *missfallen* *der Frau.*
de.NOM kinderen misvallen.3P.PL de.DAT vrouw
'De kinderen bevallen de vrouw niet.'

(2) a. *Die Kur* *hat* *mir* *gutgetan.*
de.NOM kuur heeft.3P.SG mij.DAT deugddoen.VD
'De kuur heeft me deugd gedaan.'

(3) a. *Diesem Buch* *fehlt* *eine klare Linie.*
dit.DAT boek ontbeert.3P.SG een.NOM duidelijke lijn
'Dit boek ontbeert een duidelijke lijn.'

(4) a. *Dieses Wort* *kommt* *mir* *absurd* *vor.*
dit.NOM woord komt.3P.SG mij.DAT absurd voor
'Dit woord komt mij absurd voor.'

In elk van de bovenstaande zinnen neemt de datief een rol op die in canonieke structuren door de nominatief wordt vervuld. Dat blijkt ook uit de (b)-voorbeelden hieronder, waarin het perspectief telkens wordt omgekeerd en de subjectachtige datief uit (1a)-(4a) middels een ander werkwoord in de nominatief wordt gerealiseerd:

(1) b. *Die Frau* *mag* *die Kinder* *nicht.*
de.NOM vrouw mag.3P.SG de.ACC kinderen niet
'De vrouw houdt niet van de kinderen.'

(2) b. *Ich* *habe* *die Kur* *genossen.*
ik.NOM heb.1P.SG de.ACC kuur genieten.VD
'Ik heb van de kuur genoten.'

(3) b. *Dieses Buch* *hat* *keine klare Linie.*
dit.NOM boek heeft.3P.SG geen.ACC duidelijke lijn
'Dit boek ontbeert een duidelijke lijn.'

[1] Ik ben me ervan bewust dat de term 'onpersoonlijk' ook wordt gebruikt om te verwijzen naar uitingen met een loos onderwerp (bv. Ndl. *het regent*) of zinnen zonder nominatiefargument (bv. Du. *mir ist kalt* 'ik heb het koud'). Bij gebrek aan een andere term, en in navolging van Barðdal (2004), zal ik hem echter gebruiken om te verwijzen naar nominatief-datiefwerkwoorden met een subjectachtige datief.

(4) b. *Ich* *finde* *dieses Wort* *absurd.*
ik.NOM vind.1P.SG dit.ACC woord absurd
'Ik vind dit woord absurd.'

Het meest wezenlijke kenmerk van onpersoonlijke datief-nominatiefstructuren is dat ze laag-agentief zijn (Barðdal 2001: 103, op basis van Barðdal 2004; Barðdal et al. 2016: 50): ze drukken onder meer ondervinding uit (zoals in 1a), baat (zoals in 2a), bezit (zoals in 3a) of oordeel (zoals in 4a). Subjectachtige datieven zijn dus niet enkel geassocieerd met de semantische rol van de ondervinder (Eng. *experiencer*), hoewel die in vele talen wel niet-canonieke casusmarkering draagt (zie bv. Haspelmath 2001). Ik sluit me zo aan bij auteurs als Jónsson (1997–98), Barðdal (2004), Barðdal et al. (2012) en Van Valin (2018), die er eerder al op wezen dat datieve casusmarkering een veel breder semantisch domein bestrijkt dan enkel de signalisatie van ontvangers (Eng. *recipients*), beneficiënten (Eng. *beneficiaries*) en ondervinders. Voorts is het ook niet zo dat subjectachtige datieven per se naar bezielde entiteiten dienen te verwijzen: vaak is dat wel het geval, maar zoals (3a) illustreert, is dat geenszins een wetmatigheid. Dat geldt, mutatis mutandis, ook voor de nominatief, die vaak bezield is, maar dat evengoed niet kan zijn (cf. voorbeeld 1a).

In het verleden werd al meermaals attent gemaakt op de syntactische eigenschappen die niet-canoniek gemarkeerde onderwerpen delen met canonieke (nominatieve) onderwerpen. Eén eigenschap in het bijzonder is het vermogen van subjectachtige datieven om de eerste zinsplaats te bezetten in mededelende zinnen, zonder dat die positie als gemarkeerd wordt ervaren (zie voorbeeld 5a; Barðdal et al. 2019: 131). Barðdal et al. (2019) argumenteren dat onpersoonlijke datief-nominatiefwerkwoorden daarnaast ook een tweede casusframe toelaten, namelijk een constructie met zinsinitiële nominatief (zie voorbeeld 5b). De volgende voorbeelden, die ik in licht gewijzigde vorm overnam uit Barðdal et al. (2019), illustreren die eigenschap:

(5) a. *Mir* *hat* *das Hotel* *gut* *gefallen.*
mij.DAT heeft.3P.SG het.NOM hotel goed bevallen.VD
'Ik vond het hotel een meevaller.'

(5) b. *Dieses Hotel* *hat* *mir* *gut* *gefallen.*
dit.NOM hotel heeft.3P.SG mij.DAT goed bevallen.VD
'Het hotel beviel me.'

De variatie in woordvolgorde die onpersoonlijke datief-nominatiefwerkwoorden toelaten, is de reden waarom ze ook wel als 'alternerende werkwoorden' worden aangeduid (Barðdal et al. 2019, *inter alia*). Andere eigenschappen die subjectachtige datieven delen met canonieke onderwerpen, zijn reflexivisering, voegwoordelijke reductie en controle-infinitieven. Die syntactische eigenaardigheden vallen buiten het blikveld van deze bijdrage, maar voor een uitgebreide bespreking kan de lezer terecht bij Barðdal 2006, Eyþórsson & Barðdal 2005 en Barðdal et al. 2019.

2.2 De reikwijdte van de onpersoonlijke datief-nominatiefconstructie in het Hedendaagse Duits

Het hedendaagse Duits kent heel wat structuren die raakvlakken vertonen met onpersoonlijke datief-nominatiefwerkwoorden, maar die er om diverse redenen geen zijn. In deze sectie zal ik daarom trachten mijn uitgangspunt verder duidelijk te maken, en de subjectachtige datief af te zetten tegen andere gebruikswijzen van de datief, in de eerste plaats tegen de zogenoemde 'vrije datief':

(6) *Da* *brach* *ihm* *die Stimme.*
toen brak.3P.SG hem.DAT de.NOM stem
'Toen brak zijn stem.'

(7) *Mir* *ist* *das Glas* *gebrochen.*
mij.DAT is.3P.SG het.NOM glas breken.VD
'Ik heb het glas gebroken.'

(8) *Die Übung* *ist* *dem Mädchen* *zu schwierig.*
de.NOM oefening is.3P.SG het.DAT meisje te moeilijk
'Het meisje vindt de oefening te moeilijk.'

(9) *Komm* *mir* *doch* *nicht* *immer* *mit Ausreden!*
kom.IMP me.DAT toch niet steeds met smoesjes
'Kom (me) toch niet steeds met smoesjes af!'

Alle bovenstaande voorbeeldzinnen bevatten datieven die in de literatuur als zogenoemde 'vrije datieven' worden omschreven, wat betekent dat ze niet in het valentieschema van het hoofdwerkwoord vervat zitten.[2] Dat is meteen ook de reden waarom ze buiten het blikveld van deze bijdrage vallen: subjectachtige datieven zijn immers geen adjuncten, maar volwaardige argumenten.

De datief in (6) wordt ook wel de possessieve datief (of 'Pertinenzdativ') genoemd. Hij codeert een innige relatie tussen het nominatiefargument (dat vaak een lichaamsdeel of een kledingstuk aanduidt) en het datiefargument (dat de bezitter aanduidt) die zich het best laat vatten als een van onvervreemdbaar bezit (Lamiroy & Delbecque 1998: 29). Die possessieve betekeniscomponent delen constructies met een possessieve datief met een aantal onpersoonlijke datief-nominatiefwerkwoorden die gebrek of bezit uitdrukken, zoals *abgehen* 'ontberen; gevoelens van gemis ervaren (regionaal)', *fehlen* 'missen, ontberen; missen (emotioneel); niet gezond zijn (regionaal); niet voorhanden zijn, verloren gegaan

[2] Zie echter Draye 1996: 175 en Duden (2016: 829), die alle datieven als complementen beschouwen, met uitzondering van de ethische datief.

zijn; nodig hebben' of *bleiben* 'blijven, resteren'. Het verschil is echter dat possessieve datieven vrije toevoegingen zijn, terwijl onpersoonlijke datief-nominatiefwerkwoorden vaste datiefargumenten vergen. De status van de possessieve datief als vrije toevoeging is echter niet helemaal onomstreden: sommige bronnen, zoals Duden (2016: 831) beschouwen hem als een object, andere, zoals Lamiroy & Delbecque (1998: 30), argumenteren dat hij net niet in het subcategorisatieframe van het werkwoord vervat zit. De analyse van Hentschel & Weydt (2021: 176) is wat dat betreft verhelderend: zij stellen dat er weliswaar een innige relatie bestaat tussen werkwoord en bezitsdatief, maar beklemtonen dat dat er een is van louter semantische aard.

Voorbeeld (7) bevat een zogenoemde datief van voor- of nadeel, ook wel *dativus commodi/incommodi*. Hij duidt de persoon aan die als gevolg van een bepaalde actie baat of schade ondervindt (Hentschel & Weydt 2021: 172). Specifiek in combinatie met passissieve of anticausatieve werkwoorden (zoals in voorbeeld 7) kan hij tot verwarring leiden met onpersoonlijke datief-nominatiefwerkwoorden, aangezien de datief van voor- of nadeel, weliswaar als perifere constituent, een rol codeert die traditioneel met het onderwerp geassocieerd wordt, namelijk die van de instigator.[3] Bijgevolg is er, net als bij onpersoonlijke datief-nominatiefwerkwoorden, een conversie mogelijk waarbij de medespelers aan hun prototypische naamval worden toegewezen. De anticausatieve datief-nominatiefstructuur uit voorbeeld (7), hier herhaald als (10a), kan dus worden omgevormd tot een causatieve nominatief-accusatiefstructuur zoals in (10b):

(10)	a.	*Mir*	*ist*	*das Glas*	*gebrochen.*
		mij.DAT	is.3P.SG	het.NOM glas	breken.VD
		'Ik heb het glas gebroken.'			
(10)	b.	*Ich*	*habe*	*das Glas*	*gebrochen.*
		ik.NOM	heb.3P.SG	het.ACC glas	breken.VD
		'Ik heb het glas gebroken.'			

In totaal zijn er volgens Wegener (1985: 200–201) een 65-tal werkwoorden die in een anticausatieve datief-nominatiefstructuur (of "ergatiefconstructie") gebruikt kunnen worden. In de praktijk is het onderscheid met onpersoonlijke datief-nominatiefwerkwoorden echter niet altijd eenvoudig om te maken. De keuze om een welbepaalde datieve constituent al dan niet als vast argument te beschouwen, berust daarom op een lexicografisch criterium: zodra het *Duden Universalwörterbuch* of het DWDS (*Digitales Wörterbuch der deutschen Sprache*) een datiefargument als inherent deel van de valentiestructuur van een werkwoord beschouwt, wordt het opgenomen. Die strategie illustreert overigens dat Wegeners 'ergatieve' werkwoorden een categorie in flux zijn: van

[3] Passissieve (of anticausatieve) werkwoorden zijn (intransitief gebruikte) werkwoorden waarvan het onderwerp een patiens is, zoals *breken*, *afkoelen* of *scheuren* (Vandeweghe 2015: 54).

de 65 werkwoorden die Wegener goed 35 jaar geleden als ergatief classificeerde, verbinden er zich nu vier met een vast datiefargument. Het betreft, in alfabetische volgorde, de volgende lemma's:

> *entfallen* 'ontglippen, laten vallen; vergeten' – *passieren* 'overkomen' – *vergehen* 'voorbijgaan (tijd); verliezen, afnemen' – *verlorengehen* 'verlorengaan'

De vrije datief in (8) wordt ook wel de *dativus iudicantis* genoemd. Hij codeert volgens Hentschel & Weydt (2021: 177) een 'oordelende instantie' en komt vaak voor in combinatie met adjectieven die gemodificeerd worden door de bijwoorden *zu* 'te' en *genug* 'voldoende'.

Voorbeeld (9), ten slotte, bevat een (erg zeldzame) ethische datief. Hij komt enkel voor als persoonlijk voornaamwoord in de eerste persoon en verwijst volgens Hentschel & Weydt (2021: 176) naar iemand die een emotioneel belang heeft bij een bepaalde handeling.

In tegenstelling tot de hierboven behandelde voorbeelden behoren subjectachtige datieven steevast tot de zinskern. Dat betekent echter niet dat elk vast datiefargument sowieso als subjectachtig geanalyseerd kan worden. Voorbeelden (11) en (12) mogen dat punt verder illustreren:

(11) *Der Schüler* *widerspricht* *dem Lehrer.*
de.NOM leerling spreekt.3P.SG tegen de.DAT leraar
'De leerling spreekt de leraar tegen.'

(12) *Das hübsche Mädchen* *gefällt* *dem Jungen.*
het.NOM mooie meisje bevalt.3P.SG de.DAT jongen
'Het mooie meisje bevalt de jongen.'

De datiefargumenten in de bovenstaande zinnen zijn beide noodzakelijke aanvullingen. Op grond van hun (niet-nominatieve) casusmarkering worden ze daarom traditioneel als objecten beschouwd. Toch vertonen ze een wezenlijk semantisch onderscheid: terwijl het werkwoord in (11) een actie codeert die in wezen uitgaat van de leerling, codeert het (laag-agentieve) werkwoord in (12) een gewaarwording die uitgaat van de jongen, die in de datief wordt gemarkeerd. Los van het feit dat hij bezield is, vertoont de datief in (11) dus veel meer overlap met een klassiek object dan die in (12), die veeleer 'subjectachtig' is.

Overigens is het niet zo dat subjectachtige datieven als inherente eigenheid van een welbepaald werkwoord gelden: ze zijn hoogstens geassocieerd met bepaalde betekenisschakeringen (Eng. *senses*). Bij monoseme werkwoorden, zoals in (13), is de situatie eenduidig, aangezien ze slechts één gebruik kennen, maar zodra een werkwoord polyseem is, ligt dat enigszins anders. In dat geval zijn er twee mogelijkheden: ofwel vergt een werkwoord een subjectachtige datief in de ene betekenis, maar een objectachtige

datief in de andere (zie voorbeelden 14a en 14b), ofwel vergt het een subjectachtige datief in meerdere van zijn betekenissen (zie voorbeelden 15a en 15b).

(13) *Mir* *ist* *ein Unfall* *passiert.*
mij.DAT is.3P.SG een.NOM ongeluk overkomen.VD
'Ik heb een ongeluk gehad.'

(14) a. *Niemand* *kann* *seinem Geschick* *entgehen.*
niemand.NOM kan.3P.SG zijn.DAT lot ontsnappen.INF
'Niemand kan aan zijn lot ontsnappen.'

(14) b. *Sein Name* *ist* *mir* *leider* *entgangen.*
zijn.NOM naam is.3P.SG mij.DAT helaas ontgaan.INF
'Zijn naam is mij ontgaan.'

(15) a. *Die Suppe* *schmeckt* *mir* *nicht.*
de.NOM soep smaakt.3P.SG mij.DAT niet
'De soep smaakt me niet.'

(15) b. *Der Bundesregierung* *schmecken* *diese Reformen* *nicht.*
de.DAT Bondsregering smaken.3P.PL deze.NOM hervormingen niet
'Deze hervormingen vallen niet in de smaak bij de Bondsregering.'

Beide polyseme paren hierboven illustreren overigens ook de rol van metaforiek bij het ontstaan en de evolutie van onpersoonlijke werkwoorden. Haspelmath (2001: 79) wijst erop dat onpersoonlijke werkwoorden die ondervinding uitdrukken oorspronkelijk een concrete betekenis hadden. Hij geeft onder andere het voorbeeld van het Engelse werkwoord *worry* 'zich zorgen maken'. Aanvankelijk betekende dat 'wurgen' of 'bij de keel grijpen', maar later ontwikkelde het een metaforisch gebruik dat een tijd lang naast het letterlijke gebruik bestond, maar dat mettertijd ging verdringen: waar men vroeger dus letterlijk bij de keel gegrepen werd, is dat nu enkel nog in overdrachtelijke zin het geval.

Een in oorsprong objectachtig argument evolueerde zo tot een subjectachtig argument met niet-canonieke casusmarkering. Het werkwoord *entgehen* uit voorbeeldpaar (14) bevindt zich precies in het zojuist beschreven tussenstadium waarbij een letterlijk en een figuurlijk gebruik naast elkaar bestaan. Interessant is evenwel het betekeniscontrast dat door voorbeeldpaar (15) wordt geïllustreerd. Daaruit blijkt namelijk dat werkwoorden die al een subjectachtig obliek argument verlangen aan dezelfde evolutie onderhevig kunnen zijn, en eveneens een meer abstracte, metaforische betekenis kunnen verwerven. Hoewel Haspelmaths claim dus op zijn minst een deel van de data kan verklaren, lijkt er geen algemene diachrone geldigheid vanuit te gaan: hij erkent immers niet dat niet-canoniek gemarkeerde structuren ook uit andere niet-canoniek gemarkeerde structuren kunnen ontstaan.

2.3 Naar een inventaris van alle onpersoonlijke datief-nominatiefwerkwoorden

In deze sectie schets ik de methode aan de hand waarvan ik de verschillende onpersoonlijke datief-nominatiefwerkwoorden in het hedendaagse Duits in kaart breng. De Nederlandse parafrasen vermelden altijd alle mogelijke betekenisschakeringen die een onpersoonlijk gebruik toelaten, waarbij verschillende betekenisschakeringen van elkaar gescheiden worden door puntkomma's. Het onderscheid tussen de betekenissen is gebaseerd op lexicografische informatie: het weerspiegelt de betekenisschakeringen die het *Duden Universalwörterbuch* en het *Digitales Wörterbuch der deutschen Sprache* als geconventionaliseerd beschouwen. Als beide bronnen verschilden in het aantal betekenissen dat ze onderscheidden, heb ik telkens voor de meest fijnmazige semantische opdeling geopteerd. Ook gebruikswijzen die als eerder regionaal of spreektalig gelden, worden in de lijst vermeld.

2.3.1 Het uitgangspunt

Het vertrekpunt van deze studie is Barðdal (2004), waarin een lijst van 56 onpersoonlijke werkwoorden in het hedendaagse Duits gepubliceerd staat. Het betreft bij haar niet enkel onpersoonlijke datief-nominatiefwerkwoorden, maar ook (al dan niet monovalente) werkwoorden met subjectachtige accusatieven (monovalente werkwoorden met een subjectachtige datief komen in het moderne Duits niet voor, behalve in koppelwerkwoordconstructies als *mir ist kalt*). Barðdals lijst bevat dus werkwoorden van de volgende drie types: intransitieve werkwoorden met een subjectachtige accusatief (zoals in voorbeeld 16), twee- of meerplaatsige werkwoorden met een subjectachtige accusatief en een objectachtige nominatief (zoals in voorbeeld 17), en twee- of meerplaatsige werkwoorden met een subjectachtige datief en een objectachtige nominatief (zoals in voorbeeld 18).

(16) *Mich* *dürstet.*
mij.acc dorst.3p.sg
'Ik heb dorst.'

(17) *Ihn* *ekelt* *es* *davor.*[4]
hem.ACC huivert.3P.SG het.NOM daarvoor
'Hij huivert daarvoor.'

(18) *Dieses Buch* *gefällt* *mir.*
dit.NOM boek bevalt.3P.SG mij.DAT
'Dit boek bevalt mij.'

Aan de hand van de onlineversie van het *Duden Universalwörterbuch* werd voor elk van Barðdals werkwoorden het casusframe bepaald. In 45 gevallen bleek het om een onpersoonlijk datief-nominatiefwerkwoord te gaan. Het betreft de volgende lemma's:

> *anliegen* 'van belang zijn' – *behagen* 'behagen' – *bekommen* 'bekomen, bevallen' – *eilen* 'dringend zijn' – *einfallen* 'te binnen schieten (van een idee); zich herinneren' – *einleuchten* 'duidelijk zijn' – *ekeln* 'doen walgen' – *fehlen* 'missen, ontberen; missen (emotioneel); niet gezond zijn (regionaal); niet voorhanden zijn, verloren gegaan zijn; nodig hebben' – *frommen* 'tot nut strekken' – *gefallen* 'bevallen' – *gehören* 'toebehoren (let.); toebehoren (fig.), behoren aan; toekomen, verdienen (Zuid-Duits, Oostenrijks, spreektalig)' – *gelingen* 'lukken' – *genügen* 'volstaan' – *geziemen* 'betamen, horen' – *glücken* 'lukken' – *grauen* 'beangstigen' – *gutgehen* 'goed gaan (emotioneel); goed gaan (economisch)' – *kommen* 'komen (gelegen, ongelegen, onverwacht …); overvallen' – *leichtfallen* 'makkelijk zijn' – *leidtun* 'medelijden hebben met; spijten' – *liegen* 'liggen, goed zijn in; geven om, waarderen; liggen, het kunnen vinden met (spreektalig)' – *mangeln* 'ontbreken' – *mangeln an* 'ontbreken aan' – *missfallen* 'misvallen' – *missglücken* 'mislukken' – *misslingen* 'mislukken' – *munden* 'smaken' – *nutzen* 'nuttig zijn (voor)' – *nützen* 'nuttig zijn (voor)' – *obliegen* 'moeten, als plicht toekomen' – *passen* 'passen (van kleren); goed uitkomen (spreektalig)' – *passieren* 'overkomen' – *reichen* 'volstaan (spreektalig)' – *schaudern* 'huiveren (let.); huiveren (fig.), vrezen' – *scheinen* 'schijnen' – *schlechtgehen* 'slecht gaan (emotioneel); slecht gaan (economisch)' – *schmecken* 'een bepaalde smaak hebben (let.); een bepaalde smaak hebben (fig.); goed smaken (let.); bevallen' – *schwanen* 'een voorgevoel hebben' – *schwerfallen* 'moeilijk zijn (voor)' – *schwindeln* 'duizelen' – *stinken* 'kwaad maken' – *träumen* 'dromen' – *vorkommen* 'overkomen (gebeurtenis); ervaren; voorkomen, treffen (als)' – *vorschweben* 'voor ogen hebben' – *wehtun* 'pijn doen (fysiek)'

Daarnaast lijst Barðdal (2004) ook een aantal zogenoemde complexe predicaten met *sein* op, zoals *egal sein* 'niet uitmaken, eender zijn' of *angeboren sein* 'aangeboren zijn', die eveneens een subjectachtige datief selecteren. Om de haalbaarheid van het onderzoek te

[4] De subjectachtige accusatief kan bij het werkwoord *ekeln* ook als datief gerealiseerd worden.

garanderen, werden dergelijke structuren echter geweerd.[5] Werkwoorden die zowel in één als in twee woorden gespeld kunnen worden, werden in de dataset opgenomen in hun geüniverbeerde vorm. Dat is bijvoorbeeld het geval voor *wehtun* 'pijn doen' of *verlorengehen* 'verloren gaan'.

2.3.2 Lexicografische zoekopdracht

Om de oorspronkelijke lijst uit te breiden, werd een zoekopdracht gelanceerd in de digitale versie van het *Duden Universalwörterbuch*. Het doel daarvan was om aan de hand van de bekende onpersoonlijke werkwoorden andere op te sporen die Duden als parafrasen voor de gerunde werkwoorden aanbood. De idee achter deze werkwijze is dat men van semantisch gelijkaardige werkwoorden mag verwachten dat ze ook een gelijkaardig casusframe aannemen (cf. Barðdal 2004: 114, Barðdal 1999a, Barðdal 1999b, Barðdal & Eyþórsson 2020). Als input voor deze zoekopdracht dienden alle 56 werkwoorden uit Barðdals oorspronkelijke lijst, dus ook diegene die een subjectachtige accusatief selecteren (cf. voorbeelden 16 en 17). Het was immers niet ondenkbaar dat werkwoorden met een subjectachtige accusatief uit de oorspronkelijke lijst een synoniem zouden genereren met een subjectachtige datief. In totaal leverde de lexicografische zoekopdracht 90 onpersoonlijke datief-nominatiefwerkwoorden op, waarvan er 48 niet in de oorspronkelijke lijst voorkwamen. Het gaat om de volgende lemma's:

> *abgehen* 'ontberen; gevoelens van gemis ervaren (regionaal)' – *ahnen* 'een voorgevoel hebben (spreektalig); weten van, kennis hebben van' – *angehören* 'behoren tot (een bepaalde groep); behoren tot (een bepaalde periode); nauw verbonden zijn met iemand' – *ankommen* 'overvallen, overkomen; zwaar vallen; van belang zijn; aankomen op' – *anstehen* 'bevallen, aanstaan; genoeg hebben van (Oostenrijks)' – *auffallen* 'opvallen' – *aufgehen* 'zich openbaren aan' – *ausgehen* 'verminderen, op geraken; tekortkomen (adem)' – *ausreichen* 'genoeg hebben van' – *begegnen* 'overkomen; (toevallig) aantreffen'[6] – *beifallen* 'te binnen schieten' – *beikommen* 'opkomen bij, te binnen schieten' – *blühen* 'te wachten staan' – *danebengeraten* 'mislukken' – *dazwischenkommen* 'tussenkommen' – *dienen* 'dienen' – *eignen* 'typeren, eigen zijn

[5] Een anonieme reviewer merkt op dat het hier geen onpersoonlijke datief-nominatief*werkwoorden* betreft, maar hoogstens onpersoonlijke datief-nominatief*structuren*, aangezien de subjectachtige datief strikt genomen niet wordt geselecteerd door het werkwoord, maar door het adjectief.

[6] De betekenis '(toevallig) aantreffen' staat noch in het DWDS, noch in Duden apart vermeld, maar moet wel van de betekenis 'ontmoeten' onderscheiden worden. Vergelijk daartoe de voorbeelden *Wir waren Hundeschlitten fahren und sind tatsächlich Wölfen begegnet* (*Bild am Sonntag*, 13/01/2019, n° 2; via DWDS, lemma *begegnen*) en *Und sollte ihm in seiner Wohnung einmal ein Juchtenkäfer begegnen, dann hätte er sofort ein Insektenvernichtungsspray zur Hand* (*Der Spiegel*, 31/12/2010, online; via DWDS, lemma *begegnen*). In de eerste zin is de datief objectachtig, in de tweede zin is hij subjectachtig: hij codeert immers de persoon die (per toeval) de kever aantreft.

(aan)' – *einkommen* 'te binnen schieten' – *erbarmen* 'spijten (Oostenrijks)' – *ermangeln* 'ontbreken' – *ermangeln an* 'ontbreken aan' – *erscheinen* 'voor iemand verschijnen (let.); voorkomen als'– *gebrechen* 'ontbreken' – *gebrechen an* 'ontbreken aan' – *gebühren* 'toekomen' – *geraten* 'lukken; voor iemand op een bepaalde manier aflopen' – *geschehen* 'gebeuren, overkomen' – *graulen* 'doen huiveren' – *grausen* 'beangstigen' – *gruseln* 'beangstigen' – *imponieren* 'imponeren' – *konvenieren* 'uitkomen, passen (Oostenrijks)' – *missraten* 'mislukken' – *nachgehen* 'bijblijven' – *pressieren* 'dringend zijn' – *schauern* 'huiveren (let.); huiveren (fig.), vrezen' – *widerfahren* 'overkomen' – *widerstehen* 'mishagen; niet smaken (eten)' – *widerstreben* 'met tegenzin vervullen' – *zufallen* 'in het bezit komen; te beurt vallen' – *zugehören* 'behoren (tot)' – *zugutekommen* 'ten goede komen' – *zukommen* 'ontvangen; toegestuurd krijgen (bv. van een bericht); toekomen, mogen; krijgen (fig., zoals van een bepaalde rol); toe te meten zijn' – *zusagen* 'bevallen' – *zusetzen* '(mentaal) kwellen; (fysiek) kwellen' – *zustattenkommen* 'van pas komen' – *zustoßen* 'overkomen' – *zuteilwerden* 'te beurt vallen'

Deze zoekstrategie had tot gevolg dat het aantal datief-nominatiefwerkwoorden meer dan verdubbelde (van 45 naar 93). 42 van de 90 gevonden werkwoorden waren al bekend uit Barðdals lijst, wat de sterke semantische samenhang tussen de verschillende oorspronkelijke lemma's illustreert. Heel wat werkwoorden genereerden bij de lexicografische zoekopdracht immers elkaar.

Omgekeerd waren er wel drie werkwoorden uit de oorspronkelijke lijst die niet in de nieuwe voorkwamen, nl. *ekeln* 'doen walgen', *scheinen* 'schijnen' en *träumen* 'dromen'. Dat aantal is misschien gering, maar het laat wel zien dat niet *alle* onpersoonlijke datief-nominatiefwerkwoorden een semantische link delen met een ander soortgelijk werkwoord. Als dat wel zo zou zijn, zouden alle 45 werkwoorden uit Barðdals oorspronkelijke lijst zonder uitzondering ook in de nieuwe voorkomen. Het is uiteraard ook denkbaar dat het systeem op basis waarvan het *Duden Universalwörterbuch* synoniemen voor bepaalde lemma's genereert niet helemaal sluitend is – een woordenboek blijft immers mensenwerk – maar dat verandert niets aan de conclusie waartoe deze observatie noopt, namelijk dat de drie werkwoorden die onder de waterlijn bleven symptomatisch zouden kunnen zijn voor een grotere tot nog toe onontgonnen groep. De lexicografische zoekopdracht diende dus te worden aangevuld met andere methodes.

2.3.3 Aanvullende zoekstrategieën

De eerste alternatieve zoekstrategie bestond uit een corpusstudie in het *Deutsches Referenzkorpus* (DeReKo), meer bepaald in het Tagged-M-subcorpus, dat uit circa 30 miljoen woorden lopende tekst bestaat, waarvan een kleine 20 miljoen openbaar te raadplegen zijn. Alle woorden zijn voorzien van morfosyntactische annotatie, wat het

mogelijk maakte een zoekopdracht te formuleren die casuele informatie mee in beschouwing neemt. Concreet werd de volgende syntaxis gerund: MORPH(NOU nom) /+w1:3,s0 MORPH(VRB) /+w1:3,s0 MORPH(NOU dat).

De syntaxis ging dus op zoek naar een verbaal element dat geflankeerd werd door naamwoordgroepen die respectievelijk nominatieve en datieve casusmarkering droegen. De naamwoorden die als kern van deze nominale groepen fungeerden, konden daarbij tot drie woorden van het werkwoord verwijderd staan. Dat leverde in totaal 32.503 treffers op, die vervolgens manueel werden opgeschoond. In totaal konden er op die manier 59 werkwoorden worden vergaard, waarvan 24 nieuwe. Dat aandeel ligt gevoelig lager dan het aantal werkwoorden dat via de lexicografische studie kon worden verzameld, maar vormde nog steeds een wezenlijke aanvulling op de al gekende werkwoorden. Het gaat om de volgende lemma's:

> *anhängen* 'op iemand wegen' – *aufstoßen* 'opvallen (spreektalig); oprispen; tegensteken' – *beiwohnen* 'eigen zijn (aan)' – *bevorstehen* 'te wachten staan; opzien tegen (Noord-Duits)' – *bleiben* 'blijven, resten' – *drohen* 'dreigen' – *entgegenschlagen* '(met heftigheid) gewaarworden (let.); (met heftigheid) gewaarworden (fig.)' – *entgegentreten* 'opmerken, stoten op' – *entschwinden* 'mentaal ontgaan, vergeten; uit iemands blikveld verdwijnen'– *entstehen* 'ontstaan (voor iemand)' – *erwachsen* 'ontstaan' – *fernliegen* 'niet in iemand opkomen; niet iemands intentie zijn' – *folgen* 'volgen op' – *gegenübertreten* 'voorkomen, treffen als' – *gelten* 'waard zijn; gelden (als), beschouwen' – *guttun* 'goeddoen' – *nachhängen* 'blijven hangen' – *unterlaufen* 'overkomen; overkomen (van een fout)' – *winken* 'in het vooruitzicht staan' – *wohltun* 'goeddoen' – *zufliegen* 'te binnen schieten' – *zufließen* 'ontvangen' – *zustehen* 'het recht hebben, mogen; verdienen' – *zuwachsen* 'krijgen'

In de tweede plaats werd er een beroep gedaan op het elektronische valentiewoordenboek van het *Leibniz-Institut für deutsche Sprache*, dat voor 677 werkwoorden de complementstructuur in kaart brengt. Via de applicatie werd een zoekopdracht opgezet naar alle werkwoorden die op zijn minst een nominatiefcomplement en een datiefcomplement vorderden, wat in totaal 112 treffers opleverde. Na manuele doorzoeking konden 31 onpersoonlijke werkwoorden worden geïdentificeerd, waarvan vijf nieuwe lemma's:

> *ausmachen* 'uitmaken' – *bedeuten* 'betekenen voor; zijn' – *lachen* 'toelachen' – *sein* 'toebehoren (regionaal, spreektalig)' – *stehen* 'staan (van een kledingstuk); staan (fig.), afgaan'

Vervolgens werd er een beroep gedaan op Wegener (1985), een referentiewerk over de datief in het hedendaagse Duits, en Engelen (1975), die een index van werkwoorden presenteert op basis van hun respectievelijke complementstructuren. Op die manier

konden vijftien werkwoorden aan de lijst worden toegevoegd. Voor Wegener (1985) gaat het om de volgende dertien lemma's:

> *abhandenkommen* 'verliezen' – *belieben* 'believen' – *dämmern* 'dagen' – *dünken* 'dunken' – *eingehen* 'begrijpen; bevallen' – *entfallen* 'ontglippen, laten vallen; vergeten' – *entgehen* 'zich laten ontglippen, voorbij laten gaan; niet opgemerkt worden, ontgaan' – *entgleiten* 'ontglippen, laten vallen; ontglippen (fig.), verliezen; ontsnappen (woorden)' – *klingen* 'klinken (als)' – *langen* 'volstaan (spreektalig)' – *nahegehen* 'bedroeven' – *vergehen* 'voorbijgaan (tijd); verliezen, afnemen' – *verlorengehen* 'verloren gaan'

De overige twee werkwoorden stammen uit Engelen (1975):

> *klarwerden* 'duidelijk worden' – *verbleiben* 'resteren'

Om de laatste gaten te stoppen, werd er specifiek aandacht besteed aan samengestelde datief-nominatiefwerkwoorden op basis van de prefixen *ent-*, *entgegen-*, *gegenüber-*, *nach-* en *zu-*. Duden (2016: 398) wijst erop dat het aantal werkwoordstammen dat zich met deze prefixen verbindt, stelselmatig uitdijt, en dat de vorming van dergelijke werkwoorden bijgevolg een productief proces vormt. Ook partikelwerkwoorden met *hinterher-* en *voran-* werden verder onder de loep genomen. Zes lemma's konden zo aan de lijst worden toegevoegd:

> *entfahren* 'ontsnappen (van woorden)' – *entgegenschallen* 'in iemands richting weerklinken' – *entschlüpfen* 'ontsnappen (van woorden)' – *hinterhereilen* 'achtervolgen (bv. van succes)' – *zuhandenkommen* 'krijgen' – *zupasskommen* 'van pas komen'

Ondanks alle inspanningen waren er drie werkwoorden die veeleer op toevallige wijze werden verworven.[7] Het betreft de volgende lemma's:

> *aufdämmern* 'dagen' – *innewohnen* 'eigen zijn (aan)' – *rauchen* 'irritatie ondervinden, boos worden (spreektalig)'

2.3.4 Besluit

Het aantal werkwoorden dat via de hierboven besproken methodologie kon worden vergaard, bedraagt 146. Onpersoonlijke datief-nominatiefwerkwoorden blijken daarmee veel talrijker te zijn dan aanvankelijk werd vermoed. Barðdal et al. (2016: 56), bijvoorbeeld, schatte het *totale* aantal onpersoonlijke datiefwerkwoorden (dus ook de predicaten met

[7] Het gaat hier om werkwoorden die bij een vervolgonderzoek in het German Web 2013-corpus (deTenTen13) werden aangetroffen.

sein) voor het hedendaagse Duits op slechts 75. Daarnaast illustreert de doorgevoerde zoekopdracht het in Sectie 2.2 besproken belang van polysemie: van de in totaal 146 werkwoorden bleken er 40 een subjectachtige datief te regeren in meerdere van hun betekenissen. Meestal betrof het werkwoorden met twee gebruikswijzen (30 lemma's), maar zes werkwoorden (*angehören*, *aufstoßen*, *entgleiten*, *gehören*, *liegen* en *vorkommen*) laten drie lezingen toe, twee werkwoorden (*ankommen* en *schmecken*) laten vier lezingen toe en twee werkwoorden (*fehlen* en *zukommen*) zelfs vijf lezingen. Toekomstig onderzoek zou op basis van corpusdata kunnen achterhalen in welke mate de lexicografisch geponeerde betekenissen ook voorkomen in het taalgebruik, en of het semantische bereik van de onderzochte werkwoorden uitdijt, inkrimpt, dan wel constant blijft.

2.4 Semantische classificatie

2.4.1 Clustering in zestien klassen

Nadat alle werkwoorden waren vergaard, konden ze semantisch worden geclusterd. Dat gebeurde op basis van een typologie door Barðdal et al. (2016), waarin vijftien lexicale velden worden beschreven waarover de onderwerpsdatief in het Germaans zich uitstrekt.[8] Het gaat de onderzoekers daarbij zowel om werkwoorden in monovalente constructies van het type *mir schwindelt* 'ik duizel', als om een van vijf meerplaatsige constructies waarin een onderwerpsdatief gecombineerd wordt met een nominatief-, accusatief- of genitiefobject, een voorzetselcomplement of een zinscomplement. Barðdal et al. (2016) nemen dus meer types constructies in beschouwing dan enkel onpersoonlijke datief-nominatiefstructuren. Bovendien vermelden zij voor het moderne Duits slechts 75 werkwoorden, maar toch valt het op hoe goed hun classificatie zich leent tot de beschrijving van de veel uitgebreidere dataset uit mijn eigen onderzoek.

Er hoefde slechts één lexicaal veld aan de classificatie worden toegevoegd, met name dat van de werkwoorden van toebehoren. Merk op dat ik in mijn eigen classificatie bij polyseme werkwoorden een onderscheid maak tussen de verschillende betekenissen en betekenisschakeringen, waardoor het totale aantal lemma's in het overzicht hieronder op 202 komt (i.e. de som van 86 monoseme werkwoorden, vermeerderd met 40 werkwoorden met elk twee betekenissen, zes werkwoorden met elk drie betekenissen,

[8] De typologie van Barðdal et al. (2016) is specifiek op de volgende talen gebaseerd: IJslands, Faeröers en Duits wat betreft de moderne Germaanse talen, en Gotisch, Oudengels, Oudsaksisch, Oudhoogduits, Middelengels, Middelnederlands, Middelhoogduits, Oudijslands en Oudzweeds wat betreft de historische Germaanse talen.

twee werkwoorden met elk vier betekenissen en twee werkwoorden met vijf betekenissen). Verschillende betekenissen worden telkens onderscheiden aan de hand van Romeinse cijfers. Alle categorieën, i.e. zowel de zestien macrocategorieën als hun respectievelijke microcategorieën, worden gevolgd door een aanduiding van het aantal werkwoorden of betekenissen dat ze bevatten. De classificatie wordt geordend volgens aflopend aantal lemma's.

1. Werkwoorden van emotie (54)
    - **Welbehagen**: *anstehen* I 'bevallen, aanstaan', *behagen* 'behagen', *bekommen* 'bekomen, bevallen', *belieben* 'believen', *eingehen* II 'bevallen', *gefallen* 'bevallen', *guttun* 'goeddoen', *liegen* III 'liggen, het kunnen vinden met (spreektalig)', *schmecken* IV 'bevallen', *wohltun* 'goeddoen', *zusagen* 'bevallen' (11)
    - **Geven om**: *angehören* III 'nauw verbonden zijn met iemand'; *ankommen* III 'van belang zijn', *ankommen* IV 'aankomen op', *anliegen* 'van belang zijn', *ausmachen* 'uitmaken', *bedeuten* I 'betekenen voor', *gelten* I 'waard zijn', *leidtun* I 'medelijden hebben met', *liegen* II 'geven om, waarderen' (9)
    - **Angst/gevaar**: *grauen* 'beangstigen', *graulen* 'doen huiveren', *grausen* 'beangstigen', *gruseln* 'beangstigen', *schaudern* II 'huiveren (fig.), vrezen', *schauern* II 'huiveren (fig.), vrezen' (6)
    - **Last**: *anhängen* 'op iemand wegen', *ankommen* II 'zwaar vallen', *eilen* 'dringend zijn', *nachhängen* 'blijven hangen', *pressieren* 'dringend zijn', *schwerfallen* 'moeilijk zijn (voor)' (6)
    - **Verdriet/droefenis**: *abgehen* II 'gevoelens van gemis ervaren (regionaal)', *erbarmen* 'spijten (Oostenrijks)', *fehlen* II 'missen (emotioneel)', *leidtun* II 'spijten', *nahegehen* 'bedroeven' (5)
    - **Voelen/ervaren**: *ankommen* I 'overvallen, overkomen', *entgegenschlagen* II '(met heftigheid) gewaarworden (fig.)', *schmecken* II 'een bepaalde smaak hebben (fig.)', *hinterhereilen* 'achtervolgen (bv. van succes)', *vorkommen* II 'ervaren' (5)
    - **Boosheid/irritatie**: *anstehen* II 'genoeg hebben van (Oostenrijks)', *aufstoßen* I 'tegensteken', *rauchen* 'irritatie ondervinden, boos worden (spreektalig)', *stinken* 'kwaad maken' (4)
    - **Mishagen**: *bevorstehen* II 'opzien tegen (Noord-Duits)', *missfallen* 'misvallen', *widerstehen* I 'mishagen', *widerstreben* 'met tegenzin vervullen' (4)
    - **Genot/blijdschap**: *gutgehen* I 'goed gaan (emotioneel)' (1)
    - **Opluchting/gemak**: *leichtfallen* 'makkelijk zijn' (1)
    - **Leed**: *schlechtgehen* I 'slecht gaan (emotioneel)' (1)
    - **Bitterheid/haat**: *ekeln* 'doen walgen' (1)

2. Werkwoorden van cognitie (24)
    - **Denken**: *aufgehen* 'zich openbaren aan', *beikommen* 'opkomen bij, te binnen schieten', *dünken* 'dunken', *einfallen* I 'te binnen schieten (van een idee)', *eingehen* I 'begrijpen', *einkommen* 'te binnen schieten', *einleuchten* 'duidelijk zijn', *fernliegen* I 'niet in iemand opkomen', *gelten* II 'gelden (als), beschouwen', *klarwerden* 'duidelijk worden', *vorschweben* 'voor ogen hebben', *zufliegen* 'te binnen schieten' (12)
    - **Weten**: *ahnen* II 'weten van, kennis hebben van', *aufdämmern* 'dagen', *beifallen* 'te binnen schieten', *dämmern* 'dagen', *einfallen* II 'zich herinneren', *entfallen* II 'vergeten', *entschwinden* I 'mentaal ontgaan, vergeten', *nachgehen* 'bijblijven' (8)
    - **Argwaan**: *ahnen* I 'een voorgevoel hebben (spreektalig)', *schwanen* 'een voorgevoel hebben' (2)
    - **Verrassing/verwarring**: *imponieren* 'imponeren' (1)
    - **Vastberadenheid/twijfel**: *fernliegen* II 'iemands intentie zijn' (1)

3. Werkwoorden van winst en voordeel (20)
    - **Baat**: *dienen* 'dienen', *entstehen* 'ontstaan (voor iemand)', *erwachsen* 'ontstaan', *lachen* 'toelachen', *nutzen* 'nuttig zijn (voor)', *nützen* 'nuttig zijn (voor)', *winken* 'in het vooruitzicht staan', *zufallen* I 'in het bezit komen', *zufallen* II 'te beurt vallen', *zufließen* 'ontvangen', *zugutekommen* 'ten goede komen', *zuhandenkommen* 'krijgen', *zukommen* I 'ontvangen', *zukommen* II 'toegestuurd krijgen (bv. van een bericht)', *zukommen* IV 'krijgen (fig., zoals van een bepaalde rol)', *zupasskommen* 'van pas komen', *zustattenkommen* 'van pas komen', *zustehen* I 'verdienen', *zuteilwerden* 'te beurt vallen', *zuwachsen* 'krijgen' (20)

4. Werkwoorden van gebeuren (17)
    - **Gebeuren**: *begegnen* I 'overkomen', *begegnen* II '(toevallig) aantreffen', *bevorstehen* I 'te wachten staan', *blühen* 'te wachten staan', *dazwischenkommen* 'tussenkomen', *drohen* 'dreigen', *erscheinen* I 'voor iemand verschijnen (let.)', *folgen* 'volgen op', *geraten* II 'voor iemand op een bepaalde manier aflopen' *geschehen* 'gebeuren, overkomen', *kommen* I 'komen (gelegen, ongelegen, onverwacht …)', *kommen* II 'overvallen', *passieren* 'overkomen', *unterlaufen* I 'overkomen', *vorkommen* I 'overkomen (gebeurtenis)', *widerfahren* 'overkomen', *zustoßen* 'overkomen' (17)

5. Werkwoorden van obstructie en belemmering (14)
    - **Verlies en ontglippen**: *abhandenkommen* 'verliezen', *ausgehen* I 'verminderen, op geraken', *entfallen* I 'ontglippen, laten vallen', *entgehen* I

'zich laten ontglippen, voorbij laten gaan' *entgleiten* I 'ontglippen, laten vallen', *entgleiten* II 'ontglippen (fig.), verliezen', *vergehen* II 'verliezen, afnemen', *verlorengehen* 'verloren gaan' (8)

- **Falen**: *danebengeraten* 'mislukken', *missglücken* 'mislukken', *misslingen* 'mislukken', *missraten* 'mislukken', *schlechtgehen* II 'slecht gaan (economisch)', *unterlaufen* II 'overkomen (van een fout)' (6)

6. Werkwoorden van attitude (13)
   - **Gepastheid**: *gebühren* 'toekomen', *gehören* III 'toekomen, verdienen (Zuid-Duits, Oostenrijks, spreektalig), *geziemen* 'betamen, horen', *konvenieren* 'uitkomen, passen (Oostenrijks)', *passen* I 'passen (van kleren)', *passen* II 'goed uitkomen (spreektalig)', *stehen* I 'staan (van een kledingstuk)', *stehen* II 'staan (fig.), afgaan' (8)
   - **Volstaan/nuttig zijn**: *ausreichen* 'genoeg hebben van', *frommen* 'tot nut strekken', *genügen* 'volstaan', *langen* 'volstaan (spreektalig)', *reichen* 'volstaan (spreektalig)' (5)

7. Werkwoorden van bezit (13)
   - **Gebrek**: *abgehen* I 'ontberen', *ermangeln* 'ontbreken', *fehlen* I 'missen, ontberen', *fehlen* IV 'niet voorhanden zijn, verloren gegaan zijn', *fehlen an* 'ontbreken aan', *gebrechen* 'ontbreken', *gebrechen an* 'ontbreken aan' *mangeln* 'ontbreken', *mangeln an* 'ontbreken aan' (9)
   - **Bezit**: *bleiben* 'blijven, resteren', *gehören* I 'toebehoren (let.)', *sein* 'toebehoren (regionaal, spreektalig)', *verbleiben* 'resteren' (4)

8. Werkwoorden van perceptie (13)
   - **Perceptie**: *auffallen* 'opvallen', *aufstoßen* II 'opvallen (spreektalig)', *entgegenschallen* 'in iemands richting weerklinken', *entgegenschlagen* I '(met heftigheid) gewaarworden (let.)', *entgegentreten* 'opmerken, stoten op', *entgehen* II 'niet opgemerkt worden, ontgaan', *entschwinden* II 'uit iemands blikveld verdwijnen', *klingen* 'klinken (als)', *munden* 'smaken', *schmecken* I 'een bepaalde smaak hebben (let.)', *schmecken* III 'goed smaken (let.)', *träumen* 'dromen', *widerstehen* II 'niet smaken (eten)' (13)

9. Werkwoorden van lichamelijkheid (8)
   - **Ziektebeelden**: *aufstoßen* III 'oprispen', *ausgehen* II 'tekortkomen (adem)', *fehlen* III 'niet gezond zijn (regionaal)', *schaudern* I 'huiveren (let.)', *schauern* I 'huiveren (let.)', *schwindeln* 'duizelen', *wehtun* 'pijn doen', *zusetzen* 'kwellen' (8)

10. Werkwoorden van modaliteit (5)
    - **Toelating**: *zukommen* III 'toekomen, mogen', *zustehen* II 'toekomen, mogen' (2)
    - **Noodzaak**: *fehlen* V 'nodig hebben', *zukommen* V 'toe te meten zijn' (2)
    - **Verplichting**: *obliegen* 'moeten, als plicht toekomen' (1)

11. Werkwoorden van ontologische staat (5)
    - **Eigenschap/vermogen**: *bedeuten* II 'zijn', *beiwohnen* 'eigen zijn (aan)', *eignen* 'typeren, eigen zijn (aan)', *innewohnen* 'eigen zijn (aan)', *liegen* I 'liggen, goed zijn in' (5)

12. Werkwoorden van evidentialiteit (4)
    - **Schijnen/blijken**: *erscheinen* II 'voorkomen als', *gegenübertreten* 'voorkomen, treffen als', *scheinen* 'schijnen', *vorkommen* III 'voorkomen, treffen (als)' (4)

13. Werkwoorden van succes (4)
    - **Succes**: *gelingen* 'lukken', *geraten* I 'lukken', *glücken* 'lukken', *gutgehen* II 'goed gaan (economisch)' (4)

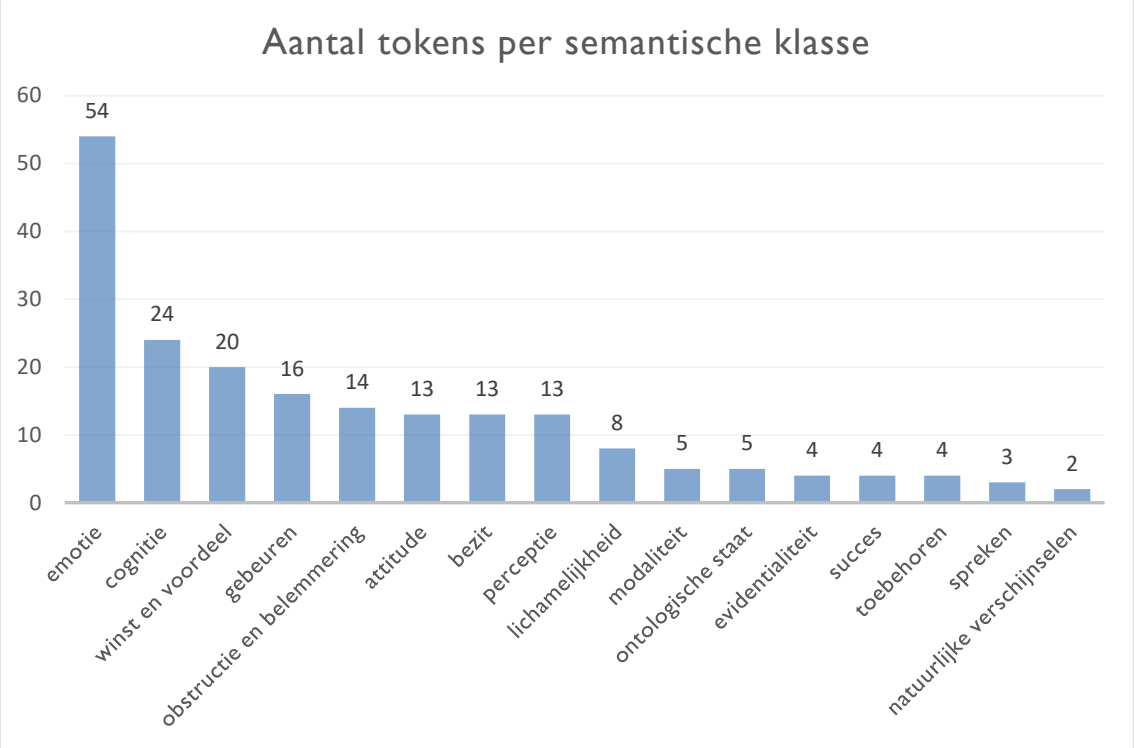
14. Werkwoorden van toebehoren (4): *angehören* I 'behoren tot (een bepaalde groep)', *angehören* II 'behoren tot (een bepaalde periode), *gehören* II 'toebehoren (fig.), behoren aan', *zugehören* 'behoren (tot)'

15. Werkwoorden van spreken (3): *entfahren* 'ontsnappen (van woorden)', *entgleiten* III 'ontsnappen (van woorden)', *entschlüpfen* 'ontsnappen (van woorden)'

16. Werkwoorden van natuurlijke verschijnselen (1):
    - **Beweging in de tijd**: *vergehen* I 'voorbijgaan (tijd)' (1)

2.4.2 Algemene tendensen

Uit de bovenstaande classificatie kunnen een aantal tendensen worden afgelezen. Om te beginnen worden alle vijftien door Barðdal et al. (2016) oorspronkelijk geponeerde klassen door minstens één werkwoord vertegenwoordigd. Dat is opmerkelijk, aangezien bijvoorbeeld werkwoorden van ontologische staat voor het Duits vroeger nog onvermeld bleven (Barðdal 2004; 2006). Ten tweede valt het op dat de verschillende klassen sterk variëren in het aantal werkwoorden dat ze bevatten, zoals ook blijkt uit Figuur 2.1.



Figuur 2 Aantal lemma's per semantische klasse (op het niveau van de individuele betekenisschakeringen)

Uit de verdeling van het aantal lemma's per semantische klasse, zoals die in Figuur 2.1 wordt gevisualiseerd, spreekt een opmerkelijke exponentiële tendens: de helft van de lemma's behoort tot de klassen emotie, cognitie, en winst en voordeel, terwijl de andere helft zich over de dertien resterende klassen verdeelt. Een dergelijke, steil aflopende curve is in het taalkundige onderzoek bepaald geen nieuwigheid. Het was George Zipf (1935: 25) die er voor het eerst op wees dat slechts een fractie van de woordenschat vaak voorkomt, en dat lexicale verscheidenheid geassocieerd is met een lage gebruiksfrequentie: het gros van de woorden wordt dus maar sporadisch gebruikt. Later werd de naar hem vernoemde Zipfiaanse distributie ook toegepast op onderzoek naar taalproductiviteit, onder andere om de verzadigingsgraad van de woordenschat ten aanzien van een welbepaald proces te bepalen. Dat laatste gebeurt bijvoorbeeld in Zeldes (2012: 69), die aantoont dat voor productieve processen de curve veel trager afvlakt dan voor niet-productieve.

De data die ik presenteer, laten, voor zover ik weet, voor het eerst ook een Zipfiaanse correlatie zien tussen een welbepaald morfologisch procedé (i.e. niet-canonieke casusmarkering) en het semantische bereik daarvan. In feite is ook dat een vorm van taalproductiviteit. Met name betreft het hier wat Baayen (2009: 901) omschrijft als *realized productivity*, i.e. de som van het aantal verschillende vormen (ook wel 'types' genoemd) dat een bepaalde taalregel volgt. Daarnaast onderscheidt Baayen een vorm van productiviteit die hij *expanding productivity* noemt, waarmee hij bedoelt dat een bepaald fenomeen steeds verder uitdijt en zo nieuwe vormen aantrekt (Baayen 2009: 7). Waar *realized productivity* dus verwijst naar een toestand, verwijst *expanding productivity* naar een proces.

Of onpersoonlijke datief-nominatiefwerkwoorden ook productief zijn in de zin dat de klasse steeds verder uitdijt, is moeilijk te zeggen, maar polyseme werkwoorden die een subjectachtige datief verlangen in de ene betekenisschakering, maar een subjectachtige nominatief in de andere, doen dat wel vermoeden. Voorbeelden van dat proces, dat in Sectie 2.2 al werd besproken, zijn *begegnen* 'ontmoeten' > *begegnen* 'overkomen', of *beiwohnen* 'bijwonen' > *beiwohnen* 'eigen zijn (aan); ervaren'. Ook metaforische extensies van werkwoorden die *an sich* al een subjectachtige datief vergen, wijzen in de richting van een productief proces (bv. *schmecken* 'goed smaken' > *schmecken* 'bevallen'). Anderzijds is het wel zo dat een aantal van de werkwoorden in mijn lijst in toenemende mate in onbruik geraken of hun niet-canonieke casusmarkering inruilen voor canonieke (nominatieve) casusmarkering, zoals *mir träumt* > *ich träume* of *mir ekelt vor* > *ich ekele mich vor*.[9] Een productief en een antiproductief proces lijken zich dus tegelijkertijd te voltrekken, maar onderzoek zal moeten uitwijzen hoe ze zich ten opzichte van elkaar verhouden.

2.4.3 Bespreking per klasse

De rijkst gevulde klasse is die van de werkwoorden van emotie (54 lemma's), en ook haar subdomeinen zijn goed vertegenwoordigd: enkel werkwoorden die te maken hebben met verlangen, verveling/moeheid, wensen/hopen, schaamte en mentale pijn komen in het moderne Duits niet in de onpersoonlijke datief-nominatiefconstructie voor. De andere subdomeinen bevatten minstens één lemma, maar het gros van de werkwoorden clustert wel rond de subdomeinen welbehagen (elf lemma's), geven om (negen lemma's), angst/gevaar (zes lemma's) en last (zes lemma's). Dat deze klasse, althans voor het moderne Duits, zo omvangrijk is, zou wel eens mee aan de basis kunnen liggen van de hardnekkige misvatting dat het oblieke argument van onpersoonlijke constructies steeds een zogenoemde ondervinder (Eng. *experiencer*) zou betreffen (zie ook Sectie 2.1). Verder hoeft het niet te verbazen dat ook polyseme werkwoorden deze klasse talrijk bevolken: twintig van de in totaal 40 meerduidige werkwoorden hebben te maken met emotie. Vaak betreft het werkwoorden die een metaforische betekenis ontwikkelden uit een meer letterlijke gebruikswijze, zoals lichamelijkheid (bv. *aufstoßen* 'oprispen' > *aufstoßen* 'tegensteken'), perceptie (bv. *widerstehen* 'niet smaken (eten)' > *widerstehen* 'mishagen') of bezit (bv. *abgehen* 'ontberen' > *abgehen* 'gevoelens van gemis ervaren').

Met 24 lemma's zijn ook de werkwoorden van cognitie goed vertegenwoordigd. Ze bestrijken vooral de subdomeinen 'denken' (twaalf lemma's) en 'weten' (acht lemma's), en in mindere mate ook 'argwaan' (twee lemma's), 'verrassing/verwarring' (één lemma) en

[9] Dank aan beide anonieme reviewers, die me op deze tendensen attent maakten.

vastberadenheid/twijfel (één lemma). Werkwoorden of betekenissen die te maken hebben met overeenstemming of onenigheid komen echter niet voor.

De categorie winst en voordeel telt in totaal twintig werkwoorden of gebruikswijzen, en allemaal vallen ze onder het subdomein 'baat'. Barðdal et al. (2016) onderscheiden ook nog het subdomein 'groei', maar daarvan komen in het hedendaagse Duits geen voorbeelden voor.

Vijf categorieën die tot de kwantitatieve middenmoot behoren, zijn de werkwoorden van gebeuren (zeventien lemma's), obstructie en belemmering (veertien lemma's), attitude (dertien lemma's), bezit (dertien lemma's), en perceptie (dertien lemma's). Opvallend is wel dat hun dekking nagenoeg absoluut is, aangezien haast alle door Barðdal et al. (2016) geponeerde subdomeinen bezet worden: enkel werkwoorden van obstructie en belemmering in de enge zin komen in het hedendaagse Duits niet voor.

Tot slot zijn er acht categorieën die de laatste 34 lemma's onder zich verdelen. Wat de werkwoorden van evidentialiteit (vier lemma's), succes (vier lemma's), en spreken (drie lemma's) betreft, is de bezetting absoluut, maar aangezien ze alle slechts één subcategorie tellen, is dat resultaat weinig informatief. Voor de werkwoorden van lichamelijkheid (acht lemma's), ontologische staat (vijf lemma's), modaliteit (vijf lemma's) en natuurlijke verschijnselen (één lemma) is de bezetting veel fragmentarischer.

De werkwoorden van lichamelijkheid en ontologische staat komen elk zelfs maar in één subcategorie voor, nl. die van 'eigenschap/vermogen' voor werkwoorden van ontologische staat, en 'ziektebeelden' voor werkwoorden van lichamelijkheid. Ook de modale werkwoorden zijn dun gezaaid, maar ze komen wel voor in drie van de vier door Barðdal et al. (2016) geponeerde subcategorieën: enkel het bereik 'mogelijkheid' komt niet voor. Onpersoonlijke datief-nominatiefwerkwoorden die natuurlijke verschijnselen aanduiden, zijn nagenoeg onbestaande in het hedendaagse Duits: enkel *vergehen* 'voorbijgaan (van tijd)' past in deze categorie. Het Duits kent geen onpersoonlijke datief-nominatiefwerkwoorden die te maken hebben met daglicht, weer/wind, beginnen/eindigen, of beweging.

2.4.4 Besluit

De in deze sectie besproken semantische classificatie laat zien dat het moderne Duits over een brede waaier aan onpersoonlijke datief-nominatiefwerkwoorden beschikt. Het valt op dat de verdeling van de data Zipfiaans is van aard en dat de helft van de lemma's tot de categorieën 'emotie', 'cognitie', of 'winst en voordeel' behoort: denken, voelen en krijgen zijn processen die zich blijkbaar bij uitstek lenen tot een conceptualisering door onpersoonlijke datief-nominatiefwerkwoorden.

Toch lijkt het aangewezen de voorliggende typologie computationeel te toetsen, zodat kan worden achterhaald in welke mate de geponeerde klassen ook aan objectief-semantische criteria voldoen. Werkwoorden van falen, bijvoorbeeld, worden momenteel

behandeld als een subgroep binnen de categorie obstructie en belemmering. Ze zijn semantisch echter nauw verwant met werkwoorden van succes, waarvan ze morfologisch vaak zijn afgeleid. In een toekomstige bijdrage zal ik daarom de eerder impressionistische methode van de huidige studie trachten te falsifiëren aan de hand van VSS (*Vector Space Semantics*).

2.5 Slotbemerking

In deze bijdrage heb ik het begrip 'subjectachtige datief' geïntroduceerd en behandeld in de context van wat ik 'onpersoonlijke datief-nominatiefwerkwoorden' noem. Subjectachtige datieven zijn vaste argumenten van laag-agentieve werkwoorden en vervullen een rol die doorgaans met het onderwerp geassocieerd wordt. Ze kunnen vooral verward worden met possessieve datieven en datieven van voor- of nadeel in anticausatieve structuren. Daarnaast werd duidelijk dat subjectachtige datieven zijn niet zozeer geassocieerd zijn met een bepaald werkwoord, maar met welbepaalde verbale betekenisschakeringen (Eng. 'senses').

Aan de hand van een niet-exhaustieve lijst in Barðdal (2004), die 45 onpersoonlijke datief-nominatiefwerkwoorden noemt, kon worden aangetoond dat het hedendaagse Duits over 146 verschillende werkwoorden en 202 betekenissen beschikt die zich tot een gebruik in de onpersoonlijke datief-nominatiefconstructie lenen. Een semantische clustering gebaseerd op Barðdal et al. (2016) blijft ook voor het hedendaagse Duits overeind en kon zelfs met één categorie worden uitgebreid, namelijk met werkwoorden van toebehoren. De typologie toont dat meeste onpersoonlijke datief-nominatiefwerkwoorden te maken hebben met emotie (54 betekenissen), cognitie (24 betekenissen) en winst en voordeel (20 betekenissen) en dat de algehele verdeling een Zipfiaanse trend vertoont. Dat resultaat is opvallend, aangezien de Zipfcurve in taalkundig onderzoek vooral wordt geassocieerd met woordfrequenties en niet met de bezetting van een semantisch spectrum.

Toekomstig onderzoek zou zich kunnen richten op de productiviteit van onpersoonlijke datief-nominatiefwerkwoorden door te bestuderen hoe frequent dit soort werkwoord in het hedendaagse Duits precies is. In Sectie 2.4.2 raakte ik kort de idee aan dat dit type werkwoord enerzijds op zijn retour is, maar dat er anderzijds ook nieuwe onpersoonlijke datief-nominatiefwerkwoorden lijken te ontstaan. Daarnaast zou het zinvol zijn om de classificatie uit Sectie 2.4.1 te onderwerpen aan een computationele methode, zodat ze gefalsifieerd kan worden en de samenhang tussen de verschillende klassen nauwkeuriger in kaart kan worden gebracht.

Referenties

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

Baayen, Harald. 2009. Corpus linguistics in morphology: Morphological productivity. In Anke Lüdeling & Merja Kytö (eds.), *Corpus Linguistics: An International Handbook. Volume 2*, 899–919. Berlin: De Gruyter-Mouton.

Barðdal, Jóhanna. 1999a. Case in Icelandic: A construction-grammar approach. *TijdSchrift voor Skandinavistiek* 20(2): 65–100.

Barðdal, Jóhanna. 1999b. Case and argument structure of some loan verbs in 15th-century Icelandic. In Inger Haskå & Carin Sandqvist (eds.), *Alla Tiders Språk: En Vänskrift till Gertrud Pettersson November 1999*, 9–23. Lund: Institutionen för nordiska språk.

Barðdal, Jóhanna. 2001. *Case in Icelandic: A Synchronic, Diachronic and Comparative Approach*. Lund: Department of Scandinavian Languages, Lund University.

Barðdal, Jóhanna. 2004. The semantics of the impersonal construction in Icelandic, German and Faroese: Beyond thematic roles. In Werner Abraham (ed.), *Focus on Germanic typology*, 105–137. Berlin: Akademie Verlag.

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

Barðdal, Jóhanna, Thomas Smitherman, Valgerður Bjarnadóttir, Serena Danesi, Gard B. Jenset & Barbara McGillivray. 2012. Reconstructing constructional semantics: The dative-subject construction in Old-Norse Icelandic, Latin, Ancient Greek, Old Russian and Old Lithuanian. *Studies in Language* 36(3): 511–547.

Barðdal, Jóhanna, Carlee Arnett, Stephen Mark Carey, Þórhallur Eyþórsson, Gard B. Jenset, Guus Kroonen & Adam Oberlin. 2016. Dative subjects in Germanic: A computational analysis of lexical semantic verb classes across time and space. *Language Typology and Universals* 69(1): 49–84.

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

Barðdal, Jóhanna & Þórhallur Eyþórsson. 2020. How to identify cognates in syntax: Taking Watkins' legacy one step further. In Jóhanna Barðdal, Spike Gildea & Eugenio R. Luján (eds.), *Reconstructing syntax*, 197–238. Leiden: Brill.

Blake, Barry J. 2001. *Case*. Cambridge: University Press.

Cosmas II. 2018. Textorganisation. Tagged-M. Mannheim: Institut für Deutsche Sprache.

Draye, Luk. 1996. The German dative. In William Van Belle & Willy Van Langendonck (eds.), *The Dative. 1: Descriptive studies*, 155–215. Amsterdam: Benjamins.

Dudenredaktion (ed.). 2015. *Duden – Deutsches Universalwörterbuch: Das umfassende Bedeutungswörterbuch der deutschen Gegenwartssprache*.

Dudenredaktion (ed.). 2016. *Die Grammatik*. Mannheim, Leipzig, Wien & Zürich: Dudenverlag. Berlin: Bibliographisches Institut.

Dudenredaktion (ed.). z.J. *Duden online*. https://www.duden.de.

DWDS – Digitales Wörterbuch der deutschen Sprache. Das Wortauskunftssystem zur deutschen Sprache in Geschichte und Gegenwart, uitgegeven door de *Berlin-Brandenburgischen Akademie der Wissenschaften*

Engelen, Bernhard. 1975. *Untersuchungen zu Satzbauplan und Wortfeld in der geschriebenen deutschen Sprache der Gegenwart*. München: Hüber.

Eyþórsson, Þórhallur & Jóhanna Barðdal. 2005. Oblique subjects: A common Germanic inheritance. *Language* 81(4): 824–881.

Haspelmath, Martin. 2001. Non-canonical marking of core arguments in European languages. In Alexandra Aikhenvald, Robert M. W. Dixon, Masayuki Onishi (eds.), *Non-canonical subjects and objects*, 53–83. Amsterdam: Benjamins.

Hentschel, Elke & Harald Weydt. 2021. *Handbuch der deutschen Grammatik*. Berlin/Boston: De Gruyter.

Lamiroy, Béatrice & Nicole Delbecque. 1998. The possessive dative in Romance and Germanic languages. In William Van Belle & Willy Van Langendonck (eds.), *The Dative. 2: Theoretical and contrastive studies*, 29–74. Amsterdam: Benjamins.

Lehmann, Christiann. 2015. *Thoughts on grammaticalization*. Berlin: Language Science Press.

Jónsson, Jóhannes Gísli. 1997–98. Sagnir með aukafallsfrumlagi. *Íslenskt mák og almenn málfræði* 19–20: 11–43.

Leibniz-Institut für Deutsche Sprache. *Wörterbuch zur Verbvalenz*. Grammatisches Informationssystem grammis.

Sigurðsson, Halldór Ármann. 1989. *Verbal Syntax and Case in Icelandic*. Doctoraatsverhandeling. Universiteit Lund.

Van Valin, Robert D. Jr. 2018. Dative case and oblique subjects. In Jóhanna Barðdal, Stephen Mark Carey, Na'ama Pat-el (eds.), *Non-Canonically Case-Marked Subjects. The Reykjavík-Eyjafjallajökull papers*, 115–131. Amsterdam: John Benjamins.

Vandeweghe, Willy. 2015. *Grammatica van de Nederlandse zin*. Antwerpen/Apeldoorn: Garant.

Wegener, Heide.1985. *Der Dativ im heutigen Deutsch*. Tübingen: Günther Narr.

Zeldes, Amir. 2012. *Productivity in Argument Selection: From Morphology to Syntax*. Berlin: De Gruyter-Mouton.

Zaenen, Annie, Joan Maling and Höskuldur Þráinsson. 1985. Case and grammatical functions: The Icelandic passive. *Natural Language and Linguistic Theory* 3: 441–483.

Zipf, George Kingsley. 1935. *The Psycho-Biology of Language: An Introduction to Dynamic Philology*. Cambridge (MA): MIT Press.

Þráinsson, Höskuldur. 1979. *On Complementation in Icelandic*. New York: Garland Publishing.

Appendix I

Werkwoord	Betekenis
abgehen	ontberen; gevoelens van gemis ervaren (regionaal)
abhandenkommen	verliezen
ahnen	een voorgevoel hebben (spreektalig); weten van, kennis hebben van
angehören	behoren tot (een bepaalde groep); behoren tot (een bepaalde periode); nauw verbonden zijn met iemand
anhängen	op iemand wegen
ankommen	overvallen, overkomen; zwaar vallen; van belang zijn; aankomen op
anliegen	van belang zijn
anstehen	bevallen, aanstaan; genoeg hebben van (Oostenrijks)
aufdämmern	dagen
auffallen	opvallen
aufgehen	zich openbaren aan
aufstoßen	opvallen (spreektalig); oprispen; tegensteken
ausgehen	verminderen, op geraken; tekortkomen (adem)
ausmachen	uitmaken
ausreichen	genoeg hebben van
bedeuten	betekenen voor; zijn
begegnen	overkomen; (toevallig) aantreffen
behagen	behagen
beifallen	te binnen schieten
beikommen	opkomen bij, te binnen schieten
beiwohnen	eigen zijn (aan)
bekommen	bekomen, bevallen
belieben	believen
bevorstehen	te wachten staan; opzien tegen (Noord-Duits)
bleiben	blijven, resteren
blühen	te wachten staan

dämmern	dagen
danebengeraten	mislukken
dazwischenkommen	tussenkomen
dienen	dienen
drohen	dreigen
dünken	dunken
eignen	typeren, eigen zijn (aan)
eilen	dringend zijn
einfallen	te binnen schieten (van een idee); zich herinneren
eingehen	begrijpen; bevallen
einkommen	te binnen schieten
einleuchten	duidelijk zijn
ekeln	doen walgen
entfahren	ontsnappen (van woorden)
entfallen	ontglippen, laten vallen; vergeten
entgegenschallen	in iemands richting weerklinken
entgegenschlagen	(met heftigheid) gewaarworden (let.); (met heftigheid) gewaarworden (fig.)
entgegentreten	opmerken, stoten op
entgehen	zich laten ontglippen, voorbij laten gaan; niet opgemerkt worden, ontgaan
entgleiten	ontglippen, laten vallen; ontglippen (fig.), verliezen; ontsnappen (van woorden)
entschlüpfen	ontsnappen (van woorden)
entschwinden	mentaal ontgaan, vergeten; uit iemands blikveld verdwijnen
entstehen	ontstaan (voor iemand)
erbarmen	spijten (Oostenrijks)
ermangeln	ontbreken
erscheinen	voor iemand verschijnen (let.); voorkomen als
erwachsen	ontstaan
fehlen	missen, ontberen; missen (emotioneel); niet gezond zijn (regionaal); niet voorhanden zijn, verloren gegaan zijn; nodig hebben
fehlen an	ontbreken aan
fernliegen	niet in iemand opkomen; niet iemands intentie zijn
folgen	volgen op
frommen	tot nut strekken
gebrechen	ontbreken

gebrechen an	ontbreken aan
gebühren	toekomen
gefallen	bevallen
gegenübertreten	voorkomen, treffen als
gehören	toebehoren (let.); toebehoren (fig.), behoren aan; toekomen, verdienen (Zuid-Duits, Oostenrijks, spreektalig)
gelingen	lukken
gelten	waard zijn; gelden (als), beschouwen
genügen	volstaan
geraten	lukken; voor iemand op een bepaalde manier aflopen
geschehen	gebeuren, overkomen
geziemen	betamen, horen
glücken	lukken
grauen	beangstigen
graulen	doen huiveren
grausen	beangstigen
gruseln	beangstigen
gutgehen	goed gaan (emotioneel); goed gaan (economisch)
guttun	goeddoen
hinterhereilen	achtervolgen (bv. van succes)
imponieren	imponeren
innewohnen	eigen zijn (aan)
klarwerden	duidelijk worden
klingen	klinken (als)
kommen	komen (gelegen, ongelegen, onverwacht …); overvallen
konvenieren	uitkomen, passen (Oostenrijks)
lachen	toelachen
langen	volstaan (spreektalig)
leichtfallen	makkelijk zijn
leidtun	medelijden hebben met; spijten
liegen	liggen, goed zijn in; geven om, waarderen; liggen, het kunnen vinden met (spreektalig)
mangeln	ontbreken
mangeln an	ontbreken aan
missfallen	misvallen
missglücken	mislukken
misslingen	mislukken

missraten	mislukken
munden	smaken
nachgehen	bijblijven
nachhängen	blijven hangen
nahegehen	bedroeven
nutzen	nuttig zijn (voor)
nützen	nuttig zijn (voor)
obliegen	moeten, als plicht toekomen
passen	passen (van kleren); goed uitkomen (spreektalig)
passieren	overkomen
pressieren	dringend zijn
rauchen	irritatie ondervinden, boos worden (spreektalig)
reichen	volstaan (spreektalig)
schaudern	huiveren (let.); huiveren (fig.), vrezen
schauern	huiveren (let.); huiveren (fig.), vrezen
scheinen	schijnen
schlechtgehen	slecht gaan (emotioneel); slecht gaan (economisch)
schmecken	een bepaalde smaak hebben (let.); een bepaalde smaak hebben (fig.); goed smaken (let.); bevallen
schwanen	een voorgevoel hebben
schwerfallen	moeilijk zijn (voor)
schwindeln	duizelen
sein	toebehoren (regionaal, spreektalig)
stehen	staan (van een kledingstuk); staan (fig.), afgaan
stinken	kwaad maken
träumen	dromen
unterlaufen	overkomen; overkomen (van een fout)
verbleiben	resteren
vergehen	voorbijgaan (tijd); verliezen, afnemen
verlorengehen	verloren gaan
vorkommen	overkomen (gebeurtenis); ervaren; voorkomen, treffen (als)
vorschweben	voor ogen hebben
wehtun	pijn doen
widerfahren	overkomen
widerstehen	mishagen; niet smaken (eten)
widerstreben	met tegenzin vervullen
winken	in het vooruitzicht staan
wohltun	goeddoen

zufallen	in het bezit komen; te beurt vallen
zufliegen	te binnen schieten
zufließen	ontvangen
zugehören	behoren (tot)
zugutekommen	ten goede komen
zuhandenkommen	krijgen
zukommen	ontvangen; toegestuurd krijgen (bv. van een bericht); toekomen, mogen; krijgen (fig., zoals van een bepaalde rol); toe te meten zijn
zupasskommen	van pas komen
zusagen	bevallen
zusetzen	kwellen
zustattenkommen	van pas komen
zustehen	verdienen; toekomen, mogen
zustoßen	overkomen
zuteilwerden	te beurt vallen
zuwachsen	krijgen

Appendix 2

The current Appendix provides a set of 16 examples for each of the semantic classes presented in Chapter 2. Each set comprises an (a) example and a (b) example, with the former illustrating the Dat-Nom argument structure and the latter the Nom-Dat argument structure.

Emotion

Dat-Nom

(1) a. *Den Schülern* *fiel*
the.DAT pupils.DAT fell
die Entscheidung zwischen diesen interessanten Themen *schwer.*
the.NOM decision between these interesting topics hard
'The pupils found it difficult to choose between these interesting topics.'

Nom-Dat

(1) b. *Der Abschied* *dürfte* *den wenigsten Fahrgästen*
the.NOM goodbye could the.DAT fewest.DAT passengers.DAT
schwerfallen.
hard.fall
'Saying goodbye is likely to be difficult for very few passengers indeed.'

Cognition

Dat-Nom

(2) a. *Den drei Männern* *imponierten* *seine Worte.*
the.DAT three men.DAT impressed his.NOM words
'The three men were impressed with his words.'

Nom-Dat

(2) b. *Macht* *imponiert* *den Menschen* *immer.*
power.NOM impresses the.DAT people always
'Power always impresses people.'

Gain and benefit

Dat-Nom

(3) a. *Der besten Einsendung* *winkt* *ein Traumurlaub nach Brasilien.*
the.DAT best.DAT entry beckons a.NOM dream.holiday to Brazil
'The best entry will be treated to a dream holiday to Brazil.'

Nom-Dat

(3) b. *Vielfältige Aufgaben* *winken* *den Bachelorabsolventen.*
various.NOM tasks beckon the.DAT bachelor.graduates
'A wide range of tasks await Bachelor graduates.'

Happening

Dat-Nom

(4) a. *Jung-Akademikern* *droht* *die Privatinsolvenz.*
young.academics.DAT threatens the.NOM personal.insolvency
'Early-career academics face personal insolvency.'

Nom-Dat

(4) b. *Das Gleiche* *droht* *Deutschland*
the.NOM same threatens Germany.DAT
bei der europäischen Wirtschaftsregierung.
by the European economic government
'The same fate threatens Germany from the European economic government.'

Hindrance

Dat-Nom

(5) a. *Einem Großteil der Radiosender* *sind* *in den letzten Jahren*
a.DAT large.share the.GEN radio.stations are in the last years
die Hörer *abhandengekommen.*
the.NOM listeners lost
'A majority of radio stations have lost their listeners in recent years.'

Nom-Dat

(5) b. *Der gesunde Menschenverstand* *ist* *den Herrschenden*
the.NOM healthy.NOM people's.sense is the.DAT rulers.DAT
abhandengekommen.
lost
'Common sense has deserted rulers.'

Attitude

Dat-Nom

(6) a. *Der deutschen Mannschaft* *genügt* *zum Weiterkommen*
the.DAT German.DAT football.team suffices to progress
ein Unentschieden.
a.NOM draw
'The German team needs no more than a draw to advance.'

Nom-Dat

(6) b. *Dieser kurze Augenblick* *genügte* *unbekannten Langfingern,*
this.NOM short.NOM moment sufficed unknown.DAT thieves.DAT
um das Lauflernrad im Wert von 35 Euro zu stehlen.
to steal the baby walker worth 35 euros
'This brief moment was enough for unknown thieves to steal the baby walker worth 35 euros.'

Possession

Dat-Nom

(7) a. *Den Byzantinern* *verblieb* *im äußersten Nordosten Siziliens*
the.DAT Byzantines.DAT were.left.with in the far north.east Sicily.GEN
das Festungsdreieck Taormina-Rometta-Demanna.
the.NOM fortress.triangle Taormina-Rometta-Demanna
'The Byzantines were left in the far north-east of Sicily with the fortress triangle Taormina-Rometta-Demanna.'

Nom-Dat

(7) b. *Zwei Drittel* *verbleiben* *den Beschäftigungsträgern.*
two.NOM thirds remain.with the.DAT employment.organisations.DAT
'Two thirds remain with the employment organisations.

Perception

Dat-Nom

(8) a. *Auch meiner Frau* *schien* *das Essen* *zu munden.*
also my.DAT wife seemed the.NOM food to taste
'My wife also seemed to enjoy the food.'

Nom-Dat

(8) b. *Ein Gläschen Wein* *mundet* *hier* *keinesfalls*
a.NOM little.glass wine tastes here by.no.means
nur dem Weinkenner.
just the.DAT wine.connoisseur
'A glass of wine is not just for the connoisseur to enjoy here.'

Bodily states

Dat-Nom

(9) a. *Dem jungen Kind* *stieß* *die herbe Speise* *auf.*
the.DAT young.DAT child repeated the.NOM bitter.NOM food on
'The young child found the bitter food difficult to digest.'

Nom-Dat

(9) b. *Zu viel Obst* *könnte* *auf Dauer* *dem Pferd* *sauer*
too.NOM much fruit could in run the.DAT horse sour
aufstoßen.
repeat.on
'Too much fruit could disagree with the horse's stomach in the long run.'

Modality

Dat-Nom

(10) a. *Dem Geschäftsführer* *obliegt*
the.DAT Managing.Director is.responsible.for
die Verwaltung der üblichen Geschäfte.
the.NOM administration the.GEN ordinary.GEN business
'The Managing Director is charged with business administration as usual.'

Nom-Dat

(10) b. *Die Geschäftsführung des Vereins* *obliegt*
the.NOM company.running the.GEN organisation.GEN is.incumbent.upon
dem Vorstand.
the.DAT management
'The running of the association is incumbent upon the secretariat.'

Ontological state

Dat-Nom

(11) a. *Der Sozialdemokratie* *wohnt* *ein starker zentralistischer Zug*
the.DAT social.democracy lives a.NOM strong.NOM centralist.NOM streak
inne.
in
'Social democracy has a strong centralist streak.'

Nom-Dat

(11) b. *Der Wechsel zwischen Fülle und Leere* *wohnt*
the.NOM change between fullness and emptiness lives
den Prinzipien unseres Kosmos *inne.*
the.DAT principles our.GEN cosmos in
'The alternation between fullness and emptiness is inherent to the principles of our cosmos.'

Evidentiality

Dat-Nom

(12) a. *Den Entdeckern der Region* *erschien*
the.DAT discoverers.DAT the.GEN region seemed
der fruchtbare Landstrich *wie der Garten Eden.*
the.NOM fertile. NOM stretch.of.land like the garden Eden
'To the discoverers of the region the fertile stretch of land seemed like the garden of Eden.'

Nom-Dat

(12) b. *Der Start der Raumfähre* *erschien* *den Technikern*
the.NOM start the.GEN space.shuttle seemed the.DAT technicians.DAT
zu riskant.
too risky

'The launch of the space shuttle was deemed too risky by the technicians.'

Success

Dat-Nom

(13) a. *Den römischen Truppen* *gelang*
the.DAT Roman.DAT troops succeeded
der Durchbruch ins Lager und dessen Verteidigung.
the.NOM breakthrough into the camp and its defence
'The Roman troops managed to break through into the camp and its defences.'

Nom-Dat

(13) b. *Ein spektakulärer Angriff* *gelang* *Stefan*
a.NOM spectacular.NOM attack succeeded Stefan.DAT
kurz vor der 1. Sprintwertung.
shortly before the first sprint classification
'A spectacular attack was successfully launched by Stefan shortly before the first sprint classification.'

Belonging

Dat-Nom

(14) a. *Der Oberschicht* *sollten* *nach Meinung dieser User*
the.DAT upper.class should after opinion these.GEN users
mehr Leute *zugehören.*
more people.NOM belong.to
'The upper class should comprise more people in the opinion of these users.'

Nom-Dat

(14) b. *Drei Milliarden mehr Menschen* *werden* *der Mittelklasse*
three billion more people.NOM will the.DAT middle.class
zugehören.
belong.to
'Three billion more people will belong to the middle class.'

Speaking

Dat-Nom

(15) a. *Auch dem coolsten Fahrgast* *entfuhr* *da* *ein erschreckter Aufschrei.*
also the.DAT coolest.DAT passenger let.out there a.NOM startled.NOM cry
'Even the coolest passenger let out a startled cry at that moment.'

Nom-Dat

(15) b. *Ein leises Seufzen* *entfuhr* *dem Piraten.*
a.NOM soft.NOM sigh escaped the.DAT pirate.DAT
'A soft sigh escaped the pirate.'

Natural occurrences

Dat-Nom

(16) a. *So manchem Besucher* *verging* *die Zeit* *am Sonntagabend*
so many.a.DAT visitor went.by the.NOM time on Sunday.evening
zu langsam.
too slowly
'For many a visitor, time went by too slowly on Sunday evening.'

Nom-Dat

(16) b. *Die Zeit* *verging* *uns* *nur so* *im Fluge.*
the.NOM time went.by us.DAT just so in flight
'Time just flew by for us.'

Chapter 3
Linearization and unaccusativity: The relative order of dative and nominative arguments with five German verbs of success and failure

The present study maps out the word order distributions in verb-second clauses of five German verbs of success and failure: *gelingen* 'to succeed', *glücken* 'to succeed', *missglücken* 'to fail', *misslingen* 'to fail', and *missraten* 'to fail'. Given their status as unaccusatives, they display an asymmetric mapping between case marking and thematic roles. As a consequence, several researchers have argued that they sanction an inversion of the canonical nominative-before-dative word order pattern. Two positions emerge: those scholars who associate these verbs with a dative-before-nominative order and those who claim that they license both a dative-before-nominative order as well as a nominative-before-dative order. A corpus study of 982 tokens principally shows the latter claim to be true in that, in a configuration with double NPs, the Nom-Dat order is attested 200 times (or 40%), and the Dat-Nom order 299 times (or 60%). Remarkably, when both arguments are realised as pronouns, the variation in word order is neutralised almost entirely in favour of the nominative-before-dative order.

3.1 Introduction

Scholars of German have long been intrigued by the interplay between verbal syntax and semantics, on the one hand, and their influence on word order patterns, on the other. With regard to predicates licensing both a nominative and a dative argument, it has been

argued that not all of them are associated with a nominative-before-dative base order. Among these are verbs of success and failure: as unaccusative verbs, they display an asymmetric mapping between case marking and thematic roles. This is because with these verbs it is the dative, and not the nominative, that takes on the thematic role normally associated with the subject. It is precisely this mismatch that facilitates an inversion of the canonical word order pattern. Examples (1a–b) below illustrate this phenomenon for the verbs *gelingen* 'to succeed' and *missraten* 'to fail':

(1) a. *WhatsApp* *gelang* *der Durchbruch.*
WhatsApp.DAT succeeded the.NOM breakthrough
'WhatsApp managed to break through.'

(1) b. *Einem minderbegabten Schriftsteller* *wäre* *der Roman* *missraten.*
a.DAT less.talented.DAT author were the.NOM novel failed
'A less talented writer would have failed (with) the novel.'

The aim of this paper is to lend corpus-based support to the above-mentioned claim that verbs of success and failure allow for the canonical order of constituents to be inverted. I do this by mapping out the word order distributions of five relevant verbs: *gelingen* 'to succeed', *glücken* 'to succeed', *missglücken* 'to fail', *misslingen* 'to fail', and *missraten* 'to fail'. These verbs are either synonyms or antonyms to one another, which in turn means that there is a high degree of internal semantic coherence between the members of the dataset. The idea behind this method is that verbs that are highly similar in meaning may be assumed to also select for the same argument structure (cf. Barðdal 2001a, 2004: 114, 2008, 2012; Barðdal & Eythórsson 2020: 211–216, 223–229).

This paper is organised as follows. In Section 3.2, I present the two positions that scholars of German have taken regarding linearization with verbs of success and failure. Several researchers have claimed that these are associated with a dative-before-nominative base order, but others have argued that they license two base orders: a dative-before-nominative order and a nominative-before-dative order. I also discuss previous corpus work tying in with either of these positions. Section 3.3 formulates the hypotheses which build the foundations of this study. Section 3.4 lays out the methodological background, detailing the processes of verb selection, data cleaning, and data annotation. The analysis and the discussion of the obtained results are jointly subsumed under Section 3.5. Section 3.6 concludes this paper and provides a brief outlook for future work.

3.2 Word order patterns with verbs of success and failure

In this section, I review the two positions the German scholarship has taken regarding word order preferences with verbs of success and failure. On the one hand, there are those researchers who have associated this class of verbs with a dative-before-nominative base order. Their work is discussed in Section 3.2.1. Others have argued that verbs of success and failure correlate with two base orders: a dative-before-nominative order and a nominative-before-dative order. This position is outlined in Section 3.2.2.

3.2.1 Dative-before-nominative

First, there are those researchers who assume verbs of success and failure to be associated with a dative-before-nominative base order. This claim is sustained by syntactic as well as by semantic observations, even though these partially imply each other.

Syntactically speaking, verbs of success and failure are so-called 'unaccusative' verbs. Unaccusatives are a special class of intransitives which encode a prototypical patient in their subject slot. In some languages, like German, they build perfect tenses with the verb *sein* 'to be' (Duden 2016: 419; Eisenberg 2013: 75, *inter alia*). The oblique argument, which invariably bears dative case marking, is ranked thematically higher than the nominative and due to this thematic asymmetry, the dative is assumed to precede the nominative in the linear order. This view is adopted by Fanselow (2002: 231; 2003: 38), Bader & Häussler (2010: 751), Haider (2010: 254, 260), Eisenberg (2013: 74), and Bader (2020: 1088), *inter alia*.

One source that has lent corpus-based support to the claim that unaccusative verbs indeed prefer the dative-before-nominative order is Bader (2020). His study on linearization in the middlefield with dative pronouns and nominative full NPs has shown that, with unaccusatives, the dative precedes the nominative 91% of the time (out of a total of 636 tokens). Note that the German middlefield spans all nominal arguments following the finite verb in verb-first and verb-second clauses. In verb-final clauses, the middlefield begins after the complementizer.

Bader's findings are very convincing, although these should nevertheless be taken with caution, as object pronouns are inclined to precede full NPs in the middlefield anyway. Disregarding reflexives, research has also shown that this tendency is even stronger for dative pronouns than it is for accusative pronouns (Shannon 2000; Heylen 2005). An example of object-fronting in the middlefield is given under (2):

(2) *Nun sieht er ein, dass* ***ihm*** *kein Lehrer helfen konnte.*
now sees he in that him.DAT no.NOM teacher help.INF could
'Now he realises that no teacher could help him.'

Another study worth mentioning is Kurz (2000). She, too, investigates word order variation in the middlefield, but her study only comprises tokens containing two full NPs. For the verb *gelingen*, she finds that the dative-before-nominative order is attested 257 times out of a total of 260 observations (or 99%), of the type shown in (3):

(3) *Im Februar gelang Evelyn der Berufseinstieg.*
in February succeeded Evelyn.DAT the.NOM career.start
'In February, Evelyn managed to start her career.'

Kurz's results convincingly show that *gelingen* indeed instantiates a dative-before-nominative order in the middlefield, but it remains to be seen whether this tendency is equally strong in verb-second clauses, i.e. when one constituent is in the prefield and the other is in the middlefield.

Other scholars have a broader conception of verbs correlating with a dative-before-nominative base order. Lenerz (1977: 112–120), for instance, states that this order typically occurs with psychological verbs, like *gefallen* 'to please' or *auffallen* 'to strike'. Nevertheless, he observes that a dative-before-nominative order may also occur with verbs not necessarily referring to emotion or cognition, like *fehlen* 'to lack' or *gelingen* 'to succeed'. Hoberg (1981: 75), too, lists *gelingen* as a verb with a dative-before-nominative base order. Neither Lenerz nor Hoberg mention any other verbs of success and failure, but given the shared syntactic and semantic nature of these verbs, I assume that their observations also hold for the verbs *glücken* 'to succeed', *missglücken* 'to fail', *misslingen* 'to fail', and *missraten* 'to fail'.

The synchronic position outlined above finds a diachronic counterpart in Johnson et al. (2019), who reconstruct a dative-subject construction for verbs of success and failure not only in Proto-Germanic, but also in Proto-Indo-European. These scholars define the notion of subject as the first argument in the argument structure, which means that they essentially agree with the aforementioned sources that the default position of the dative for verbs of success and failure is, at least historically, to the left of the nominative (Johnson et al. 2019: 487).

3.2.2 Dative-before-nominative and nominative-before-dative

There are also researchers who associate verbs of success and failure with two base orders: a dative-before-nominative order and a nominative-before-dative order. One of the first to put forward this argument, albeit indirectly, were Belletti & Rizzi (1988), as they famously introduced the distinction between three classes of experiencer verbs in

Present-Day Italian: *temere* verbs, *preoccupare* verbs, and *piacere* verbs. *Piacere* verbs bear a striking resemblance to the verbs under study here: they code the experiencer in the dative case, they select for the perfect tense with *essere* 'to be', thus confirming their status as unaccusatives, and, crucially, they allow both verbal arguments to take clause-initial position without one order being more marked than the other (Belletti & Rizzi 1988: 334). According to Belletti & Rizzi, such unmarked inversion of the dative and the nominative in unaccusative verbs is by no means a peculiarity of Italian, as it can also be found in other languages (Belletti & Rizzi 1988: 292).

Among the first researchers to claim that German is one such language was Primus (1999: 156). She maintains that, whenever the case hierarchy, presented under (4), and the thematic hierarchy, presented under (5), are in conflict with one another, both the dative and the nominative may take initial position, because speakers cannot possibly comply with both hierarchies at the same time.

(4) **Case hierarchy**: nominative > accusative > dative
(5) **Thematic hierarchy**: proto-agent > proto-recipient > proto-patient

As mentioned above, verbs of success and failure, being unaccusative verbs, exhibit an asymmetric mapping between case marking and thematic roles. For the proponents of the unaccusative approach, the thematic hierarchy clearly trumps the case hierarchy, so that unaccusative verbs are argued to correlate with a dative-before-nominative base order only (cf. Section 3.2.1). For Primus, by contrast, both hierarchies balance each other out, thus facilitating two base orders: a dative-before-nominative order and a nominative-before-dative order. Nevertheless, Primus concludes that (5) is somewhat stronger than (4) for many German speakers and that dative-before-nominative order is therefore slightly, but not significantly, more acceptable than nominative-before-dative order (Primus 1999: 157).

Another scholar who has claimed that verbs of success and failure exhibit two base orders is Haspelmath (2001: 67). He argues that they share this property with other verbs displaying non-canonical case marking, like verbs of emotion (e.g. *gefallen* 'to please'), verbs of possession (e.g. *fehlen* 'to lack'), and happenstance verbs (e.g. *passieren* 'to happen to').

The same reasoning may be found in Barðdal (2004, 2023: Ch 3), Eythórsson & Barðdal (2005), and Barðdal et al. (2019). Couching their work in a Construction Grammar framework, they argue that German verbs with subject-like datives and object-like nominatives (which includes verbs of success and failure) may be used in two diametrically opposed case frames: a Dat-Nom case frame and a Nom-Dat case frame. This is a property they share with certain Dat-Nom verbs in Icelandic, which also select for two different argument structure constructions and have therefore been termed 'alternating predicates' or Dat-Nom/Nom-Dat predicates in the literature (cf. Barðdal 2001b; Rott 2016).

Two studies that have been able to lend corpus-based support to the positions discussed above are Bader & Häussler (2010) and Verhoeven (2015). In contexts with

double NPs in which either argument is adjacent to the conjugated verb, Bader & Häussler have found verbs with a nominative and a dative argument to attest the dative-before-nominative order and the nominative-before-dative order approximately equally often; that is, out of 42 attestations, the dative precedes the nominative 19 times (or 45%), and the nominative precedes the dative 23 times (or 55%). However, there are two reasons for why these results should be received with caution. First, Bader & Häussler's dataset is not limited to unaccusatives alone: it also contains an unknown number of observations with verbs that strongly correlate with a nominative-before-dative order, like *gehorchen* 'to obey' or *vertrauen* 'to trust'. As a result, the number of tokens attesting the nominative-before-dative order is potentially greatly inflated. Second, as an artefact of the data collection process, Bader & Häussler's dataset only contains tokens with definite datives. The nominative, by contrast, is not subject to any such limitations, which means that it may or may not be definite. The dative-before-nominative order is thus facilitated by a potential asymmetry in definiteness.

The second limitation is inherent to the data retrieval process, but the first is not. In order to make up for it, Bader & Häussler zoom in on the word order distributions of four subsets of verbs that favour an oblique-before-nominative order: verbs in the passive voice (dative before nominative), unaccusative verbs (dative before nominative), so-called "active dative *haben* verbs" like *gefallen* 'to please' (dative before nominative), and so-called "active accusative *haben* verbs" like *interessieren* 'to interest' (accusative before nominative). However, these results, too, should be taken with caution, since Bader & Häussler amalgamate all four subsets into one, thereby obscuring the case information. Bearing this limitation in mind, Bader & Häussler find that, in contexts where both arguments are adjacent to the conjugated verb, the oblique argument precedes the nominative 16 times (or 47%), and the nominative precedes the oblique argument 18 times (or 53%). This distribution ties in well with the studies reviewed above.

One final study to be mentioned in this context is Verhoeven (2015). She collects data for ten so-called 'dative-experiencer' verbs in verb-second clauses in which both arguments are realised as full NPs. As many as 448 tokens (or 38%) have the dative preceding the nominative, whereas 716 tokens (or 62%) attest the reverse order. Unfortunately, Verhoeven's sample does not contain any verbs of success and failure, but her results do show that thematic asymmetries in general indeed bring about an inversion of the canonical order of constituents.

3.3 Hypotheses

The positions outlined in the previous section give rise to two linearization hypotheses for German verbs of success and failure. Hypothesis 1, which is inspired by the scholarly work discussed in Section 3.2.1, may be formulated as follows:

H1 German verbs of success and failure are associated with a dative-before-nominative base order. Hence, they are hypothesised to overwhelmingly realise the dative in initial position in declarative clauses.

The second hypothesis, which builds on the work discussed in Section 3.2.2, reads thus:

H2 German verbs of success and failure are associated with two base orders: a dative-before-nominative order and a nominative-before-dative order. In declarative clauses, they are consequently hypothesised to realise either argument in initial position approximately equally often.

I now turn to the methodology of this study before analysing and discussing the obtained results.

3.4 Methodology

The present section introduces the methodology of this study. The first subsection discusses which verbs have been included in this study and the criteria for doing so. The second subsection outlines the data harvesting and cleaning process. The third subsection deals with the annotation variables.

3.4.1 Verb selection

As is already stated above, this paper zooms in on five German verbs of success and failure: *gelingen* 'to succeed', *glücken* 'to succeed', *missglücken* 'to fail', *misslingen* 'to fail', and *missraten* 'to fail'. As is documented in Somers (2021), Present-Day German possesses more dative-nominative verbs than these five to express success and failure, but these have not been retained for the following reasons.

First, the verbs under study all license a dative as part of the main verb's subcategorization frame. Potential candidate verbs like *danebengehen* 'to miss', *fehlschlagen*

'to go wrong, to fail', or *schiefgehen* 'to go wrong' may also occur with a dative, but theirs is arguably a so-called free dative: neither the *Duden* dictionary nor the *Digitales Wörterbuch der deutschen Sprache* explicitly acknowledge these datives as being part of the verbal case frame. Hence, these verbs have not been included for study.

Second, every verb type has to yield a sufficient number of hits in the deTenTen13 corpus (Kilgarriff et al. 2004, 2014; 16.5 billion words), which is the corpus selected for the present study. In order to gauge the prevalence of the [Nom-V-Dat] order and the [Dat-V-Nom] order, I have scrutinised the first 300 randomised tokens per verb type. As soon as this 300-word sample returned seven or more eligible tokens for both configurations combined, the verb in question was retained. If it did not yield seven or more eligible tokens the verb was discarded. Both *geraten* 'to succeed' and the particle verb *danebengeraten* 'to fail' failed to reach this threshold and were consequently eliminated from the study. The main reason for introducing this threshold was to safeguard the study's feasibility: in case a sample returned six (or fewer) eligible tokens, more than 10,000 tokens would have had to be cleaned in order to reach the aim of 200 eligible tokens per type, and that was deemed unmanageable.

Third, in order to keep the dataset semantically coherent, certain polysemous types were also barred from the study. These include *gutgehen* 'to thrive (economically); to thrive (emotionally)', *schlechtgehen* 'to go badly (economically); to go badly (emotionally)', and *unterlaufen* 'to make a mistake; to happen to'. Each of these verbs has exactly two senses, but only the first of these senses is compatible with the semantic field of success and failure. Because the dataset has not been annotated for sense, it was decided to exclude these verbs altogether.

3.4.2 Data cleaning

The present study endeavours to shed light on word order distribution in German verb-second clauses. All eligible tokens therefore contain (pro)nominal arguments that are situated on either side of the finite verb, as opposed to contexts in which they both follow the finite verb and, thus, occupy the middlefield.

As mentioned earlier, the data have been collected from the deTenTen13 corpus. For each verb, I have run a lemmatized search query using the Sketch Engine interface. The data were subsequently downloaded in sets of 10,000 randomised tokens per type, unless the search returned fewer than 10,000 hits. For each verb, I collected the first 200 eligible tokens per type. Data sparsity was an issue only for *missraten*: for this verb, the corpus only contains 182 instantiations of the two word orders. Thus, the total number of eligible tokens for all verbs combined equals 982, not 1,000.

Table 3.1 presents an overview of the five verbs under study (column 1), the number of randomised tokens extracted per type (column 2), the number of tokens checked

before the target of 200 was reached, or, in case of *missraten*, before the available data ran out (column 3), as well as the number of tokens ultimately included for study (column 4).

Table 3.1 Overview per verb of the total number of tokens extracted, checked and retained

Verb	Tokens extracted	Tokens checked	Tokens retained
gelingen	10,000	1,011	200
glücken	10,000	3,871	200
missglücken	5,930	2,652	200
misslingen	10,000	3,308	200
missraten	1,896	1,896	182
Total	37,826	12,738	982

As shown in Table 3.1, some verbs reveal a wide discrepancy between the number of tokens checked and the number of tokens ultimately retained. This is especially relevant for *glücken*, *missglücken*, and *misslingen*. As it turns out, all three are particularly common as intransitives. Moreover, the results for *glücken* are obscured by a number of false positives containing the noun *Glück* 'luck', which is homonymous with the stem of the verb.

3.4.3 Data annotation

All tokens were annotated for the following variables: case, constituent order, (pro)nominality, pronoun type (if applicable), referentiality, person, number, definiteness, animacy, and length. For reasons of space, the present study mainly focuses on the first three variables, which were annotated according to one of two values:

- Case: nominative or dative
- Constituent order: nominative-dative or dative-nominative
- (Pro)nominality: **pronoun** or **full NP**

3.5 Results and discussion

In this section, I first present a general overview of word order distributions across configurations, that is: without distinguishing between pronouns and full NPs (subsection 3.5.1). The four following subsections each highlight word order patterns in a specific configuration. The competition between two full NPs is the subject of subsection 3.5.2, whereas word order distributions in the double-pronoun configuration are discussed in

subsection 3.5.3. Subsections 3.5.4 and 3.5.5 deal with the interplay between full NPs and pronouns: they respectively discuss the competition between nominative NPs and dative pronouns (subsection 3.5.4), and dative NPs and nominative pronouns (subsection 3.5.5).

3.5.1 General frequencies

Table 3.2 presents an overview of the word order distributions within the dataset as a whole. It shows that both word order patterns are attested approximately equally often: the Nom-Dat pattern occurs 542 times (or 55%), and the Dat-Nom pattern 440 times (or 45%). The intra-class differences are also relatively small: the two most extreme verbs in the sample, i.e. *gelingen* and *missraten*, only show a 19.5% interval between their preference for the Nom-Dat order and Dat-Nom order, respectively. The three remaining verbs are situated between these two, with *glücken* and *missglücken* approximating a 50–50 distribution.

Table 3.2 Word order distributions across configurations

	Nom-Dat		**Dat-Nom**	
Verb	**N**	**f**	**N**	**f**
gelingen	91	45.5%	109	54.5%
glücken	98	49%	102	51%
missglücken	106	53%	94	47%
misslingen	128	64%	72	36%
missraten	119	65%	63	35%
Total	542	55%	440	45%

The numbers in Table 3.2 corroborate Hypothesis 2, which assumes that verbs of success and failure license two base orders: a dative-before-nominative order and a nominative-before-dative order. However, it is important to bear in mind that these frequencies are only rough tallies, and that they disregard any effects of (pro)nominality. The following subsections therefore explore word order variation in configurations where lexical specifications are held constant.

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

Table 3.3 shows word order distributions for double NPs. What is especially striking, is that the Dat-Nom order is consistently more felicitous than it is across configurations (cf. Table 3.2): its prevalence increases from 45% across configurations to 60% when both arguments are realised as full NPs.

Table 3.3 Word order distributions in the [NP-V-NP] configuration

	Nom-Dat		Dat-Nom	
Verb	**N**	**f**	**N**	**f**
gelingen	17	28%	43	72%
glücken	42	33%	84	67%
missglücken	55	42%	77	58%
misslingen	34	42%	47	58%
missraten	52	52%	48	48%
Total	200	40%	299	60%

The results presented in Table 3.3 again confirm the alternating nature of the verbs under scrutiny: word order distributions, both for the class as a whole, as well as for each verb individually, show that either nominal argument may take initial position in declarative clauses approximately equally often. Thus, Hypothesis 2 is once more borne out. An example of each word order pattern, Dat-Nom and Nom-Dat, is given in (6a–b) below:

(6) a. *Dem Spieler* *missglückt* *ein Eckstoß.*
the.DAT player fails a.NOM corner.kick
'The player misses a corner kick.'

(6) b. *Die Revanche* *ist* *unserem Team* *geglückt.*
the.NOM revenge is our.DAT team succeeded
'Our team succeeded in taking revenge.'

Furthermore, it is striking that these findings reflect Primus's (1999: 157) prediction about the interplay between the case hierarchy and the thematic hierarchy. She argues that the latter hierarchy overrules the former, and that many German speakers consequently perceive of the dative-before-nominative order as slightly more felicitous than the nominative-before-dative-order. The fact that 60% of all tokens containing double NPs realise the dative in clause-initial position corroborates this claim.

The numbers in Table 3.3 also present the mirror image of Verhoeven's (2015) numbers. In the present study, the dative-before-nominative order outmatches the nominative-before-dative order by 20%, but in Verhoeven's, the nominative-before-dative order outmatches the dative-before-nominative order by 24%. It has already been mentioned that Verhoeven's sample comprises different verbs than those presented here, and that she does not include any verbs of success and failure in her dataset. At the same time, Verhoeven does not specify how each individual verb weighs in on the distribution across word order patterns, which means that it is difficult to assess what lies behind her numbers.

Note that the verb *gelingen*, despite its propensity for the dative-before-nominative order (43 tokens, or 72%), still instantiates the nominative-before-dative order with

reasonable frequency (17 tokens, or 28%). This is particularly interesting given that Kurz's (2000) study on word order distributions in the middlefield found this verb to overwhelmingly tend to the dative-before-nominative order (257 tokens out of a total of 260, or 99%). Recall that Kurz only included tokens with full NPs, which makes her numbers perfectly comparable to the ones presented here. Thus, even with full NPs, the topological field seems to exert a major influence on the resulting word order pattern.

Finally, it is striking that the dative-before-nominative order is slightly more felicitous for the two verbs of success (*gelingen* and *glücken*) than it is for the three verbs of failure (*missglücken*, *misslingen*, and *missraten*). The difference between both semantic groups is also statistically informative: comparing each to a chance distribution, the result is significant only for verbs of success (χ^2 = 24.86; df = 1; $p_{two\text{-}tailed} < 0.0001$), but not for verbs of failure (χ^2 = 3.07; df = 1: $p_{two\text{-}tailed}$ = 0.08). Thus, the null hypothesis may be refuted for the former, but not for the latter. What this means is that verbs of success have a much stronger preference for the dative-before-nominative order than verbs of failure, whose frequencies are much more evenly distributed across word order patterns.

It remains to be investigated whether these differences may indeed be attributed to a semantic effect or whether there are other factors (e.g. definiteness, animacy, or length) that are perhaps unevenly distributed across the dataset, thus steering these frequencies into a particular direction.

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

Table 3.4 presents a radically different picture from Tables 3.2 and 3.3. When both arguments are pronouns, linearization almost uniquely tends towards the nominative-before-dative order: in total, 172 tokens (or 91%) instantiate the nominative-before-dative order, as opposed to a mere 16 tokens (or 9%) instantiating the dative-before-nominative order.

Table 3.4 Word order distributions in the [Pro-V-Pro] configuration

	Nom-Dat		**Dat-Nom**	
Verb	**N**	**f**	**N**	**f**
gelingen	53	96%	2	4%
glücken	30	94%	2	6%
missglücken	19	76%	6	24%
misslingen	54	92%	5	8%
missraten	16	94%	1	6%
Total	172	91%	16	9%

Two example sentences containing double pronouns are given in (7a–b):

(7) a. *Das* *gelingt* *ihm* *problemlos.*
that.NOM succeeds him.DAT effortlessly
'He succeeds effortlessly.'

(7) b. *Ihr* *missglückte* *alles.*
her.DAT failed everything.NOM
'She failed at everything.'

In light of the hypotheses formulated in Section 3.3, the obtained findings seem puzzling. However, a synchronic comparison with Icelandic shows that this is not so unexpected after all. As mentioned earlier (cf. Section 3.2.2), scholars of Icelandic posit the existence of a class of 'alternating Dat-Nom verbs'. These are a group of two-place predicates which allow both the dative as well as the nominative to take on the role of subject, albeit not at the same time, of course. As a consequence, both arguments are equally likely to occur in clause-initial position. An exploratory corpus study by Somers & Barðdal (2022) shows that four out of five such verbs indeed alternate quite freely between the nominative-dative and the dative-nominative order in the double-NP configuration while in the double-pronoun configuration, the variation in word order is almost entirely neutralised in favour of the nominative-before-dative pattern (see also Allen 1995: 107–109 for similar findings for alternating verbs in Old English). The results in Table 3.4 therefore lend further (albeit indirect) support to Hypothesis 2.

3.5.4 Competition between nominative NPs and dative pronouns

When a nominative NP enters into competition with a dative pronoun, an interesting picture emerges, which is laid out in Table 3.5. The general tendency is for the dative pronoun to take postverbal position (126 tokens in total, or 74%), although 45 tokens (or 26%) still prefer the pronoun to occupy the preverbal slot. Two relevant examples, one for each word order pattern, are given in (8a–b):

(8) a. *Ihr Auftrag* *misslingt* *ihr* *jedoch* *gründlich.*
her.NOM mission fails her.DAT however completely
'However, her mission fails completely.'

(8) b. *Mir* *missglückten* *zwei weitere Versuche* *knapp.*
me.DAT failed two further.NOM attempts.NOM narrowly
'I narrowly failed at two further attempts.'

Thus, the variation in word order previously observed in the double-NP configuration remains largely intact, even though it is considerably more skewed towards the nominative-before-dative order under the present circumstances.

Table 3.5 Word order distributions for combinations of nominative NPs and dative pronouns

	Nom_{NP}-Dat_{PRO}		Dat_{PRO}-Nom_{NP}	
Verb	**N**	**f**	**N**	**f**
gelingen	4	57%	3	43%
glücken	14	64%	8	36%
missglücken	25	69%	11	31%
misslingen	35	76%	11	24%
missraten	48	80%	12	20%
Total	126	74%	45	26%

The findings presented in Table 3.5 constitute the mirror image of what Bader (2020) finds in his investigation of word order distributions in German embedded clauses (i.e. middlefield contexts) with nominative full NPs and dative pronouns. Recall that the dative pronoun precedes the nominative full NP in 91% of cases with unaccusative verbs in his study. However, the German middlefield is prone to fronting object pronouns anyway. Thus, once again, the topological field is shown to exert a major influence on word order distributions.

3.5.5 Competition between dative NPs and nominative pronouns

Table 3.6 shows that dative full NPs generally precede nominative pronouns (80 tokens, or 65%), but nominative pronouns still precede dative NPs with reasonable frequency (44 tokens, or 35%). Again, the variation in word order is largely preserved, but now it is skewed more towards the dative-before-nominative order. The intra-class variation is also rather substantial: some verbs, like *gelingen*, mostly tend towards the dative-before-nominative order, whereas others, like *missglücken*, overwhelmingly prefer the nominative-before-dative order.

Table 3.6 Word order distributions for combinations of dative NPs and nominative pronouns

	Nom_{PRO}-Dat_{NP}		Dat_{NP}-Nom_{PRO}	
Verb	**N**	**f**	**N**	**f**
gelingen	17	22%	61	78%
glücken	12	60%	8	40%
missglücken	7	100%	0	0%
misslingen	5	36%	9	64%
missraten	3	60%	2	40%
Total	44	35%	80	65%

One factor with considerable explanatory power is the referential status of the personal pronoun *es* 'it'. Out of a total of 124 observations in the present configuration, 77 tokens contain *es* as a clause-anticipating pronoun. Of these, 13 clause-anticipating pronouns (or 17%) occur preverbally, but 64 (or 83%) occupy the postverbal slot. An example of such a postverbal clause-anticipating pronoun is given in (9):

(9) *Den Bewohnern* *gelang* *es,* *die Kleinstadt*
the.DAT inhabitants.DAT succeeded it.NOM the small.town
gegen einen übermächtigen Gegner *zu verteidigen.*
against a far.superior opponent to defend
'The inhabitants succeeded in defending the small town against a far superior opponent.'

Thus, in competition with a dative NP, clause-anticipating pronouns overwhelmingly tend to follow the dative NP. The logic behind this behaviour no doubt lies in the impoverished semantic status of the correlative pronoun, which merely functions as a placeholder for a subclause, rather than as a referential element. Still, it is remarkable that this inherently light, pronominal constituent so consistently gives way to a full NP.

3.6 Conclusions and outlook

In this paper, I have presented corpus-based support for the claim that German verbs of success and failure in verb-second clauses are associated with two base orders: a dative-before-nominative order and a nominative-before-dative order. Word order variation has been shown to be greatest in the double-NP configuration and smallest in the double-pronoun configuration, as the latter almost uniquely tends towards the nominative-before-dative order. However, a comparison with Icelandic shows that such a swing towards nominative-first may be an epiphenomenon of pronominality.

The main question prompted by this study is how to further interpret the alternating numbers that have been found. A multivariate analysis factoring in all variables mentioned in Section 3.4.3 might aid in uncovering some tendencies that currently remain hidden. One factor that is particularly promising is definiteness, as it may serve as a proxy for topicality. As argued by Barðdal (1999, 2001b) and Barðdal, Eythórsson & Dewey (2014, 2019), it is the discourse-prominence of either the dative or the nominative that steers the word order variation. A comparison with other unaccusative verbs as well as with other verbs generally displaying a thematic asymmetry between the dative and the nominative would also be instructive.

Another question is how topological fields may influence word order preferences. Kurz's (2000) numbers for *gelingen* indicate that double NPs show an almost absolute

tendency towards a dative-before-nominative order if both nominal arguments are placed in the middlefield. A comparison of word order patterns in contexts where both arguments flank the finite verb with contexts in which both arguments are in the middlefield for the remaining four verbs in this study only including double NPs might help shed light on this issue.

References

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

Bader, Markus. 2020. The position of object pronouns in the German middlefield. *Linguistics* 58(4): 1059–1115.

Bader, Markus & Häussler, Jana. 2010. Word order in German: A corpus study. *Lingua* 120: 717–762.

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

Barðdal, Jóhanna. 2001a. *Case in Icelandic: A Synchronic, Diachronic and Comparative Approach*. Lund: Department of Scandinavian Languages, Lund University.

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

Barðdal, Jóhanna. 2004. The semantics of the impersonal construction in Icelandic, German and Faroese: Beyond thematic roles. In Werner Abraham (ed.), *Focus on Germanic Typology*, 105–137. Berlin: Akademie Verlag.

Barðdal, Jóhanna. 2008. *Productivity: Evidence from Case and Argument Structure in Icelandic*. Amsterdam: John Benjamins.

Barðdal, Jóhanna. (2012). Predicting the productivity of argument structure constructions. *Berkeley Linguistics Society* 32(2006): 467–478.

Barðdal, Jóhanna. 2023. *Oblique Subjects in Germanic: Their Status, History and Reconstruction*. Berlin: Mouton de Gruyter.

Barðdal, Jóhanna & Thórhallur Eyþórsson. 2020. How to identify cognates in syntax: Taking Watkins' legacy one step further. In Jóhanna Barðdal, Spike Gildea & Eugenio R. Luján (eds.), *Reconstructing syntax*, 197–238. Leiden: Brill.

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.

Belletti, Adriana & Rizzi, Luigi. 1988. Psych-verbs and θ–theory. *Natural Language and Linguistic Theory 6*: 291–352.

Duden = *Die Grammatik*. (2016). Berlin: Dudenverlag.

Eisenberg, Peter. 2013. *Grundriss der deutschen Grammatik. 2: Der Satz*. Stuttgart: Metzler.

Eythórsson, Thórhallur & Jóhanna Barðdal. 2005. Oblique subjects: A common Germanic inheritance. *Language* 81(4): 824–881.

Fanselow, Gisbert. 2002. Quirky subjects and other specifiers. In Ingrid Kaufmann & Barbara Stiebels (eds.), *More than Words: A Festschrift for Dieter Wunderlich*, 227–250. Berlin: Akademie Verlag.

Fanselow, Gisbert. 2003. Zur Generierung der Abfolge der Satzglieder im Deutschen. *Neue Beiträge zur Germanistik* 112: 9–47.

Haider, Hubert. 2010. *The Syntax of German*. Cambridge: Cambridge University Press.

Haspelmath, Martin. 2001. Non-canonical marking of core arguments in European languages. In Alexandra Aikhenvald, Robert M. W. Dixon & Masayuki Onishi (eds.), *Non-Canonical Subjects and Objects*, 53–83. Amsterdam: John Benjamins.

Heylen, Kris. 2005. *Zur Abfolge (pro)nominaler Satzglieder im Deutschen: Eine korpusbasierte Analyse der relativen Abfolge von nominalen Subjekt und pronominalem Objekt im Mittelfeld*. Ph.D. dissertation, Katholieke Universiteit Leuven, Leuven.

Hoberg, Ursula. 1981. *Die Wortstellung in der geschriebenen deutschen Gegenwartssprache*. München: Hüber.

Johnson, Cynthia A., Peter Alexander Kerkhof, Leonid Kulikov, Esther Le Mair & Jóhanna Barðdal. 2019. Argument structure, conceptual metaphor and semantic change: How to succeed in Indo-European without really trying. *Diachronica* 36(4): 463–508.

Kilgarriff, Adam, Pavel Rychlý, Pavel Smrž & David Tugwell. 2004. The Sketch Engine. *Proceedings of the 11th EURALEX International Congress*. 105–116.

Kilgarriff, Adam, Vít Baisa, Jan Bušta, Miloš Jakubíček, Vojtěch Kovář, Jan Michelfeit, Pavel Rychlý & Vít Suchomel. 2014. The Sketch Engine: Ten years on. *Lexicography* 1: 7–36.

Kurz, Daniela. 2000. A statistical account on word order variation in German. In Anne Abeillé, Thorsten Brants & Hans Uszkoreit (eds.), *Proceedings of the COLING-2000 Workshop on Linguistically Interpreted Corpora*, 35–41. Luxemburg: Centre Universitaire.

Lenerz, Jürgen. 1977. *Zur Abfolge nominaler Satzglieder im Deutschen*. Tübingen: Gunter Narr.

Primus, Beatrice. 1999. *Cases and Thematic Roles: Ergative, Accusative and Active*. Tübingen: Max Niemayer.

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.

Shannon, Thomas F. 2000. On the order of (pro)nominal arguments in Dutch and German. In Thomas F. Shannon & Johan P. Snapper (eds.), *The Berkeley Conference on Dutch Linguistics 1997: The Dutch Language at the Millennium*, 145–196. New York: University Press of America.

Somers, Joren. 2021. Onpersoonlijke datief-nominatiefwerkwoorden in het hedendaagse Duits: Definitie, inventaris en semantische classificatie. *Handelingen – Koninklijke Zuid-Nederlandse Maatschappij Voor Taal- en Letterkunde en Geschiedenis (KZM)* 75: 211–238.

Somers, Joren & Barðdal, Jóhanna. 2022. Alternating Dat-Nom/Nom-Dat verbs in Icelandic: An exploratory corpus-based analysis. *Working Papers in Scandinavian Syntax* 107: 83–110.

Verhoeven, Elisabeth. 2015. Thematic asymmetries do matter! A corpus study of German word order. *Journal of Germanic Linguistics* 27(1): 45–104.

Chapter 4
Argument structure constructions in competition: The Dat-Nom/Nom-Dat alternation in Icelandic

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.

4.1 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; 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 4.2 below).

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, 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 Section 4.4.3 and 4.5.3 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 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 that 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 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 4.2 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 4.3 gives an overview of the methodology applied, whereas Section 4.4 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 4.5, we single out a set of four alternating verbs, whose behaviour we describe using a logistic regression model. Section 4.6 summarises the main content and conclusions of the article.

4.2 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 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 either refer to them 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 and 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.

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 topicalizing 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'. 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:

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.[1]

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) below:

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).

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

[1] 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.

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 it 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 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.

Another study worth mentioning in this respect is 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 Callegari & Ingason's. The main difference between the two versions is the size of the corpus, with the more recent version containing one third 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 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.[2] 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 & Ingasson's (2021) ditransitives. For that reason, we map out word order preferences for both the *hjálpa-*

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

and the *líka*-classes and use these counts as a baseline against which word order preferences for the *nægja*-class will be measured. We now turn to a description of our methodology, before we present our findings in Sections 4.4–4.5 below.

4.3 Methodology

This study is based on 15 simple verbs that fall into one of three categories: (i) ordinary Nom-Dat verbs (the *hjálþa*-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:

1) Ordinary Nom-Dat verbs: *hjálþa* 'help', *líkjast* 'resemble', *mótmæla* 'contradict', *treysta* 'trust' and *þakka* 'thank'
2) 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/…'
3) Alternating Dat-Nom/Nom-Dat verbs: *duga* 'suffice, be enough', *dyljast* 'be hidden to sby, 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 sby, 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; 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 4.3 and 5.3 below). For each variable, the relevant values are rendered in boldface, followed by examples in brackets. All nine variables are discussed below:

1) 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).
2) (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).
3) 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 post-verbal slot. In line with Heylen (2005:103), conjoined pronouns are also excluded, as they arguably lose their pronominal status.
4) 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.

5) 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).
6) 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).
7) 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[3] (*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), places (*Keflavík* nom.sg), and population groups (*Reykvíkingum* 'the people of Reykjavík' dat.pl) are inherently definite. In case 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 that is the label that captures the definiteness status of the first conjunct (*4x4 klúbburinn* 'the 4x4 club'), which is definite.
8) 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[4] (*myglusveppum* 'mould' dat.pl), and dead animals, are labelled 'inanimate'. In case animacy cannot unequivocally be determined, we have resorted to the label 'non-inferable'. This, for instance, applies to *lánatakanda* '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. In case

[3] 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 teachers' book'.

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

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'), which indeed refers to individuals.

9) Length: constituent weight measured in **words**.

We now turn to our findings and a discussion thereof.

4.4 Results and discussion

The current section details the results of our study. Section 4.4.1 establishes a baseline by mapping out word order preferences for both ordinary Nom-Dat verbs, i.e. verbs of the *hjálþa*-type, and non-alternating Dat-Nom verbs, or *líka*-verbs. In Section 4.4.2 we compare the statistics for alternating *nægja*-verbs with the baseline established for *hjálþa*- and *líka*-verbs in Icelandic. Section 4.4.3 discusses the main implications and conclusions.

4.4.1 Establishing a baseline: *hjálþa*- and *líka*-verbs

In this section, we discuss our findings for both *hjálþa*- and *líka*-verbs. We single out two configurations, i.e. contexts in which both arguments are full NPs (Section 4.4.1.1) and contexts in which both arguments are pronouns (Section 4.4.1.2). 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 4.4.1.3.

4.4.1.1 Word order in the [NP-V-NP] configuration

Table 4.1 presents an overview of the word order distributions for both *hjálþa*- and *líka*-verbs in clauses where both arguments are full NPs.

Table 4.1 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**
hjálþa	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%

Starting with *hjálþa*-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 *hjálþa* 'help' and one with *líkjast* 'resemble', are given in (8a–b) below:

Nom-Dat

(8) a. *Listamaður* *hjálþaði* *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 *hjálþa*-verb that (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*
this.DAT intended.DAT procedure.DAT opposed
ýmsir þingmenn …
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 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 *hjálþa*-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 4.1, 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 *hjálþa*- and *líka*-verbs in the double-NP configuration can be considered to be very robust.

Finally, both *hjálþa*- and *líka*-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 either the Nom-Dat or the Dat-Nom order in equal manner.

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

Table 4.2 summarises the results for *hjálþa*- and *líka*-verbs in the [Pro-V-Pro] configuration.

Table 4.2 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**
hjálþa	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%

For *hjálþa*-verbs, word order preferences in the [Pro-V-Pro] configuration constitute a near-perfect copy of the results presented in Table 4.1 above. With the exception of *mótmæla*, all *hjálþa*-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) below:

Topicalisation

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

Topicalisation

(12) b. *En hitt, að þetta hafi verið gjört*
but the.other that this had been done
í fullkomnu óþakklæti skólastjóra, því mótmæli ég
in perfect ingratitude headmaster's that.DAT oppose I.NOM
algjörlega.
entirely
'But the other [option], that this was done in a total ingratitude of the head master, 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 *hjálþa*-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* 'that'. This finding is reminiscent of the tendency discussed above for the verb *mótmæla* 'contradict', 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).

4.4.1.3 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.

4.4.2 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 four configurations (4.4.2.1), after which we turn to word order variation in the [NP-V-NP] configuration (4.4.2.2), the [Pro-V-Pro] configuration (4.4.2.3) and finally we discuss configurations where one of the arguments is a pronoun (4.4.2.4). The results are compared to the baseline set by Nom-Dat *hjálpa*-verbs and Dat-Nom *líka*-verbs.

4.4.2.1 General findings

The results for the class of *nægja*-verbs, which are presented in Table 4.3, 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 4.2.

Table 4.3 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%

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, 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 4.1 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.

4.4.2.2 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 4.4.

Table 4.4 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%

The frequencies in Table 4.4 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 both word orders attested with *nægja*-verbs across all five verbs yields a highly significant result with a large effect size (χ^2 = 80.14; df = 4; $p_{two\text{-}tailed}$ < .001; Cramér's V = .61), which should be interpreted as a statistical indication that the distribution of both word orders cannot be attributed to chance. A more in-depth analysis of the factors driving the alternation is presented in Section 4.5.

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) below:

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 …'

Nom-Dat

(14) b. *En skotfærasafnið hans hefði nægt*
but munition.collection.NOM his had sufficed
hverri meðal herdeild.
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 4.3 are evidently replicated in the [NP-V-NP] configuration in Table 4.4, 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 4.2, 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 4.3. 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 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 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.

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

Table 4.5 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.

Table 4.5 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%

Some examples of Dat-Nom word orders involving pronouns are given in (15a–c) below, 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 be.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 be.aware it.NOM not that
'She was not aware that …'

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 4.5 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 biassed 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 4.5). 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 (3 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).

4.4.2.4 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 4.6.

Table 4.6 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 4.6 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 4.6 are highly reminiscent of our findings for alternating verbs in the double-pronoun configuration (cf. Section 4.4.2.3 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 beg 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 4.6 is a general effect of pronouns, not specific for *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%). 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 for *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 4.7.

Table 4.7 Alternating verbs in configurations with a dative pronoun and a nominative full NP

	$\textbf{Nom}_{\textbf{NP}}\textbf{-Dat}_{\textbf{PRO}}$		$\textbf{Dat}_{\textbf{PRO}}\textbf{-Nom}_{\textbf{NP}}$	
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%

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 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.

4.4.3 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 *hjálþa*- 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 4.8.

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 4.8 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.

Table 4.8 Proportional prevalence of the Nom-Dat vs. Dat-Nom linear order in the [NP-V-NP] configuration for *hjálþa*-, *líka*- and *nægja*-verbs, and for *nægja*-verbs excluding *henta*

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

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 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.

4.5 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 2013; 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 4.3, 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 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 e.g. medicine 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 4.5.1 presents the results of the logistic regression analysis modelling all the data obtained across configurations, whereas Section 4.5.2 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 4.5.3 we discuss how our findings tie in with the concept of topicality.

4.5.1 Across configurations

The output of the first logistic regression model, which builds on all 200 observations per verb type, is presented in Table 4.9. 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 4.9 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**

Table 4.9 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 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 4.4.2.1 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 4.5.2 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 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 considerably smaller (a mere 250 across argument positions). One reason why local pronouns are so rare is because these are never nominative in our dataset (see however, 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 of taking first position anyway.

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.'					

Intended neutral Dat-Nom word order

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

Recall that in 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%). 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, 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.

4.5.2 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 4.10. 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.

Table 4.10 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

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 do not appear to have any predictive power anymore 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 (1) whether a different sample would equally single out nominative case marking as a significant predictor, and (2) whether other alternating verbs are equally sensitive to the effect of nominative case marking.

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.

4.5.3 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 4.3 and 4.4.3 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, like, 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 anyway (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:

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.

4.6 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 *hjálþa* '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, because 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 *hjálþa*-verbs, and non-alternating Dat-Nom verbs, or *líka*-verbs, in configurations with two full NPs. It turns out that both these verb classes, *hjálþa*- 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 *hjálþa*- 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 *hjálþa*- 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 *hjálþa*- 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 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 like 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.

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. *North Eastern Linguistic Society* 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: 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: 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. 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]. Ms., University of Iceland.
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.
Cristofaro, Sonia. 2019. Taking diachronic evidence seriously: Result-oriented vs. source-oriented explanations of typological universals. In Karsten Schmidtke-Bode, Natalia Levshina, Susanne 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*. 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 Ph.D. 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, N., S. Kim, K. He & J.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 Ph.D. 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, Ph.D. Dissertation.
Jónsson, Jóhannes Gísli. 1997–1998. 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 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: 103–116.
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.
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.
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 Ph.D. 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 Ph.D. 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, Elizabeth. 2009. Subjects, agents, experiencers, and animates in competition: Modern Greek argument order. *Linguistische Berichte* 2019: 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): 1–10.
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.

Chapter 5
Subjecthood and argument structure of synonymous Dat-Nom/Nom-Dat verbs across German and Icelandic

This paper investigates a set of 15 Icelandic verbs licensing both a nominative and a dative argument, recently analysed in the literature, comparing them with a corresponding set of 15 German verbs. The Icelandic dataset consists of verbs selecting for three different argument structures: a) ordinary Nom-Dat verbs, b) non-alternating Dat-Nom verbs and c) alternating Dat-Nom/Nom-Dat verbs. The German dataset contains either (near-)synonyms or cognates to the Icelandic verbs. One of our most important findings is that apparent Dat-Nom verbs in German, like *gefallen* 'please, like', are in fact alternating Dat-Nom/Nom-Dat verbs in that language. That is, these verbs can either instantiate the Dat-Nom or the Nom-Dat argument structure. This conclusion is supported by word order counts, which show a major difference between alternating Dat-Nom/Nom-Dat verbs and ordinary Nom-Dat verbs across both German and Icelandic. We also examine the set of alternating verbs across both languages statistically using a conditional inference tree. The differences between Icelandic and German especially play out in configurations involving pronouns, whereas, for double NPs, we identify the same additional factors as have already been identified for the monolingual (Icelandic) dataset, namely animacy, length and (in)definiteness.

5.1 Introduction

It is by now well known in the scholarship on argument structure in the Germanic languages that there are verbs that select for two arguments, dative and a nominative, which alternate between two diametrically opposed argument structure constructions without any change in the assignment of semantic roles. In other words, the semantic roles are constant across the morphological cases, with an experiencer in the dative case and a stimulus in the nominative case. Such alternating structures have been discussed by Barnes (1986) and Barðdal (2023: 94–97) for Faroese, Allen (1995: 107–109, 114–117) and Barðdal (2023: 97–99, 125–133) for Old/Middle English, Barðdal (1998, 2023: 99–112) for Old Swedish and Old Danish, Lenerz (1977: 112–116), Primus (1994, 2012), Haspelmath (2001), Eythórsson & Barðdal (2005), Barðdal, Eythórsson & Dewey (2014, 2019), Rott (2016), Barðdal (2023: Ch. 6), Somers (2023) and Somers et al. (2024) for Modern German, and last but not least, Bernódusson (1982), Sigurðsson (1991), Jónsson (1997–98), Barðdal (1998, 1999, 2001, 2023), Platzack (1999), Wunderlich (2009), Rott (2013, 2016), Barðdal, Eythórsson & Dewey (2014, 2019), Wood & Sigurðsson (2014), Somers & Barðdal (2022) and Somers, Jenset & Barðdal (2024) for Modern Icelandic.

Examples of this type are shown in (1–3) below for Icelandic, Faroese and German, the three modern Germanic languages still exhibiting such alternating structures, with *falla vel* 'like, be to sb's liking' in Icelandic (1) and Faroese (2) and their German cognate *gefallen* with the same meaning (3):

Modern Icelandic

Dat-Nom

(1)	a.	***Ungu fólki***	*hefur*	*ávallt*	*fallið*	***þessi tíska***
		young.DAT people.DAT	has.3SG	always	fallen	this.NOM fashion.NOM
		vel.				
		well				
		'Young people have always liked this fashion.'				

Nom-Dat

(1)	b.	***Þessi tíska***	*hefur*	*ávallt*	*fallið*
		this.NOM fashion.NOM	has.3SG	always	fallen
		ungu fólki	*vel.*		
		young.DAT people.DAT	well		
		'This fashion has always been to the liking of young people.'			

Modern Faroese

Dat-Nom (here realised as Dat-Acc)

(2) a. ***Føroyingum*** *hevur* *altíð* *fallið*
Faroese.people.DAT has.3SG always fallen
fermenteraðan mat *væl.*
fermented.ACC food.ACC well
'The Faroese people have always liked fermented food.'

Nom-Dat

(2) b. ***Tað fermenteraða*** *hevur* *altíð* *fallið*
the fermented.NOM has.3SG always fallen
føroyingum *væl.*
Faroese.people.DAT well
'Fermented food has always been to the Faroese people's liking.'

Modern German

Dat-Nom

(3) a. ***Den Kunden*** *haben* ***die Autos*** *nicht* *gefallen.*
the.DAT customers.DAT have.3PL the.NOM cars not liked
'The customers did not like the cars.'

Nom-Dat

(3) b. ***Die Autos*** *haben* ***den Kunden*** *nicht* *gefallen.*
the.NOM cars have.3PL the.DAT customers.DAT not liked
'The cars were not to the customers' liking.'

Observe that the nominative in the Dat-Nom construction in Faroese has changed into accusative during the course of time, but this has only taken place in the Dat-Nom construction and not in the Nom-Dat construction; there, the nominative in first position is intact. That nominative objects change to accusative objects is a well-known change in the syntax and argument structure of Modern Faroese (cf. Barnes 1986; Petersen 2002; Thráinsson et al. 2012: 229, 314; Barðdal 2023: 94–97).

For Modern Icelandic and Faroese there is a consensus in the literature that it is the dative that is the syntactic subject in Dat-Nom constructions, while the nominative is the subject in Nom-Dat constructions. This has been shown with a range of syntactic tests teasing apart subjects from objects (Barnes 1986; Barðdal 1999, 2001, 2023: 81–97; Rott 2016; Barðdal, Eythórsson & Dewey 2014, 2019). We have identified alternating predicates on the basis of the acceptability of the two word orders in the sample pairs above; that is, native speakers have evaluated both word orders as equally neutral.

For German, however, there is no such consensus in the field. The traditional view claims that the relevant verbs are Nom-Dat verbs with the Dat-Nom surface order turning up due to a topicalisation of the dative to initial position (Haspelmath 2001; Kempen & Harbusch 2005; Bader & Häussler 2010; Verhoeven 2015, among many others). According

to the most recent approach, these verbs are indeed assumed to be Dat-Nom verbs, in the sense that the dative is the first argument of the argument structure and the nominative the second argument (Bayer 2004; Haider 2005; Schlesewsky & Bornkessel 2006; Wunderlich 2009). Despite acknowledging the dative as the first argument of the argument structure, the reasoning is still that the nominative, i.e. the second argument of the argument structure, is the subject in such structures in German (Bayer 2004: 70; Haider 2005: 23–24; Wunderlich 2009: 592). Contra these approaches, it has been argued by Eythórsson & Barðdal (2005), Barðdal, Eythórsson & Dewey (2014, 2019), Rott (2016), Barðdal (2023: Ch. 6), Somers (2023) and Somers & Barðdal (2023) that the relevant German verbs, corresponding to the Icelandic and Faroese verbs above, also alternate between two diametrically opposed argument structure constructions. The analysis is based on the following subject tests:

- First position in declarative clauses
- Subject-verb inversion
- First position in subordinate clauses
- Conjunction reduction
- Clause-bound reflexivization
- Raising-to-subject
- Raising-to-object
- Control infinitives

For an overview of how the two arguments, the dative and the nominative, fare with regard to the subject tests in both Icelandic and German, see the references cited immediately above. For the most recent examples of either argument of alternating verbs in German, the dative or the nominative, passing the subject tests listed above, see Barðdal, Eythórsson & Dewey (2014, 2019), Rott (2016: 239–249) and Barðdal (2023: Ch. 6).

In addition, the two linear word orders in (1–3) above appear to be equally neutral in the sense that speakers do not view either one as being an instance of contrastive focus or topicalisation in any of the three languages. Therefore, given the apparent systematic alternation between these two argument structure constructions, Dat-Nom and Nom-Dat, the question arises which factors decide on why speakers choose to use one of the constructions over the other. That is, when do speakers use the Dat-Nom construction and when do they prefer to use the Nom-Dat construction in their language? To our knowledge, the only explanation found in the literature so far is that the Dat-Nom construction in Icelandic is used when the dative is the more topical argument, while the Nom-Dat construction is used when the nominative is the more topical argument (cf. Barðdal 1999, 2001; Rott 2013; Barðdal, Eythórsson & Dewey 2014, 2019; Somers, Jenset & Barðdal 2024). No explanation for the choices of German speakers exists in the literature, most likely because these have been viewed as non-alternating Dat-Nom verbs in that language.

In this article, we focus on German data of the type shown in (3a–b) above, comparing it with corresponding data from Icelandic, with the aim to a) document statistically the word order variation shown above for a set of candidate verbs, and b) uncover the factors motivating the choice of the two word orders, dative-before-nominative and nominative-before-dative, by analysing an extracted set of corpus data involving the relevant candidate verbs. In Section 2, we describe our methodology, including how the dataset was extracted, cleaned and annotated. In Section 3, we introduce the three verb classes, established by Somers & Barðdal (2022) for Icelandic, namely i) ordinary Nom-Dat verbs, ii) non-alternating Dat-Nom verbs and iii) alternating Dat-Nom/Nom-Dat verbs. We then compare word order preferences of these Icelandic verbs with their German counterparts by scrutinising both configurations involving two full NPs as well as configurations involving at least one pronoun. We also discuss word order preferences when the two arguments are referential personal pronouns. We show that the factors we documented for Icelandic in Somers & Barðdal (2022) and Somers, Jenset & Barðdal (2024) are also valid for German, namely animacy, length and (in)definiteness. Section 4 zooms in on the set of alternating verbs across both Icelandic and German and analyses them statistically with the help of a conditional inference tree, again teasing apart configurations involving at least one pronoun and configurations in which neither argument is pronominal. Section 5 summarises and concludes our discussion.

5.2 Methodology

5.2.1 Dataset

The point of departure for this study is a set of fifteen Icelandic verbs first explored by Somers & Barðdal (2022) and Somers, Jenset & Barðdal (2024). These verbs crucially divide into one of three classes: (1) ordinary Nom-Dat verbs, (2) non-alternating Dat-Nom verbs, and (3) alternating Dat-Nom/Nom-Dat verbs. The three verb classes comprise five verbs each:

1) Ordinary Nom-Dat verbs: *hjálpa* ‘help’, *líkjast* ‘resemble’, *mótmæla* ‘contradict’, *treysta* ‘trust’, and *þakka* ‘thank’
2) Non-alternating Dat-Nom verbs: *áskotnast* ‘receive’, *blöskra* ‘be shocked, be horrified’, *leiðast* ‘be bored’, *líka* ‘like’, and *þykja* ‘think, find, seem’
3) Alternating Dat-Nom/Nom-Dat verbs: *duga* ‘suffice, be enough’, *dyljast* ‘be hidden to sby, be aware’, *endast* ‘last’, *henta* ‘suit, befit’; and *nægja* ‘be enough, be sufficient’

The Icelandic Dat-Nom verbs in Somers & Barðdal (2022) and Somers, Jenset & Barðdal (2024), alternating and non-alternating verbs alike, were randomly selected on the basis of information in the earlier literature (Bernódusson 1982; Jónsson 1997–1998; Barðdal 1999, 2001, 2023; Rott 2013). As a matter of fact, Rott (2013) provides a comparison between a modest set of four non-alternating Dat-Nom verbs and four alternating Dat-Nom/Nom-Dat verbs in Icelandic, although using a very small corpus. Given the availability of gargantuan corpora today, a larger-scale study is certainly called for.

In addition, we wish to expand upon existing research by introducing a second control set of ordinary Nom-Dat verbs, in order to provide a baseline against which to compare both of the aforementioned sets, i.e. both Dat-Nom and Dat-Nom/Nom-Dat verbs. This, in essence, entails establishing how frequent object topicalisation is with ordinary Nom-Dat verbs yielding a dative-before-nominative order, which may then serve as a touchstone for evaluating the word order distributions found with either of the two remaining verb classes.

The current study takes the analyses by Somers & Barðdal (2022) and Somers, Jenset & Barðdal (2024) one step further by matching each of the fifteen Icelandic verbs to a German cognate or (near-)synonym, which allows us to adopt a cross-Germanic approach to a set of syntactically and semantically highly similar verbs. The analysis is confined to instances of the Dat-Nom and Nom-Dat argument structure constructions and does not involve any other potential argument structure constructions the relevant 30 verbs might otherwise instantiate in their respective languages, as this lies beyond the scope of the present study.

For the selection of the German verbs, we bank on Somers' (2021) study of Dat-Nom verbs in Present-Day German, which takes stock of 146 verbs that licence a subject-like dative and an object-like nominative in one or more of their senses. From Somers' dataset, we have selected the best semantic and/or etymological fits to the Icelandic verbs cited above. Table 1 reiterates the Icelandic verbs that served as our starting point (column 1), their respective argument structures (column 2), the German verbs matched to the Icelandic input verbs (column 3), a gloss for each German verb (column 4), and a description of the relation holding between each cross-linguistic pair, i.e. cognate or synonym (column 5).

Table 5.1 Icelandic source verbs, their respective argument structures, their German correspondences and the criterion according to which the German verbs were matched to the Icelandic types

Icelandic verb	**Argument structure**	**German match**	**Gloss**	**Matching criterion**
hjálþa	Nom-Dat	*helfen*	'help'	*cognate*
líkjast	Nom-Dat	*ähneln*	'resemble'	*synonym*
mótmæla	Nom-Dat	*widersprechen*	'contradict'	*synonym*
treysta	Nom-Dat	*vertrauen*	'trust'	*cognate*
þakka	Nom-Dat	*danken*	'thank'	*cognate*
áskotnast	Dat-Nom	*zufallen*	'fall to, receive (lit.); fall to, receive (fig.)'	*synonym*
blöskra	Dat-Nom	*grauen*	'dread, be afraid of'	*synonym*
leiðast	Dat-Nom	*leidtun*	'take pity; be sorry'	*cognate*
líka	Dat-Nom	*gefallen*	'please, like'	*synonym*
þykja	Dat-Nom	*dünken*	'seem, appear'	*cognate*
duga	Dat-Nom/Nom-Dat	*nützen*	'be of use'	*synonym*
dyljast	Dat-Nom/Nom-Dat	*entgehen*	'miss out on; fail to notice'	*synonym*
endast	Dat-Nom/Nom-Dat	*reichen*	'suffice'	*synonym*
henta	Dat-Nom/Nom-Dat	*geziemen*	'befit'	*synonym*
nægja	Dat-Nom/Nom-Dat	*genügen*	'be enough, be sufficient'	*cognate*

The German types have been selected as follows. First, an etymological link between an Icelandic verb and a German candidate verb has been established, as cognate verbs across related languages may inherit their argument structures from a shared proto-stage. In six out of fifteen cases, we managed to garner German cognates with the same or a very similar meaning and the same case frame as the Icelandic input verbs: *hjálþa* and *helfen*, *treysta* and *vertrauen*, *þakka* and *danken*, *leiðast* and *leidtun*, *þykja* and *dünken*, and *nægja* and *genügen*. This means that the remaining nine German verbs only share a semantic link, and the same case frame, with their Icelandic counterparts and no etymological link. The semantic pairs are *líkjast* and *ähneln*, *mótmæla* and *widersprechen*, *áskotnast* and *zufallen*, *blöskra* and *grauen*, *líka and gefallen*, *duga* and *nützen*, *dyljast* and *entgehen*, *endast* and *reichen*, and *henta* and *geziemen*.

Verbs exhibiting only one sense within their Dat-Nom/Nom-Dat use have been prioritised over verbs displaying multiple such senses, in order to minimise a potential effect of different verb senses. In fact, only two of the German verb types turn out to have more than one Dat-Nom/Nom-Dat sense, i.e. *entgehen* and *zufallen*. As for *entgehen*, it can mean both 'fail to notice', which is congruent with Ice. *dyljast*, as well as 'miss out on', which is incongruent with Ice. *dyljast*. Likewise, *zufallen* can mean both 'fall to, receive' in its literal sense, as well as 'fall to, receive' in its metaphorical sense. Only the former dovetails with Ice. *áskotnast*, but not the latter. We are aware that this creates a slight incongruity in the dataset, but since both senses of each of these two verbs, *zufallen* and

entgehen, are potentially alternating, we believe that the benefits outweigh the disadvantages. Future research will have to determine whether the different Dat-Nom/Nom-Dat senses with *zufallen* and *entgehen* alternate to different degrees.

The decision to include a verb for study, Icelandic or German, also depends on the frequency with which it occurs in either the nominative-before-dative or the dative-before-nominative word order pattern in the corpus we employ (cf. below). More specifically, if the first 300 randomised tokens for any given verb yield fewer than seven eligible tokens, the verb was excluded. That is the reason why certain types, like *passen* 'suit', have not been examined any further here.

In order to ensure maximal comparability with Somers & Barðdal's (2022) and Somers, Jenset & Barðal's (2024) dataset, the German data were retrieved from the German Web 2013 corpus, also referred to as deTenTen13 (Killgarriff et al. 2004, 2014). This is the German counterpart of the Icelandic Web 2020 corpus, also referred to as isTenTen20 (Jakubíček et al. 2013), which was used for extracting the Icelandic dataset. The deTenTen13 corpus comprises more than 16.5 billion words and has been accessed through the Sketch Engine interface. For each of the German and Icelandic verbs, we have run lemmatized search queries.

Subsequently, and also in accordance with Somers & Barðdal (2022) and Somers, Jenset & Barðdal (2024), we have downloaded one or more files of 10,000 randomised tokens per verb, depending on how abundant the data are. The first 200 eligible tokens of each verb type have been retained. Hence, the total number of tokens for German equals 3,000, which is identical to the 3,000 tokens Somers & Barðdal (2022) and Somers, Jenset & Barðdal (2024) retrieved for Icelandic using exactly the same method. How exactly the data have been cleaned is accounted for in the next section.

5.2.2 Data cleaning

It has already been mentioned that the data for the present study have been retrieved through lemmatized corpus queries. All tokens have subsequently been cleaned manually. Which tokens have been retained, and which ones were barred from study is outlined in the remainder of the present section.

First, only tokens in which the main verb is flanked by either a nominal or a pronominal element have been included for study. Thus, all eligible tokens correspond to a template of the type [Nom-V-Dat], or [Dat-V-Nom], as opposed to strings in which both (pro)nominal arguments follow the finite verb. The reason for this approach is twofold. First, by restricting our study to prefield structures we avoid mixing different syntactic conditions. At least in German, word order preferences in the so-called middle field may deviate significantly from the canonical order of constituents, especially when pronouns are involved (Hoberg 1981: 62; Primus 1994; Shannon 2000; Kempen & Harbusch 2004;

Heylen 2005; Bader & Häussler 2010; Bader 2020, *inter alia*). Moreover, the templates in question capture a word order pattern that is common to both Icelandic and German. Thus, we ensure maximal comparability between the two datasets, both within languages as well as between languages.

Secondly, both argument slots are required to be filled by either a pronoun or a full NP. Tokens containing clausal constituents have been excluded because they cannot bear morphological case marking, and because clausal arguments are usually considerably longer than (pro)nominal arguments, which, in turn, makes them more prone to occupying the post-verbal position. Somers & Barðdal (2022) as well as Somers, Jenset & Barðdal (2024) point out that the high number of clausal constituents in Rott (2013) is probably the reason why he manages to collect such a high number of Dat-Nom attestations for Icelandic alternating Dat-Nom/Nom-Dat verbs, since 82 out of 94 Dat-Nom attestations for that verb class contain clausal nominatives in his study.

It is also worth mentioning that two German verbs in our dataset, viz. *grauen* 'dread' and *dünken* 'seem, appear', allow for their oblique argument to be realised either in the dative or the accusative. For these verbs, all tokens with accusative obliques have been excluded, as well as tokens with oblique arguments that are structurally ambiguous between accusative and dative case marking, such as personal names, or pronouns that only employ one levelled oblique form, like *uns* 'us' 1p.acc/dat.pl or *euch* 'you' 2p.acc/dat.pl.

Finally, we decided only to include tokens that, at least theoretically, allow for alternation. This means that questions introduced by question words or interrogative phrases have been considered non-eligible (cf. also Bader & Häussler 2010: 717). The same goes for reflexive pronouns, which are amongst the few linguistic units that cannot occupy preverbal position, at least in German (see Duden 2016: 893–894). Following Verhoeven (2015) and Somers et al. (2024), tokens containing elided constituents have equally been excluded.

5.2.3 Data annotation

The annotation process of the present study is modelled on the procedure adopted by Somers, Jenset & Barðdal (2024) for Icelandic as well as Somers et al. (2024) for German. All tokens, Icelandic and German alike, have been annotated for ten variables: language, case, (pro)nominality, pronoun type (if applicable), referentiality, person, number, definiteness, animacy, and length. Each of these is discussed below. All values are illustrated using examples from the current dataset.

1) Language: **Icelandic** or **German**.
2) Case: **nominative** (Icel. *orkan* 'the energy' nom.sg, G. *Schimpansen* 'chimpanzees' nom.pl) or **dative** (Icel. *nemendum* 'students' dat.pl, G. *einem Konzert* 'a concert' dat.sg).

3) (Pro)nominality: **pronoun** (Icel. *hún* 'she' 3p.nom.sg, G. *niemandem* 'nobody' 3p.dat.sg) or full **NP** (Icel. *fullorðnum* 'adults' dat.pl, G. *seiner Mutter* 'his mother' dat.sg).
4) Pronoun type: **personal** (Icel. *okkur* 'us' 1p.dat.pl, G. *ich* 'I' 1p.nom.sg), **demonstrative** (Icel. *þetta* 'that' nom.sg.n, G. *dem* 'this' dat.sg), **indefinite** (Icel. *einhverjum* 'someone' dat.sg, G. *nichts* 'nothing') or **reciprocal** (Icel. *hverjir öðrum* 'each other' dat.pl, G. *einander* 'each other'). Conjoined pronouns have been excluded, as they arguably lose their pronominal status (cf. Heylen 2005: 103).
5) Referentiality: **referential**, **correlate** or **expletive**. The label 'expletive' signals all semantically void third-person pronouns. Coincidentally, these are only attested for German, where they take the form of personal *es* 'it', as in *Manchen graut es regelrecht davor* 'Many people are utterly appalled by that'. Correlates, which are pronouns serving as placeholders for a subclause, are attested for both languages in our dataset. Two relevant examples include Icel. *Mér er farið að leiðast það að allt sem ég geri er litið hornauga* 'I am starting to get annoyed by the fact that everything I do is viewed with suspicion' and G. *Es genügt uns schon vollkommen, einfach bessere Menschen zu sein* 'It is already entirely enough for us to simply be better people'. In line with Siewierska (1993: 831), it is hypothesised that both correlates as well as expletives, given their impoverished semantic status, are inclined to occupy the less prominent postverbal slot, rather than the preverbal slot. Any remaining pronouns, as well as all full NPs, are coded as 'referential'.
6) Person: **first** person (Icel. *ég* 'I' nom.sg, G. *uns* 'us' dat.pl), **second** person (Icel. *ykkur* 'you' dat.pl, G. *dir* 'you' dat.sg) or **third** person (Icel. *hann* 'he' nom.sg , G. *der Vorteil* 'the advantage' nom.sg).
7) Number: **singular** (Icel. *árangurinn* 'the success' nom.sg, G. *meinem Bruder* 'my brother' dat.sg) or **plural** (Icel. *þessar aðgerðir* 'these actions' nom.pl, G. *die Minuten* 'the minutes' nom.pl).
8) Definiteness: **definite** or **indefinite**. Icelandic and German personal and reciprocal pronouns are always definite, whereas indefinite pronouns are consistently indefinite in both languages. German demonstratives (*das* 'that' nom.sg, *diesem* 'this one' dat.sg) are invariably definite, but Icelandic demonstratives may either be definite (*þetta* 'that' nom.sg, *hinni* 'that' dat.sg) or indefinite (*slíkt* 'such a thing' nom.sg). Full NPs, in both languages, are coded for as definite if they contain a definite determiner. These include definite articles (Icel. *strákunum* 'the boys' dat.pl, G. *der Gegner* 'the adversary' nom.sg), possessive pronouns (Icel. *sín þjáning* 'his suffering' nom.sg, G. *Ihrem Unternehmen* 'your business' dat.sg) and (definite) demonstratives (Icel. *þessum konum* 'these women' dat.pl, G. *diese Rolle* 'this role' nom.sg). Attributive genitives also render constituents definite (Icel. *slagkraftur náttúrunnar* 'the power of nature' nom.sg, G. *Obamas Außenminister John Kerry* 'Obama's foreign secretary John Kerry' dat.sg). Proper names are inherently definite (Icel. *Birni* 'Björn' dat.sg, G. *Charlotte* 'id.' nom.sg). Conjoined constituents exhibiting conflicting definiteness are coded for their first conjunct (cf. Somers, Jenset & Barðdal 2024; Somers et al. 2024).
9) Animacy: **individual**, **collective**, **inanimate**, **non-inferable** or **NA**. The label 'individual' captures all constituents referring to animate entities. These include humans (Icel. *Magnúsi* 'Magnus' dat.sg, G. *dem Komponisten* 'the composer' dat.sg), animals (Icel. *hornsíli* 'stickleback' dat.sg, G. *diese Hunde* 'these dogs' nom.pl) and so-called humanoids, as Bresnan & Ford (2010: 175) call them (Icel. *Guð* 'God' nom.sg,

G. *die Sims* 'the Sims' nom.pl. [from a video game]). Collectives are those entities referring to a group of individuals (Icel. *alþýðunni* 'the common people' dat.sg, G. *dem Staat* 'the state' dat.sg). Cases where the referent could either by individual or collective have been labelled 'non-inferable'. The tag 'NA' comprises all third-person pronouns which do not have an extra-linguistic referent, as they are either used expletively or coreferentially with a subclause. All remaining tokens have been considered to be inanimate. Conjoined constituents displaying conflicting animacy are labelled in accordance with the first conjunct (cf. supra for definiteness; see also Somers, Jenset & Barðdal 2024; Somers et al. 2024).

10) Length: measured in **words**.

Note that the term *expletive* is sometimes also used to refer to presentative pronouns in existential constructions, as in G. *Es steht ein Mann vor der Tür* 'There is a man at the door'. However, such pronouns are excluded from the present study, as they are not syntactically mandatory, i.e. they are not a part of the verb's argument structure. This is evident from the fact that they do not show up when the word order is inverted: *Ein Mann steht vor der Tür*. This is true for both German and Icelandic.

5.3 Findings

The current section discusses raw frequencies for *hjálþa-* and *helfen*-verbs (Section 5.3.1), *líka-* and *gefallen*-verbs (Section 5.3.2), and *nægja-* and *genügen*-verbs (Section 5.3.3). Each subsection first examines word order patterns in the double-NP configuration. The figures for Icelandic double NPs have been adopted from Somers & Barðdal (2022) and Somers, Jenset & Barðdal (2024), while the counts for German are new. The discussion of the double-NP configuration is followed by an exploratory analysis of each of the three remaining configurations, which all involve at least one pronoun. In order to facilitate a between-language comparison, all verbs are presented alongside their semantic counterparts.

5.3.1 Verbs of the *hjálþa-* and *helfen*-type

Section 5.3.1 compares Icelandic *hjálþa*-verbs to German *helfen*-verbs. We first zoom in on word order variation in the double-NP configuration (Section 5.3.1.1), after which we shift our focus to configurations involving at least one pronoun (Section 5.3.1.2). Section 5.3.1.3 presents a first set of interim conclusions.

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

The results for *hjálþa-* and *helfen*-verbs in the double-NP configuration are presented in Table 5.2, which allows for two observations. First, all verbs in either language show an overwhelming preference for the nominative-before-dative order, with 99% of Icelandic tokens and 96% of German tokens instantiating a nominative-first structure. Note that the synonymous verbs *mótmæla* and *widersprechen* exhibit remarkably analogous behaviour, as these account for the largest share of topicalisations in Icelandic and German, respectively: as many as two out of two Icelandic dative-before-nominative attestations and ten out of 15 German dative-before-nominative attestations are found with these verbs.

Table 5.2 *Hjálþa-* and *helfen*-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**
hjálþa	25	100%	0	0%	*helfen*	53	96%	2	4%
líkjast	125	100%	0	0%	*ähneln*	132	100%	0	0%
mótmæla	98	98%	2	2%	*widersprechen*	95	90%	10	10%
treysta	31	100%	0	0%	*vertrauen*	46	98%	1	2%
þakka	55	100%	0	0%	*danken*	49	96%	2	4%
Total	334	99%	2	1%	**Total**	375	96%	15	4%

Icelandic

Neutral nominative-before-dative order

(4) a. *Kennarar og nemendur* *mótmæla*
teachers.NOM and students.NOM oppose
breytingum á stundatöflu *í grunnskólum Frakklands.*
changes.DAT to timetable in primary schools France.GEN
'Teachers and students are objecting to timetable changes in primary schools in France.'

Topicalised dative-before-nominative order

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

German

Neutral nominative-before-dative order

(5) a. *Die Grünen* *widersprechen* *dieser Begründung* *vehement.*
the.NOM Greens object.to this.DAT rationale vehemently
'The Greens vehemently object to this rationale.'

Topicalised dative-before-nominative order

(5) b. ***Dieser Aussage*** *widersprachen* *jedoch*
this.DAT statement contradicted.3PL however
Vertreter der Zivilgesellschaft.
representatives.NOM the.GEN civil.society
'This statement, however, representatives of civil society contradicted.'

Secondly, it is striking that the total number of observations in the [NP-V-NP] configuration across languages is highly similar, even though the Icelandic dataset contains as many as 336 double-NP tokens, and the German dataset contains 390 double-NP tokens. Moreover, the number of tokens in the present configuration generated by each German–Icelandic pair is highly comparable as well. The only verb type that yields considerably more [NP-V-NP] tokens in German is *helfen* (55 tokens, compared to a mere 25 for Icelandic). The four remaining verbs all show very similar token counts: 125 for *líkjast* compared to 132 for *ähneln*, 100 for *mótmæla* compared to 105 for *widersprechen*, 31 for *treysta* compared to 47 for *vertrauen*, and, finally, 55 tokens for *þakka* compared to 51 for *danken*.

5.3.1.2 Word order in configurations involving at least one pronoun

Table 3 shows the results for configurations involving at least one pronoun, i.e. [Pro-V-Pro], [Pro-V-NP], and [NP-V-Pro], each of which is exemplified in (6–8) below:

[Pro-V-Pro]

(6) a. *Ég* *þakka* *ykkur* *mikið* *fyrir það.*
I.NOM thank you.DAT.PL a.lot for that
'I thank you very much for that.'

(6) b. *Ich* *danke* *euch* *sehr* *für die Antworten!*
I.NOM thank you.DAT.PL a.lot for the answers
'I thank you very much for the answers!'

[Pro-V-NP]

(7) a. *Enginn* *mótmælti* *lögmæti fundarins.*
no.one.NOM opposed legitimacy.DAT meeting.the.GEN
'No one challenged the legitimacy of the meeting.'

(7) b. *Wir* *vertrauen* *unseren Kunden.*
we.NOM trust our.DAT customers
'We trust our customers.'

[NP-V-Pro]

(8) a. *Fólk* *treystir* *ekki* *þeim sem treysta sér ekki sjálfir.*
people.NOM trust not them.DAT that trust REFL not selves
'People do not trust those who do not trust themselves.'

(8) b. *Das Prinzip* *ähnelt* *dem eines Lügendetektors.*
the.NOM principle be.similar.to the.one.DAT a.GEN lie.detector.GEN
'The principle is similar to that of a lie detector.'

The results in Table 5.3 are very much in line with those presented in Table 5.2 above, as both Icelandic and German Nom-Dat verbs retain their inclination towards the nominative-before-dative pattern: as many as 655 Icelandic tokens, or 99%, and 553 German tokens, or 91%, contain a clause-initial nominative. Nevertheless, German tokens containing at least one pronoun exhibit topicalisation more than twice as often as tokens involving double NPs in that language (9% vs. 4%). Most dative-before-nominative instances are found with *widersprechen* (19 tokens), *vertrauen* (15 tokens) and *danken* (14 tokens). Approximately half of all German topicalisations, i.e. 29 out of 57 tokens, or 51%, contain a dative pronoun. For Icelandic, the association between a topicalisation structure and a dative pronoun is even stronger (six out of nine tokens, or 67%), but absolute numbers are markedly lower.

Table 5.3 *Hjálþa-* and *helfen*-verbs in configurations involving at least one pronoun

	Nom-Dat		**Dat-Nom**			**Nom-Dat**		**Dat-Nom**	
Verb	**N**	**f**	**N**	**f**	**Verb**	**N**	**f**	**N**	**f**
hjálþa	174	99%	1	1%	*helfen*	136	94%	9	6%
líkjast	75	100%	0	0%	*ähneln*	68	100%	0	0%
mótmæla	92	92%	8	8%	*widersprechen*	76	80%	19	20%
treysta	169	100%	0	0%	*vertrauen*	138	90%	15	10%
þakka	145	100%	0	0%	*danken*	135	91%	14	9%
Total	655	99%	9	1%	**Total**	553	91%	57	9%

5.3.1.3 Interim conclusions

The current section has shown that Modern German, like Modern Icelandic, possesses a class of non-alternating Nom-Dat verbs, since all ten verbs across both languages show a very clear preference for the Nom-Dat order regardless of lexical specifications. This means that the effect of (pro)nominality is fairly limited, as both *hjálþa-* and *helfen*-verbs show an overwhelming preference for the Nom-Dat order even in configurations involving at least one pronoun (cf. Table 3). Evidently, the attested numbers come as no surprise, as it is generally assumed in the literature that verbs like 'help' in German and Icelandic take a nominative subject and a dative object. For our purposes, however, it is important to establish this with frequency counts of the type we have presented above, as these counts will now serve as a baseline for our comparison with non-alternating Dat-Nom verbs and alternating Dat-Nom/Nom-Dat verbs below.

The most salient difference between the languages under study is the extent to which they licence topicalisation. For Icelandic, topicalisation only affects 1% of the tokens, both

in the [NP-V-NP] configuration as well as in configurations involving at least one pronoun. For German, dative fronting is somewhat more frequent, affecting 9% of tokens in configurations involving at least one pronoun and 4% of tokens when both arguments are full NPs. We now proceed to take a closer look at the word order distributions with *líka*- and *gefallen*-verbs.

5.3.2 Verbs of the *líka*- and *gefallen* type

Section 5.3.2 investigates word order variation with *líka*- and *gefallen*-verbs. Sigurðsson (2002), Barðdal & Eythórsson (2003), Eythórsson & Barðdal (2005), Barðdal (2006), Rott (2013), Barðdal, Eythórsson & Dewey (2014, 2019), Wood & Sigurðsson (2014), Somers & Barðdal (2022) and Somers, Jenset & Barðdal (2024) have shown that the relevant Icelandic verb types are principally associated with a Dat-Nom case frame and do not alternate systematically between a dative-before-nominative and a nominative-before-dative order. In German, in contrast, it appears that so-called Dat-Nom verbs are always alternating verbs (cf. Barðdal 2023: Ch. 6; Somers 2023; Somers & Barðdal 2023; Somers et al. 2024).

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

Table 5.4 presents the results for *líka*- and *gefallen*-verbs in the double-NP configuration. The results for Icelandic *líka*-verbs in the present configuration convincingly confirm their status as non-alternating Dat-Nom verbs: out of 194 observations, only one token (or 1%) instantiates a nominative-first structure. The remaining 193 tokens (or 99%) all instantiate the dative-before nominative order.

Below we show first the dative-before-nominative order with *þykja* in (9a), before introducing the one token found in our dataset with the nominative-before-dative order in (9b).

Icelandic

Neutral dative-before-nominative order

(9)	a.	*Stúlkunum*	*þótti*	*Bandaríkjamennirnir*	*afskaplega spennandi.*
		girls.the.DAT	thought	Americans.the.NOM	extremely exciting
		'The girls found the Americans extremely exciting.'			

Topicalised nominative-before-dative order

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

Table 5.4 *Líka*- and *gefallen*-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**
áskotnast	0	0%	48	100%	*zufallen*	71	62%	44	38%
blöskra	0	0%	68	100%	*grauen*	–	–	–	–
leiðast	0	0%	26	100%	*leidtun*	–	–	–	–
líka	0	0%	28	100%	*gefallen*	7	47%	8	53%
þykja	1	4%	23	96%	*dünken*	10	45%	12	55%
Total	1	1%	193	99%	**Total**	88	58%	64	42%

For German, it may first be observed that the verbs *grauen* and *leidtun* do not yield a single token with double NPs. For *grauen*, the lack of double NPs is hardly surprising, as its nominative slot is nearly always filled by an expletive, which is pronominal by definition. For *leidtun*, however, the lack of double NPs seems to be an epiphenomenon of the verb's prototypical usage, which strongly favours pronominal constituents.

The results for the three remaining verbs in the current configuration convincingly confirm their status as alternating verbs, with a total of 88 tokens, or 58%, occurring in the nominative-first pattern and the remaining 64 tokens, or 42%, in the converse dative-first pattern. For *gefallen* and *dünken*, the obtained frequencies for the nominative-first and the dative-first order approximate a 50–50 distribution, but it should be mentioned that the total number of tokens for both of these verbs in the present configuration with two full NPs is quite low, i.e. 15 for *gefallen* and 22 for *dünken*. Examples (10–12) illustrate both word order patterns for all three verbs, *zufallen*, *gefallen* and *dünken*:

German

Dat-Nom

(10) a. *Dem Layouter* *wird* *nämlich* *eine wichtige Rolle*
the.DAT layout artist will namely an.NOM important.NOM role
zufallen.
receive
'The reason is that the layout artist will have an important role to play.'

Nom-Dat

(10) b. *Das Königtum* *ist* *meinem Bruder* *zugefallen.*
the.NOM kingship is my.DAT brother fallen.to
'The kingship has fallen to my brother.'

Dat-Nom

(11) a. *Unserem 7-jährigen Sohn* *gefällt* *das Buch* *auch*
our.DAT seven.year.old.DAT son pleases.3SG the.NOM book also
sehr gut.
very well
'Our seven-year-old son also likes the book a lot.'

(11) b. *Negative Publicity für Anwälte* *scheint* *zumindest*
negative publicity.NOM for lawyers seems.3SG at.least
den Lesern *zu gefallen.*
the.DAT readers.DAT to please
'Negative publicity for lawyers seems at least to be to the readers' liking.'

Dat-Nom

(12) a. *Dem Komponist* *dünkt* *dezent aufgetragene Resignation*
the.DAT composer thinks subtly applied.NOM resignation
wahrhaftiger *als verordnetes sich aufplusterndes Pathos.*
more.truthful as prescribed REFL pompous pathos
'The composer considers subtly applied resignation more truthful than prescribed pompous pathos.'

Nom-Dat

(12) b. *Dieser bleischwere Blick* *dünkte* *dem Baron* *wie ein Blick Gottes.*
this.NOM leaden.NOM gaze thought the.DAT baron as a gaze God.GEN
'This leaden gaze seemed to the baron like a gaze from God.'

As for *zufallen*, there is also considerable word order variation found, with the nominative-before-dative order being attested 71 times, or 62%, and the Dat-Nom order 44 times, or 38%. As will become evident in Section 5.3.3, and as established by Somers et al. (2024), this kind of variation is quite common with alternating Dat-Nom/Nom-Dat verbs in general in German.

5.3.2.2 Word order in configurations involving at least one pronoun

The statistics for both *líka*- and *gefallen*-verbs in contexts with at least one pronoun are laid out in Table 5.5. Interestingly, the overall picture we have uncovered for the double-NP configuration (cf. Table 5.4) remains largely intact in configurations involving at least one pronoun. That is, across verbs, *líka*-verbs retain their overall inclination towards the dative-before-nominative order, whereas *gefallen*-verbs are generally permissive of both orders. Still, the share of nominative-first attestations, for both languages, is somewhat higher under the present circumstances.

Table 5.5 *Líka*- and *gefallen*-verbs in configurations involving at least one pronoun

	Nom-Dat		**Dat-Nom**			**Nom-Dat**		**Dat-Nom**	
Verb	**N**	**f**	**N**	**f**	**Verb**	**N**	**f**	**N**	**f**
áskotnast	3	2%	149	98%	*zufallen*	52	61%	33	39%
blöskra	1	1%	131	99%	*grauen*	33	16.5%	167	83.5%
leiðast	7	4%	167	96%	*leidtun*	187	93.5%	13	6.5%
líka	7	4%	165	96%	*gefallen*	115	62%	70	38%

þykja	50	28%	126	72%	*dünken*	140	79%	38	21%
Total	68	8%	738	92%	**Total**	527	62%	321	38%

As for the Icelandic verbs, it is notable that *þykja* accounts for most instances involving topicalisation, of which two examples are presented in (13) below:

Icelandic

Topicalised nominative-before-dative order

(13) a. *Það* *þótti* *honum* *æðislegt.*
that.NOM thought him.DAT fantastic
'That he found fantastic.'

(13) b. *Landslagið neðan fossins* *þykir* *mér*
landscape.the.NOM below waterfall.the.GEN thinks me
fegurra *austan árinnar.*
more.beautiful east.of river.the
'The landscape below the waterfall I consider more beautiful east of the river.'

It has been argued that *þykja* is a verb that is especially sensitive to pronominal influence, and particularly to that of nominative demonstratives (Somers & Barðdal 2022: 98–99; Somers, Jenset & Barðdal 2024). That tendency may explain why topicalisation numbers for *þykja* amount to 28%, compared to a mere 1–4% for the four remaining *líka*-verbs.

Zooming in on the set of German *gefallen*-verbs, it may first and foremost be observed that *grauen* and *leidtun*, neither of which have yielded any tokens in the [NP-V-NP] configuration, now account for 200 tokens each. Both of these verbs show heavy skewness, with *grauen* tending very heavily towards the dative-before-nominative order and *leidtun* towards the nominative-before-dative order. Consider the example with *grauen* in (14) below:

German

Dative-before-nominative order

(14) *Mir* *graut* *es* *schon* *vor dem Wetter.*
me.DAT dreads it.NOM already before the weather
'I am already dreading the weather.'

Note that *grauen*'s inclination towards a dative-first structure is probably due to the fact that it nearly always has its nominative slot filled by a dummy pronoun *es* 'it'. Dummy pronouns, or expletives, are non-referential slot fillers. As is illustrated by example (14) above, these prefer the less prominent post-verbal slot to the more prominent preverbal slot. For the class of *gefallen*-verbs as a whole, expletives occur 51 times preverbally, or 23%, and 171 times postverbally, or 77%.

Turning now to German *leidtun*, it is notable that it is strongly inclined towards the nominative-before-dative order, with as many as 187 tokens, or 93.5%, instantiating a

nominative-first structure. As such, this potentially alternating verb is more strongly drawn to a nominative-before-dative pattern than the Nom-Dat verbs *widersprechen*, *vertrauen* and *danken*, equally in configurations involving at least one pronoun (cf. Table 5.3 above).

However, it turns out that *leidtun* is exceptionally common with dative indefinite pronouns, the bulk of which occur in postverbal position (137 out of 140). The nominative-first bias *leidtun* exhibits may thus, at least in part, be attributed to secondary factors involving indefinites. Two relevant examples with this verb, one exhibiting a dative-before-nominative pattern and another exhibiting the reverse nominative-before-dative pattern with a dative indefinite pronoun, are shown in (15a–b) below:

German

Dat-Nom

(15) a. *Mir* *haben* *die Zuschauer* *leidgetan.*
me.DAT have the.NOM spectators pitied
'I took pity on the spectators.'

Nom-Dat

(15) b. *Die Gegner* *konnten* *einem* *fast* *leidtun.*
the.NOM opponents could one.DAT almost pity
'The opponents just about aroused one's pity for them.'

Secondly, it is striking that the number of pronominal datives in our German dataset, irrespective of argument position, is considerably higher than the number of pronominal nominatives: in total, there are 774 tokens containing a dative pronoun as opposed to 491 tokens containing a nominative pronoun. As shown in Table 5.6, dative pronouns have a slight predilection for the second slot, while nominatives are fairly evenly distributed across argument positions. A chi-square test shows the distribution to be significant ($\chi^2 = 33.34$; df = 1; $p_{\text{two-tailed}} < .001$), but the concomitant effect size is so small that it is difficult to attach any significance to these findings (Cramér's V = .16).

Table 5.6 Distribution of pronominal arguments across argument positions for German *gefallen*-verbs

	1st position	**2nd position**	**Subtotal**
Nominative pronoun	258	233	491
Dative pronoun	278	496	774
Subtotal	536	729	

5.3.2.3 Interim conclusions

In this section we have shown that Icelandic *líka*-verbs and German *gefallen*-verbs do not behave in the same manner with regard to argument structure. The Icelandic *líka*-verbs are unmistakably non-alternating Dat-Nom verbs, as the statistics exactly mirror the

statistics presented for *hjálpa*-verbs in the previous section; the nominative-before-dative linear order is maximally found in 4% of the cases with two full NPs, while the dative-before-nominative word order is found in 96–100% of the cases. This supports our analysis that the nominative-first order is a topicalisation, while the dative-first order represents the default word order for these verbs. Hence, *líka*-verbs are clearly non-alternating Dat-Nom verbs in Icelandic.

The situation with the "corresponding" German verbs is radically different, as *gefallen*-verbs appear to be alternating Dat-Nom/Nom-Dat verbs. This is evident from the statistics, which range from 45–62% Nom-Dat to 38–55% Dat-Nom, depending on the verb, again when none of the arguments is a pronoun. This finding is perhaps not altogether unexpected, as it has been argued that German Dat-Nom verbs are in fact alternating Dat-Nom/Nom-Dat verbs, and not non-alternating Dat-Nom verbs (Barðdal & Eythórsson 2003; Eythórsson & Barðdal 2005: 868; Barðdal, Eythórsson & Dewey 2014, 2019: 131–148; Rott 2016: 239–249; Barðdal 2023: Ch. 6; Somers 2023; Somers & Barðdal 2023). We have found two outliers in the German dataset, *grauen* and *leidtun*, although their behaviour appears to be explained by secondary factors: *grauen* is nearly always attested with non-referential expletive nominatives which clearly prefer the second argument slot with this verb, while *leidtun* is nearly always attested with dative indefinite pronouns which generally prefer the second argument slot cross-linguistically.

5.3.3 Verbs of the *nægja*- and *genügen*-type

We now turn to a comparison of Icelandic *nægja*- and German *genügen*-verbs. The results for the [NP-V-NP] configuration are presented in Section 5.3.3.1, whereas the remaining part of the dataset is discussed in Section 5.3.3.2. Section 5.3.3.3 is devoted to an interim conclusion.

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

Table 5.7 yields the results for *nægja*- and *genügen*-verbs for tokens with two full NPs. It may first and foremost be observed that Icelandic *nægja*-verbs quite readily allow for both word order patterns, dative-before-nominative and nominative-before-dative, as opposed to *líka*-verbs. The only verb that is at variance with this trend is *henta* 'suit, befit', as it does not yield a single Dat-Nom token. Two such nominative-before-dative examples of *henta* 'suit befit' are given in (16) below:

Icelandic
Nom-Dat

(16)	a.	*Meðferðin*	*hentar*	*öllum aldri og húðgerðum.*
		treatment.the.NOM	suits	all.DAT age.DAT and skin.types.DAT
		'The treatment is suitable for all ages and skin types.'		

(16) b. *Bækurnar* *henta* *börnum á aldri 6–8 ára.*
books.the.NOM suit children.DAT in age 6–8 years
'The books are suitable for children aged 6–8.'

For a more in-depth discussion of this verb as an outlier, due to its lack of occurrences with the dative-before-nominative order, the reader is referred to Somers & Barðdal (2022: 105–107) and Somers, Jenset & Barðdal (2024). Still, it should be mentioned here that, once *henta* is removed from the analysis, the summed frequencies for the remaining four Icelandic verbs display a much more even distribution, i.e. 54% Nom-Dat and 46% Dat-Nom (Somers & Barðdal 2022: 105, Somers, Jenset & Barðdal 2024).

Secondly, it is remarkable that every Icelandic–German verb pair displays roughly the same frequency distributions. Only *henta* shows a radical preference for the nominative-before-dative order whereas its German match *geziemen* has a considerably more evenly distributed frequency pattern, with 56% nominative-before-dative and 44% dative-before-nominative orders. For illustration, we present one example of each of the two word orders with *geziemen* in (17) below:

German
Dat-Nom
(17) a. *Dem Hause Gottes* *geziemet* *Heiligkeit.*
the.DAT house.DAT God.GEN befits holiness.NOM
'The house of God is befitting of holiness.'

Nom-Dat
(17) b. *Der bloße nackte Instinkt* *geziemt* *nicht*
the.NOM mere.NOM naked.NOM instinct befits not
dem Menschen.
the.DAT human.DAT
'The mere naked instinct does not befit man.'

Table 5.7 *Nægja-* and *genügen*-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**
duga	33	79%	9	21%	*nützen*	53	83%	11	17%
dyljast	2	25%	6	75%	*entgehen*	23	39%	36	61%
endast	9	30%	22	70%	*reichen*	15	38%	24	62%
henta	86	100%	0	0%	*geziemen*	20	56%	16	44%
nægja	27	54%	23	46%	*genügen*	36	55%	29	45%
Total	157	72%	60	28%	**Total**	147	56%	116	44%

The German dataset reveals that, across verbs, the nominative-before-dative order is attested 147 times, or 56%, and the dative-before-nominative order 116 times, or 44%.

Examples involving *nützen* 'be of use' and *genügen* 'suffice' with both word orders are given in (18–19) below for illustration purposes:

German

Dat-Nom

(18) a. *Vor allem Jüngeren* *nützt* *der Abschluss.*
especially younger.people.DAT benefits the.NOM qualification
'Especially younger people benefit from the qualification.'

Nom-Dat

(18) b. *Hilfe* *muss* *den Ärmsten* *nützen.*
aid.NOM must the.DAT poorest.DAT benefit
'Aid must benefit the poorest.'

Dat-Nom

(19) a. *Der Stadt* *genügte* *die bloße Mitteilung* *nicht.*
the.DAT city sufficed the.NOM mere.NOM announcement not
'The city did not find the mere announcement sufficient.'

Nom-Dat

(19) b. *Die erste Lösung* *hat* *Ihrem Türsteher* *nicht* *genügt.*
the.NOM first.NOM solution has your.DAT doorman not sufficed
'The first solution was not enough for your doorman.'

The even frequencies in Table 5.7 are very much in line with the counts obtained for *gefallen*-verbs in the double-NP configuration, as described in Section 5.3.2.1 above. Recall that these occur in a nominative-first structure 58% of the time, and in dative-first structure 42% of the time. This raises the question whether *gefallen*- and *genügen*-verbs build one class in German, instead of two. We therefore carry out a chi-square test comparing the prevalence of the dative-before-nominative and the nominative-before-dative orders in configurations involving two full NPs for both *gefallen*- and *genügen*-verbs. The relevant numbers are shown in Table 5.8 below:

Table 5.8 Word order distributions between *gefallen*- and *genügen*-verbs in the [NP-V-NP] configuration

	Dat-Nom	**Nom-Dat**	**Subtotal**
gefallen-verbs	64	88	152
genügen-verbs	116	147	263
Subtotal	180	235	

The test shows that a distinction between *gefallen*- and *genügen*-verbs is not statistically justified ($\chi^2 = 0.09$; df = 1; $p_{\text{two-tailed}} = 0.77$). In other words, there is no evidence to consider *gefallen*- and *genügen*-verbs two separate classes. Instead, these frequencies and statistical

tests confirm the conjecture expressed in Section 5.3.2.1 above, that *gefallen*-verbs are also alternating Dat-Nom/Nom-Dat verbs in German.

All of this means that the three-part distinction in Icelandic between Nom-Dat verbs, Dat-Nom verbs, and Dat-Nom/Nom-Dat verbs may correspond to a two-part distinction in German, as the latter language only seems to be endowed with ordinary Nom-Dat verbs of the 'help' type and alternating Dat-Nom/Nom-Dat verbs, but not with non-alternating Dat-Nom verbs. However, these findings are only based on ten out of the 146 potentially alternating verbs in German documented by Somers (2021). Therefore, future research will have to confirm whether non-alternating Dat-Nom verbs are non-existent in that language or not.

5.3.3.2 Word order in configurations involving at least one pronoun

Let us now examine how *nægja*- and *genügen*-verbs behave in contexts involving at least one pronoun. The relevant numbers are presented in Table 5.9 below. For Icelandic, 75% of tokens, i.e. 590 out of 783, instantiate the nominative-before-dative pattern. The remaining 25% of tokens, or 193, evidently display the reverse dative-before-nominative pattern. The relative frequencies presented here constitute a near-perfect copy of the results laid out in Table 5.7 above, which shows that *nægja*-verbs display a 72%–28% distribution between the Nom-Dat and the Dat-Nom linear orders, respectively. Still, the nominative-first bias observed for pronominal configurations is more robust that the one observed for double NPs, as it is based on more than three times as many tokens. Additionally, as has already been mentioned in Section 5.3.3.1 above, the nominative-before-dative bias uncovered for the double-NP configuration is mainly due to the skewed behaviour of *henta*.

Table 5.9 *Nægja*- and *genügen*-verbs in configurations involving at least one pronoun

	Nom-Dat		**Dat-Nom**			**Nom-Dat**		**Dat-Nom**	
Verb	**N**	**f**	**N**	**f**	**Verb**	**N**	**f**	**N**	**f**
duga	147	93%	11	7%	*nützen*	127	93%	9	7%
dyljast	148	77%	44	23%	*entgehen*	74	52%	67	48%
endast	69	41%	100	59%	*reichen*	99	61%	62	39%
henta	114	100%	0	0%	*geziemen*	98	60%	66	40%
nægja	112	75%	38	25%	*genügen*	73	54%	62	46%
Total	590	75%	193	25%	**Total**	471	64%	266	36%

In configurations involving at least one pronoun, the nominative-first bias observed for Icelandic is to a large degree the result of nominative pronouns, which are heavily inclined to take first rather than second position: out of a total 416 nominative pronouns, as many

as 391, or 94%, occupy the preverbal slot. Two such examples are given in (20) below, with the verbs *dyljast* 'miss out on' and *endast* 'suffice':

Icelandic

Nom-Dat triggered by a nominative pronoun

(20) a. *Það* *dylst* *engum sem hana þekkti.*
that.NOM be.hidden.to no.one.DAT who her knew
'That did not remain hidden to anyone who knew her.'

(20) b. *Hún* *entist* *mér* *út ferðalagið.*
she.NOM lasted me.DAT out journey
'It lasted me throughout the journey.'

In fact, the tendency for alternating verbs to allocate nominative pronouns to the first slot regardless of the (pro)nominal status of the dative, has been observed before (cf. Somers & Barðdal 2022 for Icelandic; Somers 2023: 68–69 for German; Cluyse, Somers & Barðdal 2024 for Latin). The same trend holds, mutatis mutandis, for the dative-before-nominative order, which very strongly correlates with dative pronouns (187 out of 193 tokens, 97%), again shown with two examples of *dyljast* and *endast* in (21) below.

Icelandic

Dat-Nom triggered by a dative pronoun

(21) a. *Engum* *dylst*
no.one.DAT be.hidden.to
mikilvægi forvarna gegn sjúkdómum.
importance.NOM prevention.GEN against diseases
'No one fails to acknowledge the importance of disease prevention.'

(21) b. *Honum* *entist* *ekki* *gæfa þar til.*
him.DAT lasted not good.fortune.NOM there to
'He outlasted his good fortune for that.'

The German data also show an increase in the number of Nom-Dat tokens, but the alternating behaviour of the verbs under study remains largely intact, with 471, or 64% instantiating the nominative-before-dative order and 266, or 36%, the dative-before-nominative order. The difference in behaviour between Icelandic and German nominative pronouns is largely due to the effect of correlates, i.e. third-person pronouns functioning as stand-ins for subclauses. Two examples of tokens featuring such correlates, one for Icelandic and one for German, are provided under (22a–b) below:

(22) a. *Það* *dugar* *mér* *að eiga einn bíl í einu.*
it.NOM be.enough me.DAT to have one car at one
'It is enough for me to have one car at a time.'

(22) b. *Es* *nützt* *ihnen* *nichts*

it.NOM be.of.use them.DAT nothing
laktosefreie Produkte zu konsumieren.
lactose.free products to consume
'It is of no use to them to consume lactose-free products.'

Now, the situation is such that Icelandic correlates with *nægja*-verbs pattern with other nominative pronouns, in that they categorically prefer the first argument position (156 tokens or 95%) to the second (eight tokens or 5%). For German, however, correlates may be realised on either side of the conjugated verb, i.e. preverbally (121 tokens 58%) as well as postverbally (88 tokens or 42%).

5.3.3.3 Interim conclusions

In this section, we have shown that German *genügen*-verbs, exactly like Icelandic *nægja*-verbs, are in fact alternating predicates, i.e. they vacillate between the two argument structures, Dat-Nom and Nom-Dat, in both languages. At first blush, alternation seems to be more pervasive in German than it is in Icelandic, as the results for *genügen*-verbs in the double-NP configuration are much closer to a 50–50 distribution than they are for *nægja*-verbs. However, as soon as the outlier verb *henta* is excluded from the statistical analysis, the resulting frequencies for the remaining four Icelandic verbs included in the investigation reveal a considerably more even distribution.

Yet another result our analysis has brought to light relates to the effect of correlative *es* and *það* on word order distributions in German and Icelandic. As it turns out, nominative correlates radically swing Icelandic alternating verbs towards the nominative-before-dative order. Alternating verbs in German, by contrast, retain their alternating behaviour. This shows that referentiality is a factor steering word order variation in German, but not in Icelandic. Why exactly German *nützen* goes against the general trend uncovered for that language remains at present unknown.

5.3.4 Personal pronouns

In her work on case marking and grammatical relations in Old and Early Middle English, Allen (1995: 109) observes that when the two arguments of alternating verbs, the dative and the nominative, are full NPs, their word order distributions are relatively even, i.e. 19 examples exhibit the dative-first pattern and 21 the nominative-first pattern. However, in cases where both arguments are personal pronouns, only the nominative-before-dative order is found in Old English. Allen documents this with 48 examples. This suggests a major asymmetry between the two argument structures when personal pronouns are involved.

The question arises whether a similar effect for two referential personal pronouns may also be found in Icelandic and German. Consider, first, Table 5.10, which shows both the raw frequencies and percentages for four of the Icelandic *nægja*-verbs when occurring with referential personal pronouns. This excludes tokens containing expletives and clause-anticipating correlatives of the type discussed in Section 5.3.3.3 above. For *duga*, *endast* and *henta*, the percentages are 100% nominative-before-dative or close to that. Below we give one example of each of these three verbs with two personal pronouns:

Icelandic

Nom-Dat with two personal pronouns

(23) a. *Keyptu nýjar ónotaðar patrónur, þær duga þér best.*
buy new unused cartridges they.NOM suffice you.DAT best
'Buy yourself new unused cartridges, they will last you the longest.'

(23) b. *Sjáum til hvað hún endist okkur lengi.*
see.1PL to what she.NOM lasts us.DAT long
'Let's see how long it will last us.'

(23) c. *… þeir henta þeim ekki.*
they.NOM suit them.DAT not
'... they are not suitable for them.'

For *nægja*, in contrast, the percentage is 71% nominative-first, although the reader should keep in mind that *nægja* only yields seven instances in total. To some degree, therefore, it seems that Allen's findings are also valid for Icelandic, where the total numbers are 94% nominative-first and 6% dative-first.

Table 5.10 *Nægja*-verbs with two referential personal pronouns

	Nom-Dat		**Dat-Nom**	
Verb	**N**	**f**	**N**	**f**
duga	15	100%	0	0%
dyljast	–	–	–	–
endast	18	95%	1	5%
henta	10	100%	0	0%
nægja	5	71%	2	29%
Total	48	94%	3	6%

Turning to German, four out of ten alternating verbs, *zufallen*, *leidtun*, *nützen* and *entgehen*, show a 100% nominative-before-dative distribution, as is evident from Table 5.11, when the two arguments are referential personal pronouns. One example of each of these four verbs is given in (24) below:

German

Nom-Dat with two personal pronouns

(24) a. *Es* *muss* *Ihnen* *schon* *zufallen,*
it.NOM must you.HON.DAT already fall.to
bevor Sie es aufheben.
before you it pick.up
'It has to come to you before you pick it up.'

(24) b. *Er* *könnte* *mir* *glatt* *leidtun.*
he.NOM could me.DAT really be.pitied
'He might just about arouse pity in me.'

(24) c. *Sie* *nützen* *mir* *nichts.*
they.NOM be.of.use me.DAT nothing
'They are of no benefit to me.'

(24) d. *Es* *ist* *Ihnen* *entgangen.*
it.NOM is you.HON.DAT failed.to.notice
'It escaped your attention.'

Three additional alternating verbs, *reichen*, *geziemen* and *genügen*, only occur two, three and four times, respectively, with two referential personal pronouns; thus, their totals are simply too low to draw any statistical conclusions from. The two remaining alternating verbs, *gefallen* and *dünken*, both occur in the nominative-first order in the majority of cases, although the numbers are somewhat lower than for *zufallen*, *leidtun*, *nützen* and *entgehen*, or 61 vs. 78% respectively. While it is obvious that more research is needed on German to confirm these numbers, there is clearly a tendency for alternating verbs to show up with the nominative-before-dative word order when the two arguments are referential personal pronouns.

Table 5.11 *Gefallen*- and *genügen*-verbs with two referential personal pronouns

	Nom-Dat		**Dat-Nom**	
Verb	**N**	**f**	**N**	**f**
zufallen	7	100%	0	0%
grauen	–	–	–	–
leidtun	8	100%	0	0%
gefallen	17	61%	11	39%
dünken	7	78%	2	22%
nützen	16	100%	0	0%
entgehen	2	100%	0	0%
reichen	2	50%	2	50%
geziemen	1	50%	1	50%
genügen	1	33%	2	67%

Total	61	77%	18	23%

5.4 Conditional inference tree

The current section scrutinises the entire set of alternating verbs, Icelandic and German alike, through a conditional inference tree. For Icelandic, the analysis is based on the verbs *duga*, *dyljast*, *endast* and *nægja*. In accordance with Somers, Jenset & Barðdal (2024), we leave out *henta*, as it does not pass the word order test. For German, the tree builds on a total of ten verbs: the *gefallen*-verbs *dünken*, *gefallen*, *grauen*, *leidtun* and *zufallen* and the *genügen*-verbs *entgehen*, *genügen*, *geziemen*, *nützen* and *reichen*. All in all, we feed into the analysis 200 tokens for each of the 14 alternating verbs, thus amounting to a dataset of 2,800 tokens. As Section 5.3 has shown that word order variation is sensitive to pronominal influence, we introduce a new variable which distinguishes tokens involving two full NPs from tokens involving at least one pronoun.

Conditional inference trees can be viewed as non-parametric regression models (Everitt & Hothorn 2006). They offer an attractive alternative to regular regression models, as they visualise interactions between predictor variables in a comprehensive way. They can also handle interactions between variables better than logistic regression models (Tagliamonte & Baayen 2012). Concretely, conditional inference trees can be read as decision trees or flow charts, starting with the most important variable at the top, with the binary splits representing statistically significant classification threshold values. The final leaf nodes at the bottom of the tree represent the classification probability of the response for the preceding combination of variable values. The tree, which yields an accuracy of just under 70%, has been modelled using the *party* package in R.

The model, which is shown in Figure 5.1, identifies Case as the most important variable, with a leftward split for dative and a rightward split for nominative case. The fact that the case variable is allocated the top node of the tree ties in well with Somers, Jenset & Barðdal (2024), who have studied the word order variation of the Icelandic alternating verbs *duga*, *dyjast*, *endast* and *nægja* through a logistic regression analysis. Their model singles out nominative case as the most important predictor, yielding a logistic regression coefficient of 2.26, which corresponds to a 56.5% increase in the likelihood of switching from second to first position. Importantly, Somers, Jenset & Barðdal (2024) builds on exactly the same 800 Icelandic tokens as the ones that have been included here. Our current study, however, additionally considers 2,000 German tokens. It is remarkable that case presently retains its predictive power, even though the dataset has been substantially expanded upon to also include data from another language, viz. German.

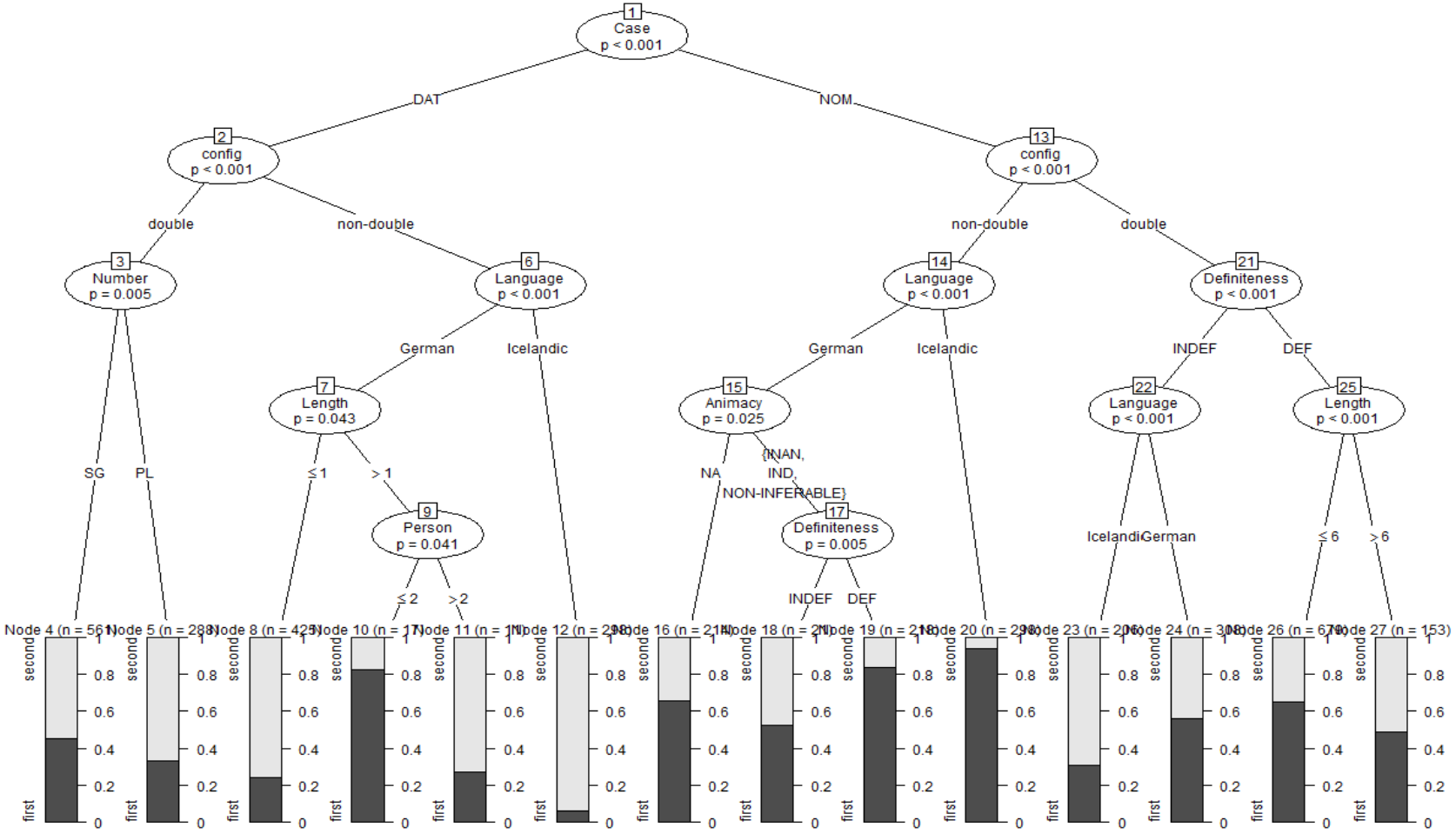


Figure 5.1 Conditional inference tree modelling the variation in word order for the set of alternating verbs

Either branch of the Case node interacts with Configuration, which splits off into configurations involving double NPs and configurations involving at least one pronoun. The fact that lexical specifications exert a major influence on word order with alternating verbs has been observed before, e.g. by Allen (1995) for Old English (see discussion above), Somers & Barðdal (2022) and Somers, Jenset & Barðdal (2024) for Icelandic, Somers (2023) for German, and Cluyse, Somers & Barðdal (2024) for Latin. The prominent position of Configuration in the tree once again underscores the importance of (pro)nominality in the study of alternating verbs.

Configuration, in turn, interacts with three additional variables: Number, Definiteness and Language. The last variable, Language, on either side of the tree, interacts with the pronominal branch, thereby demonstrating that pronouns have different inclinations in Icelandic compared to German, at least for the verbs under study here, if not in general. For Icelandic, configurations involving at least one pronoun almost categorically prefer the nominative-dative order (cf. leaf nodes six and ten). This effect is to a large degree due to the preponderance of nominative pronouns in our dataset, which not only tend to precede dative full NPs, but also dative pronouns. The implicational hierarchy under (25) captures linearisation in Icelandic alternating predicates in configurations involving at least one pronoun:

(25) nominative pronoun > dative pronoun > full NP

The hierarchy above stipulates that Icelandic nominative pronouns always tend to come first, regardless of the lexical specifications of the second argument. Dative pronouns also tend to take the preverbal slot, but only in the absence of nominative pronouns.

As for the double-NP configuration, the dative branch of the tree shows an interaction with Number, whereby singular datives tend to occur in first position more readily than plural datives. Number was also identified as a factor in Somers, Jenset & Barðdal (2024), but it only turned out to be significant across configurations, and not in configurations with double NPs only. However, Somers et al.'s (2024) study of 76 potentially alternating verbs in German found an effect for number both across configurations as well as in configurations involving at least one pronoun, whereby singular (dative or nominative) constituents consistently allowed for alternation more than plurals. The fact that singular forms are more prone to taking first position than plurals may relate to the fact that singular referents are more identifiable (cf. Bickel & Nichols 2007: 225).

Turning now to double NPs in the Nominative branch of the tree, there is a split for Definiteness, with definite nominatives showing an interaction with Length and indefinite nominatives showing an interaction with Language. The Length node creates a pathway for shorter constituents (counting six words or fewer) and a pathway for longer ones (counting seven words or more), where the former take first position more promptly than the latter (see leaf nodes 13 and 14). In other words, longer nominatives tend to come

second rather than first. This result is evidently in tune with the Behaghelian principle that shorter constituents tend to precede longer ones (Behaghel 1909/10; see also Allan 1987; Arnold et al. 2000; Hawkins 1992; Siewierska 1993; Jónsson 2018; Somers, Jenset & Barðdal 2024, *inter alia*).

Leaf nodes 11 and 12 reveal that indefinite nominative NPs in the [NP-V-NP] configuration in Icelandic are more strongly confined to the second argument slot than in German. The role of indefiniteness in the word order distributions of Icelandic alternating verbs has also been identified by Somers, Jenset & Barðdal (2024). For German, however, indefinite nominatives in the double-NP configuration do not have a clear preference for either the first or the second argument slot. This is therefore a clear difference between Icelandic and German. Yet, note that nominative indefiniteness is also a fairly weak predictor in German configurations involving at least one pronoun (leaf node 8).

In conclusion, it is worth mentioning that the Language variable only shows up once in the branches pertaining to double NPs and that it only surfaces on the fourth and lowest level of the tree. This affirms that alternating verbs in the double-NP configuration, in Icelandic and German alike, generally behave in the same way, except with regard to indefinite nominatives, as discussed above. Furthermore, the leaf nodes are generally less skewed for those branches pertaining to double NPs, which confirms that alternation is most pervasive in contexts not involving pronouns in both languages.

Let us finally consider those branches that represent the German configurations involving at least one pronoun. First, the model shows that datives in such a configuration tend to come second, whether they be short or long (cf. leaf nodes 3 and 5). The only exception to this dative-second trend are dative local pronouns consisting of more than one word, which generally come first (leaf node 4), irrespective of whether the second argument is a pronoun or a full NP. Two relevant examples are shown under (26a–b) below:

(26) a. *Mir persönlich* *reicht* *das* *als Quelle.*
me.DAT personally suffices that.NOM as source
'For me personally, that is satisfactory as a source.'

(26) b. *Uns normalen Arbeitern* *graut* *es*
us.DAT normal.DAT labourers.DAT dreads it.NOM
vor solchen Zuständen.
before such conditions
'We normal workers dread such conditions!'

It is clear that neither (26a) nor (26b) represents neutral word order. That is, the dative takes sentence-initial position because it is in focus. Inspection of the remaining tokens featuring clause-initial multi-word dative pronouns reveals that these too are examples of focus rather than neutral word order. Thus, the dative-first effect captured by leaf node 4

may clearly be attributed to pragmatic reasons. No such effect is found for Icelandic by the model.

Secondly, the model shows that German nominatives in configurations involving at least one pronoun come first if they are animate, inanimate or non-inferable as well as definite (leaf node 9), but that these very same nominatives allow for considerably more variation if they are indefinite (cf. supra). Comparable variation may be observed for German nominatives which denote either a collective, an expletive or a correlate. Recall that the latter two, which in our dataset always take the form of third-person personal *es* 'it', do not have an extralinguistic referent, as they are either placeholders for subclauses, in the case of correlates, or dummy pronouns, in the case of expletives. It has already been mentioned in Section 5.3.2.2 that expletives are strongly attracted to the postverbal slot, which illustrates that German is a language in which referential elements clearly take precedence over non-referential ones. A similar trend has been observed for correlates in Section 5.3.3.2, even though these exhibit slightly more balanced frequencies between the preverbal and the postverbal slot. By assigning them to the same branch, the conditional inference tree acknowledges that German expletives and correlates behave similarly. The same holds for nominative collectives, but these are altogether quite rare.

5.5 Summary and conclusions

In the present paper we have compared the word order distributions of 30 verbs, 15 from German and 15 from Icelandic, which licence both a nominative and a dative argument. Crucially, the Icelandic verbs divide into one of three classes (cf. Somers, Jenset & Barðdal 2024): ordinary Nom-Dat verbs, here referred to as *hjálþa*-verbs, non-alternating Dat-Nom verbs, here referred to as *líka*-verbs, and, finally, alternating Dat-Nom/Nom-Dat verbs, which are here referred to as *nægja*-verbs. In this study, we have gathered 15 German verbs, which are semantic corollaries of the Icelandic verbs in the three classes listed above, which we have here labelled *helfen*-, *gefallen*-, and *genügen*-verbs.

The German dataset was extracted in an exactly parallel manner to the Icelandic dataset from 2022, i.e. through a corpus of web texts, the deTenTen13 corpus, which contains more than 16.5 billion words, while the Icelandic dataset was gathered through the more recent isTenTen20 corpus, which contains 520 million words. The two datasets have been annotated in the same way, on the basis of exactly the same variables, i.e. language, case marking, (pro)nominality, pronoun type, referentiality, person, number, definiteness, animacy and length.

We have shown that the similarities between *helfen*- and *hjálþa*-verbs in German and Icelandic are striking in that the nominative-first linear word order is found in 96% of the

cases in German when both arguments are full NPs, while the corresponding number is 99% for Icelandic. This is, of course, what one would expect, since it is entirely uncontroversial that 'help' verbs take a nominative subject and a dative object in the Germanic languages which still have morphological case marking. Yet, this is of particular importance here since it aids in establishing a baseline for how frequent topicalisation is in German and Icelandic. Thus, we have here demonstrated that for verbs of the 'help' type, topicalisation of the dative object to preverbal position is found in ca. 4% of the cases in German but only in ca. 1% of the cases in the Icelandic dataset when the two arguments are full NPs.

Turning to *gefallen*- and *líka*-verbs in the two languages, i.e. the subset of verbs in Icelandic which systematically occur only in the Dat-Nom argument structure, these were hypothesised to be Dat-Nom verbs, thus showing no inclination towards being alternating Dat-Nom/Nom-Dat verbs. As expected, the frequencies for the Icelandic *líka*-verbs exactly mirror the frequencies for *hjálþa*-verbs in Icelandic in that the dative-before-nominative order occurs in 96–100% of the cases, again when both arguments are full NPs, while the nominative-before-dative linear order is found in maximally 4% of cases. These numbers show unambiguously that *líka*-verbs in Icelandic are non-alternating Dat-Nom verbs. However, our findings for the German *gefallen*-verbs are not in accordance with the behaviour of what we have assumed to be their Icelandic counterparts. Instead, German *gefallen*-verbs turn out to behave as alternating verbs, systematically instantiating the Dat-Nom and the Nom-Dat argument structure, ranging from 45–55% Nom-Dat vs. Dat-Nom to 62–38% Nom-Dat vs. Dat-Nom, depending on the verb. Future research (cf. Somers et al. 2024) will have to determine whether or not non-alternating Dat-Nom verbs exist in Present-Day German, as our dataset here consists of only ten potential candidates for this alternation.

One of the most important results of our study is that German, precisely like Icelandic, possesses a class of alternating verbs, here referred to as *genügen*- and *nægja*-verbs. Crucially, as may be deduced from Table 5.12, the alternating Dat-Nom/Nom-Dat argument structure is adopted by a larger set of verbs in German than in Icelandic. This means that whereas Icelandic shows a three-part distinction between Nom-Dat, Dat-Nom, and Dat-Nom/Nom-Dat verbs, German seems to possess only two verb classes: Nom-Dat verbs, and Dat-Nom/Nom-Dat verbs. In other words, all ten Dat-Nom verbs in German, under investigation here, turn out to be alternating Dat-Nom/Nom-Dat verbs. This rhymes well with what has been argued in recent literature, i.e. that German Dat-Nom predicates are in fact alternating predicates, as originally argued by Barðdal & Eythórsson (2003), Eythórsson & Barðdal (2005: 868) and later established by Barðdal, Eythórsson & Dewey (2014, 2019: 131–148), Rott (2016: 239–249) and Barðdal (2023: Ch. 6) on the basis of a series of subject tests available for that language.

Table 5.12 Distribution of Nom-Dat, Dat-Nom and Dat-Nom/Nom-Dat argument structures in Icelandic and German across verb classes

	Icelandic	**German**
Nom-Dat	*hjálþa*-verbs	*helfen*-verbs
Dat-Nom	*líka*-verbs	–
Dat-Nom/Nom-Dat	*nægja*-verbs	*gefallen*-verbs
		genügen-verbs

We have also briefly addressed the issue of word order with alternating Dat-Nom/Nom-Dat predicates when the two arguments are referential personal pronouns. It has been shown in the literature on Old English that such lexical-syntactic configurations, in fact, favour the Nom-Dat word order (Allen 1995). We have found some effect of this for both Icelandic and German, although it appears to be considerably stronger in Icelandic than in German. However, due to how few the relevant verbs are in number, additional research is required to throw further light on this issue. In any case, it is remarkable that a configuration with two referential personal pronouns in Icelandic, German and Old English alike exhibits such congruent behaviour. We believe this to be a reflection of the common origin of these predicates and these argument structure constructions, as has been argued for in more detail by Eythórsson & Barðdal (2005) and Barðdal (2006, 2023: 81–112, 125–137, 225–247).

With the aid of a conditional inference tree, we have attempted to lay bare the variables correlating with the choice between the Dat-Nom and the Nom-Dat alternants for the set of alternating verbs across both languages, Icelandic and German. For Icelandic, we have included the verbs *duga*, *dyljast*, *endast* and *nægja* and for German, the analysis builds on the verbs *dünken*, *gefallen*, *grauen*, *leidtun*, *zufallen*, *entgehen*, *genügen*, *geziemen*, *nützen* and *reichen*. The model has shown that variation in the dataset is best explained through an interaction of case marking and the different configurations of nouns and pronouns, with tertiary, quaternary and quinary factors again associating with those.

One difference between Icelandic and German that we have detected in our dataset relates to the configurations involving at least one pronoun: these show very heavy skewness towards the nominative-first order in Icelandic, whereas the same configurations in German are permissive of more variation. Moreover, the inference tree identifies an interaction with animacy, definiteness, length and person in German. The fact that nominative pronouns with Icelandic alternating verbs so strongly tend towards first position has also been observed by Somers, Jenset & Barðdal (2024: 22, 31). They present evidence that the nominative-first order in the double-pronoun configuration in Icelandic, in particular, is an epiphenomenon of length. For the double NP configuration, by contrast, the nominative-first order in Icelandic appears to be an effect of definiteness and topicality. All of this shows that there are similar factors at play in both German and Icelandic, even though these are undetected for Icelandic in the current inference tree.

Another important finding is that language is only of minor importance in configurations involving two full NPs. It should be reiterated, though, that the conditional inference tree builds on more data points for German than for Icelandic, but this skewness is mitigated in the double-NP configuration, as the German verbs *grauen* and *leidtun* do not yield any tokens with two full NPs. Still, when pronouns are excluded, alternating verbs across Icelandic and German behave in a highly similar fashion, thus confirming that these are parallel structures which should not be analysed in distinctive ways. In other words, the similarities between the two languages clearly outweigh the differences.

One of our goals here has been to investigate to which degree word order may be used as a subject test in German, despite the consensus in the German scholarship that the word order is too free in that language to be usable to tease apart syntactic subjects from syntactic objects. Contrary to this, we have here shown that neutral word order may be established in German, as well as in Icelandic, through comparing corpus-based usage frequencies for different word order constellations for distinct sets of verbs. We have found, for instance, that the neutral word order for Nom-Dat verbs like *hjálþa* and *helfen* is nominative-before-dative, while the reverse dative-before-nominative order with these verbs is much rarer, i.e. 96–99 vs. 1–4% of the cases in the two languages when both arguments are full NPs.

Due to German's general sensitivity to pronouns when it comes to word order distributions, and especially so in the middle field, we have chosen to only work with full NPs when establishing neutral word order. Thus, the numbers which we have obtained for *hjálþa*- and *helfen*-verbs show beyond any doubt that the nominative with these verbs occupies the first position in neutral word order, thus corroborating our claims that word order may be used as a subject test in German. Hence, the nominative of *helfen*-verbs is the syntactic subject and the dative the syntactic object, as is also expected on the assumption that these are Nom-Dat verbs in that language.

For our hypothesised alternating verbs, in contrast, we have established that the distribution between the dative-before-nominative and the nominative-before dative word orders are relatively even for the two languages when both arguments are full NPs, that is 46 vs. 54% in Icelandic (excluding *henta*) and 57 vs. 43% in German (excluding *grauen* and *leidtun*). Recall that these frequencies are extracted from a larger dataset involving the same number of instances per verb for both ordinary Nom-Dat verbs like *hjálþa* and *helfen* and alternating Dat-Nom/Nom-Dat verbs. As such, the distribution of the two word orders, dative-before-nominative and nominative-before dative, supports the assumption that both word orders are neutral with these verbs, again corroborating our analysis that these verbs may instantiate two diametrically opposed argument structure constructions, i.e. Dat-Nom and Nom-Dat in German, exactly as in Icelandic. As a consequence, it is the dative that is the syntactic subject in Dat-Nom argument structure constructions, while it is the nominative that is the subject in Nom-Dat constructions. In conclusion, these facts entail that the word order test may easily be applied as a subject test in German, provided,

of course, that a baseline is first established for verbs which take a nominative subject, to be compared with other non-canonically case-marked argument structure constructions, exactly as has been done here.

References

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

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

Bader, Markus. 2020. The position of object pronouns in the German middlefield. *Linguistics* 58(4): 1059–1115.

Bader, Markus & Jana Häussler. 2010. Word order in German: A corpus study. *Lingua* 120: 717–762.

Barðdal, Jóhanna. 1998. Argument structure, syntactic structure and morphological case of the impersonal construction in the history of Scandinavian. *Scripta Islandica* 49: 21–33.

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. 2004. The semantics of the impersonal construction in Icelandic, German and Faroese: Beyond thematic roles. In Abraham, Werner (ed.), *Focus on Germanic typology*, 105–137. Berlin: Akademie Verlag.

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 & 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.

Barnes, Michael. 1986. Subject, nominative and oblique case in Faroese. *Scripta Islandica* 37: 13–46.

Bayer, Josef. 2004. Non-nominative subjects in comparison. In Bhaskararao, Peri & Subbarao, Karumuri V. (eds.), *Non-Nominative Subjects,* Vol I, 49–76. Amsterdam: John Benjamins.

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

Bernódusson, Helgi. 1982. *Ópersónulegar setningar* [Impersonal sentences]. Reykjavík: University of Iceland Master's thesis.

Bickel, Balthasar & Johanna Nichols. 2007. Inflectional morphology. In Shopen, Timothy (ed.), *Language Typology and Syntactic Description* (2nd edn.), 169–240. Cambridge: Cambridge University Press.

Bresnan, Joan & Marilyn Ford. 2010. Predicting syntax: Processing dative constructions in American and Australian varieties of English. *Language* 86(1): 168–213.
Cluyse, Brian, Joren Somers & Jóhanna Barðdal. 2024. Latin *placēre* as an alternating Dat-Nom/Nom-Dat verb: A radically new analysis. Submitted.
Croft, William. 2003. *Typology and Universals*. 2nd edn. Cambridge: Cambridge University Press.
Duden = *Die Grammatik*. (2016). Berlin: Dudenverlag.
Everitt, Brian S. & Torsten Hothorn. 2006. *A Handbook of Statistical Analyses Using R*. Boca Raton: Chapman & Hall/CRC.
Eythórsson, Thórhallur & Jóhanna Barðdal. 2005. Oblique subjects: A common Germanic inheritance. *Language* 81(4): 824–881.
Haider, Hubert. 2005. How to turn German into Icelandic—and derive the OV–VO contrast. *Journal of Comparative Germanic Syntax* 8: 1–53.
Haspelmath, Martin. 2001. Non-canonical marking of core arguments in European languages. In Aikhenvald, Alexandra Y. & Dixon, R. M. W. & Onishi, Masayoki (eds.), *Non-Canonical Marking of Subjects and Objects*, 53–83. Amsterdam: John Benjamins.
Hawkins, John A. 1992. Syntactic weight versus information structure in word order variation. In Jacobs, Joachim (ed.), *Informationsstruktur und Grammatik*, 196–219. Wiesbaden: Westdeutscher Verlag.
Heylen, Kris. 2005. *Zur Abfolge (pro)nominaler Satzglieder im Deutschen: Eine korpusbasierte Analyse der relativen Abfolge von nominalem Subjekt und pronominalem Objekt im Mittelfeld*. Leuven: KU Leuven Ph.D. dissertation.
Hoberg, Ursula. 1981. *Die Wortstellung in der geschriebenen deutschen Gegenwartssprache*. München: Max Hueber Verlag.
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. 1997–98. Sagnir með aukafallsfrumlagi [Verbs selecting for oblique subjects]. *Íslenskt mál* 19–20: 11–43.
Jónsson, Jóhannes Gísli. 2018. Word order as a subject test in Old Icelandic. In Barðdal, Jóhanna & Pat-El, Na'ama & Carey, Stephen Mark (eds.), *Non-Canonically Case-Marked Subjects: The Reykjavík–Eyjafjallajökull Papers*, 135–154. Amsterdam: John Benjamins.
Kempen, Gerard & Karin Harbusch. 2004. A corpus study into word order variation in German subordinate clauses: Animacy effects linearization independently of grammatical function assignment. In Pechmann, Thomas & Habel, Christopher (eds.), *Multidisciplinary Approaches to Language Production*, 173–181. Berlin: Mouton de Gruyter.
Kempen, Gerard & Karin Harbusch. 2005. The relationship between grammaticality ratings and corpus frequencies: A case study into word order variability in the midfield of German clauses. In Kepser, Stephan & Reis, Marga (eds.), *Linguistic Evidence*, 329–349. Berlin: Mouton de Gruyter.
Kilgarriff, Adam, Pavel Rychlý, Pavel Smrž & David Tugwell. 2004. The Sketch Engine. *Proceedings of the 11th EURALEX International Congress*. 105–116.
Kilgarriff, Adam, Vít Baisa, Jan Bušta, Miloš Jakubíček, Vojtěch Kovář, Jan Michelfeit, Pavel Rychlý & Vít Suchomel. 2014. The Sketch Engine: Ten years on. *Lexicography* 1: 7–36.
Lenerz, Jürgen. 1977. *Zur Abfolge nominaler Satzglieder im Deutschen*. Tübingen: Gunter Narr.
Petersen, Hjalmar P. 2002. Quirky case in Faroese. *Fróðskaparrit* 50: 63–76.
Platzack, Christer. 1999. The subject of Icelandic psych-verbs: A minimalistic account. *Working Papers in Scandinavian Syntax* 64: 103–116.
Primus, Beatrice. 1994. Grammatik und Performanz: Faktoren der Wortstellungsvariation im Mittelfeld. *Sprache und Pragmatik* 32: 39–86.
Primus, Beatrice. 2012. *Semantische Rollen*. Heidelberg: Carl Winter.
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 Stolz, Christel & Stolz, Thomas (eds.), *From Africa via the Americas to Iceland*, 215–320. Bochum: Dr. N. Brockmeyer.
Schlesewsky, Matthias & Ina Bornkessel. 2006. Context-sensitive neural responses to conflict resolution: Electrophysiological evidence from subject–object ambiguities in language comprehension. *Brain Research* 1098: 139–152.
Shannon, Thomas F. 2000. On the order of (pro)nominal arguments in Dutch and German. In Thomas F. Shannon & Johan P. Snapper (eds.), *The Berkeley Conference on Dutch Linguistics 1997: The Dutch Language at the Millennium*, 145–196. New York: University Press of America.
Siewierska, Anna. 1993. Syntactic weight vs. information structure and word order variation in Polish. *Journal of Linguistics* 29: 233–265.
Sigurðsson, Halldór Ármann. 1990–91. Beygingarsamræmi [Agreement]. *Íslenskt mál og almenn málfræði* 12–13: 31–77.
Sigurðsson, Halldór Ármann. 2002. Agree and agreement: Evidence from Germanic. *Working Papers in Scandinavian Syntax* 70: 101–156.
Somers, Joren. 2021. Onpersoonlijke datief-nominatiefwerkwoorden in het hedendaagse Duits: Definitie, inventaris en semantische classificatie. *Handelingen van de Koninklijke Zuid-Nederlandse Maatschappij voor Taal- en Letterkunde en Geschiedenis* 75: 211–238.
Somers, Joren. 2023. An evaluation of the Dat-Nom and the Dat-Nom/Nom-Dat hypotheses for five German verbs of success and failure. *Papers of the Linguistic Society of Belgium* 17: 54–75.
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.
Somers, Joren, Gard B. Jenset & Jóhanna Barðdal. 2024. Argument structure constructions in competition: The Dat-Nom/Nom-Dat alternation in Icelandic. *Nordic Journal of Linguistics*: online first article. DOI: https://doi.org/10.1017/S0332586524000039.
Somers, Joren, Torsten Leuschner, Ludovic De Cuypere & Jóhanna Barðdal. 2024. A corpus-based analysis of the Dat-Nom/Nom-Dat alternation in German. Accepted.
Tagliamonte, Sali A. & Harald R. Baayen. 2012. Models, forests, and trees of York English: Was/were variation as a case study for statistical practice. *Language Variation and Change* 24(2): 135–178.
Thráinsson, Höskuldur. 2007. *The Syntax of Icelandic*. Cambridge: Cambridge University Press.
Thráinsson, Höskuldur, Hjalmar P. Petersen, Jógvan í Lon Jacobsen & Zakaris S. Hansen. 2012. *Faroese: An Overview and Reference Grammar*. 2nd edn. Tórshavn: Faroe University Press.
Verhoeven, Elisabeth. 2015. Thematic asymmetries do matter! A corpus study of German word order. *Journal of Germanic Linguistics* 27(1): 45–104.
Wood, Jim & Halldór Ármann Sigurðsson. 2014. *Let*-causatives and (a)symmetric DAT-NOM constructions. *Syntax* 17(3): 269–298.
Wunderlich, Dieter. 2009. The force of lexical case: German and Icelandic compared. In Kristin Hanson & Sharon Inkelas (eds.), *The Nature of the Word: Studies in Honor of Paul Kiparsky*, 587–620. Cambridge, Ma: MIT Press.

Chapter 6
A corpus-based analysis of the Dat-Nom/ Nom-Dat alternation in German

A subgroup of German Nom-Dat verbs have received considerable attention in the literature due to the propensity of the dative to occur preverbally, which is unexpected on an object analysis of the dative (see references below). Here we argue for an alternative analysis, namely that the relevant verbs alternate between two different argument structures, Dat-Nom and Nom-Dat, and hence that either argument, the dative or the nominative, may be the syntactic subject. Earlier studies have shown that topicalization of direct arguments is found in ca. 4–12% of the cases in German texts. For comparison, we have extracted 13,000 tokens of 76 verbs from the deTenTen13 corpus and coded them for ten different variables. Our findings support an alternating Dat-Nom/Nom-Dat analysis for these verbs, as 42% of the tokens instantiate the Dat-Nom order and the remnant 58% instantiate the Nom-Dat order. In contexts with full NPs only, the share of Dat-Nom tokens is even higher, 46% compared to 54% Nom-Dat, which altogether excludes a topicalization analysis of the Dat-Nom word order. In order to throw further light upon this alternation, we carry out a multivariate analysis which confirms the effect of topicality, definiteness, length, the animacy of the dative and the inanimacy of the nominative.

6.1 Introduction

Ever since Primus (1994), it has been observed that certain verbs selecting for both a dative and a nominative argument may instantiate two diametrically opposite word orders, both of which appear to be equally neutral. By "equally neutral" we mean that the two word orders neither involve topicalization nor fronting for focus purposes. Typical (constructed) examples involve the German verb *gefallen* 'like, please':

(1)	a.	*Diese Idee*	*gefiel*	*dem Mann*	*richtig*	*gut.*
		this.NOM idea	liked	the.DAT man	really	well
		'This idea really was to the man's liking.'				
(1)	b.	*Dem Mann*	*gefiel*	*diese Idee*	*richtig*	*gut.*
		the.DAT man	liked	this.NOM idea	really	well
		'The man really liked this idea.'				

Corresponding verbs in Icelandic also appear to alternate between two different argument structures with an associated alternation in word order (Barðdal 1999, 2001), as shown here with the verb *falla vel* 'like, please', a cognate of German *gefallen* in (1) above:

(2)	a.	*Þessi hugmynd*	*féll*	*manninum*	*virkilega*	*vel.*
		this.NOM idea	liked	man.the.DAT	really	well
		'This idea really was to the man's liking.'				
(2)	b.	*Manninum*	*féll*	*þessi hugmynd*	*virkilega*	*vel.*
		man.the.DAT	liked	this.NOM idea	really	well
		'The man really liked this idea.'				

While the dative of the Dat-Nom argument structure and the nominative of the Nom-Dat argument structure each passes the subject tests in Icelandic (Barðdal 1999, 2001), this type of analysis has not yet gained a foothold in the German scholarship. Traditionally, verbs like in (1) have been regarded as Nom-Dat verbs with a high frequency of topicalization of the dative to first position, thus resulting in Dat-Nom word order (Bader and Häussler 2010; Haspelmath 2001; Kempen and Harbusch 2005; Verhoeven 2015, *inter alia*). Such an analysis is not satisfactory, however, since topicalization is an optional (and rare) response to information-structural pressures (Engel 1970: 90; Rosengren 1993: 274, *inter alia*), as is confirmed empirically by Sahel and Jonischkait (2008), who show that clauses containing two arguments topicalize the object (direct object, indirect object and prepositional object) in a mere 12% of the cases. More recently, Somers and Barðdal (2023) have shown that topicalization of dative objects in German is substantially lower than Sahel and Jonischkait's baseline, since verbs like *helfen* 'help' only front the dative in 4–7% of the cases.

Under more formal approaches (Bayer 2004; Fanselow 2002; Haider 2005, 2010; Schlesewsky and Bornkessel 2006; Wunderlich 2009), it has become customary to view these verbs as selecting for the Dat-Nom argument structure and to label these as *Dat-Nom* verbs. Still, the second argument of the argument structure, i.e. the nominative, has been analysed as the syntactic subject of such verbs (see references above). Clearly, such an analysis is not satisfactory either, since the subject tests do not target the second argument, but the first one (cf. Barðdal 2023: Ch. 2).

In the present paper, we argue in favour of a third hypothesis, namely that these verbs really alternate between two diametrically opposite argument structures, Dat-Nom and Nom-Dat. In other words: when the relevant verbs instantiate the Dat-Nom word order, the dative is the first argument of the argument structure and the nominative the second argument, whereas the nominative is the first argument of the argument structure and the dative the second argument when these verbs instantiate the Nom-Dat word order. Such an analysis is implicit in Primus's (1994, 2012) account, on the assumption that both word orders are indeed equally neutral. The relevant alternation has also been observed by Haspelmath (2001) for German and several other modern Indo-European languages. An analysis in terms of an alternation between Dat-Nom and Nom-Dat has also been suggested for German in Eythórsson and Barðdal (2005), and specifically argued for by Barðdal, Eythórsson and Dewey (2014, 2019), Rott (2016), Barðdal (2023: Ch. 6) and Somers and Barðdal (2023). The analysis of Barðdal, Eythórsson & Dewey (2014, 2019) is based on several subject tests such as reflexivization, conjunction reduction, raising-to-subject, raising-to-object and control infinitives.

However, what is missing from the literature is a systematic study of an additional subject test, namely preverbal position of the subject in verb-second clauses. We aim to bridge this gap by investigating the word order variation attested by a broad sample of 76 hypothesised alternating Dat-Nom/Nom-Dat verbs in German, selected on the basis of the research presented in Somers (2021). This involves 200 tokens of each of the relevant 76 verbs, extracted from the deTenTen13 corpus, thus amounting to a total of 13,000 observations. In addition, we subject our material to a multivariate analysis in order to uncover the factors controlling the alternation.

We start in Section 6.2 below with a description of the state-of-the-art in the field, including a comprehensive overview of the relevant structures, a terminological discussion and an outline of the semantic and syntactic properties of Dat-Nom/Nom-Dat verbs in German. Section 6.3 describes the dataset, how the relevant material has been extracted from the deTenTen13 corpus, and how it has been cleaned and then annotated for ten different variables. In Section 6.4 we describe our statistical findings: Section 6.4.1 discusses the correlation between case marking and six particular properties, and Section 6.4.2 presents the results of our logistic regression analysis. Section 6.5 summarises and concludes our paper.

6.2 State-of-the-art

Our research contributes to a terminologically diverse area of research. Section 6.2.1 therefore contrasts the notion of Dat-Nom/Nom-Dat verbs with other related notions in the field such as experiential verbs, psych verbs and unaccusative verbs. In Section 6.2.2 we flesh out the semantic and syntactic properties of our hypothesised Dat-Nom/Nom-Dat verbs in German.

6.2.1 The scope of Dat-Nom/Nom-Dat verbs

Research on non-canonical case marking often revolves around the study of so-called 'psych verbs', 'experiential verbs', and 'unaccusative verbs'. Several scholars have observed similarities between these classes, yet in an attempt to emphasise their commonalities, they have sometimes failed to acknowledge their differences. The present section bridges this gap. For non-canonically case-marked verbs licensing both a dative and a nominative argument, we argue that the classes of 'psych verbs', 'experiential verbs', and 'unaccusative verbs' are in a meronymic relation to each other. At the same time, we maintain that each of these verb classes may be subsumed under the umbrella of 'Dat-Nom/Nom-Dat verbs', which additionally contains verbs not belonging to any of the remaining semantic categories.

The structure of this section is as follows: Section 6.2.1.1 discusses the labels 'psych verbs' and 'experiential verbs', Section 6.2.1.2 introduces the overarching label of 'Dat-Nom/Nom-Dat verbs', and Section 6.2.1.3 examines the status of so-called 'unaccusative verbs'. Section 6.2.1.4 summarises the main findings.

6.2.1.1 Psych verbs and experiential verbs

The term 'psych verbs' has been used rather loosely in the literature to refer to verbs expressing some sort of psychological awareness. Depending on the verb in question, the psych participant may be realised either in the nominative, as in (3), the accusative, as in (4), or the dative, as in (5):

(3) *Ich* *liebe* *diesen Duft.*
I.NOM love this.ACC smell
'I love this smell.'

(4) *Mich* *wurmt* *dieses Problem* *schon* *länger.*
me.ACC bothers this.NOM problem already longer
'I have been bothered by this problem for a long time.'

(5) *Diese Idee* *gefiel* *mir* *richtig gut.*
this.NOM idea liked me.DAT really well
'I liked this idea a lot.'

Psych verbs are not limited to two-participant constructions as in (3–5). Two other construction types frequently observed in German are compositional predicates, as in (6), and constructions with a mandatory prepositional phrase, as in (7):

(6) *Mir* *ist* *angst.*
me.DAT is frightened
'I am scared.'

(7) *Mir* *graut* *vor dem Arbeitsberg.*
me.DAT dreads before the pile.of.work.DAT
'I am dreading the pile of work.'

The focus of the present paper is on verbs selecting for both a dative and a nominative argument, i.e. constructions like the one in (5). In order to set these apart from the remaining constructions containing psych verbs (as in 3–4 and 6–7), they will be referred to as **Dat-Nom/Nom-Dat psych verbs**.

Interestingly, several scholars have used the label 'psych verbs' interchangeably with the label 'experiential verbs' (Anagnostopoulou 1999; Haspelmath 2001; Verhoeven 2015, *inter alia*). This perceived synonymy is rather odd, as experience need not affect a participant mentally. Consider the Dat-Nom/Nom-Dat verbs *munden* 'taste good', which denotes perception, or *aufstoßen* 'make (sb) burp', which refers to a bodily state: these surely make reference to human experience, but are certainly not psychological verbs.

Klein and Kutscher (2002) also observe the alleged synonymy of the terms 'psych verbs' and 'experiential verbs', prompting them to distinguish between psych verbs in the narrow sense (which only comprise verbs of emotion), and psych verbs in the broad sense (which are synonymous to experiential verbs). Klein and Kutscher suspect that the frequent interchange of the terms 'psych verbs' and 'experiential verbs' is due to the fact that verbs of emotion account for the largest subgroup of experiential predicates. Returning again to verbs licensing both a dative and a nominative argument, Somers (2021) has shown for Present-Day German that this is indeed the case: nearly 50 different German Dat-Nom/Nom-Dat verbs (out of a total of 146) may be used to denote emotion in at least one of their senses. The second largest subgroup of Dat-Nom/Nom-Dat verbs, which is made up of verbs of cognition, only accounts for 21 different lemmas.

Now, if psych verbs are essentially verbs of emotion, how exactly should one define 'experiential verbs'? According to Dowty (1991), an experiencer non-volitionally engages in an act of sentience or perception. Building on his definition, experiencers are licenced by verbs of emotion (*gefallen* 'please, like', *ekeln* 'disgust, nauseate'), verbs of cognition (*klarwerden* 'become clear', *einleuchten* 'be clear'), verbs of perception (*auffallen* 'strike,

notice', *schmecken* 'taste; have a particular taste (lit.)'), verbs referring to bodily states (*aufstoßen* 'make (sb) burp', *schwindeln* 'feel dizzy'), and verbs of evidentiality (*scheinen* 'seem, appear', *erscheinen* 'seem, appear'). Observe that all of these predicates denote activities that are prototypically carried out by humans. This has led some scholars to believe that experiencers are animate *by definition* (see Kutscher 2009; Rott 2013; Verhoeven 2009, 2015). Such a presupposition is clearly an oversimplification, as it fails to acknowledge the role of creativity (and of metaphor in particular) in language use. We return to this point in Section 6.4.1.5.

The fact that experiencers are prototypically, yet not necessarily, animate has probably also facilitated the perceived synonymy between the terms 'experiential verb' and 'psych verb': the capacity to experience emotion is what sets humans, and animate beings in general, apart from objects and abstract concepts. In the current paper, we adopt Klein and Kutscher's (2002) narrow definition of psych verbs as verbs of emotion (cf. also Viti 2016). Experiential verbs, by contrast, are defined in line with Dowty's (1991) definition of the experiencer role, thus comprising verbs of emotion alongside verbs of cognition, verbs of perception, bodily state verbs, and verbs of evidentiality. As such, Dat-Nom/Nom-Dat psych verbs are fully subsumed under the category of **Dat-Nom/Nom-Dat experiential verbs**.

6.2.1.2 Dat-Nom/Nom-Dat verbs

Apart from the confusion surrounding the classes of Dat-Nom/Nom-Dat psych verbs and Dat-Nom/Nom-Dat experiential verbs, there is yet another inconsistency in the literature, namely the use of these terms by various scholars in reference to verbs that are neither experiential nor psychological. By doing so, the scholars in question acknowledge, at least implicitly, the inability of both labels to capture the richness of a category we collectively refer to as **Dat-Nom/Nom-Dat verbs**.

Within the German scholarship, one of the first to unduly stretch the term 'psych verb' was Lenerz (1977: 114), who includes in this class the following verbs: *folgen* 'succeed, follow' (a happenstance verb), *fehlen* 'lack' (a verb of possession), *gelingen* 'succeed' (a success verb), *gehören* 'belong' (a verb of belonging), and also *zukommen* 'be entitled to; be allowed to', which is both a verb of gain and benefit as well as a verb of modality. Haspelmath (2001) expands the list even further with verbs of attitude, like *genügen* 'be enough, be sufficient'.

While all the aforementioned verbs share important properties, labelling them as 'psych verbs' or 'experiential verbs' is inappropriate and confusing. In line with Barðdal (2001; 2004) and Barðdal, Eythórsson and Dewey (2019), we therefore refer to the category as a whole as 'Dat-Nom/Nom-Dat verbs'. The interrelation between the categories Dat-Nom/Nom-Dat psych verbs, Dat-Nom/Nom-Dat experiential verbs, and Dat-Nom/Nom-Dat verbs may be captured by the following implicational hierarchy:

(8) Dat-Nom/Nom-Dat psych verbs < Dat-Nom/Nom-Dat experiential verbs < Dat-Nom/Nom-Dat verbs

According to (8), all Dat-Nom/Nom-Dat psych verbs are Dat-Nom/Nom-Dat experiential verbs, and all Dat-Nom/Nom-Dat experiential verbs are Dat-Nom/Nom-Dat verbs – but not vice versa. The specific properties of German Dat-Nom/Nom-Dat verbs, including the semantic fields to which they belong, are discussed in Section 6.2.2.

6.2.1.3 Unaccusative verbs

Attempts to shed light on the syntactic properties of Dat-Nom/Nom-Dat verbs have shown that some of these verbs are in fact unaccusatives, as they form the perfect tense with *sein* 'be', as opposed to *haben* 'have'. In their seminal paper on non-canonical case marking in Present-Day Italian, Belletti and Rizzi (1988) argue that forming the perfect tense with *essere* 'be' is a core property of so-called '*piacere* verbs', i.e. Dat-Nom/Nom-Dat psych verbs. At least for Present-Day German, Belletti and Rizzi's (1988) proposal raises three issues.

First, under Belletti and Rizzi's account, all Dat-Nom/Nom-Dat verbs would seem to be psych verbs in the narrow sense, i.e. verbs of emotion. The preceding sections have shown that this is not the case: some are actually Dat-Nom/Nom-Dat experiential verbs, whereas others, like verbs of success, verbs of speaking, or happenstance verbs, solely belong to the overarching class of Dat-Nom/Nom-Dat verbs.

Secondly, Belletti and Rizzi claim that Dat-Nom/Nom-Dat psych verbs are invariably unaccusative verbs. At least for German, such a characterisation is too reductive. Consider verbs like *wohltun* 'do sb. good', *grauen* 'dread, be afraid of', or *stinken* 'enrage': these pertain to emotion, but they are unergatives, not unaccusatives. The fact that German psych verbs do not always form the perfect tense with *sein* has also been observed by den Besten (1985: 46). He sidesteps this incongruity by stating that auxiliary selection as a diagnostic for unaccusativity is less dependable for German than for Italian. As a result, some verbs forming the perfect tense with *haben* 'have' (like *gefallen* 'please, like') may still be considered unaccusatives. Den Besten does not elaborate on the purported invalidity of the auxiliary test in German, which is peculiar, because auxiliary selection has been considered one of the core properties of unaccusativity (Alexiadou et al. 2004: 5; Haider 1985: 236; Kallulli 2007: 777, *inter alia*).

Third, not every dative unaccusative verb is a psych verb. Consider the verbs *auffallen* 'strike, notice' and *entschlüpfen* 'escape (words)'. Both are unaccusatives, yet *auffallen* is a verb of cognition and *entschlüpfen* in this particular sense concerns speaking. As such, the former belongs to the class of Dat-Nom/Nom-Dat experiential verbs and the latter to the overarching class of Dat-Nom/Nom-Dat verbs.

Thus, the relation between unaccusative verbs on the one hand, and Dat-Nom/Nom-Dat psych verbs, Dat-Nom/Nom-Dat experiential verbs, and Dat-Nom/Nom-Dat verbs, on the other, is one of partial overlap: the class of unaccusatives cross-cuts each of the remaining three categories, but it does not fully coincide with any of these.

Finally, it should be mentioned that certain German unaccusatives allow for their case frame to be expanded upon by a free dative (Grimm 1837; Gutzmann 2007; Hole 2014; Schmid 1988, *inter alia*). Wegener (1985: 200–201) shows that this so-called 'ergative construction' may be filled by over 50 verbs, which all denote an unfavourable change of state. Two examples include:

(9)	*Mir*	*ist*	*der Kuchen*	*angebrannt.*
	me.DAT	is	the.NOM cake	burnt
	'The cake burnt on me.'			
(10)	*Ihm*	*ist*	*die Tante*	*gestorben.*
	him.DAT	is	the.NOM aunt	died
	'His aunt died on him.'			

In line with Somers (2021), the present paper disregards such ergative structures, as the dative is not part of the respective main verb's subcategorization frame. Hence, the verbs in question are not subsumed under the label **dative unaccusative verbs**, which we take to refer only to unaccusative verbs that license both a nominative as well as a dative.

6.2.1.4 Summary

The present section has argued that the classes of Dat-Nom/Nom-Dat psych verbs, Dat-Nom/Nom-Dat experiential verbs, dative unaccusative verbs and the semantically more general Dat-Nom/Nom-Dat verbs share a meronymic relation. At the heart of these four categories lies the class of *piacere* verbs, which displays all properties exhibited by each of the remaining classes. Figure 6.1 visualises the interrelation between the classes:

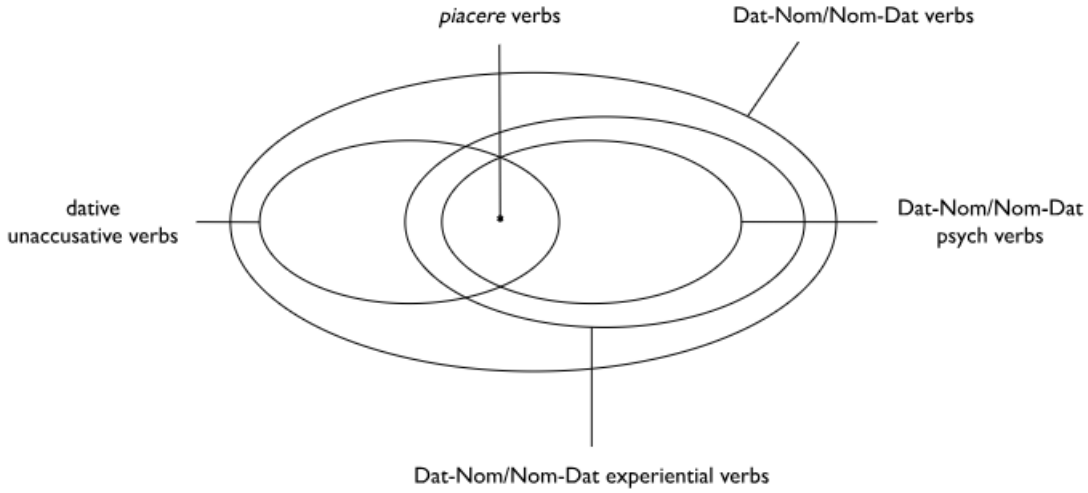


Figure 6.1 The interrelation between Dat-Nom/Nom-Dat psych verbs, Dat-Nom/Nom-Dat experiential verbs, dative unaccusative verbs and *piacere* verbs. Diagram not to scale

Now that we have cleared the terminological fog, we move on to Section 6.2.2, which examines the semantic and syntactic properties of the overarching class of Dat-Nom/Nom-Dat verbs.

6.2.2 Dat-Nom/Nom-Dat verbs: definition

Dat-Nom/Nom-Dat verbs display a number of properties that set them apart from other verbs, in particular from ordinary Nom-Dat verbs like *gehorchen* 'obey', *vertrauen* 'have faith in', or *glauben* 'trust'. Section 6.2.2.1 provides an overview of the semantic properties of Dat-Nom/Nom-Dat verbs while Section 6.2.2.2 elaborates on their syntactic peculiarities.

6.2.2.1 Semantic properties

Semantically, German Dat-Nom/Nom-Dat verbs exhibit three fundamental properties: (1) they are low in agentivity, (2) they display a thematic asymmetry between the dative and the nominative, and (3) they express a meaning that is compatible with one of 16 semantic domains. Each of these properties will be discussed in turn.

First, various scholars have noted that Dat-Nom/Nom-Dat verbs are low in agentivity, as they constitute a deviation from the transitive prototype, which postulates that a highly individuated agent willingly targets a non-individuated, inert patient (Barðdal 2004; Barðdal and Eythórsson 2009; Barðdal et al. 2016; Hopper and Thompson 1980; Somers 2021, *inter alia*). This is a property that Dat-Nom/Nom-Dat verbs share with all other predicates with non-canonical case marking (Haspelmath 2001). Hence, it qualifies as a necessary, but not sufficient, diagnostic for Dat-Nom/Nom-Dat verbs.

Second, Dat-Nom/Nom-Dat verbs display a mismatch between case marking and thematic roles. According to Primus (1999), cases and thematic roles may each be ranked hierarchically, which gives rise to the case hierarchy, presented under (11), and the thematic hierarchy, presented under (12):

(11) **Case hierarchy**: nominative > accusative > dative
(12) **Thematic hierarchy**: proto-agent > proto-recipient > proto-patient

In the case of Dat-Nom/Nom-Dat verbs, the two hierarchies are in conflict with one another. This means that the dative, being the lower-ranked case, is mapped onto the argument corresponding with the proto-agent, while the nominative, being the higher-ranked case, is mapped onto the argument bearing the lower-ranked proto-patient role. Such a mismatch between case marking and thematic roles is not a commonality of all German verbs licencing a dative and a nominative argument. Thus, this form-meaning mismatch clearly sets Dat-Nom/Nom-Dat verbs apart from Nom-Dat verbs like *gehorchen* 'obey', *vertrauen* 'have faith in', or *glauben* 'trust'. As a consequence, the mismatch between case marking and thematic roles may also be considered a necessary diagnostic for Dat-Nom/Nom-Dat verbs.

Third, German Dat-Nom/Nom-Dat verbs pertain to one of the following 16 semantic domains: emotion, cognition, gain and benefit, happening, hindrance, attitude, possession, perception, bodily states, modality, ontological states, evidentiality, success, belonging, speaking, and natural occurrences. These domains were first discussed by Barðdal et al. (2012) in the context of Proto-Indo-European, by Barðdal et al. (2016) in the context of Proto-Germanic (an early attempt at a classification can be found already in Barðdal 2004), and by Somers (2021) to Present-Day German.

Table 6.1 presents an overview of these semantic domains (column 1), each of which is illustrated by means of an example verb (column 2) and followed by a translation of the intended verb sense (column 3). The domains pertaining to bodily states (domain 11) and natural occurrences (domain 16) are asterisked, as the present study coincidentally does not include any such verbs. Nevertheless, they have been listed for the sake of completeness.

Table 6.1 Typology comprising 16 semantic domains covered by German Dat-Nom/Nom-Dat verbs. Verbs pertaining to the asterisked domains are not represented in the dataset of the present study

	semantic class	**example**	**gloss**
(1)	emotion	*gefallen*	'please, like'
(2)	cognition	*dämmern*	'dawn on'
(3)	gain and benefit	*zuwachsen*	'receive'
(4)	happening	*passieren*	'happen to'
(5)	hindrance	*abhandenkommen*	'lose'
(6)	attitude	*geziemen*	'befit'
(7)	possession	*fehlen*	'lack'
(8)	perception	*munden*	'taste good'
(9)	bodily states	*aufstoßen*	'make (sb) burp'
(10)	modality	*obliegen*	'be someone's duty'
(11)	ontological states*	*eignen*	'be peculiar to'
(12)	evidentiality	*scheinen*	'seem, appear'
(13)	success	*gelingen*	'succeed'
(14)	belonging	*zugehören*	'belong to'
(15)	speaking	*entfahren*	'escape (words)'
(16)	natural occurrences*	*vergehen*	'pass by (time)'

6.2.2.2 Syntactic properties

Apart from the aforementioned semantic properties, several scholars have drawn attention to the syntactic peculiarities of German Dat-Nom/Nom-Dat verbs. These essentially demonstrate that such verbs are in fact 'alternating predicates', which means that either the dative or the nominative may be analysed as the subject, but obviously not at the same time (cf. Barðdal 2006, 2023: Ch. 6; Barðdal, Eythórsson and Dewey 2014, 2019; Eythórsson and Barðdal 2005; Somers and Barðdal 2023). This means that the nominative and the dative independently pass an array of subject tests, namely: first position in declarative clauses, subject-verb inversion, conjunction reduction, clause-bound reflexivisation, raising-to-subject, raising-to-object, and control infinitives.

The most widely discussed fact is that either argument of a Dat-Nom/Nom-Dat verb, i.e. the dative or the nominative, may take initial position in declarative clauses without one order being more marked than the other (Barðdal 2001a, 2004, 2006, 2023; Barðdal, Eythórsson and Dewey 2014, 2019; Eythórsson and Barðdal 2005; Haspelmath 2001; Primus 1994, 2012; Rott 2016, *inter alia*). The following examples illustrate this phenomenon for the verbs *gelingen* 'succeed' (13a–b), *auffallen* 'strike, notice' (14a–b), and *schmecken* 'taste' (15a–b).

(13) a. *Der A-Jugend* *gelingt* *der Aufstieg in die Sonderklasse.*
the.DAT A-youth succeeds the.NOM rise into the top class
'The A-youth has been successfully promoted to rank A.'

(13) b. *Der Start ins zweite Tagesspiel* *gelang*
the.NOM start into the second match of the day succeeded
den Damen *besser.*
the.DAT ladies better
'The ladies got off to a better start into the second match of the day.'

(14) a. *Auch der Dame mit der Feile* *war* *der Fleck* *aufgefallen.*
also the.DAT lady with the file was the.NOM stain noticed
'The lady with the file ahd also noticed the stain.'

(14) b. *Der 15-jährige* *ist* *der Polizei* *bereits* *aufgefallen.*
the.NOM 15-year-old is the.DAT police already noticed
'The police have already noticed the 15-year-old.'

(15) a. *Den Jungs* *schmeckte* *das traditionalle Schaschlik.*
the.DAT boys tasted the.NOM traditional.NOM satay
'The boys enjoyed the traditional satay.'

(15) b. *Dieser Espresso* *schmeckt* *meinem Mann* *am besten.*
this.NOM espresso tastes my.DAT husband to best
'My husband likes this espresso best.'

A second subject test which either argument of Dat-Nom/Nom-Dat verbs has been shown to pass is subject-verb inversion (Barðdal 2006, 2023: Ch. 6.):

(16) a. *Gelingt* *der A-Jugend* *der Aufstieg in die Sonderklasse?*
succeeds the.DAT A-youth the.NOM rise into the top class
'Has the A-youth been successfully promoted to rank A?'

(16) b. *Gelingt* *der Aufstieg in die Sonderklasse* *der A-Jugend?*
succeeds the.NOM rise into the top class the.DAT A-youth
'Has the A-youth been successfully promoted to rank A?'

(17) a. *War* *der Dame mit der Feile* *der Fleck* *aufgefallen?*
was the.DAT lady with the file the.NOM stain noticed
'Had the lady with the file noticed the stain?'

(17) b. *War* *der Fleck* *der Dame mit der Feile* *aufgefallen?*
was the.NOM stain the.DAT lady with the file noticed
'Had the lady with the file noticed the stain?'

(18) a. *Schmeckte* *den Jungs* *das traditionelle Schaschlik?*
tasted the.DAT boys the.NOM traditional.NOM satay
'Did the boys enjoy the traditional satay?'

(18) b. *Schmeckte* *das traditionelle Schaschlik* *den Jungs?*
tasted the.NOM traditional.NOM satay the.DAT boys
'Did the boys enjoy the traditional satay?'

However, the word order test has been systematically ignored in research on subjecthood in German (Reis 1982), which has instead focused on conjunction reduction and control infinitives (Bayer 2004; Fanselow 2002; Haider 2005, 2010; Pankau 2016; Sigurðsson 1989, 1992; Wunderlich 2009; Zaenen, Maling and Thráinsson 1985). We therefore aim to document that (a) word order may be used to distinguish between subjects and objects in German, and (b) that either argument of Dat-Nom/Nom-Dat verbs, the dative and the nominative, pass this test.

In (13–15) above, we show with two sets of examples that either argument, the dative or the nominative, of the verbs *gelingen* 'succeed', *auffallen* 'strike, notice' and *schmecken* 'taste', may occur in first position in ordinary declarative clauses. Such clauses represent neutral word order, with no other clausal element being topicalized to first position. The same is true for the yes/no questions in (16–18), where either argument may invert with the verb, attesting to the alternating nature of these verbs. Since our goal is first and foremost to argue for the alternating status of these verbs and to unearth the factors relevant for speakers' choices when using Dat-Nom/Nom-Dat verbs in Present-Day German, we do not pursue any of the remaining subject tests in this article.

6.3 Methodology

This study takes as its starting point the list of 145 German Dat-Nom/Nom-Dat verbs compiled by Somers (2021). For each of these verbs, a lemmatised search was run in the German Web 2013, i.e. deTenTen13 corpus, accessed through the Sketch Engine interface (Kilgarriff et al. 2004; Kilgarriff et al. 2014).[1] Data inspection, cleaning and annotation proceeded based on samples of 10,000 randomised tokens per verb, which were downloaded and subsequently accessed offline. Each of these procedural stages is discussed in turn below.

[1] Available at www.sketchengine.eu/detenten-german-corpus.

6.3.1 Data inspection

In order to estimate the total frequencies of the Dat-V-Nom and the Nom-V-Dat word order patterns, each 10,000 token verb-sample was first subjected to a data inspection stage. This entails that the first 300 tokens were checked for at least seven eligible instantiations of either the Dat-V-Nom or the Nom-V-Dat word order pattern. For reasons of feasibility, if fewer than seven eligible tokens were found, the verb was excluded from the present study in order to avoid having to clean all 10,000 tokens to reach the ultimate goal of 200 eligible tokens per verb type. However, if the frequency threshold is set at seven tokens, the expected frequency of both word order patterns combined increases considerably, so that only 8,500 tokens have to be cleaned at most. That is still a significant amount, but it is nevertheless a reasonable compromise between feasibility, on the one hand, and generating an analysis with sufficient statistical power, on the other. Based on the methodological choices discussed above, the following 76 verbs were retained (see Appendix 2 for translations of the following verbs into English):

> *abhandenkommen – angehören – aufdämmern – auffallen – aufstoßen – begegnen – beifallen – dämmern – dazwischenkommen – drohen – dünken – einfallen – einleuchten – entfahren – entgegenschallen – entgegenschlagen – entgehen – entgleiten – entschlüpfen – ermangeln (an) – erscheinen – fehlen – fernliegen – folgen – gebrechen (an) – gebühren – gefallen – gelingen – gelten – genügen – geziemen – glücken – grauen – grausen – guttun – imponieren – klarwerden – konvenieren – leichtfallen – leidtun – mangeln (an) – missfallen – missglücken – misslingen – missraten – munden – nahegehen – nützen – obliegen – passieren – pressieren – reichen – scheinen – schlechtgehen – schmecken – schwanen – schwerfallen – stinken – unterlaufen – verlorengehen – vorkommen – vorschweben – widerfahren – widerstreben – winken – wohltun – zufallen – zufließen – zugehören – zugutekommen – zupasskommen – zusetzen – zustoßen – zuteilwerden – zuwachsen*

For each of the 76 types cited above, we endeavoured to collect a total of 200 eligible tokens per type. For 60 verbs, this threshold proved to be unproblematic, but for the remaining 16 verbs, we ran into data sparsity issues. The respective number of tokens which each of these verbs returned is presented in Figure 6.2.

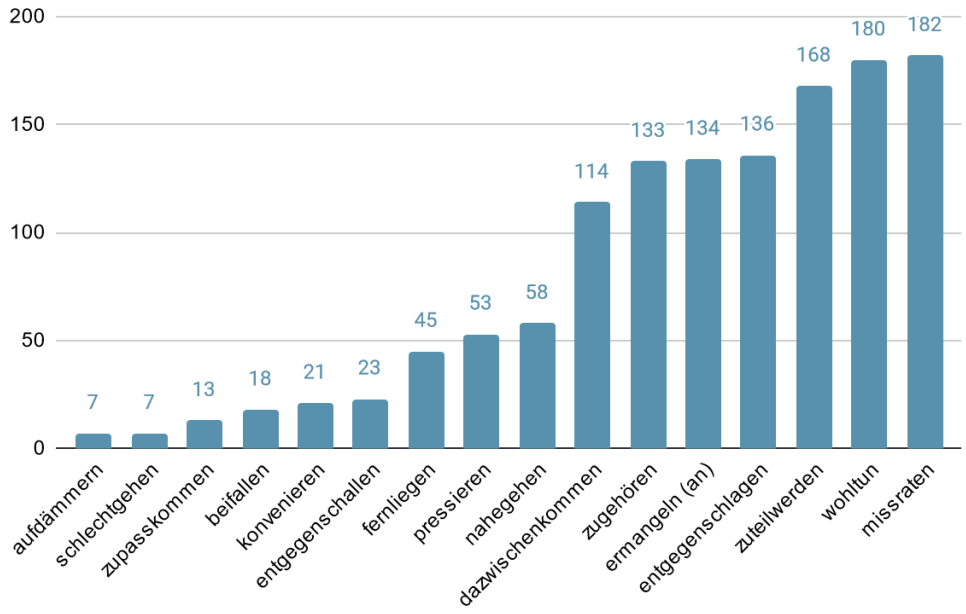


Figure 6.2 Number of data points for verbs yielding fewer than 200 eligible tokens

Thus, the full dataset comprises 60 times 200 tokens plus the remaining 1,292 tokens for the 16 verbs presented in Figure 6.2. This equals a total of 13,292 tokens.

6.3.2 Data cleaning

This study focuses on the influence of case marking on linearisation in German prefield contexts, i.e. contexts in which the main verb is flanked by two nominal arguments. Each of the argument slots may consequently be filled by either a pronoun or a full NP. Tokens with clausal constituents have been barred from study. As a result, tokens may instantiate one of four potential configurations: NP-V-NP, Pro-V-Pro, NP-V-Pro, and Pro-V-NP. Following Heylen (2005: 103), tokens containing a pronoun conjoined with either another pronoun, as in (19), or a full noun phrase, as in (20), have been excluded. The rationale behind this decision is that a conjoined pronoun loses its pronominal status (Heylen 2005: 103).

(19) *Lennox* *ist* *ein echter Multitasker,*
Lennox is a real multitasker
dem *entgeht* *nichts und niemand.*
the.DAT misses nothing and nobody.NOM
'Lennox is a veritable multitasker, he misses nothing and nobody.'

(20) *Die Auszeit,* *sagte* *Moutinho,* *habe* *ihm und dem Klub*
the.NOM time-out said Mourinho have.SBJV him.DAT and the.DAT club
gutgetan.
'Mourinho said that the time-out had benefited both him and the club.'

Another prerequisite for a token to be included is that it represents an eligible verb sense. As argued by Somers (2021), subject-like datives are not an inherent property of a particular verb. Instead, they are associated with specific verb senses. Consider the following examples, which both contain the verb *entgehen* 'escape; miss out on'. Example (21) instantiates the sense 'escape' and example (22) instantiates the sense 'miss out on':

(21) *Niemand* *kann* *seinem Geschick* *entgehen.*
nobody.NOM can his.DAT fate escape
'No one can escape his fate.'

(22) *Dem Nichtraucher* *entgeht* *diese Pause.*
the.DAT non-smoker misses.out.on this.NOM break
'The non-smoker misses out on this break.'

Examples (21) and (22) illustrate the importance of verb semantics in assessing the eligibility of a particular token (cf. also Johnson et al. 2019). It stands to reason that (21), which is more agentive than (22), contains a dative that is more object-like than the one in (22). As a result, only (22) is potentially associated with the Dat-Nom/Nom-Dat alternation, whereas (21) is invariably associated with the non-alternating Nom-Dat argument structure. Thus, only verb senses like the one in (22) are included for study.

Some polysemous verbs like *begegnen* 'come across; experience, be confronted with' have more than one eligible sense, as illustrated by examples (23–24). If such is the case, all eligible senses are included for study.

(23) *Neue Tier- und Pflanzearten* *begegnen* *uns* *heute*
new.NOM animal and plant species.NOM meet us.DAT today
fast überall.
almost everywhere
'Today, we come across new animal and plant species almost everywhere.'

(24) *Langeweile* *wird* *euch* *hier* *bestimmt* *nicht* *begegnen.*
boredom.NOM will you.3PL.DAT here definitely not meet
'You will definitely not feel bored here.'

Third, a token is required to contain a dative argument that is part of the main verb's subcategorisation frame. This is an important proviso, since certain verbs, also known as half-modals, allow for an optional dative argument to be added to the argument structure if they are used as auxiliaries (Zifonun et al. 1997: 1282). Our dataset contains three such verbs: *scheinen* 'seem, appear', *erscheinen* 'seem, appear', and *dünken* 'seem, appear'. Only tokens in which these verbs function as main verbs are included (as in 25a, 26a, and 27a), not tokens in which they are used as auxiliaries (as in 25b, 26b, and 27b).

(25) a. *Aber das scheint den Inszenatoren nicht wichtig.*
but that.NOM seems the.DAT directors.DAT not important
'But the directors do not seem to attach any importance to that.'

(25) b. *Den vielen älteren Stammgästen schien dies allerdings nicht im Geringsten zu stören.*
the.DAT many.DAT older.DAT regulars.DAT seemed this.NOM surely not at all to bother
'Many of the older regulars did not seem to be disturbed by this at all.'

(26) a. *Uns erscheint es sinnvoller, direct vor Ort zu helfen.*
us.DAT seems it.NOM more.sensible directly on site to help
'We find it more sensible to provide aid directly on site.'

(26) b. *Das erscheint mir in diesem Kontext das Richtige zu sein.*
that.NOM seems me.DAT in this context the right.thing to be

(27) a. *Denn sie dünkte ihm ein Zeichen der Genesung.*
because she.NOM seemed him.DAT a sign of recovery
'Because that appeared to him as a sign of recovery.'

(27) b. *Der kleinen Seejungfrau dünkte es eine recht lustige Fahrt zu sein.*
the.DAT little.DAT mermaid seemed it.NOM a really enjoyable journey to be
'The little mermaid found it a really enjoyable journey.'

Fourth, tokens displaying incorrect case-marking, as in (28), or erroneous subject-verb agreement, as in (29), have been excluded as well.

(28) **Kurz gesagt, es genügt meinen Ansprüche voll und ganz …*
short said it.NOM satisfies my.DAT demands fully and completely
'In short, it fully satisfies my demands …'

(29) **Allen Teilnehmern, die genügend Schätze eingesammelt haben, winken ein Quest-Heldenpaket …*
all.DAT participants.DAT that enough treasure collected have wave a.NOM quest hero package
'All participants who have collected sufficient treasure will receive a quest hero package …'

In line with Verhoeven (2015), tokens containing elided constituents, as in (30), as well as tokens with question words, as in (31), have also been excluded.

(30) *Die Apothekerin* *hat* *mir* *dann* *Kijmea* *empfohlen,*
the.NOM pharmacist.F has me.DAT then Kijmea recommended
___ *ist* *mir* *auch* *schon* *bekannt* *vorgekommen*
Ø.NOM is me.DAT also already familiar come.across
durch die Werbung.
through the ad
'The pharmacist subsequently recommended Kijmea to me, [she] also looked familiar to me from ads.'

(31) *Fragt* *der Arzt:* *"Na,* *was* *fehlt* *Ihnen* *denn?"*
asks the.NOM doctor well what.NOM lacks you.DAT then
'And so the doctor asks: "So, what's the matter?"'

Finally, it should be pointed out that a number of verbs in the dataset allow for their oblique argument to be marked either in the dative or in the accusative. One example of such a 'wavering verb' is *grauen* 'dread, be afraid of'. As the present study focuses on Dat-Nom/Nom-Dat verbs, all instantiations with accusative obliques have evidently been excluded. The same holds for 'wavering verbs' whose oblique argument is structurally ambiguous between accusative and dative case, such as proper names, or syncretised object pronouns like *uns* 'us' 1p.acc/dat.pl and *euch* 'you' 2p.acc/dat.pl.

6.3.3 Data annotation

With data collection completed, all tokens have been annotated for the following ten variables: (1) verb lemma, (2) internal order of the two arguments, (3) (pro)nominality, (4) pronoun type, (5) referentiality, (6) person, (7) number, (8) definiteness, (9) animacy, and (10) length difference between the two arguments.

1) Verb lemma: infinitival form of the main verb in question (see Section 6.3.2 for a full list of all 76 lemmas under study)
2) Order: **Dat-Nom**, or **Nom-Dat**
3) (Pro)nominality: **pronoun** (*das* 'that' nom.sg, *ihnen allen* 'them all' dat.pl) or full **NP** (*Vegetarier* 'vegetarians' nom.pl, *meinen Kindern* 'my children' dat.pl)
4) Pronoun type: **personal** (*ihm* 'him' 3p.dat.sg, *sie* 'she' 3p.nom.sg), **demonstrative** (*diese* 'these' nom.pl, *dem* 'that one' dat.sg) or **indefinite** (*etwas* 'something' nom.sg, *allen* 'all' dat.pl).
5) Referentiality: expletive, correlative or referential. The label **expletive** refers to all uses of personal *es* 'it' in which es functions as a mere syntactic filler without any

semantic content, as in *Mir persönlich graust es jetzt schon vor den Folgen* 'I am personally already dreading the consequences'. Personal *es* 'it' and, occasionally, demonstrative *das* 'that' may also be used as a placeholder for a subclause or an infinitive clause, as in *Dem anderen Vogel gelang es schließlich, sich selbst zu befreien* 'The other bird finally succeeded in freeing itself'. Hence, such uses are annotated as **correlates**. All other uses of personal *es* 'it', demonstrative *das* 'that', as well as all other pronouns and full NPs in general have been tagged as **referential**. In line with Siewierska's (1993: 831) claim that non-referential elements may be expected to follow referential elements, it is hypothesised that correlates and especially expletives are more prone to occupying the postverbal slot

6) Person: **local** (*mir* 'me' 1p.dat.sg, *wir* 'we' 1p.nom.pl, *dir* 'you' 2p.dat.sg, *euch* 'you' 2p.dat.pl) or **non-local** (*ihm* 'him' 3p.dat.sg.m, *das Schiff* 'the ship' nom.sg)
7) Number: **singular** (*dieser Film* 'this film' nom.sg, *einer jungen Frau* 'a young lady' dat.sg) or **plural** (*die Vögel* 'the birds' nom.pl, *den Texten* 'the texts' dat.pl)
8) Definiteness: definite or indefinite. With the exception of indefinites (cf. supra), German pronouns are **definite**. NPs are considered to be definite if they are fronted by a definite article (*der Satz* 'the sentence' nom.sg), a possessive pronoun (*seinem Trainer* 'his trainer' dat.sg), or a demonstrative pronoun (*diese Aussage* 'this statement' nom.sg). In line with Abbott (2004: 122), a preposed genitive (*Tacitus' ausführlicher Bericht* 'Tacitus' detailed account' nom.sg) is also considered to signal definiteness. Proper names (*Andreas*, *Emmy*), including place names (*Toronto*, *Europa*), are considered to be inherently definite. In contrast, NPs fronted by a zero determiner (*Ärzten* 'doctors' dat.pl), an indefinite article (*ein Schrei* 'a cry' nom.sg), or an indefinite pronoun (*allen Musikern* 'all musicians' dat.pl) are coded for as **indefinite**. The same holds for constituents containing numerals (*acht Jahre* 'eight years' nom.pl; see Lyons 1999: 33). Conjoined constituents displaying a conflict in definiteness, in that one of the conjuncts is definite, but the other is indefinite, are coded for the first conjunct. Thus, the constituent *Journalisten, Schülern, Studenten, und zunehmend auch den Lehrern selbst* 'journalists, pupils, students and increasingly the teachers themselves' is considered to be indefinite, since the initial conjunct *Journalisten* is indefinite.
9) Animacy: individual, collective, inanimate, non-inferable or NA. The label **individual** is used to tag humans (*der Sportler* 'the athlete' nom.sg), animals (*meinem Hund* 'my dog' dat.sg), and what Bresnan and Ford (2010: 175) term "humanoid beings" (*den Elfen* 'the elves' dat.pl). Viruses (*dem Virus* 'the virus' dat.sg) are also subsumed under this category. Groups of individuals are annotated as **collective** (*der Partei* 'the party' dat.sg). Inanimate entities (*die Jakke* 'the jacket' nom.sg) and abstract concepts (*das Glück* 'luck' nom.sg) are tagged as **inanimate**. The same is true for plants (*den Bäumen* 'the trees' dat.pl). For a number of tokens, the animacy status cannot be established beyond reasonable doubt. This is mainly the case for

constituents that may both be interpreted as denoting an individual as well as a collective, e.g. *den Sponsoren* 'the sponsors' dat.pl. These doubtful cases are all subsumed under the label **non-inferable**. Finally, there are tokens for which animacy is irrelevant as a defining feature, because they do not denote extralinguistic entities or concepts. The label **inan_corr/expl** has therefore been applied to all expletives and correlates. Conjoined constituents displaying a conflict in animacy have been coded for the first conjunct. Hence, the constituent *den Tieren und Pflanzen* 'the animals and plants' has been interpreted as signalling an individual, because that is the label capturing the animacy status of the initial conjunct *den Tieren*.

10) Length difference: calculated by subtracting the length of the dative from the length of the nominative in **number of words**. Importantly, if a constituent is broken up and partly right-dislocated (G. *ausgeklammert*, *rechtsversetzt*), only the length of the phrase that is positioned within the verbal bracket is counted (cf. also Heylen 2005: 103). Consider the following example:

(32) *Eine besondere Aufmerksamkeit* *sollte* *den Geschenken*
a.NOM special.NOM attention should the.DAT gifts.DAT
zuteilwerden, *die von dem Händler angeboten werden.*
be.bestowed.upon that by the trader offered are
'Special attention should be paid to the gifts offered by the trader.'

Here, the dative constituent is broken up into two parts, i.e. *den Geschenken*, and the right dislocated relative clause *die von dem Händler angeboten werden*. Since the relative clause has been moved into the postfield, it is discounted in determining the constituent's length, which is therefore set at two words, and not at eight. Furthermore, dislocated relative clauses are altogether quite rare in the dataset, so disregarding them is not expected to substantially influence the effect of this variable.

6.4 Results and discussion

The present section provides an overview of the results of the statistical data analysis. Section 6.4.1 examines the annotated features of both arguments, the dative and the nominative, irrespective of position and Section 6.4.2 investigates the alternation between the two arguments in terms of the relevant predictor variables, based on a mixed-effects logistic regression model.

6.4.1 Case marking

This section discusses the correlation between case marking and (pro)nominality (6.4.1.1), pronoun type (6.4.1.2), person (6.4.1.3), definiteness (6.4.1.4), animacy (6.4.1.5), and number (6.4.1.6). The main trends as well as their implications are discussed in the interim conclusions (6.4.1.7).

6.4.1.1 (Pro)nominality

Table 6.2 presents the differences in pronominality between the dative and the nominative. It shows that (1) the dative is realised as a pronoun (6,996 attestations, or 53%) and as a full NP (6,296 attestations, or 47%) approximately equally often and that (2) the nominative is slightly more often realised as a full NP (8,287 attestations, or 62%) than as a pronoun (5,005 attestations, or 38%). A chi-squared test confirms that this distribution is statistically significant (χ^2 = 601; df = 1; p < 0.0001), although the effect size is weak (Cramér's V = 0.15).

Table 6.2 (Pro)nominality of the dative and the nominative

	Dat	**Nom**
(Pro)nominality		
NP	6,296	8,287
Pronoun	6,996	5,005
Subtotal	13,292	13,292

In her study of dative and accusative "experiencer-object" verbs in Modern German, Verhoeven (2015) also investigates the extent to which either the nominative stimulus or the accusative/dative experiencer is realised pronominally. She hypothesises that accusative/dative experiencers are more prone to pronominalisation than nominative stimuli, as the former tend to be more discourse-prominent than the latter (cf. also Bickel 2004: 77). In order to test this hypothesis, Verhoeven compares the number of pronominally realised accusative and datives to the number of pronominally realised nominatives, excluding both local and non-referential pronouns. Her results show that "object-experiencers" are indeed more often realised as pronouns than nominative stimuli (Verhoeven 2015: 71).

As for the current dataset, the correlation between case marking and pronominality does not hold up to scrutiny. Counting all instances of pronominal nominatives and datives, excluding both local pronouns as well as correlates and expletives, the total number of pronominal datives amounts to 2,884 and the total number of pronominal nominatives to 2,675. The former accounts for 52% of the tokens and the latter for 48% of the tokens,

clearly showing that both arguments are equally prone to pronominalisation. This result is fully in line with the alternating hypothesis laid out in Section 6.1, as pronominality is also a property associated with subjects. Since both the dative as well as the nominative of Dat-Nom/Nom-Dat verbs independently pass the subject tests (see Barðdal 2006, 2023: Ch. 6; Barðdal Eythórsson & Dewey 2014, 2019), it is not surprising that both arguments are equally prone to pronominalisation.

The two following sections provide a further breakdown of the numbers presented in Table 6.2. These zoom in, respectively, on the association between case marking and pronoun type (Section 6.4.1.2) and on case marking and local vs. non-local pronouns (Section 6.4.1.3).

6.4.1.2 Pronoun type

Table 6.3 gives an overview of the distribution of the different types of pronouns. It reveals that both datives and nominatives are strongly inclined to be realised as personal pronouns: as many as 2,965 nominative pronouns (or 59%) and 6,154 dative pronouns (or 88%) are personal pronouns. The second most common pronoun type found in the dataset are demonstratives. These are most often realised in the nominative (1,503 tokens), but dative demonstratives are not altogether excluded (265 tokens). The last type of pronoun are indefinites, which are relatively uncommon, even though they are quite evenly distributed across both cases (577 dative tokens vs. 537 nominative tokens). According to a chi-squared test, the distribution is statistically significant (χ^2 = 1,700; df = 2; p < 0.0001). The effect size is moderately strong (Cramér's V = 0.38).

Table 6.3 Pronoun type according to case marking

	Dat	**Nom**
Pronoun type		
Personal	6,154	2,965
Demonstrative	265	1,503
Indefinite	577	537
Subtotal	6,996	5,005

One trend not presented in Table 6.3 is that the bulk of nominative personal pronouns are either expletives (1,141 attestations, or 38%) or correlates (1,163 attestations, or 39%). Expletives take the form of third person *es* 'it', which serves as a syntactic placeholder but is semantically void:

(33) *So manchem braven Limerick* *gebricht* *es* *deutlich*
so many.DAT good.DAT limerick lacks it.NOM clearly
an Geschick.
in skill
'Many a good limerick clearly lacks skill.'

The following verbs mandatorily require expletives to fill their nominative slot: *ermangeln (an)* 'lack' (134 tokens), *gebrechen (an)* 'lack' (200 tokens), *mangeln (an)* 'lack' (200 tokens), and *schlechtgehen* 'suffer; go through economically hard times' (seven tokens). Three other verbs show a very strong, yet not an absolute tendency towards nominative expletives: *grauen* 'dread, be afraid of' (198/200 tokens), *grausen* 'dread, be afraid of' (198/200 tokens), and *pressieren* 'be urgent' (12/13 tokens).

For correlates, too, there is heavy clustering around individual verbs. Correlates serve as placeholders to subclauses in extraposition. They are rendered by third-person *es* 'it' or *das* 'that':

(34) *Es* *ist* *uns* *aufgefallen,* *dass* *die Preise für Essen*
it.NOM is us.DAT noticed that the.NOM prices for food
wenig *variieren.*
little vary
'We have noticed that food prices show little fluctuation.'

The ten most common verbs with a correlative (nominative) pronoun are *geziemen* 'befit' (116 tokens), *schwerfallen* 'be difficult' (96 tokens), *gelingen* 'succeed' (90 tokens), *widerstreben* 'be reluctant' (89 tokens), *dämmern* 'dawn on' (79 tokens), *dünken* 'seem, appear' (59 tokens), *scheinen* 'seem, appear' (55 tokens), *leichtfallen* 'be easy' (54 tokens), *einleuchten* 'be clear' (42 tokens), and *genügen* 'be enough, be sufficient' (38 tokens). Taken together, these verbs account for 718 out of 1,163 tokens (or 62%) with nominative correlates. Correlates may in principle also be dative, but such forms do not occur in the present dataset.

Finally, it is worth mentioning that the verb *dämmern* in particular occurs 44 times with an expletive and 79 times with a correlative pronoun. The remaining tokens all contain either a full NP (30 observations) as in (35), or an ordinary referential pronoun (47 observations) as in (36). This is a remarkable finding, as Klein and Kutscher (2002: 27) argue that *dämmern* solely licences expletive and correlative nominatives.

(35) *Ihm* *dämmerte* *die Absurdität des gerade Gesagten.*
him.DAT realised the.NOM absurdity the.GEN already said.GEN
'The absurdity of what had just been said dawned on him.'

(36) *Mir* *dämmerte* *etwas.*
me.DAT realises something.NOM
'Something is dawning on me.'

6.4.1.3 Person

Table 6.4 maps out the interaction between local and non-local pronouns, on the one hand, and case marking on the other. Local pronouns refer to speech-act participants, which means that they are either first or second person. Non-local pronouns refer to entities situated outside of the discourse situation. Hence, they are third person.

Table 6.4 Local vs. non-local dative and nominative pronouns

	Dat	**Nom**
Person		
Local	4,112	26
Non-local	2,884	4,979
Subtotal	6,996	5,005

The distribution laid out in Table 6.4 is statistically significant (χ^2 = 4,776; df = 1; p < 0.0001). The effect size is strong (Cramér's V = 0.42). This means that dative pronouns more often than not code discourse-immediate participants: out of a total of 6,996 tokens containing a dative pronoun, 4,112 (or 59%) are first or second person, i.e. *mir* 'me' 1p.dat.sg, *dir* 'you' 2p.dat.sg, *uns* 'us' 1p.dat.pl, or *euch* 'you' 2p.dat.pl. The remaining 2,884 tokens (or 41%) self-evidently code non-local participants. Nominative pronouns show the opposite tendency: they almost invariably code third person participants (4,979 tokens, or 99%) and hardly ever first person participants (26 tokens, or 1%). Remember that approximately half of all non-local nominatives are either expletives or correlates (cf. Section 6.4.1.2).

6.4.1.4 Definiteness

Table 6.5 shows the correlation between case marking and definiteness. The tendency for both datives and nominatives to correlate with definite constituents is striking, but it is still slightly stronger for datives: in total, the dative is definite in 87% of cases (11,616 out of 13,292 tokens) and the nominative is definite in 69% of cases (9,160 out of 13,292 tokens). The remaining 13% of dative constituents (or 1,676 tokens) as well as the remaining 31% of nominative constituents (or 4,132 tokens) evidently code indefinite constituents. The distribution laid out in Table 6.5 is statistically significant (χ^2 = 1,327; df = 1; p < 0.0001). The size of effect is moderate (Cramér's V = 0.22).

Table 6.5 Definiteness of the dative and the nominative (pronominal constituents and full NPs)

	Dat	Nom
Definiteness		
Definite	11,616	9,160
Indefinite	1,676	4,132
Subtotal	13,292	13,292

Remarkably, half the 1,676 tokens with indefinite datives are due to a mere 17 verbs. The ten most common verbs with indefinite datives are *leidtun* 'take pity; be sorry' (136 tokens), *einleuchten* 'be clear' (80 tokens), *aufstoßen* 'strike, notice; make (sb) burp; annoy' (58 tokens), *vorkommen* 'happen to; experience; strike (as)' (57 tokens), *munden* 'taste good' (53 tokens), *zugutekommen* 'benefit' (48 tokens), *drohen* 'threaten' and *zugehören* 'belong to' (47 tokens each), *passieren* 'happen to' (42 tokens), *nützen* 'be of use' (39 tokens), and *dämmern* 'dawn on' (38 tokens).

Table 6.6 sheds further light on the distribution presented in Table 6.5 by zooming in on full NPs only, thus omitting the 6,996 dative pronominal constituents and the 5,005 nominative pronominal constituents.

Table 6.6 Definiteness of dative and nominative full NPs

	Dat	Nom
Definiteness		
Definite	5,197	4,692
Indefinite	1,099	3,595
Subtotal	6,296	8,287

The most notable trend is that dative NPs are rarely indefinite: out of a total of 6,296 tokens, a mere 1,099 tokens (or 17%) code for indefinite constituents. The remnant 5,197 tokens (or 83%) self-evidently code definite constituents. For the nominative, the distribution is markedly more balanced, with definite nominatives (4,692 tokens, or 57%) being only slightly more frequent than indefinite nominatives (3,595 tokens, or 43%). A chi-squared test for the distribution laid out in Table 6.6 nevertheless reports a significant result (χ^2 = 1,100; df = 1; p < 0.0001). The effect size is moderate (Cramér's V = 0.27).

As for the 1,099 tokens with indefinite dative NPs, the data again show heavy clustering around individual verbs, as half of the indefinite NPs are found with only 18 verbs. The top ten set is similar, yet not identical to the top ten shown for indefinite constituents in general (cf. supra). It comprises the following verb types: *zugutekommen* 'benefit' (47 tokens), *drohen* 'threaten' (46 tokens), *aufstoßen* 'strike, notice; make (sb) burp; annoy' (45 tokens), *zugehören* 'belong to' (44 tokens), *vorkommen* 'happen to; experience; strike (as)' (35 tokens), *mangeln (an)* 'lack' (32 tokens), *geziemen* 'befit' (30 tokens), *einleuchten* 'be

clear' (29 tokens), *gelten* 'be worth; be considered', *widerfahren* 'happen (to), experience' and *zuwachsen* 'receive' (28 tokens each), and *schwerfallen* 'be difficult' (27 tokens).

6.4.1.5 Animacy

As may be deduced from Table 6.7, there is a clear distinction in animacy between the dative and the nominative: whereas the dative overwhelmingly correlates with animate referents, the nominative shows a clear preference for inanimate referents. Recall that the label 'animate' comprises both individual referents and collectives.

Table 6.7 Animacy of the dative and the nominative

	Dat	**Nom**
Animacy		
Animate	11,755	744
Inanimate	1,477	10,238
Subtotal	13,232	10,982

Observant readers may note that the respective numbers of dative and nominative tokens presented in Table 6.7 do not add up to 13,292. Neither do both subtotals (13,232 for datives, and 10,982 for nominatives) add up to the same number. The reason for these discrepancies is that Table 6.7 abstracts away from the six nominative constituents and the 60 dative constituents whose animacy status is non-inferable. Additionally, the dataset comprises 2,304 correlative and expletive nominatives, which have also been omitted for the present purposes, since their animacy status is NA, as they do not refer to extralinguistic entities.

The distribution shown in Table 6.7 yields a statistically highly significant result (χ^2 = 16,179; df = 1; p < 0.0001) with a strong size of effect (Cramér's V = 0.83). Thus, German Dat-Nom/Nom-Dat verbs are very strongly inclined to code an animate dative and an inanimate nominative. However, the remnant 1,477 dative tokens that have an inanimate referent and the 744 nominative tokens that have an animate referent show that this correlation is not absolute.

Zooming in on the 1,477 tokens containing inanimate datives, it is striking that more than half of them cluster around ten specific verbs, i.e. *folgen* 'succeed, follow' (166 tokens), *angehören* 'belong' (105 tokens), *zusetzen* 'affect badly' (99 tokens), *zuteilwerden* 'be bestowed upon' (83 tokens), *guttun* 'do good' (78 tokens), *gelten* 'be worth; be considered' (68 tokens), *zugutekommen* 'benefit' (65 tokens), *zugehören* 'belong to' (64 tokens), *entschlüpfen* 'escape (words)' (58 tokens), and *zufließen* 'receive' (57 tokens). Five top-ten verbs are verbs of belonging (*angehören* and *zugehören*), or verbs of gain and benefit (*zuteilwerden*, *zugutekommen*, *zufließen*). Verbs of gain and benefit are also generally very common with inanimate datives. Taken together, the verbs *zuteilwerden* 'be bestowed

upon', *zugutekommen* 'benefit', *zufließen* 'receive', *zuwachsen* 'receive', *nützen* 'be of use', *zufallen* 'fall to, receive (lit.); fall to, receive (fig.)', *winken* 'can expect, be in store for', and *zupasskommen* 'come in handy' yield 312 tokens with an inanimate dative, i.e. 21% of all tokens with inanimate datives in the dataset.

What is even more remarkable is that datives can even have inanimate referents with verbs of emotion or cognition, as is illustrated by *guttun* 'do good' (78 tokens), *wohltun* 'do sb. good' (31 tokens), *behagen* 'please' (12 tokens), or *vorschweben* 'have in mind' (five tokens). Two relevant examples are given in (37) and (38):

(37) *Der Haut* *dürfte* *der kosmetische Mineralbeauty-Trend* *behagen.*
the.DAT skin should the.NOM cosmetic.NOM mineral.beauty.trend please
'The cosmetic mineral beauty trend should please the skin.'

(38) *Die naturwissenschaftliche Abweisung alles Unbewiesenen* *hat*
the.NOM scientific.NOM rejection all.GEN unproven.GEN has
diesem Aufbau *als Methode der Darstellung* *vorzuschweben.*
this.DAT structure as method the.GEN representation have.in.mind
'One should bear in mind that the scientific rejection of all that is unproven has to be the method of representation of this structure.'

This remarkable finding invalidates previous claims that experiencers (i.e. entities participating in an act of emotion, cognition, volition, perception, or bodily sensation) are animate *by definition* (Kutscher 2009: 24; Rott 2013: 93; Verhoeven 2009: 356; Verhoeven 2015: 46). Inspection of the dataset reveals that the tendency for such verbs to attract animate referents is indeed very strong, but by no means absolute.

Yet another fascinating finding is that the dative may be inanimate with verbs of speaking, like *entschlüpfen* 'escape (words)' (58 tokens). The reason is that *entschlüpfen* very often collocates with nouns like *Kehle* 'throat', *Lippen* 'lips' or *Mund* 'mouth', as illustrated by the following example:

(39) *Ein hoher Ton* *entschlüpfte* *seiner Kehle.*
a.NOM high.NOM tone escaped his.DAT throat
'A high tone escaped from his throat.'

6.4.1.6 Number

Table 6.8 summarises the distribution of number specifications according to case marking. Both datives and nominatives alike show a very strong tendency towards singular constituents. In total, the dative is singular in 9,569 out of 13,292 cases (or 72%). For the nominative, this correlation is even stronger: it is singular in 11,432 out of 13,292 cases (or 86%). A chi-squared test shows that this distribution is statistically informative (χ^2 = 786; df = 1; p < 0.0001), even though the effect size is weak (Cramér's V = 0.17).

Table 6.8 Number specifications for dative and nominative constituents

	Dat	Nom
Number		
Singular	9,569	11,432
Plural	3,723	1,860
Subtotal	13,292	13,292

Table 6.9 zooms in on number specifications for full NPs. It shows that most plural constituents presented in Table 6.8 above are in fact full NPs, and not pronouns. Comparing the number of plural NPs presented in Table 6.9 with the number of plural constituents in general presented in Table 6.8, it may be observed that the dative is a plural full NP in 2,189 out of 3,723 cases (or 59%). For the nominative, the association between plurality and full NPs is almost absolute: it is a plural full NP in 1,721 out of 1,860 cases (or 93%). A chi-squared test again yields a significant result (χ^2 = 357; df = 1; p < 0.0001), although the effect size is weak (Cramér's V = 0.16).

Table 6.9 Number specifications for dative and nominative full NPs

	Dat	Nom
Number		
Singular	4,107	6,566
Plural	2,189	1,721
Subtotal	6,296	8,287

Table 6.10 forms the negative image of Table 6.9, as it shows the association between number and pronominality.

Table 6.10 Number specifications for dative and nominative pronouns

	Dat	Nom
Number		
Singular	5,462	4,866
Plural	1,534	139
Subtotal	6,996	5,005

The inclination towards the singular, as highlighted in Table 6.8, reappears in Table 6.10, but it is considerably more outspoken for the distribution presented here. A chi-squared test yields a highly significant result (χ^2 = 890; df = 1; p < 0.0001) and a moderately strong size of effect (Cramér's V = 0.27), confirming the impression that pronouns are associated with the singular. As already pointed out above, nominative pronouns show the strongest

association with the singular, being in the plural in only 139 out of a total of 5,005 cases (or 3%).

6.4.1.7 Interim conclusions

The current section has shed light on the correlation between case marking and six other variables: pronominality, pronoun type, person, definiteness, animacy, and number. The results show that the dative and the nominative display important converging behaviour: (1) both are approximately equally often realised as pronouns and as full NPs, (2) both are very often definite, with indefinite nominatives being slightly more common than indefinite datives, and (3) both are more often singular than plural. Recall that the first two properties are strongly correlated with subjecthood, which explains why the two arguments systematically compete for the subject slot.

As for the three remaining variables, the dative and the nominative display remarkably divergent behaviour. First, pronominal datives are nearly always realised as personal pronouns, whereas pronominal nominatives are mostly either personal pronouns or demonstrative pronouns. Second, the dative is strongly associated with local pronouns, whereas the nominative is nearly always realised as a non-local pronoun. And third, the dative is very strongly associated with animate referents, whereas the nominative is most often inanimate. Interestingly, animacy is the only property for which dative full NPs in the present study so distinctly outrank nominative full NPs. Bearing in mind that there is a crosslinguistic tendency for animacy to correlate with subjecthood (Comrie 1981; Du Bois 1987; Hopper and Thompson 1980, *inter alia*), and that Dat-Nom/Nom-Dat verbs display a mismatch between case marking and thematic roles, one can expect the dative and the nominative to heavily compete for the first position in declarative clauses when both are realised as full NPs. The findings presented in Section 6.4.2.2 show that this is indeed the case.

Another result of the current section is that datives in particular correlate very consistently with both animacy and definiteness. This finding is reminiscent of Hopper and Thompson (1980: 259), who observe a crosslinguistic tendency for indirect objects to be animate and definite. In the current dataset, 11,616 tokens contain a definite dative and 11,755 tokens contain an animate dative. As many as 10,262 tokens contain a dative that is both animate and definite at the same time, equalling 77% of all datives in the dataset.

We now move on to Section 6.4.2, which scrutinises the relative order of the two arguments in relation to the relevant predictor variables. This is different from how we have proceeded in Section 6.4.1 above, which has offered a series of univariate analyses.

6.4.2 Internal order of the two arguments

We now perform a mixed-effects logistic regression analysis using the different verbs as a random intercept. The first analysis is based on the entire dataset, yet excludes the 66 cases for which animacy was non-inferable. As such, it is based on 13,226 tokens, instead of the original 13,292. In total, 5,519 tokens, or 42%, occur in the Dat-Nom order. The remaining 7,773 tokens, or 58%, occur in the Nom-Dat order. The second analysis singles out the 4,529 tokens with double NPs, of which 2,079, or 46%, instantiate the Dat-Nom order and 2,450, or 54%, the Nom-Dat order. Section 6.4.2.1 discusses why the random intercept was created and how partially overlapping variables were merged. The model results are covered in Section 6.4.2.2.

6.4.2.1 Random intercept and variable mergers

The reason why the verbs are modelled as a random intercept is that each verb seems to alternate between both word order patterns to varying degrees: some verbs tend very heavily towards the Dat-Nom order, others tend very heavily towards the Nom-Dat order, and others still are halfway the spectrum, instantiating both word order patterns approximately equally often. Table 6.11 presents the tallies of nine verbs from the dataset to further illustrate this point. It contains i) three verbs from the Dat-Nom end of the spectrum (i.e. *drohen* 'threaten', *schwanen* 'have a sense of foreboding' and *winken* 'can expect, be in store for'), ii) three verbs from the Nom-Dat end of the spectrum (i.e. *zugutekommen* 'benefit', *zusetzen* 'affect badly' and *zugehören* 'belong to'), and iii) three verbs whose frequencies are relatively evenly distributed across the Dat-Nom and the Nom-Dat word orders (i.e. *entgehen* 'miss out on; fail to notice', *gebühren* 'be due to' and *dämmern* 'dawn on').

Table 6.11 Verb biases for nine verbs from the dataset

	Dat-Nom	**Nom-Dat**
Verb		
drohen	188 (94%)	12 (6%)
schwanen	186 (93%)	14 (7%)
winken	183 (92%)	15 (8%)
entgehen	103 (51.5%)	97 (48.5%)
gebühren	98 (50.3%)	97 (49.7%)
dämmern	97 (48.5%)	103 (51.5%)
zugutekommen	14 (7%)	186 (93%)
zusetzen	14 (7%)	186 (93%)
zugehören	8 (6%)	125 (94%)

The word order counts presented in Table 6.11 seem to imply that the Dat-Nom/Nom-Dat alternation is susceptible to a verb effect. This is exactly what Somers and Barðdal (2022) found for a set of five alternating Dat-Nom/Nom-Dat verbs in Modern Icelandic. It seems appropriate, therefore, to factor a potential verb effect into the model as a random intercept.

Next, we have combined variables that appear to be closely related upon closer inspection of the data. First, we merged the variables Nominative (Pro)nominality and Nominative Referentiality, because the latter is principally relevant for personal pronouns only. The complementarity of both variables is shown in Table 6.12, where it can be deduced that correlates and expletives are virtually always personal pronouns. Our dataset also contains seven instances of correlative demonstratives, but since these only make up 0.5% of all demonstratives, these have been merged with the category of referential demonstratives for the remainder of the analysis.

Table 6.12 Overlap between the variables Nominative (Pro)nominality and Nominative Referentiality

	Nominative Referentiality		
Nominative (Pro)nominality	correlate	expletive	referential
demonstrative pronoun	7	0	1,492
indefinite pronoun	0	0	537
full NP	0	0	8,241
personal pronoun	1,147	1,140	662

Secondly, we have also merged Nominative (Pro)nominality with Nominative Definiteness. Again, the reason for the merger is the variables' mutual complementarity, as definiteness is only relevant for full NPs. After all, pronominal constituents, with the exception of indefinite pronouns, are always definite (see Table 6.13).

Table 6.13 Overlap between the variables Nominative (Pro)nominality and Nominative Definiteness

	Nominative Definiteness	
Nominative (Pro)nominality	definite	indefinite
demonstrative pronoun	1,499	0
indefinite pronoun	0	537
full NP	4,660	3,581
personal pronoun	2,949	0

The merger of the variables Nominative (Pro)nominality, Nominative Referentiality and Nominative Definiteness has resulted in an expanded version of the Nominative (Pro)nominality variable, which presently comprises seven levels: demonstrative pronoun,

indefinite pronoun, referential personal pronoun, expletive personal pronoun, correlative personal pronoun, definite full NP, and indefinite full NP.

Third, we have also combined the variables Dative (Pro)nominality and Dative Definiteness. The logic underlying this decision is identical to what was discussed above in the context of Nominative (Pro)nominality and Nominative Definiteness: for pronominals, definiteness is implicitly captured by the type of pronoun, as only indefinite pronouns are indefinite. Thus, definiteness offers additional information in the case of full NPs only (cf. Table 6.14).

Table 6.14 Overlap between the variables Dative (Pro)nominality and Dative Definiteness

	Dative Definiteness	
Dative (Pro)nominality	definite	indefinite
demonstrative pronoun	263	0
indefinite pronoun	0	575
full NP	5,151	1,093
personal pronoun	6,144	0

As such, the newly created Dative (Pro)nominality variable consists of five values: demonstrative pronoun, indefinite pronoun, personal pronoun, definite full NP, and indefinite full NP.

6.4.2.2 Results

The model estimates have been obtained using the lme4::glmer function which fits the order of both arguments in function of Dative Animacy, Dative (Pro)nominality, Dative Person, Dative Number, Nominative Animacy, Nominative (Pro)nominality, Nominative Number, and Length Difference. As was mentioned at the beginning of Section 6.4.4.1, the different verbs were modelled as a random intercept. The formula according to which the model was generated may be rendered as follows: Order ~ DatAnimacy + DatPronominality + DatPerson + DatNumber + NomAnimacy + NomPronominality + NomNumber + LengthDifference + (1|Verb). Crucially, the likelihood of the Dat-Nom order was identified as the measure of success of the model. This means that significant p-values correlate with a higher probability of Dat-Nom, as opposed to Nom-Dat. An overview of the model estimates may be found in Appendix 3.

The quality of the model appears to be very good based on various model checks. The model has a C-index of concordance equalling 82%, which suggests a good classification accuracy. A binned residuals plot shows that about 95% of the residuals fall inside the error bound, indicating a good model fit. There is collinearity between Nominative Animacy and Nominative Pronominality. All fixed effects are significant at the 5% significance level based on a Wald test and a Type II anova (see R notebook for full details).

To interpret the model, we start by examining the random intercept Verb. Based on the Intra-Correlation Coefficient, it appears that about a quarter of the variability in the Order is related to the random intercept for Verb (adjusted ICC = 0.25). Moreover, the model has a marginal R^2 of 14.9% (the R^2 associated with the fixed effects), but a conditional R^2 of 36%, when taking into account the random intercept. In line with the bivariate results, then, it appears that even though the predictors have a significant effect, their impact is fairly low, as most of the variability is explained by the Verb. The conditional modes associated with each verb are visualised in Figure 6.3:

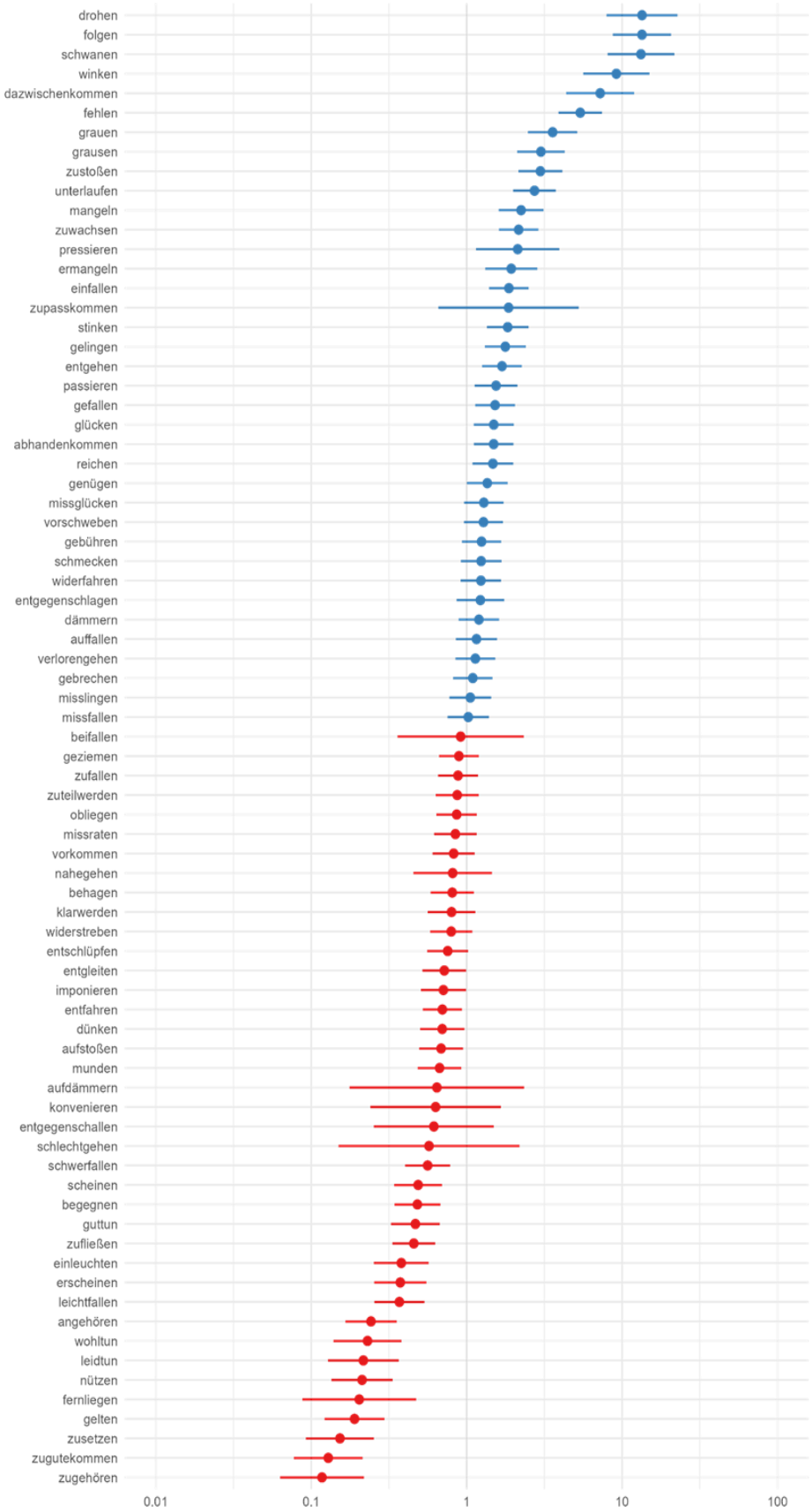


Figure 6.3 Conditional modes associated with the random intercept Verb. Conditional modes highlighted in blue indicate a preference for the Dat-Nom order, conditional modes highlighted in red a preference for the Nom-Dat order

Figure 6.3 essentially shows that the prediction made earlier, i.e. that the Dat-Nom/Nom-Dat alternation is susceptible to a verb effect, is borne out. Thus, verbs like *drohen* 'threaten', *folgen* 'succeed, follow' or *schwanen* 'have a sense of foreboding' indeed have a preference for the Dat-Nom order, whereas verbs like *zugehören* 'belong to', *zugutekommen* 'benefit' and *zusetzen* 'affect badly' show a genuine preference for the reverse Nom-Dat order. Hence, their inclination towards either order cannot be entirely reduced to secondary factors.

Apart from the verb effect, the Dat-Nom/Nom-Dat alternation is also guided by several fixed factors, whose respective effects are visualised in Figure 6.4. Each graph shows the effect of a single predictor on the alternation, i.e. Dative Animacy, Dative (Pro)nominality, Dative Person, Dative Number, Nominative Animacy, Nominative (Pro)nominality, Nominative Number, and Length Difference. Recall that the probability of the Dat-Nom order was identified as the measure of success, which means that the predicted probabilities presented in Figure 6.4 are also plotted in function of the Dat-Nom order. Steeper lines indicate stronger effects.

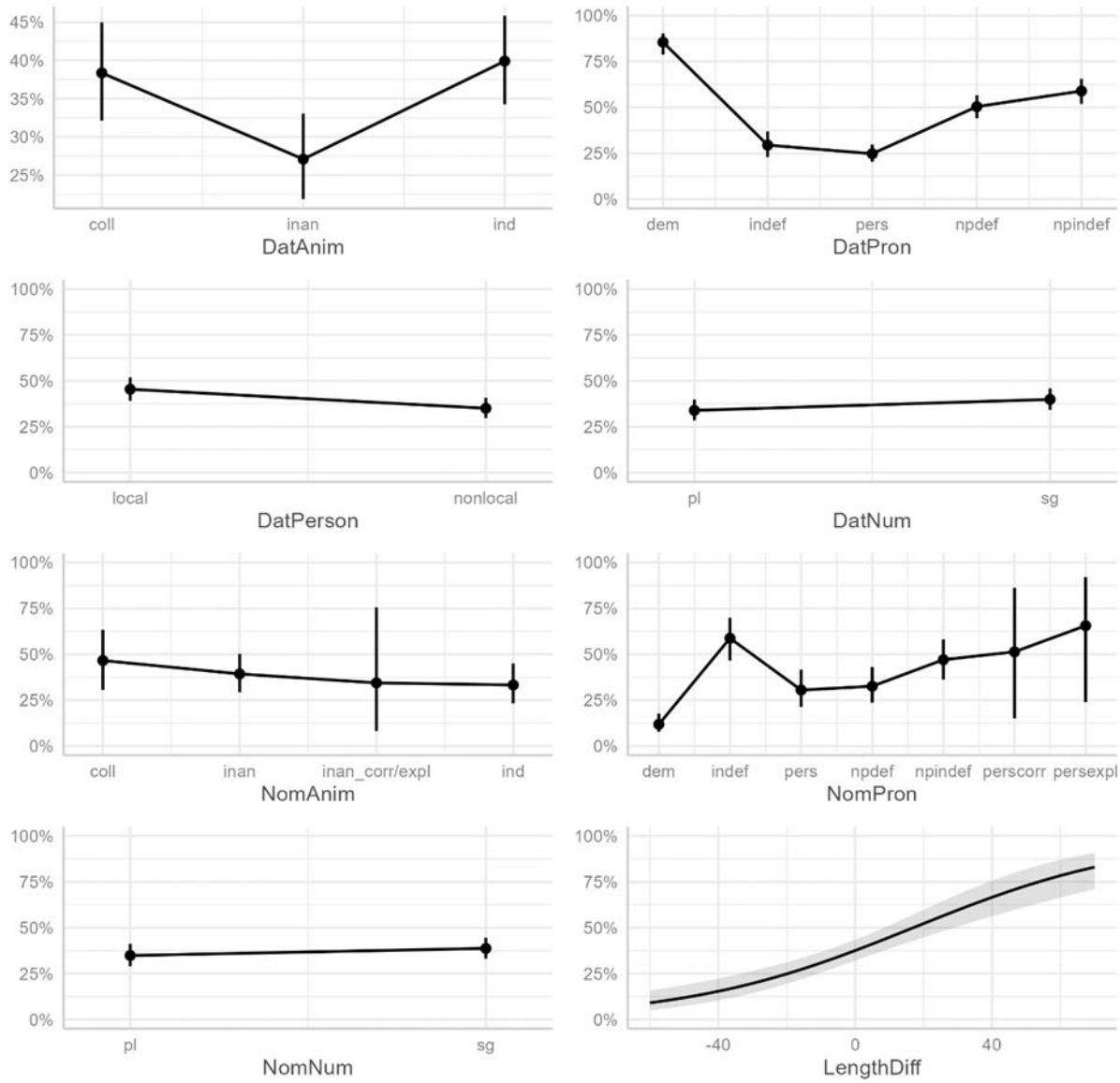


Figure 6.4 Predicted probabilities for the Dat-Nom order associated with the significant fixed effects

The variables Dative (Pro)nominality, Nominative (Pro)nominality, and Length Difference generate the strongest effects. We zoom in on the latter first, before assessing the results of the first two variables.

The difference in length between both arguments, the dative and the nominative, is calculated by subtracting the length of the dative (in number of words) from the length of the nominative (in number of words). Thus, a negative result signals a longish dative in combination with a shortish nominative, while a positive result signals a longish nominative in combination with a shortish dative. The effect plot shows that the likelihood of the Dat-Nom order increases with longer nominatives, and that both orders, Dat-Nom and Nom-Dat, are equally likely when the nominative is some 20 words longer than the dative. The observation that longer constituents tend to follow shorter ones is not new: it is at least as old as Behaghel (1909/10), and the validity of his so-called Law of Increasing Terms has

since been confirmed by various other scholars ever since (Allan 1987; Arnold et al. 2000; Hawkins 1992; Siewierska 1993; Somers, Jenset and Barðdal 2023, *inter alia*).

The Dative (Pro)nominality variable shows that especially dative demonstratives are strongly associated with the Dat-Nom order, whereas dative indefinite pronouns and dative personal pronouns are more inclined towards the Nom-Dat order. Interestingly, dative full NPs are approximately halfway along the spectrum, thus showing to alternate most freely. Zooming in even further on full NPs, it is remarkable that indefinite NPs are associated more strongly with the Dat-Nom order than definite NPs. However, closer inspection of the data suggests that this finding is an epiphenomenon of the verb effect discussed above, as the top-ten verbs associated with the Dat-Nom order (cf. Figure 6.3) account for 28% (or 157 out of 547 tokens) of all preposed indefinite dative NPs.

The results for Nominative (Pro)nominality may be summarised as follows. First, nominative demonstratives clearly favour the Nom-Dat order. This tendency is highly reminiscent of the trend we have observed earlier for dative demonstratives, which are very common in the Dat-Nom order. This means that demonstratives, regardless of case marking, strongly prefer the preverbal slot. For indefinite pronouns, the opposite holds, as nominative indefinites, precisely like dative indefinites, tend to be postposed. These findings seem to indicate that the Dat-Nom/Nom-Dat alternation is indeed guided by an effect of topicality, as has been suggested by Barðdal (2001) and Barðdal, Eythórsson and Dewey (2014, 2019).

Second, Nominative (Pro)nominality also yields strong effects for the values ‘correlative personal pronoun’ and ‘expletive personal pronoun’, even though the length of the respective error bars calls for caution. In any case, this finding confirms the prediction made in Section 6.3.3 that referentiality exerts a strong influence on the word order variation. Strikingly, the effect is also gradient, with expletives tending most strongly towards Dat-Nom, followed by correlatives, and, finally, classic coreferential personal pronouns. Thus, the poorer in semantic content the pronoun *es*, the higher the likelihood of the Dat-Nom order. As for full NPs, definite arguments pattern differently from indefinite arguments, since the latter occupy the postverbal slot more promptly than the former. This is in line with a host of previous research showing that indefinite arguments tend to give way to definite arguments (Croft 2003: 130; Siewierska 1993; Somers, Jenset and Barðdal 2023, *inter alia*).

We now turn to the effect of animacy. As for Dative Animacy, the results are entirely in line with expectations: constituents referring to individuals allow for most variation, followed closely by collectives. Inanimate datives are very rare (cf. Section 6.4.1.5), but their preference for the less prominent postverbal slot is undeniable. The output for Nominative Animacy, in contrast, appears strange, with collectives being noticeably more strongly inclined to the Nom-Dat order than individuals. However, these results build on extremely few tokens, since Dat-Nom/Nom-Dat verbs typically occur with inanimate nominatives (cf. also Section 6.4.1.5).

The remaining plots may be interpreted as follows: singulars, regardless of case marking, facilitate the choice for the Dat-Nom order. The same holds for dative local pronouns, which occupy clause-initial position more often than non-local pronouns. As such, this suggests an effect of topicality, as argued by Barðdal (2001) and Barðdal, Eythórsson & Dewey (2014, 2019) for the class of alternating Dat-Nom/Nom-Dat verbs in Icelandic and German.

Note that we have also generated a conditional inference tree and a random forest. However, since the results are in line with the mixed-effects logistic regression analysis, we refrain from discussing these analyses here. The interested reader is referred to the Notebook for further details. In what follows, we briefly elaborate on the results on configurations with double NPs. The relevant plots are shown in Figure 6.5.

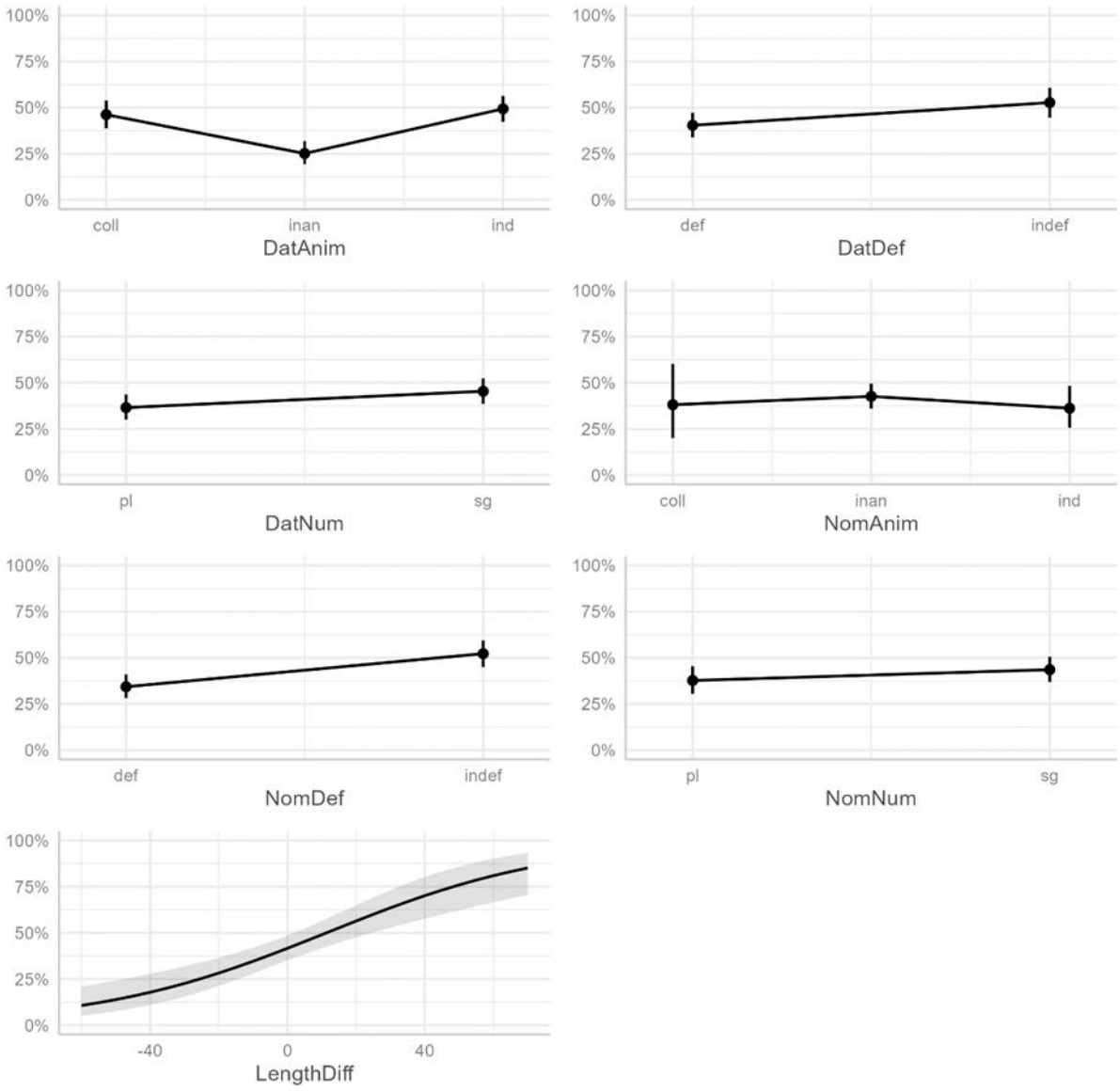


Figure 6.5 Predicted probabilities for the Dat-Nom order associated with the significant fixed effects (full NPs only)

First, the plots in Figure 6.5 are remarkably flatter than those in Figure 6.4: with the exception of the Length variable and the 'inanimate datives' value in the Dative Animacy plot, none of the explored variables yield any categorical results. We take this as evidence that configurations with pronominal arguments skew the results towards either the Dat-Nom order or the Nom-Dat order more readily than configurations with only two NPs. A similar tendency has been observed by Somers (2023) for five verbs of success and failure also included in the current dataset, and by Somers and Barðdal (2022) and Somers, Jenset and Barðdal (2023) for Modern Icelandic.

Secondly, the plots in Figure 6.5 also show a higher degree of Dat-Nom than the ones in Figure 6.4: the double-NP plots generally move up, and thereby closer towards a 50–50 distribution. This is not only true for the variables Nominative Number and Dative Number, but also for Length: the length curve is slightly higher in the current configuration and its zero point, which represents tokens with arguments of exactly equal length, is at approximately 45%. This means that the likelihood of the Dat-Nom order with equally long NPs is nearly as high as the likelihood of the Nom-Dat order. Still, Dat-Nom is more strongly associated with longer nominatives than the other way around. As for Dative Animacy, the values 'collective' and 'individual' also exhibit a higher probability of Dat-Nom, but not inanimate datives. The latter take initial position in only 25% of the cases. The results for Nominative Animacy are highly comparable to the ones presented in Figure 6.5. Nominative Definiteness shows that especially indefinite nominative NPs have a proclivity for the Dat-Nom order. Dative Definiteness again shows the counterintuitive result that indefinite datives are more prone to being fronted than definite datives, but, once again, this is a side-effect of the powerful verb effect: some verbs simply tend to allocate the dative to the preverbal slot regardless of any contending factors.

As shown by Figure 6.5, the elimination of configurations with at least one pronoun ensues in a higher share of Dat-Nom attestations. This result is consistent with earlier research, especially on Modern Icelandic, showing that nominative pronouns rigorously push alternating verbs towards the Nom-Dat pole of the spectrum. However, a comparison of Figures 6.4 and 6.5 reveals that the differences are less outspoken than originally anticipated. So why do nominative pronouns lead to less skewing in German than in Icelandic? We believe that the effect of nominative pronouns is mitigated by the effect of nominative correlates and expletives (cf. Nominative Pronominality in Figure 6.4), because German is more susceptible to the referential status of a given argument than Icelandic.

6.5 Summary and conclusions

This article has provided corpus-based arguments for an analysis of a subgroup of German Nom-Dat verbs, sometimes also referred to in the literature as "dative-object experiencer verbs", as alternating between two different argument structure constructions, Dat-Nom and Nom-Dat. Such Dat-Nom/Nom-Dat verbs constitute a syntactically and semantically homogeneous class, and one of their core properties is their ability to instantiate both a dative-first and a nominative-first structure without one order being more marked than the other. Somers and Barðdal (2023) have shown that this is not a commonality of all German verbs licencing a dative and a nominative. Rather, ordinary Nom-Dat verbs like *helfen* 'help' virtually always instantiate the Nom-Dat order. Their study of five such verbs, involving 390 tokens also harvested from deTenTen13, has revealed that *helfen* verbs allocate the preverbal slot to the nominative 96% of the time, while the remaining 4% are clear examples of topicalization. All these cases involve attested clauses where both arguments are full NPs.

The current study is based on the word order statistics of 76 German Dat-Nom/Nom-Dat verbs which have been annotated for ten different variables. A multivariate analysis has revealed that, across configurations, the variation in word order is mainly caused by the variables Nominative (Pro)nominality, Dative (Pro)nominality, and Length. Additionally, the Intra Class Correlation coefficient of the random intercept verb equals 27%, which suggests that the verb itself explains about a quarter of the word order variation. Interestingly, the extent to which the verbs in question alternate is gradient, with some verbs being strongly attracted to the Dat-Nom order, others to the Nom-Dat order, and others still to both orders approximately equally often. This begs the questions to what extent the verbs at either end of the spectrum truly alternate and whether it is possible that these are potentially non-alternating Nom-Dat verbs (like the aforementioned *help*-verbs) or non-alternating Dat-Nom verbs. The latter class has been posited for Modern Icelandic, and corpus studies by Somers and Barðdal (2022) and Somers, Jenset and Barðdal (2023) have shown that these verbs indeed correlate with a dative-before-nominative argument structure. An experimental study may help shed light on their actual status.

Contexts involving only full NPs reveal a higher share of Dat-Nom tokens than contexts involving at least one pronoun, but the difference is smaller than anticipated, because the topicalizing effect of nominative pronouns is mitigated by correlates and expletives; these alternate more readily between both word order patterns than fully referential nominative pronouns. The second multivariate analysis identifies the variables Dative Animacy, Nominative Definiteness, Nominative Animacy, and Length as the main variables influencing the choice between Dat-Nom and Nom-Dat, but, with the exception of Length, there is not a single variable that definitively determines word order variation. This finding confirms the hypothesised neutrality of both word order alternants.

We have also shown here that the distribution between Dat-Nom and Nom-Dat word orders is relatively even for full NPs, or 46% Dat-Nom vs. 54% Nom-Dat. This stands in stark contrast with 96% Nom-Dat and 4% Dat-Nom for verbs like *helfen* 'help'. Corresponding numbers for four established alternating Dat-Nom/Nom-Dat verbs in Modern Icelandic reveals a 53% Dat-Nom vs. 47% Nom-Dat distribution when both arguments are full NPs (cf. Somers and Barðdal 2022). Thus, the numbers are strikingly similar for our Icelandic and German datasets when full NPs are involved.

Turning to the distribution of Dat-Nom/Nom-Dat word orders irrespective of the (pro)nominal status of the arguments, our German dataset reveals that 42% of the examples instantiate the Dat-Nom order, while 58% instantiate the Nom-Dat order. A comparison between these numbers and the baseline established by Sahel and Jonischkait (2008) for German, where only 12% of the instances are topicalizations, shows that a topicalization analysis of the Dat-Nom alternant in German is entirely out of the question. In contrast, these numbers support the analysis suggested here, that these are alternating Dat-Nom/Nom-Dat predicates and that the dative is the subject when the relevant verbs instantiate the Dat-Nom argument structure, while the nominative is the subject when the verbs instantiate the diametrically opposite Nom-Dat argument structure.

References

Abbott, Barbara. 2004. Definiteness and indefiniteness. In Laurence R. Horn & Gregory Ward (eds.), *The Handbook of Pragmatics*, 122–149. Oxford: Blackwell.

Alexiadou, Artemis, Elena Anagnostopoulou & Martin Everaert. 2004. Introduction. In Artemis Alexiadou, Elena Anagnostopoulou & Martin Everaert (eds.), *The Unaccusativity Puzzle: Explorations of the Syntax-Lexicon Interface*, 1–21. Oxford: Oxford University Press.

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.

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.

Ashby, William J. & Paola Bentivoglio. 1993. Preferred argument structure in spoken French and Spanish. *Language Variation and Change* 5: 61–76.

Ashby, William J. & Paola Bentivoglio. 2003. Preferred Argument Structure across time and space: A comparative diachronic analysis of French and Spanish. In John W. Du Bois, Lorraine E. Kumpf & William J. Ashby (eds.), *Preferred Argument Structure: Grammar as architecture for function*, 61–80. Amsterdam: John Benjamins.

Bader, Markus & Jana Häussler. 2010. Word order in German: A corpus study. *Lingua* 120: 717–762.

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

Barðdal, Jóhanna. 2004. The semantics of the impersonal construction in Icelandic, German and Faroese: Beyond thematic roles. In Werner Abraham (ed.), *Focus on Germanic typology*, 105–137. Berlin: Akademie Verlag.
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, Carlee Arnett, Stephen Mark Carey, Thórhallur Eythórsson, Gard B. Jenset, Guus Kroonen & Adam Oberlin. 2016. Dative subjects in Germanic: A computational analysis of lexical semantic verb classes across time and space. *Language Typology and Universals* 69(1): 49–84.
Barðdal, Jóhanna & Thórhallur Eythórsson. 2009. The origin of the oblique subject construction: An Indo-European comparison. In Vit Bubenik, John Hewson & Sarah Rose (eds.), *Grammatical Change in Indo-European Languages*, 179–193. Amsterdam: John Benjamins.
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.
Barðdal, Jóhanna, Thomas Smitherman, Valgerður Bjarnadóttir, Serena Danesi, Gard B. Jenset & Barbara McGillivray. 2012. Reconstructing constructional semantics: The dative subject construction in Old-Norse Icelandic, Latin, Ancient Greek, Old Russian and Old Lithuanian. *Studies in Language* 36(3): 511–547.
Bayer, Josef. 2004. Non-nominative subjects in comparison. In Peri Bhaskararao & Karumuri V. Subbarao (eds.), *Non-Nominative Subjects*, Vol. 1, 49–76. Amsterdam: John Benjamins.
Behaghel, Otto. 1909/10. Beziehungen zwischen Umfang und Reihenfolge von Satzgliedern. *Indogermanische Forschungen* 25: 110–142.
Belletti, Adriana & Luigi Rizzi. 1988. Psych-verbs and θ–theory. *Natural Language and Linguistic Theory* 6: 291–352.
Bickel, Balthasar. 2004. The syntax of experiencers in the Himalayas. In Peri Bhaskararao & Karumuri V. Subbarao (eds.), *Non-Nominative Subjects*, Vol. 1, 77–111. Amsterdam: John Benjamins.
Bresnan, Joan & Marylin Ford. 2010. Predicting syntax: Processing dative constructions in American and Australian varieties of English. *Language* 86(1): 168–213.
Chappell, Hilary & William McGregor. 1996. Prolegomena to a theory of inalienability. In Hillary Chappell & William McGregor (eds.), *The Grammar of Inalienability: A Typological Perspective on Body Part Terms and the Part-Whole Relation*, 3–30. Berlin/New York: Mouton de Gruyter.
Comrie, Bernard. 1981. *Language Universals and Linguistic Typology: Syntax and Morphology*. Chicago: University of Chicago Press.
Croft, William. 2003. *Typology and Universals*. 2nd edn. Cambridge: Cambridge University Press.
den Besten, Hans. 1985. The ergative hypothesis and free word order in Dutch and German. In Jindrich Toman (ed.), *Studies in German Grammar*, 23–64. Dordrecht: Foris.
Dowty, David. 1991. Thematic proto-roles and argument selection. *Language* 67(3): 547–619.
Draye, Luk. 1996. The German dative. In William Van Belle & Willy Van Langendonck (eds.), *The Dative 1: Descriptive Studies*, 155–215. Amsterdam: John Benjamins.
Du Bois, John W. 1987. The discourse basis of ergativity. *Language* 63(4): 805–855.
Du Bois, John W. 2003. Argument structure: Grammar in use. In John W. Du Bois, Lorraine E. Kumpf & William J. Ashby (eds.), *Preferred Argument Structure: Grammar as Architecture for Function*, 11–60. Amsterdam: John Benjamins.
Duden = *Die Grammatik*. 2016. Berlin: Dudenverlag.

Engel, Ulrich. 1970. *Regeln zur Wortstellung*. [Forschungsberichte des Instituts für Deutsche Sprache 5.] Mannheim: Institut für Deutsche Sprache.
Eythórsson, Thórhallur & Jóhanna Barðdal. 2005. Oblique subjects: A common Germanic inheritance. *Language* 81(4): 824–881.
Fanselow, Gisbert. 2002. Quirky subjects and other specifiers. In Ingrid Kaufmann & Barbara Stiebels (eds.), *More than Words: A Festschrift for Dieter Wunderlich*, 227–250. Berlin: Akademie Verlag.
Gutzmann, Daniel. 2007. Eine Implikatur konventioneller Art: Der Dativus Ethicus. *Linguistische Berichte* 211: 277–308.
Grimm, Jacob. 1837. *Deutsche Grammatik,* Vol. IV. Göttingen: Dieterichsche Buchhandlung.
Haider, Hubert. 1985. Über *sein* oder nicht *sein*: Zur Grammatik des Pronomens *sich*. In Werner Abraham (ed.), *Erklärende Syntax des Deutschen*, 223–254. Tübingen: Gunter Narr.
Haider, Hubert. 2005. How to turn German into Icelandic – and derive the OV–VO contrast. *Journal of Comparative Germanic Syntax* 8: 1–53.
Haider, Hubert. 2010. *The Syntax of German*. Cambridge: Cambridge University Press.
Haspelmath, Martin. 2001. Non-canonical marking of core arguments in European languages. In Alexandra Aikhenvald, Robert M. W. Dixon & Masayuki Onishi (eds.), *Non-Canonical Subjects and Objects*, 53–83. Amsterdam: John Benjamins.
Hawkins, John A. 1992. Syntactic weight versus information structure in word order variation. In Joachim Jacobs, (ed.), *Informationsstruktur und Grammatik*, 196–219. Wiesbaden: Westdeutscher Verlag.
Hentschel, Elke & Harald Weydt. 2021. *Handbuch der deutschen Grammatik*. Berlin/Boston: De Gruyter.
Heylen, Kris. 2005. *Zur Abfolge (pro)nominaler Satzglieder im Deutschen: Eine korpusbasierte Analyse der relativen Abfolge von nominalem Subjekt und pronominalem Objekt im Mittelfeld*. Leuven: Katholieke Universiteit Leuven dissertation.
Hole, Daniel. 2014. *Dativ, Bindung und Diathese*. Berlin: De Gruyter.
Hopper, Paul J. & Sandra A. Thompson. 1980. Transitivity in grammar and discourse. *Language* 56(2): 251–299.
Johnson, Cynthia A., Peter Alexander Kerkhof, Leonid Kulikov, Esther Le Mair & Jóhanna Barðdal. 2019. Argument structure, conceptual metaphor and semantic change: How to succeed in Indo-European without really trying. *Diachronica* 36(4): 463–508.
Klein, Katarina & Silvia Kutscher. 2002. Psych-verbs and lexical economy. *Theorie des Lexikons* 122: 1–41.
Kallulli, Dalina. 2007. Rethinking the passive/anticausative distinction. *Linguistic Inquiry* 38(4): 770–780.
Kempen, Gerard & Karin Harbusch. 2005. The relationship between grammaticality ratings and corpus frequencies: A case study into word order variability in the midfield of German clauses. In Stephan Kepser & Marga Reis (eds.), *Linguistic evidence*, 329–349. Berlin: Mouton de Gruyter.
Kilgarriff, Adam, Pavel Rychlý, Pavel Smrž & David Tugwell. 2004. The Sketch Engine. *Proceedings of the 11th EURALEX International Congress*. 105–116.
Kilgarriff, Adam, Vít Baisa, Jan Bušta, Miloš Jakubíček, Vojtěch Kovář, Jan Michelfeit, Pavel Rychlý & Vít Suchomel. 2014. The Sketch Engine: Ten years on. *Lexicography* 1: 7–36.
Kutscher, Silvia. 2009. *Kausalität und Argumentrealisierung: Zur Konstruktionsvarianz bei Psychverben am Beispiel europäischer Sprachen*. Tübingen: Niemeyer.
Lambrecht, Knud. 1986. *Topic, focus, and the grammar of spoken French*. Berkeley, CA: University of California dissertation.
Lenerz, Jürgen. 1977. *Zur Abfolge nominaler Satzglieder im Deutschen*. Tübingen: Gunter Narr.
Lyons, Christopher. 1999. *Definiteness*. Cambridge: Cambridge University Press.
Pankau, Andreas. 2016. Quirky subjects in Icelandic, Faroese, and German: A Relational Grammar account. In Doug Arnold, Miriam Butt, Berthold Crysmann, Tracy Holloway King & Stefan

Müller (eds.), *Proceedings of the Joint 2016 Conference on Head-driven Phrase Structure Grammar and Lexical Functional Grammar*, *Polish Academy of Sciences, Warsaw, Poland*, 499–519. Stanford, Ca: CSLI Publications.
Primus, Beatrice. 1994. Grammatik und Performanz: Faktoren der Wortstellungsvariation im Mittelfeld. *Sprache und Pragmatik* 32: 39–86.
Primus, Beatrice. 2012. *Semantische Rollen*. Heidelberg: Carl Winter.
Reis, Marga. 1982. Zum Subjektbegriff im Deutschen. In Werner Abraham (ed.), *Satzglieder im Deutschen: Vorschläge zur syntaktischen, semantischen und pragmatischen Fundierung*, 171–211. Tübingen: Gunter Narr.
Rosengren, Inger. 1993. Wahlfreiheit mit Konsequenzen: Scrambling, Topikalisierung und FGH im Dienste der Informationsstrukturierung. In Marga Reis (ed.), *Wortstellung und Informationsstruktur*, 251–312. Tübingen: Niemeyer.
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.
Sahel, Said & Julia Jonischkait. 2008. Syntaktische Funktionen im Vorfeld: Eine empirische Studie. *Muttersprache* 118(4): 281–294.
Schlesewsky, Matthias & Ina Bornkessel. 2006. Context-sensitive neural responses to conflict resolution: Electrophysiological evidence from subject-object ambiguities in language comprehension. *Brain Research* 1098(1): 139–152.
Schmid Josef. 1988. *Untersuchungen zum sogenannten freien Dativ in der Gegenwartssprache und auf Vorstufen des heutigen Deutsch.* Frankfurt: Peter Lang.
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: Lund University dissertation.
Sigurðsson, Halldór Ármann. 1992. The case of quirky subjects. *Working Papers in Scandinavian Syntax* 49: 1–26.
Somers, Joren. 2021. Onpersoonlijke datief-nominatiefwerkwoorden in het hedendaagse Duits: Definitie, inventaris en semantische classificatie. *Handelingen – Koninklijke Zuid-Nederlandse Maatschappij Voor Taal- en Letterkunde en Geschiedenis (KZM)* 75: 211–238.
Somers, Joren. 2023. An evaluation of the Dat-Nom and the Dat-Nom/Nom-Dat hypotheses for five German verbs of success and failure. *Papers of the Linguistic Society of Belgium* 17: 54–75.
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.
Somers, Joren, Gard B. Jenset & Jóhanna Barðdal. 2024. Argument structure constructions in competition: The Dat-Nom/Nom-Dat alternation in Icelandic. *Nordic Journal of Linguistics*: online first article. DOI: https://doi.org/10.1017/S0332586524000039.
Verhoeven, Elizabeth. 2009. Subjects, agents, experiencers, and animates in competition: Modern Greek argument order. *Linguistische Berichte* 219: 355–376.
Verhoeven, Elisabeth. 2010. Transitivity in Chinese experiencer object verbs. In Patrick Brandt & Marco García García (eds.), *Transitivity: Form, Meaning, Acquisition and Processing*, 95–118. Amsterdam: John Benjamins.
Verhoeven, Elisabeth. 2015. Thematic asymmetries do matter! A corpus study of German word order. *Journal of Germanic Linguistics* 27(1): 45–104.

Viti, Carlotta. 2016. Areal distribution of argument marking of Indo-European experience predicates. *The Journal of Indo-European Studies* 44(1): 1–84.
Wunderlich, Dieter. 2009. The force of lexical case: German and Icelandic compared. In Kristin Hanson & Sharon Inkelas (eds.), *The Nature of the Word: Studies in Honor of Paul Kiparsky*, 587–620. Cambridge, Ma: MIT Press.
Zaenen, Annie, Joan Maling & Höskuldur Thráinsson. 1985. Case and grammatical functions: The Icelandic passive. *Natural Language and Linguistic Theory* 3: 441–483.
Zifonun, Gisela, Ludger Hoffmann & Bruno Strecker. 1997. *Grammatik der deutschen Sprache*, Vol. 2. Berlin/New York: Mouton de Gruyter.

Appendix 3

Verb	Meaning
abhandenkommen	'lose'
angehören	'belong'
aufdämmern	'dawn on'
auffallen	'strike, notice'
aufstoßen	'strike, notice; make (sb) burp; annoy'
begegnen	'come across; experience, be confronted with'
behagen	'please'
beifallen	'occur to'
dämmern	'dawn on'
dazwischenkommen	'come up'
drohen	'threaten'
dünken	'seem, appear'
einfallen	'occur to; remember'
einleuchten	'be clear'
entfahren	'escape (words)'
entgegenschallen	'hear'
entgegenschlagen	'beat against; experience intensely'
entgehen	'miss out on; fail to notice'
entgleiten	'drop; slip; escape (words)'
entschlüpfen	'escape (words)'
ermangeln (an)	'lack'
erscheinen	'appear before; seem, appear'

Verb	Meaning
fehlen	'lack; miss; be unhealthy; not be available; need'
fernliegen	'not occur to; not be someone's intention'
folgen	'succeed, follow'
gebrechen (an)	'lack'
gebühren	'be due to'
gefallen	'please, like'
gelingen	'succeed'
gelten	'be worth; be considered'
genügen	'be enough, be sufficient'
geziemen	'befit'
glücken	'succeed'
grauen	'dread, be afraid of'
grausen	'dread, be afraid of'
guttun	'do good'
imponieren	'impress'
klarwerden	'become clear'
konvenieren	'be convenient, suit'
leichtfallen	'be easy'
leidtun	'take pity; be sorry'
mangeln (an)	'lack'
missfallen	'displease'
missglücken	'fail'
misslingen	'fail'
missraten	'fail'
munden	'taste good'
nahegehen	'touch, care'

Verb	**Meaning**
nützen	'be of use'
obliegen	'be someone's duty'
passieren	'happen to'
pressieren	'be urgent'
reichen	'suffice'
scheinen	'seem, appear'
schlechtgehen	'suffer; go through economically hard times'
schmecken	'taste; have a particular taste (lit.); have a particular taste (fig.); please'
schwanen	'have a sense of foreboding'
schwerfallen	'be difficult'
stinken	'enrage'
unterlaufen	'happen to; make a mistake'
verlorengehen	'be lost'
vorkommen	'happen to; experience; strike (as)'
vorschweben	'have in mind'
widerfahren	'happen (to), experience'
widerstreben	'be reluctant'
winken	'can expect, be in store for'
wohltun	'do sb. good'
zufallen	'fall to, receive (lit.); fall to, receive (fig.)'
zufließen	'receive'
zugehören	'belong to'
zugutekommen	'benefit'
zupasskommen	'come in handy'
zusetzen	'affect badly'
zustoßen	'happen to'

Verb	**Meaning**
zuteilwerden	‘be bestowed upon’
zuwachsen	‘receive’

Appendix 4

	estimated coefficient (SE)	z-value	p-value
Intercept	0.55 (0.38)	1.43	= 0.15
Dative Animacy			
individual	0.06 (0.07)	0.86	0.39
collective			
inanimate	-0.51 (0.10)	-5.26	< 0.001
Dative (Pro)nominality			
indefinite pronoun	-2.65 (0.23)	-11.41	< 0.001
demonstrative pronoun			
personal pronoun	-2.89 (0.21)	-13.91	< 0.001
definite full NP	-1.76 (0.20)	-8.65	< 0.001
indefinite full NP	-1.41 (0.21)	-6.64	< 0.001
Dative Person			
non-local	-0.43 (0.07)	-6.16	< 0.001
Dative Number			
singular	0.26 (0.05)	5.09	< 0.001
Nominative Animacy			
inanimate	-0.30 (0.27)	-1.12	0.26
NA	-0.51 (1.12)	-0.45	0.64
individual	-0.56 (0.28)	-1.98	0.04
collective			
Nominative			

(Pro)nominality			
indefinite pronoun	2.36 (0.14)	17.16	< 0.001
demonstrative pronoun			
personal pronoun	1.18 (0.13)	8.85	< 0.001
personal pronoun (corr)	2.06 (1.09)	1.90	0.06
personal pronoun (expl)	2.65 (1.09)	2.42	0.01
definite full NP	1.28 (0.10)	13.19	< 0.001
indefinite full NP	1.88 (1.10)	18.44	< 0.001
Nominative Number			
singular	0.17 (0.07)	2.39	0.02
Length Difference	0.03 (0.005)	6.23	< 0.001
Random intercept	Var (SD)		
Verb	1.21 (1.1)		

Chapter 7
Discussion and general conclusion

7.1 Introduction

In this thesis, I have investigated alternating Dat-Nom/Nom-Dat verbs in German and Icelandic. I have compiled an exhaustive overview of these verbs in German and classified them semantically. In addition, I have carried out two parallel corpus studies, one for each language, using two sister corpora: the deTenTen13 corpus for German and the IsTenTen20 corpus for Icelandic. The study builds on a total of 81 target verbs: 76 for German and five for Icelandic. The results show that both languages possess a class of Dat-Nom/Nom-Dat verbs and that the factors correlating with the alternation are similar in both languages. In addition, both Icelandic and German, of course, possess a class of non-alternating Nom-Dat verbs of the *help*-type, whereas Icelandic also possesses a third class of non-alternating Dat-Nom verbs of the *like*-type.

The current chapter provides an overview of the main findings of this thesis. These are discussed in Section 7.2. The next section, Section 7.3, bridges the final gap in this thesis by offering a statistical comparison of the target verbs, Icelandic and in German alike, with the control set of *helfen*-verbs for German and both control sets of *hjálpa*- and *líka*-verbs for Icelandic. Such a comparison offers a tool to single out the verbs that fail to pass the word order test. The outlying verbs are subsequently discussed in more detail. Section 7.4 is dedicated to the theoretical implications of this research. The final section, Section 7.5, opens a window onto future research questions.

7.2 Main findings

This section reiterates the main findings this thesis has yielded by explicitly tying in with each of the seven research questions presented in Chapter 1.

Research question 1: How many, and which, alternating Dat-Nom/Nom-Dat verbs exist in Modern German?

The first research question is oriented towards an exhaustive overview of potentially alternating Dat-Nom/Nom-Dat verbs in German. Starting from a list of 56 such verbs first inventoried by Barðdal (2004), an additional 90 types are identified here, thus increasing the number of established Dat-Nom/Nom-Dat verbs to 146. The search builds on the *Duden Universalwörterbuch*, the *Deutsches Referenzkorpus* (DeReKo), two valency dictionaries, as well as Wegener's (1985) seminal book on the dative in Present-Day German.

Inspection of these 146 verbs has revealed the importance of polysemy for the current study. First, some verbs, like *gefallen* or *gelingen*, are monosemous Dat-Nom/Nom-Dat verbs. That is, they license a single sense, which is also compatible with the alternating Dat-Nom/Nom-Dat construction. Secondly, there are polysemous verbs which are compatible with the alternating construction in some, but not all, of their senses. One relevant example is *genügen* 'be enough, be sufficient; fulfil'. Only the former sense displays a thematic asymmetry between the dative and the nominative, not the latter. As such, only the former sense is compatible with the alternating construction. Finally, there are as many as 40 verbs licensing multiple Dat-Nom/Nom-Dat senses. One relevant example, discussed in Chapter 6, is *begegnen* 'come across; experience, be confronted with', of which both senses are compatible with the alternating construction. The issue of polysemy is also relevant in the context of research question 2, as will be elaborated on in more detail below.

As to the second part of the first research question, i.e. *which* Dat-Nom/Nom-Dat verbs Modern German possesses, it is shown in Chapter 6 that such verbs display very specific properties: they are low in agentivity, they display a thematic asymmetry between the dative and the nominative and they pertain to one of 16 semantic verb classes (cf. infra). Syntactically, they fully subsume the class of *piacere*-verbs, as well as that of unaccusative verbs. However, the opposite is not necessarily true, as some Dat-Nom/Nom-Dat verbs are neither *piacere*-verbs nor unaccusative verbs. Finally, alternating Dat-Nom/Nom-Dat verbs are bound to pass the cohort of subjecthood tests presented in Chapter 1, both for the dative and for the nominative argument, although obviously not at the same time.

Research question 2: What are the semantic domains to which German Dat-Nom/Nom-Dat verbs pertain?

Research question 2 is aimed at the semantic distribution of the Dat-Nom/Nom-Dat construction. In addressing the first research question, I have shown that the alternation is first and foremost associated with particular verb senses, and not with verbs in general. Clustering therefore proceeded on a sense-specific basis by building on a typology first put forward by Barðdal et al. (2016). Based on data for an array of historical daughter languages, Barðdal et al. show that the dative-subject construction in Proto-Germanic covered as many as 15 different semantic fields. Chapter 2 of the present thesis has confirmed the validity of their typology for Modern German, for which only one additional semantic field has to be posited, namely that of 'belonging'. A complete overview of all 16 lexical fields is presented here:

(1)	emotion	(9)	bodily states
(2)	cognition	(10)	modality
(3)	gain and benefit	(11)	ontological states
(4)	happening	(12)	evidentiality
(5)	hindrance	(13)	success
(6)	attitude	(14)	belonging
(7)	possession	(15)	speaking
(8)	perception	(16)	natural occurrences

Importantly, the extent to which the 202 senses are divided across these classes is not proportionate, but exponential, which means that every next class contains half the number of senses compared to the previous one. Such a distribution is also known as a Zipf distribution (Zipf 1935: 25).

Barðdal et al.'s (2016) typology offers a tool that is sufficiently fine-grained to distinguish between the different semantic subfields of the alternating construction. Terms like 'psych verbs' or 'experiential verbs' obscure this semantic diversity, since they create the impression that alternation is a property solely associated with verbs of emotion or verbs of experience. The latter term is often used as an umbrella term for any verb licensing non-canonical case marking, but Chapter 6 has shown that it may only be applied meaningfully to verbs of emotion, verbs of cognition, verbs of perception, verbs referring to bodily states and verbs of evidentiality.

Research question 3: Which factors in German are associated with the speaker's choice of one linearisation pattern over the other?

To answer the third research question, a maximum of 200 tokens for a total of 76 German Dat-Nom/Nom-Dat verbs were collected. This resulted in a 13,000-token dataset. The data were retrieved from the deTenTen13 corpus, manually cleaned and annotated for 10 variables. A mixed-effects logistic regression model was fitted to evaluate the simultaneous effects of the annotated predictor variables, using the verbs as a random intercept. This is the topic of Chapter 6. One of the most remarkable findings the model has yielded is that

up to 27% of the word order variation may be attributed to a verb effect. This shows that individual verbs have individual preferences, regardless of the influence of secondary factors. The model additionally revealed that dative (pro)nominality, nominative (pro)nominality and length are key factors correlating with the choice of the dative-before-nominative and the nominative-before-dative alternants across configurations. In the double-NP configuration, the strongest predictors turned out to be length and nominative indefiniteness.

Research question 4: Are all German Dat-Nom verbs veritable Dat-Nom/Nom-Dat verbs in that both arguments, dative and nominative, pass the word order test, or do some behave as either non-alternating Dat-Nom verbs or as ordinary Nom-Dat verbs?

In order to address the fourth research question, the study has made use of a control set of five German Nom-Dat verbs, also termed *helfen*-verbs. This set, which is used to set a baseline for topicalisation, comprises the verb types *ähneln*, *danken*, *helfen*, *vertrauen* and *widersprechen*. A sample containing 200 tokens for each of these five verbs has revealed that 93%, or 928 out of 1,000 tokens, allocate the preverbal slot to the nominative, with the remaining 7%, or 72 tokens, featuring the dative-before-nominative order, thus instantiating a topicalisation construction (Somers & Barðdal 2023: 9). In the double-NP configuration, numbers are even more extreme, with only 4% of tokens instantiating the dative-before-nominative order. These findings support the assumption that Nom-Dat verbs select for a nominative subject and a dative object.

Chapter 5 compares the numbers for *helfen*-verbs with those obtained for *gefallen*- and *genügen*-verbs. It turns out that both these sets of target verbs instantiated either word order pattern in the double-NP configuration considerably more often than the control set of *helfen*-verbs. *Gefallen*-verbs accounted for 64 instances of the dative-before-nominative order, i.e. 42%, and 88 instances of the converse nominative-before-dative-order, i.e. 58%. *Genügen*-verbs yielded highly similar frequencies, with 44% of tokens, or 116, instantiating a dative-first structure and 56%, or 147, a nominative-first structure. Their status as alternating verbs is thus confirmed.

One final question is how the remaining German target verbs fare with regard to the Nom-Dat baseline. As shown in Chapter 6, each target verb alternates between both word order patterns to different degrees. This begs the question which verbs do or do not deviate significantly from the control group, either because they are heavily skewed towards the dative-before-nominative order, or because they exhibit the reverse inclination towards the nominative-before-dative order. This issue is further developed in Section 7.3 below.

Research question 5: How do alternating verbs in Icelandic behave vis-à-vis the word order test?

Research question 5, which was addressed in Chapter 4, targeted a set of five Icelandic alternating verbs, namely *duga*, *dyljast*, *endast*, *henta* and *nægja*. Their results have been compared to two control sets: one comprising five Nom-Dat verbs and one comprising five Dat-Nom verbs. Exploratory analysis already shows that the five alternating verbs behave radically differently than either control group: in the NP-V-NP configuration, they instantiate the nominative-dative order 72% of the time, versus 28% for the dative-nominative order. Excluding the outlier *henta*, the prevalence of the dative-before-nominative order increases to 46%, versus 54% for the nominative-before-dative order. Both classes of non-alternating verbs realise the syntactic subject, the dative for Dat-Nom verbs and the nominative for Nom-Dat verbs, 99% of the time in initial position, again in contexts involving two full NPs. Thus, alternating verbs boast two argument structure constructions, not one.

Another significant finding is that the alternating verb *henta* does not pass the word order test, as it only occurs with the nominative-before-dative order in the corpus sample, both in the double-NP configuration (86 tokens, or 100%) as well as across configurations (200 tokens, or 100%). This issue is further addressed in Section 7.3 below.

Research question 6: Which factors in Icelandic motivate the speaker's choice of one linearisation pattern over the other?

Chapter 4 similarly used a logistic regression model to examine the set of Icelandic alternating verbs in light of nine predictor variables: case, (pro)nominality, pronoun type, referentiality, person, number, definiteness, animacy and length. The analysis only takes into account those verbs passing the word order test, thus leaving out the outlier *henta*. The data are modelled using ordinary logistic regression, as the dataset comprises too few tokens to defend the added complexity of mixed-effects modelling. The analysis revealed that alternating verbs in the double-NP configuration are skewed towards the nominative-before-dative order, but that indefinite and lengthy constituents, regardless of case marking, tend to take the second argument position. The same three variables, nominative case marking, indefiniteness and length, were also singled out as strong predictors in the across-configurations dataset, in addition to animacy, person and number. Chapter 4 further revealed that the lion's share of factors identified by the logistic regression model correlate with topicality, thus confirming a hypothesis by Barðdal (2001) as well as Barðdal, Eythórsson & Dewey (2014, 2019) that alternating verbs are, in fact, topic-first verbs.

Research question 7: How do alternating verbs behave across Icelandic and German? Are there any cross-linguistic generalisations to be made?

The last research question, as discussed in Chapter 5 focuses on a cross-linguistic comparison. To that end, each of the 15 Icelandic verb types, Dat-Nom, Nom-Dat and Dat-Nom/Nom-Dat, has been matched to a cognate or a (near)-synonym in German. Thus,

the cross-linguistic dataset comprises a total of 6,000 tokens, i.e. 3,000 each for German and Icelandic. A conditional inference tree model suggests that word order variation with alternating verbs correlates with the same factors in both languages under study, most importantly case and configuration. Furthermore, the split for configuration revealed that double-NPs are sensitive to virtually the same factors in both languages, while configurations involving at least one pronoun exhibit more language-specific tendencies.

One of these inter-language differences relating to the influence of pronouns is that Icelandic configurations involving at least one pronoun show a very strong tendency towards the nominative-first order, while German allows for more variation. However, the tree model seems to inflate the nominative-first effect for Icelandic somewhat. Recall that Icelandic alternating verbs in configurations involving at least one pronoun instantiate the nominative-before-dative order 75% of the time, with 590 tokens, and the reverse dative-before-nominative order 25% of the time, with 193 tokens. As many as 168 out of 193 dative-first tokens contain a dative pronoun followed by a nominative full NP, again showing that configuration influences word order distributions.

7.3 Comparing the targets to the baseline

One of the main findings of this study is that each Dat-Nom/Nom-Dat verb, in Icelandic and German alike, alternates to different degrees. That is, the extent varies to which the target verbs instantiate either the dative-before-nominative order or the nominative-before-dative order. This begs the question which targets do not deviate significantly from the control set of *helfen*-verbs for German or both control sets of *hjálþa-* and *líka*-verbs for Icelandic. The current section is therefore devoted to a comparison of the targets to these language-specific baselines. Section 7.3.1 zooms in on Icelandic, Section 7.3.2 is focused on German. Crucially, because word order is sensitive to pronominal influence, my conclusions are solely based on the numbers obtained for double NPs. This naturally excludes the German verbs *ermangeln (an)*, *gebrechen (an)*, *grauen*, *grausen*, *leidtun*, *mangeln (an)*, *pressieren* and *schlechtgehen*, as these do not yield any tokens involving lexical NPs only.

7.3.1 Icelandic

The set of Icelandic target verbs, here labelled *nægja*-verbs, comprises five types: *duga* 'suffice, be enough', *dyljast* 'be hidden to somebody', *endast* 'last', *henta* 'suit, befit' and *nægja* 'be enough, be sufficient'. Their word order distributions are discussed in Chapter 4. A

comparison against both baselines, the Nom-Dat baseline set by *hjálþa*-verbs and the Dat-Nom baseline set by *líka*-verbs, reveals significant results for four out of five verbs, i.e. *duga*, *dyljast*, *endast* and *nægja*. These results are presented in Table 7.1 below:

Table 7.1 Results for eight chi-square tests comparing the alternating verbs *duga*, *dyljast*, *endast* and *nægja* to the *hjálþa*-baseline and the *líka*-baseline

target	**Nom-Dat**		**Dat-Nom**		vis-à-vis *hjálþa*-verbs	vis-à-vis *líka*-verbs
	N	**f**	**N**	**f**		
duga	33	79%	9	21%	$p < 0.001$	$p < 0.001$
dyljast	2	25%	6	75%	$p < 0.001$	$p < 0.001$
endast	9	29%	22	71%	$p < 0.001$	$p < 0.001$
nægja	27	54%	23	46%	$p < 0.001$	$p < 0.001$

The verbs *duga*, *dyljast*, *endast* and *nægja*, significantly deviate from the baseline set by the class of Nom-Dat, or *hjálþa*-verbs as well as the baseline set by Dat-Nom, or *líka*-verbs. This finding is based on a comparison of the [NP-V-NP] frequencies for each individual *nægja*-verb to the [NP-V-NP] frequencies obtained for the entire class of *hjálþa*-verbs (column 2) as well as the entire class of *líka*-verbs (column 3).

Note that *dyljast* only yields eight tokens in the double-NP configuration, six instantiating the dative-before-nominative order and two instantiating the nominative-before-dative order. These numbers are too low to attach any significance to. I thus leave it to future research to provide more conclusive evidence as to whether *dyljast* passes the word order test.

Only *henta* 'suit, befit', the last member of the class of alternating verbs, fails to pass the word order test: across configurations, all 200 tokens of this verb, or 100%, instantiate the nominative-before-dative pattern and in the double-NP configuration, as many as 86 tokens, again 100%, instantiate a nominative-first structure. A statistical comparison with the class of non-alternating Nom-Dat verbs evidently yields a non-significant result ($p = 1$). Thus, the null hypothesis, assuming that *henta* belongs to a different class than the five *hjálþa*-verbs, cannot be rejected.

This is a peculiar finding, as both nominal arguments of *henta* still pass the subjecthood tests (Barðdal 1999, 2001). This is also evident from (1a–b) below, in which either argument, the nominative in (1a) and the dative in (1b), occupies the subject slot, with the other argument occupying the object slot:

(1) a. *Þetta hefur alltaf hentað mér.*
that.NOM has always suited me.DAT
'This has always pleased me.'

(1) b. *Mér* *hefur* *alltaf* *hentað* *þetta.*
me.DAT has always pleased this.NOM
'I have always liked this.' (Barðdal 2001: 50)

In order to provide evidence from word order that the dative of *henta* may indeed behave syntactically as a subject, I performed a Google search for the string *hentað það* 'suited it/that'. This query searches for contexts in which the past participle *hentað* 'suited' is followed by the nominative pronoun *það* 'it/that'. As such, it singles out clauses with complex tenses in which the nominative is realised in an object position. The dative subject may either precede the auxiliary, thus occurring in clause-initial position, or it may invert with the auxiliary. Four relevant examples this search query has returned are shown in (2–3) below. Example (2a) confirms that the dative of *henta* may take on the subject role in main clauses, while (2b) illustrates that the same is true for subordinate clauses. Examples (3a–b) both contain a dative subject inverting with the finite verb.

Dat-V-Nom

(2) a. *Allt of mörgum* *hefur* *hentað* *það* *ljómandi*
all too many.DAT has suited it.NOM splendid
að ástandið breyttist ekki.
that the situation changed not.
'Too many people found it splendid that the situation did not change.'

(2) b. *… að* *þeim* *geti* *hentað*
that them.DAT could suited
það hagkerfi, sem …
the.NOM economic.system.NOM that
'… that they could find suitable the economic system that …'

V-Dat-Nom

(3) a. *… enda* *hefur* *einyrkjabændum* *aldrei* *hentað*
… after all has single.farmers.DAT never suited
það.
that.NOM
'After all, that has never suited single farmers.'

(3) b. *og þess vegna* *geti* *okkur* *ekki* *hentað*
and therefore could us.DAT not suited
það sama og þeim.
the.NOM same.NOM as them
'And therefore we could not find the same thing suitable as them.'

The evidence presented above is arguably rare, but it does confirm that the dative argument of *henta* passes the word order test. This finding evidently raises a follow-up question, namely what the relation is between frequency and acceptability in language, a

topic out of the scope of this work (but see Barðdal & Eythórsson 2006; Divjak 2017; Flach 2020, *inter alia*).

7.3.2 German

Let us now take a closer look at the set of potentially alternating verbs in German. Chapter 6 provides an in-depth analysis of 76 such types. As mentioned above, eight target verbs do not return any tokens with two full NPs. Five additional verbs, *aufdämmern* 'dawn on', *beifallen* 'occur to', *entgegenschallen* 'hear', *passieren* 'happen to' and *konvenieren* 'be convenient, suit' only yield five (or fewer) tokens in the double-NP configuration. These 13 types are therefore excluded from the remainder of this section.

Table 7.2 compares the German target verbs tending most strongly towards the nominative-before-dative order to the control set of *helfen*-verbs. In total, there are five verbs that do not deviate significantly from the Nom-Dat baseline: *zugehören* 'belong to', *zugutekommen* 'benefit', *gelten* 'be worth; be considered' *zusetzen* 'affect badly' and *fernliegen* 'not occur to; not be someone's intention'. As such, there is no statistical evidence that these five target verbs constitute a different set than the control group of *helfen*-verbs with regard to the word order test.

Table 7.2 Results for five chi-square tests comparing the alternating verbs *zugehören*, *zugutekommen*, *gelten*, *zusetzen* and *fernliegen* to the baseline of *helfen*-verbs

target	**Nom-Dat**		**Dat-Nom**		vis-à-vis *helfen*-verbs
	N	**f**	**N**	**f**	
zugehören	86	96%	4	4%	p = 1
zugutekommen	144	94%	9	6%	p = 0.42
gelten	166	92%	14	8%	p = 0.08
zusetzen	121	92%	11	8%	p = 0.07
fernliegen	8	89%	1	11%	p = 0.87

Recall from Chapter 6 that four of the five verbs presented in Table 7.2 above, i.e. *zugehören*, *zugutekommen*, *gelten* and *zusetzen*, are especially common with inanimate datives. As such, their dative arguments behave more as canonical objects than as non-canonical subjects. After all, objects also tend to be inanimate rather than animate, whereas subjects show the reverse tendency, in that they are more often animate than inanimate (Hopper & Thompson 1980; Comrie 1981; Du Bois 1987, *inter alia*). It is therefore not surprising that the datives of these verbs are more strongly drawn to the postverbal slot, given that this is an object position. Recall that the logistic regression model presented in Chapter 6 brought to light exactly this, namely that inanimate datives strongly correlate with the postverbal slot.

Another remarkable feature shared by the verbs *zugehören*, *zugutekommen*, *gelten*, *zusetzen* and *fernliegen*, is that their nominative argument is not commonly indefinite. Numbers range from five out of nine indefinite nominatives, or 56%, for *fernliegen* down to 29 out of 153 indefinite nominatives, or 19%, for *zugutekommen*. Nevertheless, as shown by the logistic regression model discussed in Chapter 6, nominative indefiniteness is one of the key factors boosting the occurrence of the dative-before-nominative order. This is hardly surprising, as indefiniteness is a key feature of objects, whereas subjects show the reverse inclination towards being definite (Keenan 1976: 319; Givón 1978: 295; Comrie 1981: 121; Du Bois 1987: 829; Mikkelsen 2002: 1, *inter alia*). Thus, judging from the default argument structure of these verbs, the dative is generally too object-like and the nominative too subject-like to trigger the dative-before-nominative pattern.

It may also be observed that three of the verbs shown in Table 7.2, *zugehören*, *zugutekommen* and *zusetzen*, feature the same particle, i.e. *zu*. One could wonder if the erstwhile semantics of this particle, which developed out of the preposition *zu* 'to, towards', exert an influence on the obtained word order frequencies, in that the nominative referent might still be interpreted as moving towards the dative referent. However, the German corpus study comprises several more *zu*-particled verbs, including *zufließen*, *zufallen* and *zuwachsen*, each of which passes the word order test. As such, the conjecture that the particle *zu* shifts linearisation towards the nominative-first order seems unlikely.

One final possibility is that the five verbs under discussion here, *zugehören*, *zugutekommen*, *gelten*, *zusetzen* and *fernliegen*, might be polysemous verbs of the *genügen*-type, i.e. verbs licensing both an agentive as well as a non-agentive sense. Perhaps linearisation with the agentive sense, which presumably correlates very strongly with the nominative-before-dative order, contaminates linearisation with the non-agentive sense? This is not the case, however, at least not across the entire quintet, since only *zusetzen* possesses both an agentive, 'get on somebody's nerves', as well as a non-agentive sense 'be hard on somebody'. Furthermore, some of the verbs tending very strongly towards the dative-before-nominative order, like *winken* and *drohen*, are also polysemous verbs of the *genügen*-type (cf. infra). This further confirms that constructional contamination across senses is not a general tendency in polysemous Dat-Nom/Nom-Dat verbs licensing both an agentive as well as a non-agentive sense. However, it cannot be excluded that linearisation in one sense influences linearisation in another on a verb-specific basis.

This issue is relevant because Verhoeven (2015) has found that the word order distributions of agentive senses do influence those of non-agentive senses in verbs licensing an accusative and a nominative. For a group of ten Acc-Nom verbs possessing both an agentive as well as a non-agentive sense, like *nerven* 'to get on someone's nerves; to be annoyed', she has confirmed that the accusative-before-nominative order is attested 95 out of 527 times, or 18.03%, for the non-agentive sense. Ten monosemous Acc-Nom verbs possessing non-agentive sense only, like *interessieren* 'interest', instantiate the accusative-

first order 101 out 418 times, or 24.2%. Thus, monosemous Acc-Nom verbs instantiate the accusative-first order more often than polysemous Acc-Nom verbs. However, Verhoeven's numbers do not distinguish between verbs, making it impossible to interpret her statistics in a more fine-grained fashion. The present study has repeatedly shown that individual verbs may deviate from verb-class effects.

In what follows, each of the five verbs presented in Table 7.2 will be scrutinised further with the aid of example sets (4–8). The a. examples feature the nominative-before-dative order, the b. examples the dative-before-nominative order. Each example is accompanied by a predicted probability score representing the likelihood of the dative-before-nominative order. Since the dative-before-nominative order serves as the baseline for the logistic regression model in the German study, the predictors identified for that analysis correlate with a dative-first structure compared to a nominative-first structure.

The first pair of examples, presented in (4a–b), contains *zugehören*, the verb most strongly drawn to the nominative-before-dative order:

(4) a. *Ein Produkt* *kann* *immer* *nur einer Gruppe* *zugehören.*
a.NOM product can always but one.DAT group belong.to
'A product can always belong to but one category.'
predicted probability of the dative-before-nominative order = 10%

(4) b. *Der Oberschicht* *sollten* *mehr Leute* *zugehören.*
the.DAT upper.class should more people.NOM belong.to
'The upper class should comprise more people.'
predicted probability of the dative-before-nominative order = 9%

It has already been mentioned that the dative of each of the five verbs presented in Table 7.2 is very often inanimate. As for *zugehören*, it may additionally be observed that, in the rare case that the dative is animate, it nearly always refers to a collective. Collectives are by definition less individuated than individuals and consequently rank lower on the animacy hierarchy (Grimm 2012). The fact that the dative does not designate a prototypically animate, individual entity, but a less prototypical, collective, entity no doubt contributes to its being positioned in the middlefield, rather than in the prefield.

Consider now the following examples (5a–b) with the verb *zugutekommen*:

(5) a. *Das Geld* *soll* *der Schule* *zugutekommen.*
the.NOM money should the.DAT school benefit
'The money should benefit the school.'
predicted probability of the dative-before-nominative order = 8%

(5) b. *Dem Verkauf von Wörterbüchern* *war und ist* *diese Tatsache*
the.DAT sale of dictionaries was and is this.NOM fact
wahrscheinlich *zugutegekommen.*
probably benefited
'The sale of dictionaries has probably benefited from this fact.'
predicted probability of the dative-before-nominative order = 3%

The verb *zugutekommen* is morphologically complex, consisting of the simplex verb *kommen* 'come' as well as the incorporated prepositional phrase *zu Gute* 'to good (lit.)'. The verb *kommen* designates movement and *zugutekommen*'s preference for the nominative-first order suggests that its semantics are not entirely bleached yet. After all, the nominative argument may still be conceived of as moving towards the dative in order to benefit it. In other words, *zugutekommen* may be perceived as too agentive because it is still too compositional, thus instantiating the nominative-before-dative order rather than alternating between both word order patterns.

The next set of examples illustrates the two word orders for the verb *gelten*:

(6) a. *Mein Dank* *gilt* *aber* *natürlich*
my.NOM gratitude be.valid.for however of.course
allen Spielern.
all.DAT players.DAT
'However, my gratitude of course extends to all players.'
predicted probability of the dative-before-nominative order = 13%

(6) b. *Den Eltern und dem Rest der Familie* *gilt*
the.DAT parents.DAT and the.DAT rest the.GEN family be.worthy.of
mein tiefstes Mitgefühl.
my.NOM deepest.NOM sympathy
'The parents and the rest of the family enjoy my deepest sympathy.'
predicted probability of the dative-before-nominative order = 8%

The pattern in (6a) is highly fixed: out of 181 tokens containing two full NPs, there are as many as 63 containing a nominative constituent with *Dank* 'gratitude' or *Dankeschön* 'thank you' as a head noun. These nominatives take clause-initial position 61 times, or 97%, as opposed to two, or 3%, taking clause-final position. As such, this 'thank you' string with the verb *gelten* is highly formulaic.

However, the remaining 118 *gelten* tokens with two full NPs also tend very heavily towards nominative-first order, with 106 nominative-before-dative tokens, or 90%, and 12 dative-before-nominative tokens, or 10%. It is my impression that the bulk of these nominative-first structures do not instantiate neutral word order, but focus. This is evident from the fact that many contain adjectives like *besonder* 'special', *zentral* 'central', *speziell* 'special' *erst* 'first' or *weiter* 'additional', or a head noun modified by *Haupt-* 'main' or *Extra-* 'additional'. These performative strategies signal the importance of the upcoming dative

constituent. Two relevant examples, one featuring *besonder* and one featuring *zentral*, are given below:

(7) a. *Besonderes Augenmerk* *gilt* *der Beleuchtung.*
special.NOM attention be.valid.for the.DAT lighting
'Special attention is to be paid to the lighting.'

(7) b. *Das spezielle Interesse* *gilt* *den Mopeds der 50er Jahre.*
the.NOM special.NOM interest be.valid.for the.DAT scooters of the 1950s
'The special interest is in scooters from the 1950s.'

Example set (8) features two tokens containing the verb *zusetzen*. Again, the first one exhibits the nominative-before-dative order and the second one the dative-before-nominative order.

(8) a. *Das milde Wetter* *hat* *der Schneedecke* *arg*
the.NOM mild.NOM weather has the.DAT snow.blanket badly
zugesetzt.
affect.badly
'The mild weather took its toll on the snow cover.'
predicted probability of the dative-before-nominative order = 6%

(8) b. *Einer solchen Uhr* *kann* *bereits* *ein Regenschauer*
a.DAT such.DAT watch can already a.NOM rain.shower
arg *zusetzen.*
badly affect.badly
'Such a watch can already take a serious blow from a rain shower.'
predicted probability of the dative-before-nominative order = 16%

It may be observed that the dative participant of *zusetzen*, precisely as in (8a–b) above, is high in affectedness: its inanimate referent is damaged by the inanimate referent coded by the nominative argument. Being high in affectedness is one of the key features of a prototypical object (Hopper & Thompson 1980). This constitutes a verb-specific reason why *zusetzen* so strongly tends towards the nominative-before-dative order, apart from the more general tendencies discussed in relation to all five Nom-Dat outliers above.

Observe, finally, two example sentences involving the verb *fernliegen*:

(9) a. *Eine Majestätsbeleidigung* *habe*
a.NOM defamation.against.the.crown have
dem Artikelschreiber *ferngelegen.*
the. DAT author.of.the.article be.not.his.intention
'Defamation against the crown could not have been further away from the intentions of the author of the article.'
predicted probability of the dative-before-nominative order = 30%

(9) b. *Dem Kläger* *habe* *aber* *eine politische Stellungnahme*
the.DAT plaintiff have however a.NOM political.NOM statement
völlig *ferngelegen.*
entirely not.occurred.to
'However, the plaintiff was not at all interested in making a political statement.'
predicted probability of the dative-before-nominative order = 31%

The verb *fernliegen* only yields ten tokens in the double-NP configuration, but it is clearly skewed towards the nominative-first order: nine tokens exhibit the nominative-before-dative pattern, and only one the dative-before-nominative pattern. Precisely as with *zugutekommen*, this skewness is probably the result of *fernliegen*'s compositional semantics. *Fernliegen* literally translates to 'lie far away'. Clearly, the nominative referent is still interpreted as lying far away from the dative referent, making the former a more natural subject than the latter.

Observe that Table 7.3 constitutes the mirror image of Table 7.2: it compares those targets for which frequencies tend most strongly to the dative-before-nominative order to the same baseline of *helfen*-verbs. Crucially, I compare the prevalence of the dative-before-nominative order of the outlying targets to the prevalence of the nominative-before-dative order of the control set of *helfen*-verbs. Three verbs yield an insignificant result: *schwanen* 'have a sense of foreboding', *drohen* 'threaten' and *winken* 'can expect, be in store for'.

Table 7.3 Results for three chi-square tests comparing the alternating verbs *schwanen*, *winken* and *drohen* to the baseline of *helfen*-verbs

target	**Nom-Dat**		**Dat-Nom**		vis-à-vis *helfen*-verbs
	N	**f**	**N**	**f**	
schwanen	2	4%	54	96%	p = 1
winken	11	6%	165	94%	p = 0.31
drohen	11	7%	151	93%	p = 0.21

The results in Table 7.3 should be interpreted in the same way as those presented in Table 7.2 above: as showing that there is no statistical indication that the targets *schwanen*, *winken* and *drohen* pass the word order test. That is, they instantiate the nominative-before-dative order roughly to the same extent as *helfen*-verbs instantiate the dative-before-nominative order, thus suggesting that their nominative-first structure represents a topicalisation construction rather than an independent argument structure construction. In other words, the targets *schwanen*, *winken* and *drohen* behave as non-alternating Dat-Nom verbs, and not as alternating Dat-Nom/Nom-Dat verbs. This finding is remarkable, as non-alternating Dat-Nom verbs are assumed to be non-existent in German (Eythórsson & Barðdal 2005:

868; Barðdal, Eythórsson & Dewey 2014, 2019: 131–148; Rott 2016: 239–249; Barðdal 2023: Ch. 6).

Inspection of the predicted probabilities of the dative-before-nominative order for *schwanen*, *winken* and *drohen* reveals that these are extremely high, with a mean of 93% for *schwanen* and *winken* and 92% for *drohen*. Across the entire dataset, the mean predicted probabilities for the dative-before-nominative order only amount to 46%. All these percentages concern configurations involving two full NPs, excluding tokens for which animacy is non-inferable. Below are three sets of examples, each featuring one of the verbs in question. The a. examples illustrate the dative-before-nominative order, while the b. examples represent the nominative-before-dative order:

(10) a. *Dem Mann* *schwant* *Böses.*
the.DAT man has.a.sense.of.foreboding evil.NOM
'The man is having a sense of foreboding.'
predicted probability of the dative-before-nominative order = 94%

(10) b. *Die Erkentniss* *schwant* *inzwischen* *auch der SPD.*
this.NOM realisation come.to.realise meanwhile also the.DAT SPD
'The SPD has now also come to realise this.'
predicted probability of the dative-before-nominative order = 86%

(11) a. *Dem Gewinner* *winkt* *eine Reise nach Australien.*
the.DAT winner can.expect a.NOM trip to Australia
'The winner can expect a trip to Australia.'
predicted probability of the dative-before-nominative order = 95%

(11) b. *Ein besonderer Preis* *winkt* *allen Gewinnern*
a.NOM special.NOM prize awaits all.DAT winners.DAT
der jeweiligen Disziplinen.
the.GEN respective disciplines
'A special prize awaits all winners of the respective disciplines.'
predicted probability of the dative-before-nominative order = 95%

(12) a. *Dem Betreiber* *drohen* *lange Haftstrafen.*
the.DAT operator faces long.NOM prison.sentences
'The operator is facing long prison sentences.'
predicted probability of the dative-before-nominative order = 94%

(12) b. *Das wohl größte Unheil* *droht* *dem Auftraggeber.*
the.NOM probably greatest.NOM disaster threatens the.DAT client
'Probably the greatest disaster threatens the client.'
predicted probability of the dative-before-nominative order = 91%

The predicted probabilities for these three verbs are so high because they match best the factors singled out by the logistic regression model discussed in Chapter 6: dative

(pro)nominality, nominative (pro)nominality and length. In addition, it is remarkable that the dative is very often both definite and animate, whereas the nominative is oftentimes indefinite and inanimate. This is shown in Table 7.4 below:

Table 7.4 Animacy and definiteness in the verbs *schwanen*, *winken* and *drohen* (double-NP configuration)

	Nominative				**Dative**			
	indefinite		*inanimate*		*definite*		*animate*	
	N	f	N	f	N	f	N	f
schwanen	55	98%	56	100%	51	91%	56	100%
winken	159	90%	176	100%	163	93%	169	96%
drohen	101	62%	162	100%	116	72%	151	93%

The discrepancy between the dative and the nominative is clearly highest with *schwanen*: in the double-NP configuration, its dative is consistently animate (56 tokens, or 100%) and nearly always definite (51 tokens, or 91%), while its nominative is consistently inanimate (56 tokens, or 100%) and nearly always indefinite (55 tokens, 98%).

Another remarkable tendency is that *schwanen* collocates with a fixed set of nominatives: the top-three nominative arguments in the double-NP configuration are *Böses* 'something evil' (24 out of 56 tokens, or 43%), *(nichts) Gutes* 'nothing good' (19 out of 54 tokens, or 34%) and *Übles* 'something bad' (six out of 54 tokens, or 11%). Together, these three lexical units account for 49 out of 54 of all nominatives, or 91%. The dative, by contrast, is considerably less constrained, with the most common dative argument, *Fan* 'fan', occurring seven times, or 13%, and the second most common dative, *Zuschauer* 'onlooker', only three times, or 6%. It is well-known that verbs and objects share a more intimate connection than verbs and subjects (Valian, Prasada & Scarpa 2006). The fact that *schwanen* selects for a restricted number of nominatives as opposed to a less restricted number of datives suggests that it is the former that behave as subjects here, not the latter.

The verb *winken* displays a highly similar distribution for animacy and definiteness. Excluding two cases for which animacy is non-inferable, the dative, again for clauses involving two full NPs, is animate in 96% of cases (169 out of 176 tokens) and definite in 93% of cases (163 out of 176 tokens). The nominative mirrors these frequencies, with 176 inanimate NPs, or 100%, and 159 indefinite NPs, or 90%. However, in contrast to *schwanen*, *winken* also yields a fair share of tokens comprising two definite NPs. Of these, 14 tokens, or 88%, allocate the preverbal slot to the dative and two, or 12%, to the nominative. Even if both NPs are definite, this verb therefore exhibits a clear tendency towards the dative-first order.

The last verb skewed towards the dative-before-nominative order is *drohen*. Numbers are markedly lower with *drohen* than with *schwanen* and *winken*, but still relatively high. As

for animacy, 93% of datives with *drohen* are animate (151 tokens) and 100% of nominatives are inanimate (162 tokens). As for definiteness, 72% of datives with *drohen* are definite (116 tokens) as opposed to 62% inanimate nominatives (101 tokens).

The reason why all of this matters is because the dative arguments of *schwanen*, *winken* and, to a lesser extent, *drohen* are natural themes, while their nominatives are natural rhemes. After all, thematic constituents tend to be both definite and animate, whereas rhematic constituents show the reverse inclination, i.e. towards indefiniteness and inanimacy. Since the semantic-pragmatic discrepancy between the dative and the nominative is so extreme, in that the dative is particularly theme-like and the nominative particularly rheme-like, linearisation for these verbs clearly tends towards the dative-before-nominative order, suggesting that *schwanen*, *winken* and *drohen* possess one Dat-Nom argument structure construction only.

7.4 Theoretical implications

In this section, I discuss the theoretical relevance of the present research. I first present six general implications (Section 7.4.1), after which I show how this study is relevant for the representation of polysemous verbs in the constructional network (Section 7.4.2). Finally, I discuss and reject the allostructional account of the relevant Dat-Nom/Nom-Dat alternation in general (Section 7.4.3).

7.4.1 General implications

The corpus studies presented in this thesis have confirmed that Modern Icelandic and Modern German possess a class of verbs for which both the dative and the nominative argument pass the word order test, as the research motivating this thesis has claimed (Eythórsson & Barðdal 2005; Barðdal, Eythórsson & Dewey 2014, 2019; Barðdal 2023: Chs. 3 and 6). These so-called Dat-Nom/Nom-Dat verbs therefore have two arguments behaving syntactically as subjects, not just one, although not at the same time, of course. This finding has a bearing on linguistic theory in multiple respects.

First, the present thesis shows that subjecthood is not necessarily associated with a fixed verbal argument. After all, some verbs feature more than one argument structure construction and, as a consequence, their subject role may be filled by more than one constituent, although self-evidently not at the same time. More specifically, the subject role is filled by the dative in the Dat-Nom construction and the nominative in the Nom-Dat construction. Taking the subject to be the first argument of the argument structure is

considerably more parsimonious, I believe, than taking the subject to be marked for nominative case, regardless of that argument's position in the clause. As is evident from the statistics presented here, the latter approach, i.e. that the subject is marked for nominative irrespective of its position in the clause, requires an analysis involving topicalisation in ca. 30–90% of all attestations of the relevant verbs. As has been pointed out by Barðdal, Eythórsson & Dewey (2014: 55) and Barðdal, Eythórsson & Dewey (2019: 116), such a near-mandatory topicalisation operation is a contradiction in terms, as topicalisation is both optional and infrequent (see also Sections 5.1 and 6.1 of the present thesis). Still, the assumption that the subject is nominative regardless of position is exactly the approach several scholars have taken so far (Bader & Häussler 2010; Haspelmath 2001; Kempen & Harbusch 2005; Verhoeven 2015, *inter alia*).

Second, this study has shown that one can make meaningful syntactic generalisations independently of morphological case marking. By associating the subject with the nominative, one denies that such a generalisation is possible, let alone desirable, since one assumes a complete overlap between morphology and syntax: the subject is taken to be case-marked in the nominative, and the argument that is case-marked in the nominative is taken to be the subject. Such a view is problematic because it is circular and, hence, unfalsifiable. It also assumes morphology to be a derivative of syntax. However, morphology may also be motivated semantically, as is the case with every structure displaying non-canonical case marking, regardless of the syntactic properties of the non-canonically case-marked argument. Finally, the assumption that the subject in Icelandic or German is only ever marked for nominative case is psycholinguistically improbable; if speakers allocate the preverbal slot with reasonable frequency to a given argument, such as the dative, that tendency must mean something.

Third, the German corpus study has shown that the lion's share of the target verbs in that language are indeed alternating Dat-Nom/Nom-Dat verbs and not non-alternating Dat-Nom verbs. These findings empirically confirm a claim first put forward by Eythórsson & Barðdal (2005: 868), namely that all German Dat-Nom verbs appear to be able to vacillate between two case frames (see also Barðdal, Eythórsson & Dewey 2014, 2019: 131–148; Rott 2016: 239–249; Barðdal 2023: Ch. 6). At the same time, the corpus study repudiates a long-lived hypothesis in the literature on German that verbs exhibiting a thematic asymmetry between the dative and the nominative are Dat-Nom verbs, i.e. verbs instantiating a dative-before-nominative base order (Fanselow 2002: 231, 2003: 38; Bader & Häussler 2010: 751; Haider 2010: 254, 260; Eisenberg 2013: 74; Bader 2020: 1088, *inter alia.*).

Fourth, the current study reveals a genealogical generalisation across German and Icelandic with respect to the status of alternating verbs. At least for clauses involving two full NPs, the statistical model presented in Chapter 5 confirms that the factors correlating with the alternation are generally identical, given that the language factor only surfaces at the bottom of the inference tree. This suggests that the phenomenon at hand is in principle

the same in both these languages, in turn motivating a unified account of alternating Dat-Nom/Nom-Dat verbs. Additionally, the synchronic similarities between German and Icelandic indicate that their alternating constructions have been inherited from a shared proto-stage of the language, as already argued by Barðdal (1998), Eythórsson & Barðdal (2005), Barðdal, Eythórsson & Dewey (2014, 2016) and Barðdal (2023: Ch. 3, Ch. 6).

Fifth, the results of this study lend support to one of the main tenets of Radical Construction Grammar (Croft 2001; Barðdal 2006; Barðdal & Gildea 2015: 30–31), namely that grammatical relations are construction-specific roles. The alternative view, which claims that grammatical relations are generalised relations existing independently of constructions, is not confirmed. The conclusion that subject is a construction-specific category is supported by the fact that arguments marked for either nominative or dative case may fulfil different grammatical roles depending on the construction they fill, with the Nom-Dat construction licensing a nominative subject and a dative object, and the Dat-Nom construction licensing a dative subject and a nominative object. German and Icelandic verbs of the *help*-type are only compatible with the former argument structure construction, whereas Icelandic verbs of the *like*-type are only compatible with the latter. Alternating verbs may instantiate either of these argument structure constructions.

Finally, this thesis has confirmed that non-canonical case marking is not restricted to *piacere*-verbs, psych-verbs, experiential verbs or unaccusative verbs. Rather, the Dat-Nom/Nom-Dat alternation may potentially be associated with any verb displaying a thematic asymmetry, regardless of any further syntactic or semantic properties.

7.4.2 Implications for the constructional representation of polysemous verbs

I now turn to another theoretical contribution of the German corpus study, which has shown that the Dat-Nom and the Nom-Dat argument structure constructions are not only associated with *verbs* displaying a thematic asymmetry between the dative and the nominative, but even with verb *senses* displaying such an asymmetry. Recall that the German corpus study is based on three types of verbs: (1) monosemous verbs, (2) polysemous verbs of the *passen*-type and (3) polysemous verbs of the *genügen*-type. Polysemous verbs of the *passen*-type possess multiple senses displaying a thematic asymmetry, whereas polysemous verbs of the *genügen*-type possess one or more agentive senses alongside one or more senses displaying a thematic asymmetry. Only senses displaying a thematic asymmetry have been included for study, whose alternating status has been confirmed (see, however, Section 7.3 above for a discussion of the outliers). Furthermore, since a control set of five *helfen*-verbs has confirmed that *helfen*-verbs are indeed unequivocal Nom-Dat verbs, one can reasonably assume for *genügen*-verbs to also associate with the Nom-Dat argument structure construction in their agentive sense(s).

The relevance of polysemy in the study of argument structure constructions has been highlighted before. Another notable study in this respect is Bernolet, Colleman & Hartsuiker (2014), who report on the results of a priming study. Their main finding is that priming in ditransitive verbs is boosted when primes and targets exhibit the same sense as opposed to a different sense. This so-called "sense boost" compelled Bernolet, Coleman & Hartsuiker to conclude that each of the senses of ditransitive verbs is stored separately.

The very same conclusion, i.e. that speakers have sense-specific representations of verbs, which are subsequently associated with either the *to*-dative or the double-object construction is supported by evidence presented in Bernolet & Colleman's (2016) study. These scholars run a distinctive collexeme analysis on corpus data retrieved for 15 ditransitive verbs licensing both a concrete and a metaphorical sense. The results show that different senses of the same verb may be differently inclined towards either the prepositional dative or the *to*-dative. Thus, Bernolet & Coleman (2016) argue, subcategorisation preferences, at least for some verbs, are associated with individual senses, and not with the verbal lexeme. Bernolet & Colleman contend that their results are perfectly compatible with usage-based approaches to language, which take the mental representation of constructions to be a hierarchical network existing at different levels of schematicity (2016: 194).

The results of the current study are evidently fully in line with both Bernolet, Coleman & Hartsuiker (2014) as well as Bernolet & Colleman (2016). After all, they confirm that argument structure constructions are indeed principally associated with verb senses, and not with verbal lexemes. Hence, I suggest an expansion of Barðdal's (2023: 121–124) formal representation of alternating verbs (see Chapter 1) to the one presented in Figure 7.1 below:

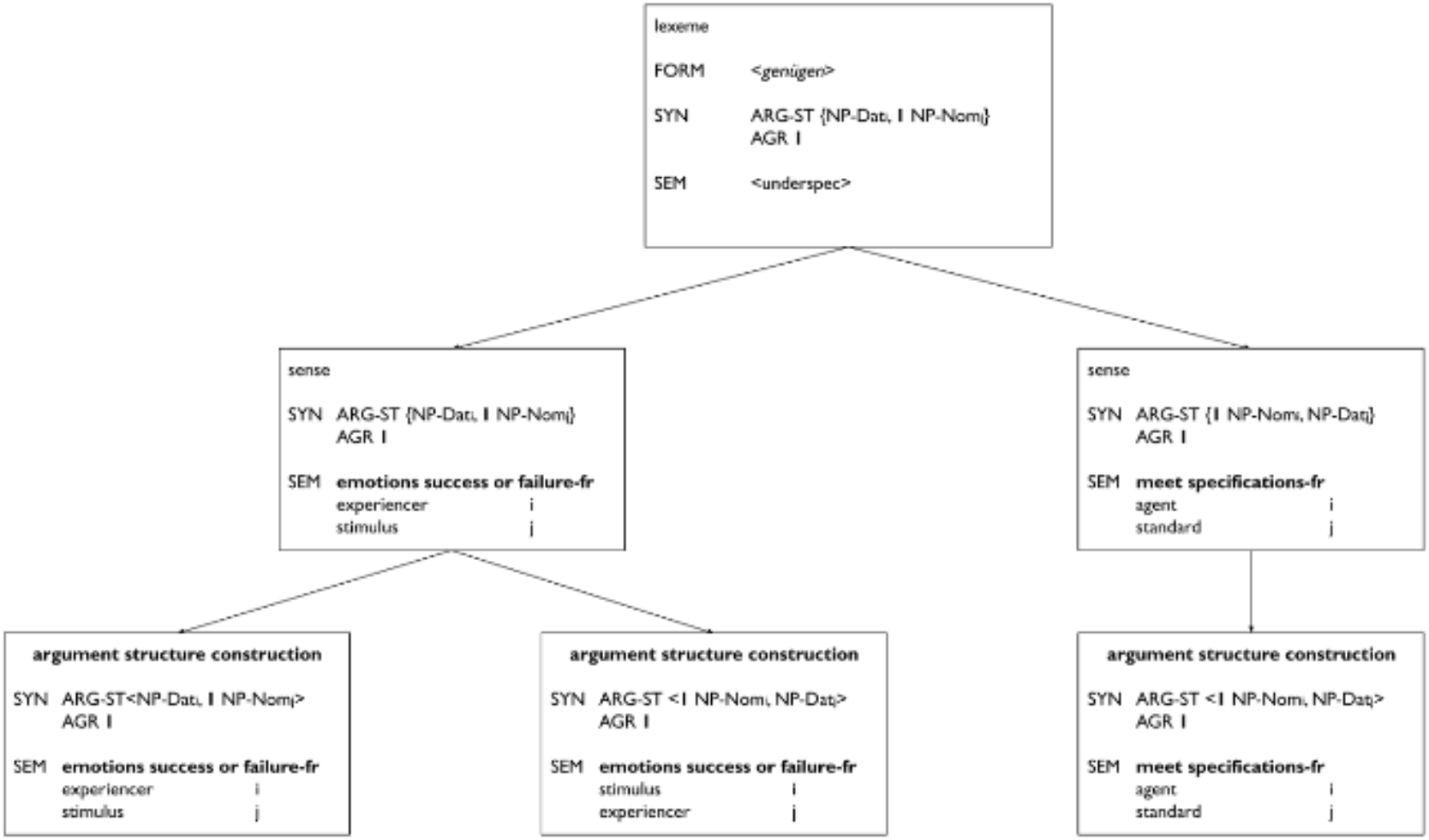


Figure 7.1 The compatibility of polysemous *genügen* with each of its senses, which, in turn, associate with their respective argument structure construction(s)

The main innovation in Figure 7.1 compared to Barðdal's model is that it posits a sense level in between the lexeme level and the level of the argument structure constructions. These senses must be stored separately by language users, as it would otherwise be impossible for them to alternate between the Dat-Nom and the Nom-Dat argument structure constructions in the thematically asymmetric sense, while adhering to a single, Nom-Dat, construction in the more agentive sense.

The model above evidently does not apply to *zusetzen*, one of the outliers discussed in Section 7.3.2 above. Precisely like *genügen*, *zusetzen* is also a polysemous verb possessing both an agentive as well as a non-agentive sense. However, word order counts for the non-agentive sense of this verb do not deviate significantly from the *helfen*-baseline. This means that language users potentially do not perceive of *zusetzen* as a polysemous verb, or that the verb is solely agentive in their mind.

Ever since Goldberg (1995: 9–10), it is well known that argument structure constructions are associated with meaning. The most famous example confirming this premise is probably the one cited below:

(13) *He sneezed the napkin off the table.* (Goldberg 1995: 9)

Examples like the one under (13) confirm that argument structure constructions indeed have meaning in themselves, as one would otherwise have to posit an additional caused-motion sense for the verb *sneeze*, which is highly implausible. By relegating the 'caused-motion' meaning to the argument structure construction, one avoids a so-called "proliferation of the senses".

However, not every instance of polysemy may be attributed to the effect of argument structure constructions. As the current study has shown, verbs like German *genügen* 'be enough, be sufficient; fulfil' are polysemous in and of themselves. Here, it is not the argument structure construction determining the sense of the verb, it is the sense that selects the appropriate argument structure construction: Nom-Dat in agentive cases and Dat-Nom/Nom-Dat in non-agentive cases. Hence, polysemy should not only be reckoned with in a top-down fashion, i.e. macroconstructions *enforcing* a particular sense onto a given verb, but also in a bottom-up fashion, i.e. verbs *selecting* a given argument structure construction based on the sense the verb instantiates (cf. also Johnson et al. 2019).

As argued by Bernolet & Colleman (2016: 194), polysemy of individual words should not be expected to play out in every single case. Precisely like the results for German *zusetzen* have shown, a verb may still display a thematic asymmetry in one of its senses while being more agentive in the other sense without selecting for different argument structure constructions in either case.

7.4.3 Against an allostructional account of the Dat-Nom/Nom-Dat alternation

One influential account dealing with syntactic alternations has come to be known as the *allostructional account* (Cappele 2006; Perek 2015; De Vaere 2020; *inter alia*). The term 'allostruction' was first suggested by Cappelle (2016), who introduced it to account for word order variation found with particle verbs where the particle may either precede or follow the direct object. The allostructional account has also been applied to the well-known *dative alternation*, i.e. the alternation between the ditransitive construction and its prepositional variant.

The main assumption has been that allostructions are synonymous yet structurally distinct manifestations of a construction that is left partially underspecified. In the case of the verb-particle construction, the assumption is that it is underspecified for word order and that language users may either home in on the continuous order (*I picked up the book*), where the particle is immediately adjacent to the main verb, or the discontinuous order (*I picked the book up*), where the direct object intervenes between the verb and the stranded particle.

The question arises whether such an analysis is also valid for the Dat-Nom/Nom-Dat alternation discussed here, as either alternant, Dat-Nom and Nom-Dat, could be regarded as an allostruction of an underlying higher-level Dat-Nom/Nom-Dat construction, which would be underspecified for word order, precisely like Cappelle's (2016) verb-particle construction. However, the present study has empirically confirmed that hypothesised alternating verbs, both in Icelandic and German, possess two veritable argument structure constructions, in that either argument, dative and nominative alike, independently passes the subject requirement of occurring in clause-initial position in contexts which are information-structurally neutral. Furthermore, this study has shown that the choice between the Dat-Nom and Nom-Dat argument structure constructions is guided by topicality, thus confirming that there is an actual difference in function between both alternants instead of a mere difference in word order.

The fact that alternating verbs deviate significantly from the *help*-baseline further confirms that we are indeed dealing with two distinct verb classes, Dat-Nom/Nom-Dat on the one hand, and Nom-Dat on the other. Additionally, both arguments of alternating verbs have been shown to satisfy any of the remaining subject criteria, albeit not at the same time, of course (see Section 1.3.1 and Barðdal 2023: Ch. 3 for Icelandic and Ch. 6 for German).

7.5 Outlook

The current study opens up a number of pathways for future research. One of the main questions remaining unanswered at present is to what extent the corpus frequencies presented here reflect a psycholinguistic reality. This question is especially relevant for the target verbs that do not pass the word order test, i.e. German *zugehören*, *zugutekommen*, *gelten*, *zusetzen*, *fernliegen*, *schwanen*, *winken* and *drohen* as well as Icelandic *henta*. An acceptability judgement study or a series of priming experiments could aid in bridging this gap.

Yet another way in which the corpus frequencies presented here could be falsified is by systematically investigating any of the remaining subjecthood diagnostics. Some of these, like control infinitives or raising-to-object structures, are notoriously hard to find, but others, like inversion or raising-to-subject, are more common. Inversion is the phenomenon whereby the subject and the finite verb switch places, either because the finite verb takes first position, as in yes-no interrogatives, or because another constituent is topicalised.

The validity of the inversion test is sometimes called into question with regard to German, as word order in that language is considered to be free rather than fixed (Reis 1982; Eisenberg 2013: 376; Hentschel & Weydt 2021: 419, *inter alia*). However, this claim is mainly based on contexts involving pronouns, which tend to defy linearisation principles across languages in general, and in German in particular (Hoberg 1981: 62; Siewierska 1993; Primus 1994; Shannon 2000; Kempen & Harbusch 2004; Heylen 2005; Bader & Häussler 2010; Bader 2020, *inter alia*).[1] As Barðdal (2006, 2023: Ch. 6) argues, there is no a priori reason for disqualifying inversion as a subject test in Modern German. Indeed, as long as one's claims about what constitutes neutral word order are based on contexts involving full NPs, any tendencies found are bound to shed light on the subject status of the inverted argument. Alternatively, one would have to give up on the analysis that there is subject-verb inversion in German in yes-no interrogatives and in contexts where a constituent other than the verb occurs in first position. This, however, would be a highly controversial claim. For alternating verbs, the expectation is that both arguments, whether they be marked in the dative or the nominative, license inversion with the finite verb.

Third, a follow-up study might also look further into sense-specific word order tendencies, either by zooming in on verbs of the *genügen*-type or by focusing on verbs of the *begegnen*-type. Recall that verbs like *genügen* 'be enough, be sufficient; fulfil' possess both an (alternating) Dat-Nom/Nom-Dat sense as well as a (non-alternating) Nom-Dat sense and that this thesis has only included potentially alternating senses, assuming that

[1] The reader should be well aware that the current study may be added tot his list.

more agentive senses correlate with a single Nom-Dat argument structure. This is also what the word order distributions for the control set of *helfen*-verbs has confirmed.

Collecting data for verbs licensing both agentive and non-agentive senses would lend further support to the alternating hypothesis, as one would be able to use the target verbs as their own control items. As an example, *genügen* 'fulfil' would be the control for *genügen* 'be enough, be sufficient'. The fact that polysemous verbs behave differently in different senses is intriguing, but not necessarily surprising, as polysemy also affects other domains of the grammar, e.g. morphology. One relevant example may be found in the diverging conjugations of polysemous verbs like German *senden* 'send; broadcast': the former sense correlates with the strong forms *sandte* 'sent' (simple past) and *gesandt* 'sent' (past participle), the latter uses the weak forms *sendete* 'broadcast' (simple past) and *gesendet* 'broadcast' (past participle). The existence of different verb forms, strong and weak, confirms that language users may be sensitive to polysemy.

A systematic investigation of verbs of the *begegnen*-type, i.e. verbs licensing multiple alternating senses, would perhaps be even more enlightening. Such a study could reveal whether different Dat-Nom/Nom-Dat senses of the same verb yield a different distribution across the dative-before-nominative and the nominative-before-dative orders. If language users can distinguish between agentive and non-agentive senses, as with *genügen*-verbs, it is not inconceivable that they are also able to distinguish between multiple non-agentive senses.

Fourth, future research could pursue further the so-called 'topicality hypothesis', which claims that the choice between the Dat-Nom and the Nom-Dat constructions for alternating verbs is a derivative of discourse prominence (Barðdal 2001; Barðdal, Eythórsson & Dewey 2014, 2019). Chapter 5 has shown that 28 Icelandic double-NP tokens instantiating the nominative-before-dative order indeed contain a more topical nominative, despite it being indefinite. Systematically coding both datasets, German and Icelandic, for topicality could lend more robust support to this hypothesis. This particularly applies to the double-NP configuration, as the logistic regression analyses for this subset do not generate particularly high C-indexes (see Chapter 4 for Icelandic and Chapter 6 for German). The same holds for the conditional inference tree presented in Chapter 5 for both languages. Clearly, the ensemble of variables included for study could not satisfactorily explain the word order variation observed within the three data subsets, (1) Icelandic, (2) German and (3) Icelandic–German, suggesting that some predictors are still missing. Therefore, expanding the number of predictor variables might yield a more fine-grained model.

Finally, the current study prompts the question how non-canonically case-marked subject verbs licensing a nominative and an accusative behave under the word order test. According to Barðdal (2023: 109), Icelandic Acc-Nom verbs alternate between two case frames, testifying to their status as Acc-Nom/Nom-Acc verbs. However, there are only two relevant types: *henda* 'happen' and *sækja syfja* 'get drowsy'. In German, this class is

more sizable, containing verbs like *interessieren* 'interest', *freuen* 'give pleasure' or *wundern* 'astonish'. A systematic study of these verbs might further improve our understanding of non-canonical subjecthood, shedding additional light on the relation between case and argument structure.

References

Bader, Markus. 2020. The position of object pronouns in the German middlefield. *Linguistics* 58(4): 1059–1115.

Bader, Markus & Häussler, Jana. 2010. Word order in German: A corpus study. *Lingua* 120. 717–762.

Barðdal, Jóhanna. 1998. Argument structure, syntactic structure and morphological case of the impersonal construction in the history of Scandinavian. *Scripta Islandica* 49: 21–33.

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

Barðdal, Jóhanna. 2004. The semantics of the impersonal construction in Icelandic, German and Faroese: Beyond thematic roles. In Werner Abraham (ed.), *Focus on Germanic Typology*, 105–137. Berlin: Akademie Verlag.

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, Carlee Arnett, Stephen Mark Carey, Thórhallur Eythórsson, Gard B. Jenset, Guus Kroonen & Adam Oberlin. 2016. Dative subjects in Germanic: A computational analysis of lexical semantic verb classes across time and space. *Language Typology and Universals* 69(1). 49–84.

Barðdal, Jóhanna & Thórhallur Eythórsson. 2006. Control infinitives and case in Germanic: 'Performance error' or marginally acceptable constructions. In Leonid Kulikov, Andrej Malchukov & Peter de Swart (eds.), *Case, Valency and Transitivity*, 147–177. Amsterdam: John Benjamins.

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. 2020. How to identify cognates in syntax: Taking Watkins' legacy one step further. In Jóhanna Barðdal, Spike Gildea & Eugenio R. Lujan (eds.), *Reconstructing Syntax*, 197–238. Leiden: Brill.

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 & Spike Gildea. 2015. Diachronic Construction Grammar: Epistemological context, basic assumptions and historical implications. In Jóhanna Barðdal, Elena Smirnova, Lotte Sommerer & Spike Gildea (eds.), *Diachronic Construction Grammar*, 1–50. Amsterdam: John Benjamins.

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.

Bernolet, Sarah & Timothy Colleman. 2016. Sense-based and lexeme-based alternation biases in the Dutch dative alternation. In Joyoung Yoon & Stefan Gries (eds.), *Corpus-Based Approaches to Construction Grammar*, 165–198. Amsterdam: John Benjamins.

Bernolet, Sarah, Timothy Colleman & Robert J. Hartsuiker. 2014. The 'sense boost' to dative priming: Evidence for sense-specific verb-structure links. *Journal of Memory and Language* 76: 113–126.

Cappelle, Bert. 2006. Particle placement and the case for 'allostructions'. *Constructions* 1: 1–28.

Comrie, Bernard. 1976. *Aspect: An Introduction to the Study of Verbal Aspect and Related Problems*. Cambridge: Cambridge University Press.

Comrie, Bernard. 1981. *Language Universals and Linguistic Typology: Syntax and Morphology*. Chicago: University of Chicago Press.

Croft, William. 2001. *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford: Oxford University Press.

De Vaere, Hilde, Julia Kolkmann & Thomas Belligh. 2020. Allostructions revisited. *Journal of Pragmatics* 170: 96–111.

Divjak, Dagmar. 2017. The role of lexical frequency in the acceptability of syntactic variants. Evidence from *that*-clauses in Polish. *Cognitive Science* 41(2): 354–382.

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

Eisenberg, Peter. 2013. *Grundriss der deutschen Grammatik. 2: Der Satz*. Stuttgart: Metzler.

Eythórsson, Thórhallur & Jóhanna Barðdal. 2005. Oblique subjects: A common Germanic inheritance. *Language* 81(4): 824–881.

Fanselow, Gisbert. 2002. Quirky subjects and other specifiers. In Ingrid Kaufmann and Barbara Stiebels (eds.), *More than Words: A Festschrift for Dieter Wunderlich*, 227–250. Berlin: Akademie Verlag.

Fanselow, Gisbert. 2003. Zur Generierung der Abfolge der Satzglieder im Deutschen. *Neue Beiträge zur Germanistik* 112: 9–47.

Flach, Susanne. 2020. Schemas and the frequency/acceptability mismatch: Corpus distribution predicts sentence judgments. *Cognitive Linguistics* 31(4): 609–645.

Givón, Talmy. 1978. Definiteness and referentiality. In Joseph H. Greenberg (ed.), *Universals of Human Language*, 149–188. Stanford: Stanford University Press.

Grimm, Scott. 2012. *Number and Individuation*. Stanford: Stanford University Ph.D. dissertation.

Haider, Hubert. 2010. *The Syntax of German*. Cambridge: Cambridge University Press.

Haspelmath, Martin. 2001. Non-canonical marking of core arguments in European languages. In Alexandra Aikhenvald, Robert M. W. Dixon & Masayuki Onishi (eds.), *Non-canonical subjects and objects*, 53–83. Amsterdam: John Benjamins.

Hentschel, Elke & Harald Weydt. 2021. *Handbuch der deutschen Grammatik*. Berlin/Boston: De Gruyter.

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

Hoberg, Ursula. 1981. *Die Wortstellung in der geschriebenen deutschen Gegenwartssprache*. München: Max Hueber Verlag.

Hopper, Paul J. & Sandra A. Thompson. 1980. Transitivity in grammar and discourse. *Language* 56(2): 251–299.

Johnson, Cynthia A., Peter Alexander Kerkhof, Leonid Kulikov, Esther Le Mair and Jóhanna Barðdal. 2019. How to succeed in Indo-European without really trying? Argument structure, conceptual metaphor and semantic change. *Diachronica* 36(4): 463–508.

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

Kempen, Gerard & Harbusch, Karin. 2004. A corpus study into word order variation in German subordinate clauses: Animacy effects linearization independently of grammatical function

assignment. In Pechmann, Thomas & Habel, Christopher (eds.), *Multidisciplinary approaches to language production*, 173–181. Berlin: Mouton de Gruyter.
Kein, Devorah E. & Gregory L. Murphy. 2001. The representation of polysemous words. *Journal of Memory and Language* 45: 259–282.
Mikkelsen, Line. 2002. Reanalyzing the definiteness effect. Evidence from Danish. *Working Papers in Scandinavian Syntax* 69: 1–75.
Perek, Florent. 2015. Argument Structure in Usage-Based Construction Grammar: Experimental and Corpus-Based Perspectives. Amsterdam: John Benjamins.
Primus, Beatrice. 1994. Grammatik und Performanz: Faktoren der Wortstellungsvariation im Mittelfeld. *Sprache und Pragmatik* 32: 39–86.
Primus, Beatrice. 1999. *Cases and Thematic Roles: Ergative, Accusative and Active*. Tübingen: Max Niemayer.
Reis, Marga. 1982. Zum Subjektbegriff im Deutschen. In Werner Abraham (ed.), Satzglieder im Deutschen: Vorschläge zur syntaktischen, semantischen und pragmatischen Fundierung, 171–211. Tübingen: Gunter Narr.
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.
Shannon, Thomas F. 2000. On the order of (pro)nominal arguments in Dutch and German. In Shannon, Thomas F. & Snapper, Johan P. (eds.), *The Berkeley conference on Dutch linguistics 1997: The Dutch language at the millennium*, 145–196. New York: University Press of America.
Siewierska, Anna. 1993. Syntactic weight vs. information structure and word order variation in Polish. *Journal of Linguistics* 29: 233–265.
Somers, Joren. 2023. Oblieke subjecten in het Nederlands? De casus wachten. *Nederlandse Taalkunde* 28(3): 350–364.
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.
Valian, Virginia, Sandeep Prasada & Jodi Scarpa. 2006. Direct object predictability: Effects on young children's imitation of sentences. *Journal of Child Language* 33(2): 247–269.
Verhoeven, Elisabeth. 2015. Thematic asymmetries do matter! A corpus study of German word order. *Journal of Germanic Linguistics* 27(1): 45–104.
Wegener, Heide. 1985. *Der Dativ im heutigen Deutsch*. Tübingen: Günther Narr.
Zipf, George Kingsley. 1935. *The Psycho-Biology of Language: An Introduction to Dynamic Philology*. Cambridge (MA): MIT Press.