



Socio-Syntactic Variation and Change in Nineteenth-Century Icelandic

The Emergence and Implementation of a National
Standard Language

Heimir Freyr van der Feest Viðarsson

Dissertation towards the degree of Doctor of Philosophy
2019



UNIVERSITY OF ICELAND
SCHOOL OF HUMANITIES

FACULTY OF ICELANDIC AND
COMPARATIVE CULTURAL STUDIES

Socio-Syntactic Variation and Change in Nineteenth-Century Icelandic

The Emergence and Implementation of a National
Standard Language

Heimir Freyr van der Feest Viðarsson

Dissertation towards the degree of Doctor of Philosophy

University of Iceland
School of Humanities
Faculty of Icelandic and Comparative Cultural Studies
December 2019

Íslensku- og menningardeild Háskóla Íslands
hefur metið ritgerð þessa hæfa til varnar
við doktorspróf í íslenski málfraði

Reykjavík, 5. nóvember 2019

Torfi Tulinius
deildarforseti

Faculty of Icelandic and Comparative Cultural Studies
at the University of Iceland
has declared this dissertation eligible for a defence
leading to a Ph.D. degree in Icelandic Linguistics

Doctoral Committee:
Þórhallur Eypórssón and Ásgrímur Angantýsson, supervisors
Leonie Cornips
Jóhannes Gísli Jónsson

Socio-Syntactic Variation and Change in Nineteenth-Century Icelandic
The Emergence and Implementation of a National Standard Language

© Heimir Freyr van der Feest Viðarsson
Reykjavík 2019

Dissertation for a doctoral degree at the University of Iceland. All rights reserved.
No part of this publication may be reproduced in any form
without written permission of the author.

ISBN 978-9935-9491-4-1

Printing: Háskólaprent

Abstract

This thesis is a study of the interface between sociolinguistics and syntax, focusing on 19th-century Icelandic and the implementation of a national standard language from a syntactic perspective. Icelandic is noted for relatively successful and far-reaching effects of prescriptive standard norms (see Ottósson 1987, 1990, 2003, 2005, Sigmundsson 1990-1991, 2002, Árnason 2003b, Thomason 1999, 2001, 2007, Kusters 2003, Friðriksson 2008, Hilmarsson-Dunn and Kristinsson 2010, Viðarsson 2017b). The main reason for this success has been claimed to lie in the fact that the standard language was based to some extent on the everyday language of (rural) commoners, which gradually replaced the Danicised written norms of previous periods: “In essence, what the majority of Icelanders did was accepting their own linguistic standard” (Friðriksson 2008:99, see also Sigmundsson 1990-1991). On the other hand, the traditional narrative also emphasises that some of these aspects involved the ‘revival’ of traits characteristic of Old Norse (see Ottósson 1987, 1990, 2003, 2005). The fact that these features include representatives from across the whole linguistic spectrum, including not only phonology, morphology and the lexicon (in addition to spelling), but also syntax, appear to put Icelandic standardisation squarely at odds with common assumptions about relatively limited awareness of abstract linguistic structure (e.g. Labov and Harris 1986, Cheshire 1987, Laycock and Mühlhäusler 1990, Cheshire et al. 2005). At the same time, there has been very little emphasis laid on investigating actual language use as found in linguistic corpora to substantiate frequently made claims with regard to standardisation.

According to Labov’s INTERFACE PRINCIPLE (cf. Labov and Harris 1986 and much later work), also referred to more specifically as the ANTISOCIAL SYNTAX HYPOTHESIS (Ingason et al. 2011, Ingason et al. 2013), the social embedding of variation is to be found primarily on the surface, particularly the words and the sounds of a language. Abstract linguistic structure, in contrast, is claimed *not* to be evaluated socially and the same has been suggested with regard to deliberate change, as these both rate low on Laycock and Mühlhäusler’s (1990) DEGREE OF INTERFERENCE HIERARCHY. With respect to variation at a (broad) syntactic level, my overarching conclusions based on the three case studies reported on in this thesis are on the one hand that the Interface Principle appears to be overstated, while on the other hand the extent of historical variation in Icelandic is grossly understated.

The present study attempts to address and to problematise common conceptions of the socio-syntactic interface insofar as the traditional Icelandic standardisation narrative is concerned (see e.g. Ottósson 1990, 2005) by studying the language use of speakers from a range of different social backgrounds and in a range of different text types. The main focus will be on three linguistic variables that differ in their level of abstractness. These variables concern the **position of the finite verb** vis-à-vis adverb(s)

in subject-initial embedded clauses (Vfin-Adv vs. Adv-Vfin), where all slots of the construction are lexically unspecified, the free form of the **definite determiner** (*hinn* vs. *sá* ‘the’), where the determiner is lexically specified but other elements are not, and the (non-standard/frowned upon) use of the **generic pronoun** *maður* ‘one’, treated here as a fully specified univariate variable. The variables are studied on the basis of 19th-century Icelandic corpora featuring three different genres, viz. private letters, newspapers/periodicals and, finally, student essays, used in and partly developed as a part of the project *Language Change and Linguistic Variation in 19th-Century Icelandic and the Emergence of a National Standard* (19LCLV).

This research was sparked in part by an accumulation of evidence that attests to much more present-day variation in the adoption of standard norms than the conventional standardisation narrative suggests (see e.g. Árnadóttir and Einarsdóttir 2007, Heimisdóttir 2008, Leonard and Árnason 2011, Angantýsson 2017b), revealed also by studies on ongoing changes that have been frowned upon such as Dative Substitution (“Dative Sickness”) and the New Passive (“Castrated Passive”) (see e.g. Svavarsdóttir 1982, Svavarsdóttir, Pálsson and Þórlindarson 1984, Jónsson and Eyþórsson 2003 and Nowenstein 2017 on the dative, and Sigurjónsdóttir and Maling 2001 and Þráinsson, Eyþórsson, Svavarsdóttir and Blöndal 2013 on the passive). The general consensus among linguists is that prescriptive dicta typically have no influence on speakers’ actual language use (cf. e.g. Anderwald 2014b) and recent years and decades have also seen critical discussion of nationalistic ideology in historiography, including and beyond language (cf. e.g. Leerksen 1999, Hálfðanarson 2005, Elspaß 2014).

What the three linguistic variables treated in the present work show is that there is a clear uptake of the standard norms in the newspapers and periodicals as well as in the student essays, where language use gradually shifts towards the codified norms. This effect is much less pronounced in the private letters, although there is also some evidence that speakers’ social status and/or sex/gender plays a role in the adoption of these norms. The results suggest that the implementation of standard norms was indeed only ‘partly successful’ (cf. Elspaß’s 2016 “successfulness measures”).

The level of abstractness of the linguistic variable does not appear to be a relevant conditioning factor. The least successful variable in terms of norm implementation appears to be the one which is the most lexically specified, the generic pronoun *maður*. This variable exhibits a temporary slowdown in the newspapers during the late 19th and early 20th century, but a steady increase throughout the 20th and 21st centuries. The definite determiner variable, argued to be intermediate between the generic pronoun and verb-adverb placement with regard to lexical specification, overwhelmingly occurs with the non-standard variant in the private letter corpus and nearly categorically with dates. The newspapers exhibit a strong standardisation effect in the latter half of the 19th century, which remains in some linguistic contexts, but not all, throughout the 20th and 21st centuries. The most abstract of the variables, verb-adverb placement, in contrast, seems to exhibit the greatest effect of standardisation, even in the private letters. However, this is presumably partly due to the different social embedding of the non-standard Adv-Vfin variant vis-à-vis the other variables. The fact that the frequency of Adv-Vfin was highest in precisely the group of speakers most likely to respond to prescriptivism in the first place, the higher echelons, highlights the interaction between prescriptive linguistic change and social factors.

Útdráttur

Viðfangsefni þessarar athugunar er snertiflötur félagsmálfræði og setningafræði þar sem 19. aldar íslenska og innleiðing opinbers málstaðals frá setningafræðilegu sjónarmiði eru í brennidepli. Íslenska er annáluð í málfræðilegri umræðu fyrir fremur víðtæk áhrif málstöðlunar/málhreinsunar (sbr. Kjartan G. Ottósson 1987, 1990, 2003, 2005, Svavar Sigmundsson 1990-1991, 2002, Kristján Árnason 2003b, Thomason 1999, 2001, 2007, Kusters 2003, Finnur Friðriksson 2008, Hilmarsson-Dunn og Ari Páll Kristinsson 2010, Heimir F. Viðarsson 2017b). Aðalástæða þess að málstýring er talin hafa borið mikinn árangur hefur verið rakin til þess að staðalmálið hafi verið byggt að hluta til á daglegu máli alþýðunnar í sveitum, sem hafi jafnt og þétt útrýmt dönskuskotnum ritmálsstíl fyrri tíðar: „Meirihluti landsmanna samþykkti í raun og veru sinn eigin málstaðal“ (Finnur Friðriksson 2008, sbr. einnig Svavar Sigmundsson 1990-1991). Á hinn bóginn hefur verið lögð áhersla á að sum þessara áhrifa varði atriði sem talin eru hafa verið endurvakin af forníslensku málskeiði (sbr. einkum Kjartan G. Ottósson 1987, 1990, 2003, 2005). Að málstýring nái til allra sviða málfræðinnar, ekki einungis hljóðkerfisfræði, beygingarfræði og orðaforða (auk stafsetningar) heldur einnig til setningafræðinnar, er að verulegu leyti gagnstæð viðteknum hugmyndum um að óthlutstæðum sviðum málfræðinnar sé almennt veitt lítil eftirtekt (sbr. t.d. Labov and Harris 1986, Cheshire 1987, Laycock and Mühlhäusler 1990, Cheshire et al. 2005)). Eins hefur lítil áhersla verið lögð á að sannreyna algengar staðhæfingar um áhrif málstýringar í íslensku byggt á mállegum gagnasöfnum.

Samkvæmt SNERTIFLATARLÖGMÁLI Labovs (sbr. Labov and Harris 1986 og síðari skrif), sem einnig hefur verið kallað TILGÁTA UM ANDFÉLAGSLEGA SETNINGAFRÆÐI (sbr. Anton Karl Ingason o.fl. 2011, Ingason et al. 2013), eru félagsleg tengsl tilbrigða fyrst og fremst bundin við yfirborðið, þ.e. orð og hljóð tungumálsins. Óthlutstæðar málfræðiformgerðir eru hins vegar ekki taldar vera til þess fallnar að tengjast félagslegum hliðum nánnum böndum og sömu sögu er að segja um skrif um meðvitaðar málbreytingar, þar sem slík atriði eru neðarlega á STIGVELDI ÍHLUTUNARHÆFIS (sjá Laycock og Mühlhäusler 1990). Almenn niðurstaða þriggja tilviksrannsókna, sem greint er frá hér og nær til tilbrigða á setningafræðilega sviðinu (í víðum skilningi), er að á annan bóginn hljóti þröng tilvísun snertiflatarlögmálsins að vera ofátæluð, en á hinn bóginn sé umfang breytileikans í íslensku í sögulegu ljósi verulega vanátælað.

Rannsóknin miðar að því að kanna og kryfja viðteknar hugmyndir um setningafræði í félags(málfræði)legu ljósi að því er varðar hina hefðbundnu frásögn af íslenskri málstöðlun (sbr. t.d. Kjartan G. Ottósson 1990, 2005). Megináhersla er lögð á þrjár málbreytur sem eru að mismiklu leyti orðasafnslega skilyrtar. Málbreyturnar snúa að **stöðu persónubeygðrar sagnar** í aukasetningum með frumlagi fremst (so.-ao. andspænis ao.-so.), **formi lauss greinis** (*hinn* andspænis *sá*) og notkun **óákveðna fornafnsins**

maður, sem löngum hefur verið amast við. Þessar málbreytur eru rannsakaðar í þremur ólíkum textategundum, þ.e. persónulegum sendibréfum, dagblöðum og tímaritum og loks skólaritgerðum, allt málheildir sem m.a. hefur verið unnið að í tengslum við verkefnið *Málbreytingar og tilbrigði í íslensku máli á 19. öld: tilurð opinbers málstaðals*.

Kveikjuna að rannsóknunum 19. aldar verkefnisins var ekki síst að finna í vísbendingum um meiri breytileika í íslensku nútímamáli en hefðbundin lýsing á íslenskri málstöðlun bendir til (sbr. Linda Ösp 2008, Leonard og Kristján Árnason 2011, Ásgrímur Angantýsson 2017b), og í athugunum á málbreytingum sem árangurslítið hefur verið reynt að sporna við eins og þágufallshneigð eða „þágufallssýki“ og nýja þolmyndin eða „geld þolmynd“ (sbr. annars vegar Ásta Svavarsdóttir 1982, Ásta Svavarsdóttir, Gísli Pálsson og Þórólfur Þórlindarson 1984, Jóhannes Gísli Jónsson og Þórhallur Eypórsson 2003, Höskuldur Þráinsson, Þórhallur Eypórsson, Ásta Svavarsdóttir og Þórunn Blöndal 2015 og Iris Edda Nowenstein 2017 varðandi tilbrigði í fallmörkun og hins vegar Sigríður Sigurjónsdóttir and Maling 2001, Höskuldur Þráinsson, Sigríður Sigurjónsdóttir, Hlíf Árnadóttir og Þórhallur Eypórsson 2013, 2015 varðandi þolmynd). Almennt eru málfræðingar á því að forskrift í málfarslegum efnun hafi lítil sem engin áhrif á daglegt mál fólks (sbr. t.d. Anderwald 2014b) og á undanförunum árum og áratugum hefur borið meira á gagnrýni á lýsingar í sagnaritun sem taldar eru litaðar af hugmyndafræðilegum þáttum, þ.m.t. í málfarslegum efnun (sbr. t.d. Leerssen 1999, Guðmundur Hálfðanarson 2005, Elspaß 2014).

Málbreyturnar þrjár eiga það sammerkt að greinileg áhrif málstöðlunar má finna í dagblöðum/tímaritum og skólaritgerðum, þar sem málnotkun færast jafnt og þétt í átt að skráðum málviðmiðum. Áhrifin eru mun síður sýnileg í persónulegu sendibréfunum, þótt þar séu einnig einhverjar vísbendingar um að félagsleg staða og/eða kyn(gervi) bréfritaranna leiki hlutverk í upptöku á viðmiðunum. Niðurstöðurnar benda til þess að málstöðlun hafi aðeins verið „árangursrík að hluta“ (sbr. árangurskvarða Elspaß 2016).

Hve óhlutstæð málbreytan er frá málfræðilegu sjónarmiði virðist ekki hafa nein teljandi áhrif. Sú málbreyta sem síst eru vísbendingar um að hafi dregið úr fyrir tilstilli málstýringar, óákveðna fornafrnið *maður*, er einmitt sú sem er mest orðasafnslega skilyrt. Á útbreiðslu þessarar málbreytu hægist tímabundið í dagblöðum/ tímaritum á ofanverðri 19. öld og byrjun 20. aldar en síðar og fram til okkar tíma færast notkun hennar þvert á móti jafnt og þétt í aukana. Lausi greinirinn, sem færa má rök fyrir að sé mitt á milli óákveðna fornafrnsins og stöðu sagnar að því er varðar hve orðasafnslega skilyrt málbreytan er, birtist í miklum meirihluta tilvika með því afbrigði sem ekki er hluti af málstaðlinum og er næstum einrátt með dagsetningum. Dagblöðin sýna greinileg áhrif málstöðlunar á síðari helmingi 19. aldar og þau áhrif haldast að hluta til út 20. öld og inn á 21. öldina. Sú breyta sem telst vera mest óhlutbundin, staða sagnar, virðist sýna mest merki um áhrif málstöðlunar, jafnvel í sendibréfunum. Ástæðu þessa má væntanlega að hluta til rekja til þess að dreifing þess afbrigðis sem ekki er hluti af málstaðlinum (ao.-so.) er ólík hinum breytunum; það afbrigði hafði einmitt hæsta tíðni hjá þeim málhöfum sem eru líklegastir til þess að fylgja forskrift í málfræðilegum efnun, úr efrilögum samfélagsins.

In memory of Piet van der Feest

Table of Contents

Abstract	iii
Útdráttur	v
Table of Contents	ix
List of Figures	xiii
List of Tables	xviii
Abbreviations	xix
Acknowledgements	xxi
Introduction	1
1 Background	9
1.1 Syntax	9
1.1.1 Structure building and basic operations	9
1.1.2 Clause structure	11
1.1.3 Nominal structure	15
1.2 Sociolinguistics	16
1.2.1 The linguistic variable	16
1.2.2 Social embedding of linguistic variation	18
1.2.3 Socially constructed identities	20
1.2.4 The socio-syntax interface	21
1.3 Standardisation	25
1.3.1 Evaluating standardisation effects	25
1.3.2 Symbolic value and power	28
1.3.3 Deliberate change	30
1.3.4 The implementation of the standard	33
1.3.5 The importance of the grammar school(s)	38
2 Methodology	43
2.1 Corpora	43
2.1.1 Private letter corpus	44
2.1.2 Newspaper corpus	45
2.1.3 Student essay corpus	46
2.2 Social structure	48
2.3 Compulsory learning and literacy	51
2.4 Statistical analysis	53
3 Verb-adverb placement	57
3.1 Codification and implementation	59
3.2 A diachronic sketch of Adv-Vfin	62
3.2.1 The Adv-Vfin order in older Icelandic	62

3.2.2	Historical variation in Danish	64
3.3	Grammatical factors	67
3.3.1	Clause type	67
3.3.2	Adverb type	71
3.3.3	Subject type	73
3.3.4	Verb type	75
3.4	Results	77
3.4.1	Newspapers and periodicals	77
3.4.1.1	Overall trends in the data	77
3.4.1.2	Modelling the variation	78
3.4.1.3	Individual variation	80
3.4.1.4	Beyond the 19 th century and hypercorrection	83
3.4.2	Private letters	85
3.4.2.1	Overall trends in the data	85
3.4.2.2	Modelling the variation	86
3.4.2.3	Inter-speaker variation	92
3.4.2.4	Intra-speaker variation	95
3.4.3	Student essays	98
3.4.3.1	Overall trends in the data	98
3.4.3.2	Modelling the variation	98
3.5	Discussion	102
3.6	Previous analyses	105
3.7	Towards an analysis	110
3.7.1	Adv-Vfin as a low Spell-out and V2 violation	110
3.7.2	Targeted change, failure and success	118
3.7.3	Stylistic Fronting, null subjects and reanalysis	121
3.7.4	Consequences for linguistic theory	126
3.8	Summary	131
4	Free definite marker <i>sá</i> vs. <i>hinn</i>	133
4.1	The standardisation of definite marking	135
4.1.1	Evidence from grammars and published material	135
4.1.2	A view from the Reykjavík Grammar School	140
4.2	Grammatical factors	142
4.3	Results	144
4.3.1	Overall trends in the data	144
4.3.2	Modelling the variation	146
4.3.3	Individual variation	156
4.3.4	Student essays	160
4.3.5	Beyond the 19 th century	162
4.4	Discussion	165
4.5	Towards an analysis	166
4.5.1	Double definiteness	169
4.5.2	Complex NP structure	171
4.5.3	Strong adjectival inflection in NP-internal definite contexts	175
4.5.4	Co-occurrence of determiners	179

4.5.5	The functional sequence: The peculiar case of NUM	180
4.6	Summary	182
5	Generic pronoun <i>maður</i>	185
5.1	Basic properties	185
5.2	Stigmatisation	189
5.3	Circumscribing the variable context	192
5.4	Results	194
5.4.1	Overall trends in the data	194
5.4.2	Newspapers and periodicals	195
5.4.3	Private letters	198
5.4.4	Student essays	202
5.4.5	Beyond the 19 th century	204
5.5	Discussion	206
5.6	Towards an analysis	207
5.6.1	Feature decomposition	207
5.6.2	Form and function	209
5.6.3	Arbitrary reference?	211
5.6.4	A note on the origin of pronominal <i>maður</i>	212
5.7	Summary	214
6	Summary and conclusions	217
	Bibliography	225
	Appendices	245

List of Figures

1.1	Forms of communication on the conceptual continuum (cf. Koch and Oesterreicher 2007:349)	28
1.2	Social background of Reykjavík Grammar School students (1846-1904) based on the occupation of the father and the occupation of the students later in life (cf. Ólafsson 2004:40,44).	40
3.3	Teacher correcting Adv-Vfin in a student essay (1861, 4 th grade).	60
3.4	A forest plot of estimates with odds ratios of fixed effects: OR>1 (blue) favours Adv-Vfin, OR<1 (red) favours Vfin-Adv. (*' $p<0.05$, '**' $p<0.01$, '***' $p<0.001$.)	79
3.5	Predicted probabilities of the fixed effect CLAUSE TYPE over time, conditioned on random effects (lower values = less Adv-Vfin).	80
3.6	Proportion Neg-Vfin in newspapers from 1850 onwards ($n=93,748$). Unanalysed data set containing clauses introduced by <i>að</i> 'that' (typically DECL) or <i>sem</i> 'which' (typically REL), immediately followed by 3P.SG. subject (<i>hann/hún/það</i>) and the negation (<i>ekki</i>) either precedes or follows an inflected form of the verb <i>hafa</i> (only 3P.SG.).	84
3.7	Stylistic Fronting: Proportion of fronting versus non-fronting of negation in raw data set ($n=89,686$) compiled from newspapers from 1850 onwards. The data set contains clauses introduced by <i>sem</i> 'which' (typically REL), followed by the negation (<i>ekki</i>) and an inflected form of the verb <i>hafa</i> (only 3P.SG.) in either order, and lacking an overt subject.	85
3.8	A forest plot of estimates with odds ratios of fixed effects: OR>1 (blue) favours Adv-Vfin, OR<1 (red) favours Vfin-Adv. (*' $p<0.05$, '**' $p<0.01$, '***' $p<0.001$.)	87
3.9	Predicted probabilities of CLAUSE TYPE over time, conditioned on random effects (lower values = less Adv-Vfin).	88
3.10	Predicted probabilities of social status across time and sex/gender, conditioned on random effects (lower values = less Adv-Vfin).	89
3.11	Verb-adverb placement across five generations of speakers (1784-1918).	91
3.12	A forest plot of estimates with odds ratios of fixed effects: OR>1 (blue) favours Adv-Vfin, OR<1 (red) favours Vfin-Adv. (*' $p<0.05$, '**' $p<0.01$, '***' $p<0.001$.)	99
3.13	Essays: Predicted probabilities of CLAUSE TYPE over time (1852-1906), conditioned on random effects (lower values = less Adv-Vfin).	100
3.14	Essays: Predicted probabilities of GRADUATION SCORE over time, conditioned on random effects (lower values = less Adv-Vfin).	101
3.15	Verb-adverb placement in 19 th -century Icelandic across four basic clause types. All data from private letters, newspapers and student essays combined (1784-1924, per 25 years).	120
4.16	Newspapers: A forest plot of estimates with odds ratios of fixed effects. OR>1 (blue) increased odds of <i>sá</i> , OR<1 (red) decreased odds of <i>sá</i> . (*' $p<0.05$, '**' $p<0.01$, '***' $p<0.001$.)	147

4.17	Letters: A forest plot of estimates with odds ratios of fixed effects. OR>1 (blue) increased odds of <i>sá</i> , OR<1 (red) decreased odds of <i>sá</i> . (*' $p<0.05$, '**' $p<0.01$, '***' $p<0.001$.)	147
4.18	Newspapers: Predicted probabilities of fixed effect adjective category over time, conditioned on random effects (higher values = greater use of non-standard <i>sá</i> norm).	149
4.19	Letters: Predicted probabilities of fixed effect adjective category over time, conditioned on random effects (higher values = greater use of non-standard <i>sá</i> norm).	149
4.20	Newspapers: Predicted probabilities of fixed effect structural phrase type over time, conditioned on random effects (lower values = lesser use of non-standard <i>sá</i> norm).	150
4.21	Letters: Predicted probabilities of fixed effect structural phrase type over time, conditioned on random effects (lower values = lesser use of non-standard <i>sá</i> norm).	150
4.22	Newspapers: Predicted probabilities of fixed effect relative modification over time, conditioned on random effects (higher values = greater use of non-standard <i>sá</i> norm).	151
4.23	Letters: Predicted probabilities of fixed effect relative modification over time, conditioned on random effects (higher values = greater use of non-standard <i>sá</i> norm).	152
4.24	Predicted probabilities of <i>sá</i> across gender over time, conditioned on random effects (lower values = greater use of standard <i>hinn</i> norm).	153
4.25	Normalised frequency of <i>hinn</i> in the unambiguous DET-ADJ-NOM pattern across SEX/GENDER over time.	154
4.26	Newspapers: A forest plot of estimates with odds ratios of fixed effects. OR>1 (blue) increased odds of <i>sá</i> , OR<1 (red) decreased odds of <i>sá</i> . (*' $p<0.05$, '**' $p<0.01$, '***' $p<0.001$.)	161
4.27	Essays: Predicted probabilities of PHRASE TYPE over time, conditioned on random effects.	162
4.28	Essays: Predicted probabilities of GRADUATION SCORE (right) over time, conditioned on random effects.	163
4.29	Proportion <i>hinn</i> in newspapers from 1840 onwards ($n=337,902$). Unanalysed data set containing the full paradigm (masc., fem., neut. in the sg. and pl.) of five distinct phrases, incl. common variants thereof (e.g. <i>sá/hinn fyr(r)()nefndi</i>).	164
5.30	Overall frequency of <i>maður</i> in newspapers and periodicals, normalised per 10,000 words—contrasted against the private letter corpus for comparison.	194
5.31	Private letters: The rate of occurrence of generic pronoun <i>maður</i> , split for time period and gender, normalised per 10,000 words.	200
5.32	Use of the generic pronoun <i>maður</i> over generations split for gender, normalised per 10,000 words, cf. Table 5.26.	201
5.33	Essays: Generic pronoun <i>maður</i> by grade (year) over six periods, cf. Table 5.28.	203

5.34	Frequency of <i>maður</i> normalised in newspapers from before 1840 onwards ($n = 511,991$) as a proportion of the raw frequency of the top 25 most frequent words (92,705,498 words in total) according to the Icelandic Frequency Dictionary. Unanalysed data set containing clauses introduced by <i>að</i> ‘that’, <i>sem</i> ‘which’, <i>hvort</i> ‘whether’, <i>ef</i> ‘if’ and <i>þegar</i> ‘when’, followed by <i>maður</i> (only NOM.SG.) and an agreeing form of the verbs <i>geta</i> ‘can’, <i>hafa</i> ‘have’, <i>koma</i> ‘come’, <i>segja</i> ‘say’, <i>vera</i> ‘be’ and <i>verða</i> ‘must; will become’ (both IND. and SUBJ. forms).	204
5.35	Frequency of <i>maður</i> normalised in selected newspapers from 1850 onwards ($n = 29,685$) as a proportion of the raw frequency of the top 25 most frequent words (51,720,463 words in total) according to the Icelandic Frequency Dictionary. Unanalysed data set containing clauses introduced by <i>að</i> ‘that’ and <i>sem</i> ‘which’, followed by <i>maður</i> (only NOM.SG.) and an agreeing form of the verbs <i>geta</i> ‘can’, <i>hafa</i> ‘have’ and <i>vera</i> ‘be’ (both IND. and SUBJ. forms).	205
6.36	Newspapers: Predicted probabilities of fixed effects over time, conditioned on random effects.	246
6.37	Letters: Predicted probabilities of language-internal fixed effects over time, conditioned on random effects.	247
6.38	Newspapers: Predicted probabilities of fixed effect adjective category over time, conditioned on random effects (higher values = greater use of non-standard <i>sá</i> norm).	251
6.39	Letters: Predicted probabilities of fixed effect adjective degree over time, conditioned on random effects (higher values = greater use of non-standard <i>sá</i> norm).	251

List of Tables

1.1	The declination of <i>ia</i> -stems as codified in 19 th -century grammars, exemplified on the basis of the noun <i>læknir</i> ‘doctor’	37
3.2	Verb-adverb placement in texts from 1600-1800 (<i>Íslenskt textasafn</i>).	64
3.3	The distribution of embedded V2 (Wiklund et al. 2009:1921).	70
3.4	Verb-adverb placement in newspapers/periodicals, per quarter century.	77
3.5	Proportion V _{fin} -Adv/Adv-V _{fin} in newspapers and periodicals (1800-1924) with a minimum of 20 examples of either variant.	81
3.6	Verb-adverb placement in the private letter corpus, per quarter century.	86
3.7	Model comparison with and without SEX/GENDER (<i>anova</i>).	89
3.8	Percentage Adv-V _{fin} across social status and gender in private letters over time.	90
3.9	Individuals with at least 20 examples of either variant, ordered by percentage of Adv-V _{fin} in descending order (<i>n</i> = total number of both variants).	93
3.10	Verb-adverb placement for data-rich letter-writers in V2 and non-V2 environments over time (highest and lowest values boldfaced).	95
3.11	Proportion of Adv-V _{fin} in Bjarni Thorarensen’s letters in V2 and non-V2 environments, per recipient (<i>n</i> = total number of both variants).	96
3.12	Proportion Adv-V _{fin} in Guðrún Jónsdóttir’s letters in V2 and non-V2 environments, per recipient (<i>n</i> = total number of both variants).	98
3.13	Proportion V _{fin} -Adv/Adv-V _{fin} orders in student essays (1852-1906).	98
3.14	Relative frequency of Adv-V _{fin} in private letters and newspapers up until 1850 vs. after 1850 across different types of declaratives.	116
3.15	Logical types of universal statements following Greenberg (from Evans and Levinson 2009:437, adapted)	129
4.16	Egilsson’s use of the free articles <i>hinn</i> and <i>sá</i> in his Odyssey translation, 1829-1840.	138
4.17	Percentage <i>sá</i> across adjective categories in private letters vs. newspapers over time (higher values = greater use of sub-standard <i>sá</i> norm).	145
4.18	Model comparison with and without SEX/GENDER (<i>anova</i>).	152
4.19	Upper: Percentage <i>sá</i> across social status and sex/gender in private letters over time using the whole data set. Lower: Alternative data set drawing a random sample of max. 30 observations per individual.	154
4.20	Letters: Generational differences among males and females in the use of <i>hinn</i> vs. <i>sá</i> , per 25 years, where birth year is known (whole data set).	156
4.21	Proportion <i>hinn/sá</i> in newspapers and periodicals (1800-1924) with a minimum of 20 examples of either variant.	157
4.22	Raw distribution of <i>hinn</i> in the adjectival and nominal pattern in Period I and II (only includes letter-writers with at least 3 occurrences of the nominal pattern).	158
4.23	The use of the free article <i>hinn/sá</i> per period, graduation score, grade and phrase type.	160
5.24	Raw and normalised frequency of the generic pronoun <i>maður</i> in periodicals and newspapers 1803-1924, per quarter century.	197

5.25	Letters: The rate of the generic pronoun <i>maður</i> in letter-writers with 10 examples of the variable or more.	199
5.26	Use of the generic pronoun <i>maður</i> over generations split for gender, normalised per 10,000 words.	200
5.27	Essays: Generic pronoun <i>maður</i> based on three variables, raw frequency and normalised frequency (per 10,000 words).	202
5.28	Essays: Generic pronoun <i>maður</i> by grade over six periods; raw frequency and normalised frequency (per 10,000 words).	203
6.29	Relative frequency of Adv-Vfin in private letters and newspapers up until 1850 vs. after 1850.	245
6.30	Relative frequency of Adv-Vfin in private letters and newspapers up until 1850 vs. after 1850 across different types of subjects.	245
6.31	Percentage Adv-Vfin for letter-writers with at least 100 examples of the linguistic variable in different types of <i>that</i> -clauses. Arranged in order of appearance in Table 3.9.	248
6.32	Top: Use of <i>maður</i> over generations, split for gender and social status. Bottom: Time period (year written). ‘♀’ refers to females, ‘♂’ to males and ‘∅’ to both sexes. Generation 1 (birth year) = <1825, G2 = 1825-49, G3 = 1850-74, G4 = >1874. Period 1 = 1784-1850, Period 2 = 1851-85, Period 3 = 1886-1918.	253

Abbreviations

19LCLV	Language Change and Linguistic Variation in 19 th -C. Icelandic
ACC	accusative
ADV	adverb(ial)
AGRP	agreement phrase
AIC	Akaike information criterion
BIC	Bayesian information criterion
C	complementiser
CMP	comparative
CP	complementiser phrase
DAT	dative
DEF	definite marker
DET	determiner (definite/demonstrative)
DP	determiner phrase
GEN	genitive
FINP	finiteness phrase
ICEPAHC	Icelandic Parsed Historical Corpus
ICENCC	Icelandic Corpus of Early Nineteenth-Century Correspondence
IP	inflectional phrase
LL	log-likelihood
MID	middle voice
MÍM	Mörkuð íslensk málheild (Tagged Icelandic Corpus)
NOM	nominative
NP	noun phrase
OR	odds ratio
PRT	particle
RAH	Rich Agreement Hypothesis
REFL	reflexive pronoun
REL	relative marker
STR	strong inflection
SUB	subjunctive
SUP	superlative
TP	tense phrase
V2/V3	finite verb in second/third position
VFIN	finite verb
VFINAL	finite verb in clause-final position
WK	weak inflection

Acknowledgements

Some say one should not thank one's supervisors in the acknowledgments. After all, they're only doing their job, right? I disagree. It is certainly not a given that a student-supervisor relationship will be a fruitful one and in this project, I've been lucky enough to have had two great supervisors. Ásgrímur Angantýsson and Þórhallur Eypórsson (Tolli), thank you for all your help and support. Thank you for giving at least some room to the Icelandic dynamic, structure-in-chaos mentality, where needed. And thank you for having faith in my project even at difficult times when I did not. I also thank my PhD committee, Leonie Cornips, Jóhannes Gísli Jónsson and my supervisors, for their continued support throughout the years and especially over the past few months as my project was finally coming to an end. I greatly appreciate all your helpful comments and pointers on dealing with language variation and change, as well as on the more global methodological aspects.

It is often hard to distinguish between work and play. However, for valuable discussion about linguistic and professional issues, big and small, I would like to thank Kristján Árnason, Kristine Bentzen, Haraldur Bernharðsson, Kristín Bjarnadóttir, Þórunn Blöndal, Úlfar Bragason, Noam Chomsky, Julia Fernández Cuesta, Hans-Martin Gärtner, Gregory R. Guy, David Hákansson, Caroline Heycock, Katja Jasinskaja, Łukasz Jeźdrzejowski, Helgi Skúli Kjartansson, Olaf Koeneman, Ari Páll Kristinsson, Katrín Lísá van der Linde Mikaelisdóttir, Alda Bryndís Möller, Brafi Þorgrímur Ólafsson, Veturliði Óskarsson, Alexander Peter Pfaff, Eiríkur Rögnvaldsson, Gijsbert Rutten, Peter Schrijver, Michael Schulte, Jóhannes B. Sigtryggsson, Halldór Ármann Sigurðsson, Sigríður Sigurjónsdóttir, Doris Stolberg, Tania Strahan, John D. Sundquist, Ásta Svavarsdóttir, Peter Trudgill, Seid Tvica, Wim Vandenbussche, Rik Vosters, Joel Wallenberg, Jan-Wouter Zwart, Guðrún Þórhallsdóttir and last but not least Höskuldur Þráinsson.

My project received funds from the University of Iceland Research Fund, for which I am deeply grateful and humbled. It is also with great pleasure and gratitude that I thank *Leiden University Centre for Linguistics* (Leiden University), *Utrecht Institute of Linguistics OTS* (Utrecht University) and *Institut für deutsche Sprache und Literatur I* (University of Cologne) for generously hosting me as an affiliated researcher for extended periods in 2014, 2015-2016 and 2018-2019, respectively, and for providing an inspiring research environment over and above the one offered by my home institutions, the Árni Magnússon Institute for Icelandic Studies and the University of Iceland.

I repeat and extend my thanks to the members of the joint 19th-century Icelandic project and other colleagues at the Árni Magnússon Institute for Icelandic Studies. Being able to benefit from your expertise (and labour) with regard to corpus building and access to a range of electronic 19th-century Icelandic corpora is what really made this project possible in the first place. Despite being without linguistic annotation and

mark-up at this time, the corpora developed as a part of the joint 19th-century Icelandic project have served really well. It has been a great pleasure for me to have been able to contribute directly to the transcription process for not only the private letters but also the student essays. I would also very much like to thank my colleagues and students while teaching at *Menntavísindasvið* for sharing with me their passion for teaching Icelandic language and literature. I also want to express my gratitude to the staff at the National Archives of Iceland, the National and University Library of Iceland and the Directorate of Education for all their effort and willingness to be of service.

For the more personal aspects, behind the scenes, socialising, sharing thoughts, mutual shipwrecks, pep-talks, “all-work-and-no-play” reminders, coffee breaks, linguistics, the universe and everything, I thank Hlíf Árnadóttir, Dominique Blok, Eefje Boef, Kees Boxma, Marta Castella, Ronald van der Fange, Moragh Gordon, Nynke de Haas, Mirjam Hachem, Björn ’t Hart, Gunnar Marel Hinriksson (apologies for the widows and orphans), Ingunn Hreinberg Indriðadóttir, Hadil Karawani, Margreet van Koert, Margot Kraaikamp, Andreas Krogull, Jón Örn Loðmfjörð, Audrey Mahn, Iris Edda Nowenstein, Mike Olson, Matthias Passer, Guido Peek, Friðgeir Pétursson, Anna Sara Hexeberg Romøren, Bob Schoemaker, Jolien Scholten, Sigríður Sæunn Sigurðardóttir, Einar Freyr Sigurðsson, Páll Sigurðsson, Martijn van Steenberg, Arngrímur Vídalín and Petra van Werven. Heartfelt thanks and gratitude to my Dutch and Icelandic family for your help and continuous support. My mother-in-law, Marrie, for your willingness to lend a helping hand and for all your help babysitting Kjartan. My father-in-law, Piet (in memory), for your inspiring perseverance. My mother, Bryndís, for always being there and for coming to visit in Utrecht more than once. My grandmother, Helga, for opening up your house to us whenever we needed. My sister, Helga, for good laughs and very welcome distractions. My father, Viðar, and my grandmother, Matta, for all the good times during much too short visits. It has been great to have you all around at various points and moments in this long learning process. For inspiration and criticism. For asking how things are going. (And sometimes for not asking.)

But most of all, I thank my wonderful wife Jolanda. Thank you for all your support through thick and thin, in sickness and in health. Thank you for putting up with me, not to mention the long and crooked road leading to this piece of work, for your help and encouragement, and clear eye for structure and symmetry. Little Kjartan, thank you for your cheerfulness, your sweet smile, your fascinating linguistic adventures, as well as your understanding and sometimes even enthusiasm for *pabbi*’s strange taste in music.

Introduction

Research topic

The case of deliberate changes in Icelandic during the 19th century onwards is a well-known topic in historical sociolinguistics, referred to by Thomason (1999:24), for instance, as “the famous archaizing of Icelandic as a part of the process of standardization”. The outcome of this process is by no means a matter of course and constitutes a problem which is right at the heart of the joint research project *Language Change and Linguistic Variation in 19th-Century Icelandic and the Emergence of a National Standard*. The overarching research question central to that project is whether or not language change can be reversed—a deep question which clearly merits a better answer than a strictly binary one.¹

The extent to which language is amenable to deliberate change is an issue of current theoretical interest, gaining momentum as a viable research topic with the advent of large and diverse diachronic corpora in recent years (see e.g. Auer 2006, 2009, 2014, Poplack and Dion 2009, Poplack et al. 2015, Hendrickx 2013, Anderwald 2014a,b, 2016, Hinrichs et al. 2015, Krogull et al. 2017, Krogull 2018, Rutten 2019). In this context, scholars broadly distinguish between two kinds of studies (see e.g. Deumert 2003b:233ff., Auer 2009:5). On the one hand, there are studies that are mainly discourse-oriented, investigating the influence of language authorities, focusing on meta-linguistic discourse emerging from grammar books or via specific individuals. On the other hand, there are use-oriented perspectives more focussed on measuring the effects of prescriptivists on language use by means of electronic corpora. As Auer (2009:5) observes, studies on prescription and praxis in this latter sense and at a macro level have emerged only recently. It is also mainly this latter type of approach that will form the backbone of the present study into the effects of prescriptivism or, more specifically, language standardisation in an Icelandic context.

Nowhere is the need for an interdisciplinary approach that applies the methodology of sociolinguistics and historical linguistics to study language variation and language change as readily apparent in the history of Icelandic as in the 19th century. Traditional language histories coupled with nationalistic ideology and a strongly perceived contrast between the Danicised language of the cultural elite and the purer rural language of peasants and labourers (see e.g. Þorkelsson 1870, Smári 1920, Ottósson 1990, 2005, Sigmundsson 1990-1991, Hróarsdóttir 1998, Kusters 2003, Viðarsson 2016) makes for a somewhat convoluted and bewildering narrative of the emerging national standard language and its implementation—an intriguing blend of fact, fiction and contradiction.

¹ ‘Can linguistic change be reversed?’ can be said to partly echo the title of one of the early publications in the field of language planning and policy research, *Can Language be Planned?* (see Rubin and Jernudd 1971).

My focus here is on the so-called “long 19th century”, beginning in late 18th and stretching into the early 20th century, a period “when mass alphabetization and a general education system contributed not only to a significant increase in literacy levels, but also facilitated access and exposure to the norms of the standard language” (Deumert and Vandebussche 2003:458). Furthermore, the 19th century is unique from a variationist perspective in that early in this time frame, the effects of language standardisation are arguably very limited. As discussed by Deumert (2003a:66f.) in a wider European context, deliberate linguistic changes generally involved in language standardisation were brought about around 1800 by a conspiracy of factors. These included the interaction of language societies and academies, intellectual elites, improved school systems and better access to various types of media, the norms of which were adopted in part as the casual code of a growing number of speakers (see e.g. Wright 2004:54ff.). An important aspect of the sort of evaluation undertaken in studies of this kind is thus to better document the steps taken and investigate “the ‘black box’ of pedagogy” (Vandebussche 2007:29).

The main research questions that arise in the present context concern the effectiveness of the implementation of norms of fundamentally different linguistic variables in likewise fundamentally different types of texts. The rationale behind the selection of variables as well as the choice of corpora is related directly to two continua: firstly, the level of LEXICAL SPECIFICITY, i.e. relatively abstract vs. more lexically specified constructions, and secondly, the level of FORMALITY/FAMILIARITY on the other, representing relative informality/closeness vs. more formal/distant text types. The objects of study belong to two well-researched domains of structure: variation at the level of the clause and the noun phrase, respectively. These variables were initially selected based on three criteria: 1) they were (in some sense) syntactic in nature, 2) they were targeted in 19th-century language commentaries and/or grammar books from early on, and 3) the suggestion has been made that they have been affected by standardisation.

The main linguistic variants are exemplified on the basis of 19th-century data below from the private letter corpus (see Section 2.1.1):

(1) VARIABLE VERB-ADVERB PLACEMENT:

- a. *eg finn eg get ei talað eða skrifað mikið um það.* (VFİN-ADV / “V2”)
 I feel I can not speak or write much about that
 ‘I feel that I cannot speak or write much about that.’
 (GudMag-1819-00-00.txt)
- b. *og gefi það nú guð, að þetta bréf ei angri þig.* (ADV-VFİN / “V3”)
 and give that now god that this letter not bothers you
 ‘And I hope to God that this letter does not bother you.’
 (GudMag-1819-00-00.txt)

(2) VARIABLE DEFINITE MARKER (*hinn* vs. *sá* ‘the’):

- a. *veduratta hefur vérid hinn æskilegasta*
 weather has been the_{HINN} desirable_{SUP}
 ‘The weather has been (the) most desirable.’ (SigPal-1830-03-07.xml)
- b. *veduráttan hefur verid sú æskilegasta i vetur*
 weather-the has been the_{SÁ} desirable_{SUP} in winter
 ‘The weather has been (the) most desirable this winter.’
 (SigPal-1841-03-07.xml)

(3) GENERIC PRONOUN (*maður* ‘one’):

En þau tímabilin eru þyngri, þegar maður getur ekki grátið.

but those time-periods_{DEF} are heavier when one can not cry

‘But the times when one cannot weep are more difficult.’

(GudMag-1844-06-13.txt)

VARIABLE VERB-ADVERB PLACEMENT is a phenomenon that has received a great deal of attention in the linguistic literature and continues to be central to linguistic theorising in general, going back to the work of 19th-century scholars such as Delbrück (1888) and Wackernagel (1892). This variable appeared in metalinguistic discussion in Iceland quite early on, at least as early as Gíslason (1844). Due to the negative evaluation of Adv-V_{fin} in Icelandic, language standardisation has been considered to be responsible for the cross-linguistic differences between Icelandic, where V_{fin}-Adv was selected as the norm, as opposed to the other Scandinavian languages, where in contrast Adv-V_{fin} was selected (cf. Pettersson 1988). However, such claims have not been substantiated on the basis of use-oriented studies, at least not beyond noting that Adv-V_{fin} had become quite common in Icelandic publications during the period from approximately 1600-1850 (cf. Smári 1920, Heycock and Wallenberg 2013).

VARIABLE DEFINITE MARKER strategies feature variation at two different levels. First of all, there is relatively abstract variation in terms of a choice between prenominal and postnominal (suffixal) determiners. As discussed at length by Pfaff (2015) and Ingason (2016), the pre- vs. postnominal determiners do not compete for the same meaning and will not be studied here. However, there is also a lexical choice within the prenominal determiner strategy between *hinn* and *sá* with the same or similar semantics, the latter of which is usually taken to be non-standard. According to the literature, the use of prenominal *sá* gradually gained terrain at the expense of prenominal *hinn* during the early modern period until *hinn* was prescribed as the norm following the purism of the 19th century and onwards (cf. Ottosson 2003). As a result, *sá* became confined to the spoken language (cf. also Ólsen 1882), where it gradually lost ground as well.

GENERIC PRONOUN *maður* is quite different from the above variables in that it is not as clear-cut what precisely the alternative variant(s) would be (see Ragnarsdóttir and Strömqvist 2005, Sigurðsson and Egerland 2009). However, similar to the above, pronominal *maður*, too, has been frowned upon and generally ascribed to foreign influences, in particular with reference to the generic (or ‘indefinite’) pronoun *man* in Danish. Overt prescriptive remarks targeting this feature appear to be non-existent in published 19th-century sources, but there is anecdotal evidence that it was a great thorn in the side of the Icelandic teacher, Halldór Kr. Friðriksson, at the Reykjavík Grammar School (cf. Ottósson 1990:96). This is confirmed here based on corrections in student assignments, although it did not prevent him from using *maður* as a generic pronoun in published as well as unpublished sources (see Section 5.2).

As discussed in more detail in Section 2.1, the corpora used in this thesis mainly consist of colloquial private letters (1.5 million words), newspapers and periodicals (2 million words) and, finally, student essays (83,000 words). For comparison, other corpora have been consulted as well, especially the Icelandic Parsed Historical Corpus (Wallenberg et al. 2011), the historical Text Collection of the Árni Magnússon Institute (*Íslenskt textasafn*), a larger corpus of newspapers and periodicals from the mid-19th century to the present day hosted by the National and University Library of Iceland

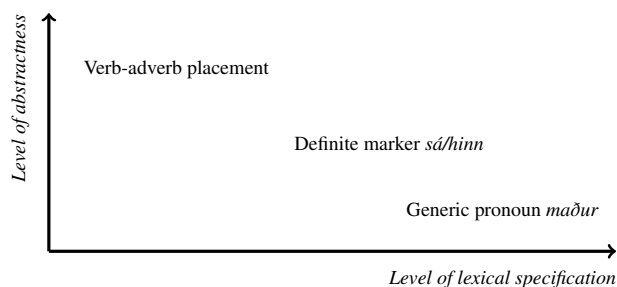
(*Tímarit.is*) and last but not least The Gigaword Corpus of 20th- and 21st-century Icelandic, which consists of about 1.3 billion words (see Steingrímsson et al. 2018).

On a more general level, the present study is an exploration into the interaction between syntactic and social aspects of language from a historical perspective, aligned with the sub-discipline of historical linguistics and sociolinguistics, collectively referred to as HISTORICAL SOCIOLINGUISTICS (see e.g. Romaine 1982, Hernández-Campoy and Conde-Silvestre 2012 and Auer et al. 2015). To underscore the role syntactic phenomena play in this work, the more specific term HISTORICAL SOCIO-SYNTAX might also be suggested, seeking inspiration in the (mainly synchronic) SOCIO-SYNTAX approach (see e.g. Cornips and Corrigan 2005a,b, Adger and Smith 2005, Cornips 2014, Freywald et al. 2015) and, more generally, the language variation and change tradition going back to the seminal work of Labov (1965, 1972a, 1972b, 1994, 2001, 2010, 2016) and variationist sociolinguistics (e.g. Tagliamonte 2012).

A common focus in traditional Labovian sociolinguistics is variation at the phonological level, whereas more abstract aspects of language subject to variation, such as syntax, are considered to be much less likely to exhibit sociolinguistic patterns (see Section 1.2.4). From the variationist/socio-syntactic perspective, however, syntactic variation, too, is considered to be determined to a large extent by social and stylistic factors (cf. Romaine 1982:206f, 167-170, Grondelaers 2009:289, Tagliamonte 2012:207, Cornips 2014:4). Much of this literature also goes against basic tenets of formalist theories of syntax, which tend to emphasise the autonomy of syntactic structure building processes. As a result, speakers may be depicted as being largely oblivious to variation at the abstract level of syntax. While it is widely acknowledged that syntax may be situated low on the awareness scale (cf. Laycock and Mühlhäusler 1990, Mair 2009), this is stated very forcefully in hypotheses such as the ANTISOCIAL SYNTAX HYPOTHESIS (cf. Ingason et al. 2012, 2013), where it is suggested that apparent socio-syntactic variation is always lexical in nature. Supposed conditions of this kind have been made explicit by Labov and Harris (1986) in what has become known as Labov's INTERFACE PRINCIPLE and relating further to constraints on the SOCIOLINGUISTIC MONITOR (see discussion in Section 1.2.4).

As briefly mentioned above, the linguistic phenomena that will form the main focus of the present study occupy different positions on a syntax-lexicon continuum, ranging from being highly abstract to being fully lexically filled:

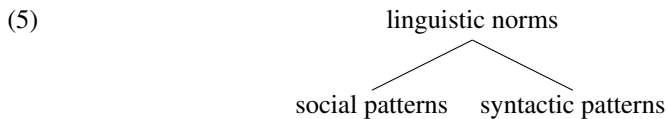
(4) *Level of schematicity of the linguistic variables under study*



The syntax-lexicon continuum is emphasised in various cognitive approaches to language such as construction grammar (see e.g. Goldberg 2006:5, 220; Boas 2010). The term is used here in a pre-theoretic sense, making no commitment to a constructional view of grammatical architecture. Highly lexically-specified constructions such as the generic pronoun are expected to be closer to the Labovian ideal of an accessible linguistic feature (“words and sounds”) and thus expected to participate more readily in sociolinguistic variation than abstract notions such as the structural position of the finite verb. The definite determiner then occupies some sort of middle-ground between the two in that the determiner is lexically specified whereas the slot for the adjective and/or nominal is not.

Similarly, the text types occupy vastly different positions on a continuum ranging from highly informal texts written to a specific, familiar individual, typically family and friends, to highly formal and distant texts such as the student essays. Again, newspapers and periodicals occupy a middle-ground, being relatively formal, addressing an abstract (national) speaker. It is particularly in the latter two text types, therefore, that we expect to find effects of standardisation and/or prescriptivism, whereas the private letters may be expected to reveal more (socially determined) variation. The student essays are expected to be subject to strong (or even ‘inflated’) normative influences, because the essays formed a part of each student’s grade at the Reykjavík Grammar School (1846-1904), taken to have been elementary in implementing and spreading the standard norms (see Sections 1.3.5 and 2.1.3).

The study can be visualised as follows, roughly corresponding to Silverstein’s (1985) ‘total linguistic fact’ taken to comprise linguistic **structure, usage** and **ideology**:²



For the sake of concreteness, the overall linguistic framework that I adopt is Minimalism (Chomsky 2000, 2001, 2008, 2013) with further elaboration along the lines of the cartographic approach (cf. Rizzi 2004, Cinque 1999, 2004, Haegeman 2004, 2012, Rizzi and Cinque 2016). The linguistic domains under study touch on aspects of embedded clause structure, the extended verbal domain and the extended noun phrase—in others words, the three phases of minimalist syntax: CP, vP and DP. Chomsky’s Merge/Agree-based framework is sufficiently rich to readily account for the syntactic patterns discussed in this thesis, further complemented by the notion of derivational (“constructional”) layers (cf. Zwart 2009, Trotzke and Zwart 2014) that may subsequently be subjected to social evaluation.

²The ‘total linguistic fact’ is sometimes presented as quadripartite with the additional factor of *domain*, referring to the recognition of an indexical link established between a linguistic form and an associated ideology by a set of speakers, subject to variation (see e.g. Wortham 2008).

Objectives

The main objective of this study is to provide an account of a range of linguistic phenomena in the history of Icelandic, focusing on the 19th-century, from a socio-historical and syntactic perspective, in a variety of different sources. The most important questions that I will focus on are given in (6) below:

- (6) a. What is the scope of variation in these three linguistic domains of study across the three text types?
- b. To what extent do language descriptions and language commentaries reflect actual usage as attested in the historical record, especially in terms of the perceived contrast between ‘educated’ vs. ‘rural’ Icelandic?
- c. How strong is the evidence that language purism, language revivalism and language planning had an active and effective role in (re)shaping the form of “Icelandic”? To what extent did the selection of a particular standard norm have an effect on actual language use?
- d. To what degree are these phenomena subject to social (e.g. in terms of sex/gender, age, education, social status) and stylistic factors (text type/medium)?
- e. To what extent are the restrictions/properties identified in the literature on Modern Icelandic relevant to 19th-century Icelandic?
- f. How can we account for social and syntactic variability in 19th-century Icelandic from a theoretical perspective?
- g. Does the social/syntactic distribution reveal anything about how each variant may have arisen (e.g. through internal or external factors)?

These questions all touch upon aspects of language variation and change discussed in the seminal paper of Weinreich et al. (1968) on the social *evaluation* of linguistic structure, the *embedding* of variation within the (socio-)linguistic system, potential *constraints* on variation and change, as well as the questions of *actuation* of change and *transition* from one stage to the next (see Section 1.2.2).

In order to be able to address such questions at all, one must approach the traditional standardisation narratives critically. In previous work, metalinguistic discourse has often simply been taken at face value, whereas a careful analysis of such descriptions might actually reveal an underlying ideological bias that the narrative does not acknowledge. Furthermore, corpus-based studies verifying commonly made claims about the standardisation of Icelandic are hitherto either very scarce or lacking completely. An important aspect of the present study is thus to try to assess such claims empirically by checking praxis against prescription in a range of different sources.

Attaining comprehensive knowledge of the attested variation at different points in time and in a broad range of sources is not only an illuminative undertaking in determining the extent of variation (including its limits) but indeed is a prerequisite to a work like the present one; during the late 18th and early 19th century, the effects of an emerging standard are unlikely to have been very strong and the efforts would not have been particularly concerted. In contrast to this, the period after 1850, once the dust had begun to settle so to speak, ought to give a more consistent view of which variants should or should not constitute the standard.

Outline

The structure of the dissertation is as follows. Chapter 1 presents the theoretical underpinnings of the study, beginning with fundamental aspects of syntactic structure in Section 1.1 and sociolinguistics in Section 1.2. Section 1.2.4 follows up on this with a discussion on the way syntax and sociolinguistics interact. The first three subsections of Section 1.2 deal with how language (including syntactic) variability is defined and operationalised in the variationist/sociolinguistic literature, focusing on the linguistic variable, the social embedding of change within society and social identities, including the complex notion of sex/gender. Sections 1.3.1 and 1.3.2 more specifically deal with aspects related to standardisation, building upon the Labovian notion of (meta)linguistic awareness and Haugen and Bourdieu's notions of symbolic value and symbolic power. Section 1.3.3 specifically deals with the issue of deliberate change in general and with regard to Icelandic standardisation in particular. This discussion continues in Sections 1.3.4 and 1.3.5 on the implementation of the standard language and the importance of the Reykjavík Grammar School, respectively.

Chapter 2 offers a brief discussion about the methodology, including the corpora used (Section 2.1), the operationalisation of social variables (Section 2.2), the basic educational background and literacy levels in 19th-century Iceland (Section 2.3) and, finally, a very brief discussion on the statistical analyses (Section 2.4), including statistical modelling used extensively in Chapters 3 and 4.

Chapter 3 presents the first case study. First, a general background to the study of Adv-Vfin/Vfin-Adv is provided, including the selection of the norm (Section 3.1) and a diachronic overview of the Adv-Vfin variant with a comparative study of historical variation in Danish (Section 3.2). Section 3.3 deals with the main grammatical factors likely to be relevant to variation in verb-adverb placement based on extensive discussion in the literature. This part of the chapter is at the same time a detailed circumscription of the variable context. Due to the complexity of this linguistic variable, readers not particularly interested in syntactic aspects might want to skip directly to the results in Section 3.4. The results section deals with both social and grammatical factors. A statistical analysis is offered based on mixed-effects statistical modelling, coupled with traditional descriptive (cross-tabs) statistics for comparison. The results are followed by a brief discussion in Section 3.5, previous accounts in Section 3.6 and analysis in Section 3.7.

Chapters 4 and 5 present the second and third case studies, respectively. Chapter 4 is a study on the use of the definite markers *hinn/sá* 'the'. The discussion in this chapter largely follows the structure of Chapter 3, presenting a statistical mixed-effects analysis of the attested variability, followed by discussion and subsequent analysis. In this respect, Chapter 5 on the generic pronoun *maður* 'one' is radically different. The study on the generic pronoun, defined here as a univariate variable, is analysed in terms of raw, normalised frequencies, as opposed to relative proportions of two (or more) variants as in the two previous chapters. The statistical analysis thus also deviates from the two previous studies, instead adopting log-likelihood tests for statistical purposes.

Finally, Chapter 6 offers a brief summary and conclusions, highlighting the main arguments presented in the preceding chapters, as well as pointers for further research. Summaries are also found at the end of each chapter.

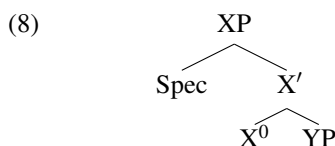
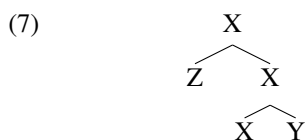
1 Background

1.1 Syntax

The syntactic framework assumed throughout this book builds on work done in the Minimalist tradition of generative grammar (cf. Chomsky 1995, 2000, 2001, 2008), with further extensions to clause and noun phrase structure as proposed on the so-called Cartographic approach to syntax (Rizzi 2004, Cinque 1999, 2004, 2013). I would like to emphasise right at the outset that I do not regard developing these sets of theories as a primary goal of my research. Instead, by adopting a concrete version of an existing approach (or set of approaches) to syntax that is reasonably well worked out with regard to the issues at hand, this framework can serve to guide or inform my research into the empirical phenomena under study.

1.1.1 Structure building and basic operations

The main ingredients of Chomsky’s generative approach are, first of all, the structure-building operation MERGE, establishing set-membership, and, second, AGREE, establishing a probe-goal relation between a set of elements. Merge combines two syntactic objects (SO), forming the unordered set {X, Y}. Projecting further, this structure receives a label necessary for computation in the course of derivation.³ By way of illustration, a basic structure of three elements, for instance a subject-predicate construction with a filled object slot (e.g. *John saw Mary*), is shown in (7), abstracting away from details as to how these elements enter the derivation, subsequent agreement relations, clause-typing and so on. The corresponding structure using the traditional primitives in the notation of X-bar theory (maximal projection, intermediate bar-level, head, specifier, complement) is given in (8):



³The specific mechanism for determining this label is irrelevant for my purposes here (for discussion, see e.g. Chomsky 2008, 2013). For concreteness, I assume that the element serving as a label is the most prominent element of the relevant piece of structure and that at least some problems arising with regard to the labelling of complex heads (Chomsky 2013) may be avoided if lexical items are treated as emergent entities as in Distributed Morphology, consisting of a category-less root merged with a category-defining head (e.g. \sqrt{kiss} merging with verbalising *v* head to derive the verb *to kiss*). These elements will then be labelled according to “the functional elements determining category” (Chomsky 2013:47) or “by selectional features at Merge” (Ingason 2016:9).

An object that is not the output of a Merge operation is a head (X^0), Merge of $\{X, Y\}$ corresponds to an intermediate bar node (X^1) and an object that does not project further is a phrase corresponding to a maximal projection (XP) in the notation in (8). When X and Y first form a set, the operation is sometimes referred to as *external Merge* (EM), whereas displacement of elements in this model obtains if an element already a part of a set is merged again (resulting in two copies), referred to as *internal Merge* (IM). When the syntactic output is externalised for communicative purposes, it is typically the structurally higher instance of an internally merged element that is pronounced, leaving the lower copy (or copies) unexpressed.

Agree establishes a relation between two syntactic objects formed by Merge, where one object contains the other, as defined by c-command (see e.g. Adger 2003, Hornstein et al. 2005). Here, structure labelling again becomes relevant: “The label selects and is selected in EM, and is the *probe* that seeks a *goal* for operations internal to the SO: Agree or IM.” (Chomsky 2008:141). Agree is a feature-matching mechanism which relates two sorts of features, interpretable [*i*F] and uninterpretable features [*u*F], mediated by a head with a probe feature seeking a matching goal. In case of a subject-predicate construction, the finite verb has uninterpretable Tense and Agreement features for person, number and gender (referred to as phi-features) that need to be valued (matched) against interpretable phi-features of the nominal. This can be achieved remotely at a distance or locally by displacement, raising the goal to the probe—an aspect in which languages may vary (see Section 3.7).

While Agree simply consists of a probe seeking a goal in its search domain, the operation is bound by intervention and locality effects that may arise in the course of a derivation. Common examples of these are intervention effects due to non-local matching of an Agree relation where a closer goal serves as an INTERVENER, blocking Agree with a structurally lower element (see e.g. Adger 2003:218f.). Intervenors will become a topic of discussion in Chapter 4, Section 4.5.

Another important (and absolute) constraint is that of PHASES, the hypothesis that the derivation proceeds in “chunks” at which point (or points) a syntactic object is passed on to the interfaces: the phonological component, mapping it to the sensorimotor interface, and the semantic component, mapping it to the semantic/conceptual-intentional interface (Chomsky 2008:142). Phases are standardly taken to consist of the CP (clause structure) and vP (argument structure) levels, but frequently also DP (nominal structure).

The derivation begins with a numeration containing the elements to be merged. For concreteness, I assume that lexical words enter the derivation as roots, attaining their word class by merging e.g. with a verbalising head *v* to yield a verb or nominalising head *n* yielding a noun. However, I consider ‘constructions’ to be an integral part of representation in the course of a derivation, yielding subderivations that may be manipulated as such but potentially also evaluated socially, much like typical lexical items or “words”. These layers of structure can feed into another derivation, “enter[ing] as atomic items in a syntactic derivation” (Trotzke and Zwart 2014:144), which means that the building blocks of a derivation may consist of phrases or even clauses. Needless to say, “functionally” atomic elements such as a constituent noun phrase, too, are composed by Merge, given the premise of *syntax all the way down*. Derivational layers will play a role in (at least partly) accounting for structural differences between different clause types, which is relevant for the verb-adverb placement variable (cf. Chapter 3).

By virtue of the ability for complex elements to serve as atomic, these elements can arguably be targeted in a constructional fashion in the case of social evaluation such that speakers may be aware of the output of a subderivation, without necessarily being aware of any interaction of that element within the larger derivational network nor, obviously, the way in which, say, a noun phrase like *the pig* is composed of a nominalised root, merged with a determiner morpheme (see Section 1.1.3). Therefore, social evaluation of externalised syntactic structure, I suggest, is still always close to the surface. When referring informally to constructions with regard to verb-adverb placement, the definite marker and the generic pronoun below, it is therefore intended in the specific meaning just outlined. The reader is, of course, free to dismiss constructions/derivational layers as theoretical notions, or, indeed, to regard constructions as a theoretical primitive in their own right.

1.1.2 Clause structure

I adopt the traditional analysis of hierarchical clausal structure as involving three basic domains or zones that can roughly be labelled CP > IP > VP. While I will use these traditional terms below, I take each of these domains to be abbreviated labels or cover terms for a more elaborated structure, the C-domain, for instance, being a shorthand for Rizzi's (2004) elaborated left periphery of the clause (cf. also Chomsky 2008:143). These domains can be taken to refer to **functional** complementiser and inflectional domains and the **lexical** domain of verbal argument structure, respectively. That being said, I take these traditional labels, assigned to structures generated by the structure-building operation Merge, to be but a notational convenience; ideally, phrase structure is 'bare' (Chomsky 1995) with labels being read off of the structure based on its composition and feature specification by minimal search (Chomsky 2013).

In line with much research in the Cartographic tradition (Rizzi 2004, Haegeman 2004, 2012, Cinque and Rizzi 2008), CP is here considered to involve a constellation of hierarchically-ordered projections denoting sentential and illocutionary force, information structure (with Topic a recursive property, indicated with '*') and finiteness:⁴

(9) [CP Sub Top* Force Top* Foc Top* Fin [IP ... [VP ...

SUB denotes sentential force in terms of an operator representing a specific clause type, e.g. conditional, relative or interrogative. FORCE expresses the semantic or illocutionary force of the clause, e.g. assertions, orders, requests and so on, and may host e.g. declarative complementisers, as required by the selectional requirements of SubP. The

⁴Force is traditionally regarded as the highest projection and Finiteness the lowest, with specialised information-structural positions relating to Topic and Focus structure occurring in-between. Subsequent work mostly on the Romance languages has revealed a potentially even more elaborated hierarchy of functional projections in the left periphery, argued to consist at least of ForceP > IntP > FocusP > ModP > QembP > FinP, with intermediate TopicP positions occurring freely in-between. The availability of these Topic-Focus heads is subject to extensive cross-linguistic variation (see e.g. Rizzi and Cinque 2016:146f., with references). The subordinator phrase (SubP) is additionally often employed to denote clause type/sentential force separately, e.g. hosting a conditional operator for conditional clauses, with an assertion operator in Force denoting illocutionary force (see e.g. Haegeman 2012:225f.), and I assume that representation here (see also Angantýsson 2011). Complex subordinating conjunctions such as *ef að* 'if (that)', *sem að* 'which (that)' and *hvort að* 'whether (that)' can then be regarded as occupying distinct positions, Sub and Force, respectively.

Force head is roughly the equivalent of C^0 in V2 languages where V2 interacts with the presence/absence of an overt complementiser (cf. den Besten 1977, 1989, Vikner 1995, with a precursor in the structuralist work of Paardekooper 1961, see also Zwart 2008). FIN encodes whether or not the clause is finite and Tense is at the border of the CP/IP domain, mediating between the left periphery and the content of the embedded IP; as such, the Fin head resembles and shares certain properties with I (or T) in non-cartographic approaches. I follow Rizzi (2004) and much subsequent work in the cartographic tradition in separating finiteness from the inflectional domain, identifying it with the C-system and not the I-system.⁵ The Split-CP analysis allows for a rather simple analysis of different V2 systems across the Germanic languages as targeting positions in the C-system, differing in whether they occur in Force, yielding the complementary distribution in asserted contexts, or in Fin (cf. Holmberg and Platzack 2005, Wiklund et al. 2007, Hrafnbjargarson and Wiklund 2010).

TOP is a position for fronted topics and is assumed, based on evidence from Romance languages, to be recursive, scattered between each of the other functional projections of the left periphery (Rizzi 2004). Finally, FOC hosts focussed elements that have been fronted, typically involving Focus-Presupposition structures but may also refer to non-presuppositional foci that introduce new information. The approach is a “syntactisation” of criterial heads in traditional Spec-head-complement configurations, guiding the interpretation of Topic-Comment, Focus-Presupposition and so on.⁶

While the split CP structure may seem overly elaborate, a richer structure is arguably necessary to account for various aspects of Icelandic syntax. A good example of this is a common particle construction, originally discussed by Rögnvaldsson and Thráinsson (1990). The examples in (10) are provided as a case of left dislocation, featuring “an “extra” XP to the left of C” in their view (i.e. traditionally in Spec,CP; from Rögnvaldsson and Thráinsson 1990:35f.):⁷

- (10) a. *Í morgun, að þá fór rafmagnið*
 this morning C_{that} then went electricity-the
 ‘This morning we had a power outage.’

⁵Finiteness (Fin) and inflection for *Tense* (T) are regarded as denoting distinct heads, where T (or usually I in earlier terminology) marks semantic tense with respect to an event, whereas Fin relates to a speech event (see e.g. Adger 2007, Sigurðsson 2011). More specifically, the temporal reference of Fin is about “topic time” that holds in relation to the illocutionary force of the clause, thus being more naturally regarded as an inherent feature of the C-level than as an inflectional/tense feature of the I-system.

⁶I am completely agnostic as to whether to regard these elements as projections in their own right as opposed to being features of C (or features of a recursive C).

⁷While this construction always features the element *þá* ‘then’, it need not have a temporal reference, also occurring in conditional statements such as *if ... then*. So as not to give the impression that this construction is only a very recent phenomenon, I point out that a range of potentially similar examples are attested in 19th-century Icelandic as well (ex. (1a) could be a verbal particle *gá* (*að*) ‘check PRT’ but (1b) is unequivocal):

- (1) a. *um daginn þegar eg fór að gá að þá Vantaði þetta*
 about day-the when I went to check C_{that} then missed this
 ‘The other day when I went to check, this was missing.’ (HalPet-1862-05-13.xml)
- b. *við hugsum að ef henni ekki batnar að þá muni Carl, fara með hana til Englands*
 we think that if her not recovers C_{that} then will_{SUB} Carl go with her to England
 ‘We think that if she does not get better, Carl will take her to England.’ (GdrJon-1891-02-02.xml)

- b. *Þegar þú kemur að þá verð ég glaður*
 when you come C_{that} then become I glad
 ‘I will be glad when you come.’

In Rögnvaldsson and Thráinsson’s analysis, *þá* ‘then’ occupies a position below the CP-level with the subject in VP. Thráinsson (2011:5), “updating” this analysis to the Split IP structure in Bobaljik and Thráinsson (1998), takes *þá* to be in the upper subject position (AgrP) and the actual subject in the lower subject position (TP) (see Section 3.6). This analysis is problematic for a number of reasons. AgrP is a *subject* position in Bobaljik and Thráinsson’s system and argued to be the locus of definite/specific phrases subjects, whereas indefinite/non-specific phrases occupy the lower subject position. More generally, fronting of topical elements to IP, also assumed in Rögnvaldsson and Thráinsson (1990), goes against the broad consensus that information structure, topicalisation in particular, is associated with the CP-domain.

Ideally, the fronted element *þá* ‘then’ occurs in the left periphery and I take it that it is in TopP, above FinP, the finite verb being in Fin as usual. Now the definite subject can be assumed to occupy the *higher* subject position, as it should if we assume the Split-IP analysis of Bobaljik and Thráinsson (1998). If the initial elements are really dislocated (cf. Rögnvaldsson and Thráinsson 1990), they are higher than Force, suggesting that the complementiser is in Force and the particle in Top. The subject is then barred from entering the C-domain (a V2 ‘bottleneck’ effect) and we have arrived at the surface word order. An alternative might also be that the initial element is fronted to Spec,Top with the complementiser in Top⁸ and that *þá* is not a particle but complementiser agreement (*að+þá*), where the complementiser agrees with the fronted element.⁹

Like CP, the IP domain is also a shorthand for richer functional structure. Following Pollock (1989) and subsequent work of Cinque (1999, 2004, 2013), I assume that the IP domain can be expanded to comprise functional projections for verbal subject AGREE-MENT, TENSE (e.g. past, future, anterior), ASPECT (e.g. habitual, perfect, completive), MOOD (e.g. evaluative, evidential, irrealis), MODALITY (e.g. epistemic, alethic, root) and VOICE (e.g. impersonal, middle, passive). The Pollockian Split-IP into Tense and Agreement (TP and AgrP) has been applied in much subsequent work on Icelandic (see e.g. Bobaljik and Thráinsson 1998, Thráinsson 2010, Angantýsson 2001, 2007, 2011). Cinque’s more elaborate decomposition of the clause takes adverbs to be functional elements that correspond directly to phenomena such as e.g. tense, aspect and mood—some languages may encode these grammatical distinctions by means of verbal morphemes, others by adverbial phrases. Cross-linguistic research reveals that each

⁸See also discussion in Section 3.6 around example (73) on page 109.

⁹Potential evidence for this might be that *þá* is obligatory if and only if the complementiser is present:

- (1) a. *Í morgun (*að) fór rafmagnið* b. *Í morgun (að) þá fór rafmagnið*
 this morning that went the.electricity this morning that then went the.electricity

Since *í morgun* is structurally higher than *þá* (or *að+þá*), these data may actually suggest upward agreement where the Goal c-commands the Probe (see e.g. Zeijlstra 2012). A way around this kind of reverse agreement would be to assume that the agreement relation is established earlier in the derivation (with a lower instance of the phrase). This does not induce a locality violation, with the subject acting as an intervener, because the fronted phrase originates in a structurally higher position than the subject. I leave the further analysis of this pattern to future research.

of these distinctions is hierarchically ordered, adverbs and different moods, tenses, modals, aspects and voices being “exponents of the same grammatical notion, merged in the same functional projection of the clause” and thus “the two sides of the same coin” (Cinque 2013:55f.). A frequently adopted alternative is the adjunction approach of adverbs, treating restrictions on the position of the adverbs mainly as arising from the semantics which are adjoined to already existing categories (see e.g. Ernst 2002, 2007). Another interesting proposal is to treat adverbs as belonging to a different domain (or “tier”) entirely which then gets interleaved with the syntactic structure much like combining two decks of cards (Bobaljik 1999). Cinque’s (1999) hierarchy of functional projections is given in (11) below:

- (11) [*frankly* Mood_{speech act} [*fortunately* Mood_{evaluative} [*allegedly* Mood_{evidential} [*probably* Mod_{epistemic} [*once* T(past) [*then* T(future) [*perhaps* Mood_{irrealis} [*necessarily* Mod_{necessity} [*possibly* Mod_{possibility} [*usually* Asp_{habitual} [*again* Asp_{repetitive(I)} [*often* Asp_{frequentative(I)} [*intentionally* Mod_{volitional} [*quickly* Asp_{celerative(I)} [*already* T(anterior) [*no longer* Asp_{terminative} [*still* Asp_{continuative} [*always* Asp_{perfect} [*just* Asp_{retrospective} [*soon* Asp_{proximative} [*briefly* Asp_{durative} [*characteristically* Asp_{generic/progressive} [*almost* Asp_{prospective} [*completely* Asp_{SgCompletive(I)} [*tutto* Asp_{PICompletive} [*well* Voice [*fast/early* Asp_{celerative(II)} [*again* Asp_{repetitive(II)} [*often* Asp_{frequentative(II)} [*completely* Asp_{SgCompletive(II)}]

From Cinque’s perspective, then, adverbs are functional elements, each adverb class belonging to a separate functional projection. The hierarchy is argued to match the attested order of verbal morphemes cross-linguistically and captures the observation that these adverbs (or classes of adverbs) must occur in this order, where adverbs located lower on the hierarchy preceding higher adverbs generally results in ungrammaticality (cf. Cinque 1999, 2004, 2013). Jónsson (2002) offers a critical assessment of Cinque’s hierarchy from the perspective of Icelandic and the hierarchy is shown to match the Icelandic data well in that “adverb orders consistent with Cinque’s hierarchy are always preferred to orders that violate it” (Jónsson 2002:84).¹⁰

As regards negation, I assume with Cinque (1999:124ff.) that it always originates in a NegP which is base-generated on top of each of the adverb-related projections he identifies, but always below ModP_{epistemic}. as epistemic modals such as *probabilmente* (*probably*, *sennilega*, etc.) are never under the scope of negation (Cinque 1999:124). From the cartographic perspective, NegP can thus be considered to be recursive similar to TopP, for which we can correspondingly adopt the asterisk notation (NegP*).

¹⁰Jónsson (2002) does remark that not all violations of the hierarchy result in absolute ungrammaticality, such as evidential > evaluative adverbs (cf. Jónsson 2002:85, his judgments):

- (1) a. *María er sem betur fer greinilega mjög ánægð* (evaluative > evidential)
 Mary is fortunately clearly very happy
 b. ? *María er greinilega sem betur fer mjög ánægð* (evidential > evaluative)
 Mary is clearly fortunately very happy

Nonetheless, the ‘?’-marked examples that Jónsson (2002) provides all sound pretty awkward. At least some of the deviant orders, to the extent that they can be used, would strike me as parenthetical, despite the fact that Jónsson (2002:84, fn. 13) explicitly states that focusing or parenthetical uses were excluded. Note moreover that his use of ‘?’ in his data judgments denotes “marginal” status (Jónsson 2002:86). See Cinque (2004) for some discussion of acceptable adverb orders in an apparent violation of the adverb hierarchy.

The finer details of Cinque’s adverb hierarchy will be relevant only insofar as finite verb placement may target a position in-between two (or more) adverbs, as there are no upper/lower bounds on how far up in the IP domain the finite verb can occur in Icelandic. In the typical case, adverbs thus serve as a diagnostic for the structural position of the verb. If the verb is realised in a position higher than the adverb (=to the left of ADV in linear order), the verb must have left its base position in the VP. Previous approaches such as Thráinsson (2010) and Angantýsson (2011), adopting the adjunction approach of Ernst (2002, 2007), do not treat Cinque’s higher adverbs (including focus adverbs such as *bara* ‘just’) as displacement diagnostics of the verb because these may break up subject-verb adjacency in main clauses; this in contrast to so-called “sentence-medial adverbs” such as negation and adverbs of quantification, corresponding to *never*, *always*, *seldom* and so on, which are taken to be more rigid in that they always occur in second position in main clauses. I try to avoid the different ramifications of these approaches with regard to the nature and attachment sites of these adverbs by focussing on the more homogeneous “sentence-medial” class (see Section 3.3.2).

1.1.3 Nominal structure

Largely corresponding to sentential structure, much previous literature has argued for extended projections of noun phrases (cf. Szabolcsi 1983, Hellan 1986, Abney 1987, Delsing 1988, 1993 and much subsequent work). The identification of separate projections for a variety of categories within this complex similarly yields a hierarchy consisting at least of quantifier > demonstrative > determiner > numeral > adjective > noun, in addition to hierarchical order within some of these categories (for an overview and references, see e.g. Rizzi and Cinque 2016:151ff.). Each level of the hierarchy can thus be conceived of as projecting a phrase:

- (12) QP > DemP > DP > NumP > AP > NP (Extended nominal domain)

A full noun phrase like *the three little pigs* could, then, be analysed as follows:

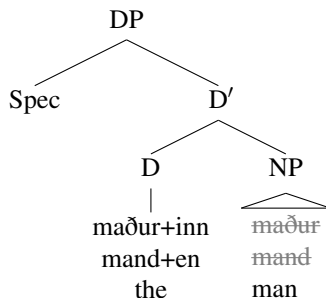
- (13) [_{DP} the [_{NUMP} three [_{AP} little [_{NP} pigs ...

This analysis raises interesting questions concerning which elements can occur within the DP and their relative order, now conceived of as representing structural differences. From an Icelandic perspective, for instance, the analysis also immediately offers a possible transformational account of the interplay between the free-standing article *hinn/hin/hið* in D and the suffixed article *-inn/-in/-ið* involving the merger or incorporation of N to D, the noun occupying the position of the determiner (cf. Hellan 1986, Delsing 1988, Sigurðsson 1993b, 2006, Vangsnes 2003, Pfaff 2015, Ingason 2016).

The analysis of the outer or extended noun phrase as a determiner phrase (DP) is often traced back to Abney (1987), but has precedence in earlier work, e.g. Hellan (1986) on Scandinavian.¹¹ Most analyses of the Scandinavian languages take the free and suffixed articles to occupy the same structural position in the nominal skeleton, the difference being due to merger of N-to-D where the two heads combine:

¹¹On Hellan’s analysis, DP is referred to as T for ‘term’, housing DET, the determiner head.

(14)



The variation found across Scandinavian can thus be represented roughly as follows (a schematic derivation of *the (three) (little) pigs*; based on Sigurðsson 2006:206):¹²

	(ADJ+NOUN) – DEF DET	NUM	ADJ	NOUN(+DEF DET)
Danish	grise – ne			←
Swedish	grisar – na			←
Icelandic	grísir – nir			←
Danish	de		små	grise
Swedish	de		små	grisar – na
(Icelandic	hinir		litlu	grísir)
Icelandic	litlu grísir – nir		←	←
Danish	de	tre	små	grise
Swedish	de	tre	små	grisar – na
(Icelandic	hinir	þrír	litlu	grísir)
Icelandic	litlu grísir – nir	þrír	←	←

Note also that Faroese and Norwegian follow Swedish above such that D+N may merge, resulting in the suffixed article as in Danish and Icelandic, and when an element such as A intervenes, we get the double definiteness pattern shown for Swedish above. As we will see in Section 4.5, 19th-century Icelandic exhibits much greater variation than attested in Modern Icelandic, for instance allowing for the double definiteness pattern (at least marginally, see Section 4.5.1), reminiscent of corresponding structures in Swedish. See further discussion on the analysis of the NP/DP level in Icelandic in Section 4.5.

1.2 Sociolinguistics

1.2.1 The linguistic variable

In the Labovian tradition of sociolinguistics, a fundamental methodological principle is the PRINCIPLE OF ACCOUNTABILITY, which dictates that the analyst must study variation in the contexts of its use in terms of “the largest homogeneous class in which all subclasses vary in the same way” (e.g. Labov 1972a:72). Most sociolinguistic research carried out to date concerns linguistic variability at the phonological level, with numerous potential issues arising when this approach is applied to grammatical

¹²Yet another strategy, not shown in the schema above, is the North-Scandinavian way of fusing together A and N with the suffixed article, i.e. *gamm-biln* (see e.g. Delsing 2003:44).

and syntactic variables (see Section 1.2.4). The variants that belong to this overarching, ‘largest homogeneous class’ are referred to as the LINGUISTIC VARIABLE and the variation attested within it the *envelope of variation* (cf. e.g. Tagliamonte 2012:10). What this principle is supposed to prevent is irrelevant or strongly biased (high- or low-frequency) subclasses distorting the regular pattern which a linguistic variable may otherwise be shown to follow or a variable being defined in a way that is too narrow.

It is necessary to identify invariant subclasses which should be excluded from the analysis. Contexts that categorically select for a particular variant introduce a confounding factor, which, while interesting in itself, especially from a purely grammatical perspective, ought to be treated separately from the sociolinguistic analysis of variation (Tagliamonte 2012:10f.). However, at the same time, it is recommended that the researcher collect data from all potential contexts in which a variable may occur, not only those contexts which follow the researcher’s own intuition or previous discussion in the literature, “to let the data provide the evidence of variability” (Tagliamonte 2012:238) and to identify categoricity in the data afterwards. A further methodological point is that the use of a specific variant should be reported not in isolation, e.g. based on the amount of text, number of words in a corpus, etc., but relative to all other variants that belong to the same variable (cf. Labov 1972a, 2016, Tagliamonte 2012). In his reflection on the history of the field, Labov (2016) refers to the latter situation, where only occurrences of a variant are counted and not the corresponding non-occurrences, as numerical studies. In compliance with the principle of accountability, the non-occurrences must be included as well.

The first case study in the present work, reported on in Chapter 3.4, goes the furthest in strictly applying this methodology. The second and the third case studies, however, deviate more from this ideal (see Sections 4 and 5); the second by admitting more ‘noise’ in the form of potential categoricity where the definite determiner may be a demonstrative;¹³ the third by treating the linguistic phenomenon not as bivariate but as a univariate variable, adopting a different method frequently used in corpus linguistics (see McEnery and Hardie 2012, Rayson 2016 and Chapter 5).

Some variationist studies that adopt mixed-effects statistical modelling (see Section 2.4) have permitted themselves to deviate from these ideals by not removing categoricity in the data beforehand. Instead, by factoring in known aspects of variation as fixed effects, in addition to a random effect for the individual, the statistical model is used as a tool to analyse the distribution regardless of potential categoricity (see Jensen and Christensen 2013). While categoricity in the data might indeed be regarded as a part of variation as well,¹⁴ I maintain that we are better off by separating non-variable contexts from variable ones that have been identified in the linguistic literature, in the context of Jensen and Christensen (2013) by distinguishing between subject-initial contexts and fronting of non-subjects (see Chapter 3), or at least demonstrate that a broader definition of the scope of variation is warranted.

¹³I do offer an evaluation of a subset of the data where this ‘noise’ is removed.

¹⁴I thank Leonie Cornips for valuable discussion on this point.

1.2.2 Social embedding of linguistic variation

The point of departure common to many (historical) sociolinguistic approaches is Weinreich et al.'s (1968) observation that linguistic variation is a matter of “structured heterogeneity” (p. 101) or, more specifically, “an orderly heterogeneous system in which the choice between linguistic alternants carries out social and stylistic functions” (p. 162). Thus, language is not only inherently variable, but also both stable variation and the diffusion of a variant to the whole population (language change) arguably follows social patterns. The issues raised above with regard to the linguistic variables and the standardisation of Icelandic all touch upon aspects (or ‘mechanisms’) of language change identified by Weinreich et al. (1968), viz. *actuation*, *constraints*, *transition*, *embedding*, and *evaluation* of change. These deal with the introduction of a new variant into the system, constraints on possible linguistic changes, the transition from one language stage to the next, how a variant is embedded both within social structure (in terms of use) and incorporated into the linguistic system itself, and, finally, the evaluation of its structural and communicative effects, e.g. in terms of language attitudes and linguistic complexity. An integrated account of language variation and change as envisioned by Weinreich et al. (1968), thus, decidedly addresses both social and linguistic factors.

When observing language variation and change, each phase in the development is considered to have a social aspect (see e.g. Tagliamonte 2012:61f.). At first, the change (or, better, the innovation) is not going to correlate strongly with any particular social group. However, when an innovation starts to spread, reaching a particular threshold in the often mentioned S-shaped curve up to its middle section, this is going to have social consequences; a change that is midway along the curve will start to lose links with factors such as age and social characteristics.

Based on insights from sociolinguistic studies, it is known that language changes usually do not originate in the top layers of society unless they involve changes ‘from above’ (see Section 1.3) or ‘targeted’ changes, i.e. external features of prestige that are borrowed from other varieties (cf. Guy 2011:180, see also Labov 1972a:290 and Section 3.7). In contrast, untargeted changes that really do bring in completely new innovations, i.e. spontaneous changes within the language community which have not been borrowed from elsewhere, tend to arise within the working classes (cf. Guy 2011:180f.).

The focus of sociolinguistics has shifted over the decades, both in terms of theory and in terms of subject matter, often identified by referring to ‘waves’. While these developments need not be regarded as chronologically ordered, they have tended to go from rather broad brush-strokes in traditional, ‘first wave’ Labovian research to ever more intricate identities, towards the individual level of the ‘third wave’ (cf. Eckert 2012). These shifts in perspective are depicted roughly in (16) below (from Tagliamonte 2012:38):

	FIRST WAVE	↔	SECOND WAVE	↔	THIRD WAVE
(16)	<i>Social groups</i>		<i>Social networks</i>		<i>Styles</i>
	<i>Sex, age, education</i>		<i>Communities of practice</i>		<i>Identities</i>
			<i>Jocks, Burnouts</i>		<i>Individuals</i>

Traditional sociolinguistic studies belonging to the first wave tend to be at a broad macro level, typically involving macrosociological categories, variables such as social status or

class, sex, age and education. Later scholarship has regarded these traditional categories, especially with regard to sex/gender, as very problematic (see e.g. Eckert 1989, 2012, Eckert and McConnell-Ginet 2013 and discussion below). Second-wave studies tend to focus on smaller units, dealing with ethnographic factors and communities of practice. Third-wave studies zoom in even further by considering individual differences and the creation of identities through language use. The third wave thus stresses not the broad social structures of traditional Labovian sociolinguistics, where linguistic variation is a reflection of a social category, but the agency of individuals as they construct their identities, shifting the focus away from (or beyond) regional or non-standard linguistic variables. These researchers seek out what makes styles distinctive (cf. Eckert 2012:96) and what sort of ideologically-related fluid meanings, or indexical fields, are conveyed by linguistic variables as they are activated (see Eckert 2008).

The present study is squarely situated in the first wave in most respects. However, a more fine-grained analysis of specific individuals will be attempted where the fragmented nature of the data allows for smaller units to be considered. Moreover, the issue of constructed identities and its potential interaction with social hierarchies and power (see Sections 1.2.3 and 1.3.2) will obviously play a role in the interpretation and analysis of linguistic variation across the social dimensions.¹⁵ The choice to categorise speakers broadly in terms of sex/gender makes it problematic, of course, to *explain* correlations between linguistic and social aspects. At the same time, scholars have also defended the view that a large-scale distinction of speaker sex/gender may be employed as a methodological, exploratory variable: “in other words, it is a purposely broad, unrefined social variable that can be easily taken into account at the data collection stage of research” (cf. Cheshire 2005a:494). The same, *mutatis mutandis*, can be said of broad categories relating to social class/status groups.

As for sex/gender, it is often claimed that females generally show more tendency to conform to conservative linguistic variants in stable environments and also to conform to change when imposed from above, but not to conform when changes arise from below (cf. Labov 2001:366f.). Labov refers to this finding as the GENDER PARADOX, whereby women may lead certain changes when not negatively evaluated socially, while conforming to norms more than men when these are overtly prescribed. He re-states this state of affairs as the CONFORMITY PARADOX, according to which women conform more to overtly prescribed norms than men but women deviate more when the deviations are not proscribed (Labov 2001:367). More recent experimental evidence based on phonological variation also suggests that women are more sensitive to deviations from overt linguistic norms than males in a moderate range of use (10-30%) but sex/gender differences disappear at higher, extreme values (cf. Labov et al. 2011). Various different explanations have been proposed for the sex/gender effect observed in sociolinguistic research, from biological to cultural (for an overview, see Tagliamonte 2012:32ff.).

From the traditional Labovian perspective, then, females are expected to be more faithful practitioners of the standard language than men, all things being equal. However as Labov stresses, this finding comes from sociolinguistic studies which mostly focus

¹⁵In theory, at least, the private letter corpus does permit second-wave applications of the data in that many speakers share ties in terms of family and friendship. I must leave most of these considerations for future research. See Conde-Silvestre (2016) for a recent attempt at a third-wave analysis of Middle English correspondence.

on modern western societies, industrial in nature, and he acknowledges that these results may not apply to more remote societies or periods of history (cf. Labov 2001:437).

1.2.3 Socially constructed identities

The basic tenet of the Labovian view of social patterns of language variation is much criticised in third-wave approaches that try to analyse and deconstruct the traditional sex/gender distinction, often presented as an opposition of sorts when in fact it is a multidimensional complex (cf. Eckert 2012) and actual differences between men and women tend not to be dichotomous but scalar (cf. Eckert and McConnell-Ginet 2013:5). As Eckert (1989) already noted, there are numerous problems that arise on the categorical/binary approach to sex/gender differences that traditional sociolinguistic studies have taken. For one, there is no constant constraint associated with gender; a speaker's sex does not have a uniform effect on linguistic variables, some changes being led by women while others may be led by men or, indeed, by groups that identify neither with males nor females. Similar findings have been reported in historical sociolinguistic research where there have been sex/gender-related differences, females leading certain changes, males others (see Nevalainen and Raumolin-Brunberg 2011).

More specifically, Eckert (1989:246) regards gender as “the social construction of sex”, the biological category of sex giving rise to distinctive roles, norms and expectations in society. From her perspective, traditional sociolinguistic findings with respect to different social patterns between men and women are “a function of gender and only indirectly a function of sex” (Eckert 1989:247). The potential interaction between sex-related differences and power also arises here in that women have to assert their status in society (or in a given group) by symbolic means, leading Eckert to state that there is really nothing which suggests it is not precisely this (i.e. power) that underlies the frequently observed sex differences in language (cf. Eckert 1989:256). Thus, with her criticism of the traditional approach, she invites the researcher not simply to fall back on traditional, unanalysed notions pertaining to sex/gender to account for apparent correlations with speakers' sex that may seem to be grounded in the data.

Another problematic aspect concerning social variables that arises with regard to the present work is that of status-class distinctions and how these (and other) social categories may interact with style. As will be discussed in more detail below, the traditional view of Icelandic language history makes a rather strict dichotomy between the language of the educated class and that of the common, peasant class, as well as the view that particular features only exist as a part of a written literary style associated with the educated elite, which are then taken to be ‘contrived’ in some sense. Such assumptions may seem rather suspect from a sociolinguistic perspective, normally viewing style as “the locus of the individual's internalization of broader social distributions of variation” (Rickford and Eckert 2001:1). Moreover, styles are considered to be hybrids, associated with a multitude of identities, personae and performances that may be a part of each speaker's social and linguistic repertoire, where speakers engage in “ongoing and lifelong projects of self-construction and differentiation” (Eckert 2012:98). There is arguably no stable ‘core’ real self independent of the outside world; instead, selves and styles are constructed, emerging and ever-changing phenomena (see Eckert and McConnell-Ginet 2013:249).

An emphasis on individual identities is somewhat at odds with the widely held view (or even tenet) that sociolinguistics is the study and analysis of variation at the level of the community. Although traditional variationist linguistics may consider individual differences as indicative of socially-determined variation, traditional sociolinguistics typically does not regard individuals as the appropriate level of analysis. Instead, it is argued that individual behaviour can only be understood through the study of social groups, leaving no room for the individual as a linguistic unit (see e.g. Labov 2010:7). This is not to say that the individual is totally meaningless or that the language of a specific person cannot be studied, say, in comparison to the average of the group to which they may be assigned by the researcher. Ideally, social grouping produces an assembly of data points from a set of individual speakers who differ only quantitatively from each other, i.e. only insofar as the rate or frequency of a construction or variant is concerned, whose underlying grammars are otherwise sufficiently similar. On traditional statistical models of variation, the individual, therefore, has typically been excluded by design on the assumption that no (relevant) individual differences exist among speakers (for discussion, see e.g. Walker and Meyerhoff 2013, Johnson 2014).

While the extent to which individual differences matter in (socio)linguistic analyses is debatable, it stands to reason that “the assumption of no individual differences is almost surely false” (Johnson 2014:4). Mixed-effects models, which add a potential random effect for individuals, offer a way to address such issues by making it harder for a statistical model to ascribe an effect to a factor by requiring it to exceed observed individual differences, thus reducing the risk of Type 1 error—a false positive, finding an effect when there is none (see Section 2.4). Due to their relative insensitivity to imbalanced data in comparison to fixed-effect models which ignore the individual, these models have been argued to be particularly useful for natural language data in sociolinguistic and/or (historical) corpus-based studies (see Baayen et al. 2008).

1.2.4 The socio-syntax interface

The socio-syntactic perspective, adopted as a point of departure in the present study (see e.g. Cornips and Corrigan 2005a, Adger and Smith 2005, 2010, Cornips 2014, 2015), assumes that syntactic variation (at least in principle) enters into social semiotic processes, where “a variant functions as a linguistic sign that is indexical of social categories” (Cornips 2014:4). The idea is that these linguistic features come to index social categories that for whatever reason become salient and “[o]nce recognized, that feature can be extracted from its linguistic surroundings and come, on its own, to index membership in that population.” (Eckert 2012:94). From the perspective of the ideological framework of Irvine and Gal (1995), a salient social opposition can thus be activated in a recursive fashion, projecting an opposition such as e.g. “native” vs. “foreign” onto linguistic aspects that are typically non-salient (see also Viðarsson 2016).

The emphasis on syntax in the socio-syntax approach may seem difficult to reconcile both with strictly usage-based approaches such as Construction Grammar, where syntax as such has no obvious place, as well as much previous work in traditional Labovian sociolinguistics, being predominantly phonological in nature. A working assumption of most constructional approaches is that “it’s constructions all the way down” (Goldberg 2006:18), with no separate level of syntax. Traditional sociolinguistics has also main-

tained that social evaluation of linguistic variation is mainly on the surface, targeting lexical and phonological aspects, and does not target abstract linguistic structure (for extensive and critical discussion, see e.g. Romaine 1984, Winford 1984, Cheshire 1987, 1999, 2005a,b, Sankoff 1988, Cheshire et al. 2005, Meyerhoff and Walker 2007, 2013, Grondelaers 2009, Cornips 2015, Levon and Buchstaller 2015, Sneller and Fisher 2015, among many others). A related concept is the differing degree in *salience* of linguistic features (Trudgill 1986). Such potential restrictions of the proposed SOCIOLINGUISTIC MONITOR (see Labov et al. 2006, 2011), a cognitive filter thought to operate on socially marked information as a part of grammatical and phonological processing, are referred to as Labov's INTERFACE PRINCIPLE:

Abstract linguistic structure has little or no social impact on members of the community. The interface of language and society is narrow, and primarily on the surface: the words and sounds of the language. (Labov and Harris 1986:21)

In a similar vein, syntax is situated at the bottom of Laycock & Mühlhäusler's (1990:849) *Degree of Interference Hierarchy*, below all other levels of linguistic representation: vocabulary, derivational morphology, inflectional morphology, phonology, and phonetics. With regard to word order, specifically, Labov's Interface Principle has been re-formulated as the ANTISOCIAL SYNTAX HYPOTHESIS:

Word order is not socially evaluated, unless it can be identified with specific phonological or lexical material. (Ingason et al. 2011, 2013:93)

More recently, however, there has been a growing appreciation for potential social effects even at the level of syntax, a case in question being the relaxing of the verb-second constraint in Norwegian, Swedish, German and Dutch by youngsters in urban areas (see e.g. Freywald et al. 2015, Cornips 2015). This variable is arguably roughly as abstract as the main variable treated in Chapter 3. Moreover, there is substantial evidence of at least partly socially determined variation in Modern Icelandic, for example in the context of passive formation, the New Passive/New Construction (cf. Sigurjónsdóttir and Maling 2001, Þráinsson et al. 2013, Þráinsson, Sigurjónsdóttir, Árnadóttir and Eypórrsson 2015), which is a highly stigmatised innovation that has been referred to as the "Castrated Passive", and also regarding variation in case-marking, known as Dative Substitution or "Dative Sickness" (cf. Svavarsdóttir 1982, Svavarsdóttir et al. 1984, Þráinsson et al. 2013, Þráinsson, Eypórrsson, Svavarsdóttir and Blöndal 2015, Nowenstein 2017).

Importantly, the literature raises numerous issues with simply applying the traditional definition of the linguistic variable to *syntactic* variation. There is a long-standing debate as to what extent the linguistic variable is applicable in this domain, if at all (see e.g. Lavandera 1978, Sankoff and Thibault 1981, Romaine 1982:31-37, Romaine 1984, Winford 1984, Cheshire 1987, Sankoff 1988, Coveney 2002, Cheshire et al. 2005, Grondelaers 2009, Tagliamonte and D'Arcy 2009, Terkourafi 2011, Nevalainen and Raumolin-Brunberg 2011, Jensen and Christensen 2013, Hasty 2014). Cheshire (1987:264f.) proposes a continuum with morpho-phonemics and syntax, respectively, at the most extreme end of the continuum, with morphology situated in-between. She further suggests that relatively clear-cut instances of syntactic variation include (i) variation involving more than one item in a construction (e.g. *he never even went down* vs. *he didn't even go down*), (ii) variation involving a change in word order, and (iii)

variation among constructions such as agentless passive/active constructions (cf. also Weiner and Labov 1983). Based on such distinctions, variation involving the *get*-passive (see Sneller and Fisher 2015), in contrast, will obviously be considered much ‘less syntactic’, leaning towards the lexical/phonological. Nonetheless, much research into syntactic variation has revealed that it, too, is determined to a large extent by social and stylistic factors (see e.g. Romaine 1982:206f., 167-170, Grondelaers 2009, Tagliamonte 2012:207, Levon and Buchstaller 2015, Sneller and Fisher 2015).

A general problem that was acknowledged early on (see e.g. Lavandera 1978, Sankoff and Thibault 1981, Sankoff 1982, 1988, Romaine 1984) is that when it comes to syntactic variants, it is almost always possible to find “some usages or contexts in which they have different meanings, or functions” (Sankoff 1988:153). In what sense can the syntactic variable be straightforwardly defined as constituting alternative ways of saying the same thing and what sort of equivalence of meaning does that refer to? Problems may even arise with lexical choices, leading Tagliamonte and D’Arcy (2009:74) to state that “[n]o two forms can have identical meaning”. In fact, language change is not necessarily expected to arise from distinct linguistic forms with an identical meaning (see also Tagliamonte 2012:16-19, with references).¹⁶ Since emphasis does not affect truth conditions, Cheshire et al. (2005:161, adapted) argue that a superficial analysis in terms of simple alternation of forms *can* be carried out and Jensen and Christensen (2013:40) also claim that variants that differ on one semantic level, such as information status, may be compared if comparable on another.

If we follow socio-syntactic approaches in assuming that syntax exhibits social effects, then we need a way to implement said effects formally as a part of the grammatical architecture. Here, not only potential cognitive limitations in the recognition of abstract linguistic structure arise, but also the way in which the formal representation of optionality with regard to the choice between two (or more) variants is accounted for in grammar—or indeed in grammars.

As Adger (2007:700) argues, the structure-building operations of syntax are blind to facts of usage, including social facts: “syntax is Socio-free and Use-free” (see also Adger and Smith 2010, Adger 2014). These aspects, then, are considered to be external to syntax and may affect selectional procedures when choosing among variants generatable by the speakers’ grammatical systems (the “Pool of Variants”). Social forces thus operate over and above (non-)categorical restrictions that arise in the grammatical structuring of the variation space. One and the same grammatical form may be the spell-out of more than one feature bundle, serving to raise the frequency of that particular variant, in competition with a less specified one (see Adger 2007, 2014 on the competition between the relativisers *that* and *who* in English). As a variant becomes increasingly disfavoured over generations of speakers, a variable grammatical distinction may eventually become a categorical category of grammar. While Adger (2006, 2007, 2014) and Adger and Smith (2005, 2010) treat these differences in terms of variable spell-out of feature bundles, i.e. variable grammars, they can also be formulated in terms of competition

¹⁶Similarly, Sankoff (1988:153ff.) specifically warns against the common suggestion among scholars that a different meaning underlies the use of a variant and maintains that these differences may often be neutralised in discourse—that is, they are not always intended by the speaker or interpreted by the interlocutor. They may be “nothing but *a posteriori* artifacts of linguistic introspection or afterthoughts inspired by linguistic norms.” (Sankoff 1988:154, see also Poplack and Dion 2009).

among two or more invariable grammars leading to variable output (cf. Kroch 1989, Yang 2010, Wallenberg 2016).

While the precise nature of syntactic variants subject(able) to social variation is still an open issue, it seems clear that variation across the social dimension was not meant to target extremely low levels of grammatical representation such as the combinatoric mechanism of syntax itself or the way atomic elements of structure interact.¹⁷ Rather, we are typically either dealing with variation bordering on the lexical or complex syntactic objects that can in some way be regarded as representing larger pieces of structure (or ‘constructions’)—thus, closer to ‘the surface’ as in the quote above.

I assume here a theoretical model of syntax that accommodates constructions as derivational LAYERS (cf. Zwart 2009, Trotzke and Zwart 2014), but departs from the constructional perspective of Construction Grammar in that syntax is taken to be elementary—more in line with the premise *syntactic hierarchical structure all the way down* (cf. Halle and Marantz 1994:276). Derivational layers are a way to syntactically implement the human cognitive capacity of (sub)categorisation which allows us to treat objects of the world as both complex and atomic (for discussion, see Hofstadter and Sander 2013). Layers allow complex elements in syntax to function as a single constituent, such as when a whole clause, a verb cluster or a complex noun phrase occurs in initial position, immediately followed by a finite verb in a verb-second language. The initial element is then opaque in terms of surface syntax, present in the numeration underlying the derivation as a monolithic item and yet it is clearly complex.

The upshot of such an approach is also that by admitting constructions (layered derivations) into the syntactic representation, we can be more precise about what it means for *syntax* to be socially evaluated, avoiding much contradiction with the widely held view that speakers allegedly lack conscious awareness of abstract syntactic operations, or, as suggested by Meyerhoff and Walker (2013), that “fundamentally grammatical” variables are in any case not as likely vary along social dimensions as lexical variables. It has, moreover, been suggested that social evaluation may be closer to the surface in another respect, viz. by not necessarily targeting such an abstract a notion as the *linguistic variable* but rather the *variant* itself (for discussion, see Maddeaux and Dinkin 2017). From the above perspective on syntactic derivations, a piece of linguistic structure (a derivational layer or a phase) can thus arguably be subjected to social evaluation by virtue of our cognitive ability to regard the complex as simple, an analysis which is largely consistent with the fact that speakers will tend to be oblivious to the combinatorial processes themselves giving rise to that structure.

The competing grammar approach accounts for both inter-speaker variation across individuals as well as intra-speaker variation within one and the same individual (see e.g. Kroch 1989, Yang 2000, 2010, Nowenstein 2017). This is what Weinreich et al. (1968) refer to as the TRANSITION PROBLEM: “the route by which a linguistic change is proceeding to completion” (1968:153) and “the transference of a linguistic form or rule

¹⁷Emphasising this point rather forcefully, Bickerton (2014) observes:

Although syntax is often regarded as part of cognition, its operations are automatic and out of reach of conscious awareness. We are no more aware of how our brains construct sentences than we are of how our stomachs digest food or our hearts circulate blood. (Bickerton 2014:74).

Writing some 150 years earlier, M. Müller (1861:36) similarly mentions our inability to control our blood circulation or adding an inch to our height as an analogy in the context of deliberate grammatical change.

from one person to another—more specifically, from one linguistic system to another” (1968:155). In my view, the grammar competition model neatly captures the idea of apparent optionality in syntax, avoiding formal optionality *within a single grammar*. As emphasised by Weinreich et al. (1968:185), the implementation of variability in use need not be regarded as being external to the system as the “control of such variation is a part of the linguistic competence of members of the speech community”. The notion of *control* is thus implemented in terms of a choice between grammars that are a part of a speaker’s linguistic repertoire.

Yang’s (2000, 2010) seminal evolutionary implementation of the competing grammars approach incorporates it with probabilistic population dynamics, viewing language acquisition “as a variational process in which the distribution of grammars changes as an adaptive response to the linguistic evidence in the environment” (Yang 2000:234). From this perspective, language acquisition is about selecting among “principled hypotheses of language” on the basis of fitness measures, where the learner assigns penalty probabilities to hypothesised grammars as they fail to match the observed input. This is formulated in terms of grammar advantage. As Yang (2000:235) explicitly states, it is possible to estimate fitness measures from (historical) texts. Grammar fitness is determined on the basis of the probability weight of unambiguous cases and the advantage value of Grammar 2 over Grammar 1 is arrived at by subtracting the fitness of Grammar 1 from Grammar 2 (see Heycock and Wallenberg 2013:136). A higher fitness value for a particular grammar will eventually result in change, all other things being equal. An insightful cognitive principle of child acquisition, the PRINCIPLE OF CONTRAST (see Clark 1987), captures the fact that grammars in competition are diachronically unstable (see Kroch 1994, Wallenberg 2016). According to this principle, a difference in form is always (expected to be) associated with a difference in meaning, where meaning is interpreted broadly so as to encompass apparent cases of synonyms (of words, structure etc.) involving dialect, register and style variation. As a result, if two forms enter into competition, the learner has an inherent bias to allot to one a different function than to the other (for further discussion, see Wallenberg 2016:e244f. and references therein).

I will return to the grammar competition approach in more detail in Section 3.7, where I will adopt such a model to capture variation attested in the 19th-century Icelandic data. Let us now move away from sociolinguistics with regard to syntactic variation in particular and move on to issues of conscious language planning.

1.3 Standardisation

1.3.1 Evaluating standardisation effects

At the heart of the present approach lies the fundamental distinction of the Labovian tradition between changes ‘from above’ vs. changes ‘from below’, representing elements imported from other systems as opposed to language-internal changes (Labov 1965, 1994, 2001, 2007). These terms are further defined below:

Changes from above are introduced by the dominant social class, often with full public awareness. Normally, they represent borrowings from other

speech communities that have higher prestige in the view of the dominant class. Such borrowings do not immediately affect the vernacular patterns of the dominant class or other social classes, but appear primarily in careful speech, reflecting a superimposed dialect learned after the vernacular is acquired.

Changes from below are systematic changes that appear first in the vernacular, and represent the operation of internal, linguistic factors. At the outset, and through most of their development, they are completely below the level of social awareness. [...] It is only when the changes are nearing completion that members of the community become aware of them. (Labov 1994:78)

Changes from above may also involve stigmatisation, possible “[o]nce a linguistic feature has risen to a sufficiently high level of social awareness” (Labov 2010:186), leading in the context of language standardisation to the rejection of variants in the standard variety. The above vs. below distinction is reprojected onto the textual dimension in the critical approach to language historiography initiated by Elspaß (2005b), correspondingly taking a view of language ‘from above’ and ‘from below’. This conceptual framework has been applied in much subsequent work on language histories, standardisation and norms vs. usage within historical sociolinguistics (see Elspaß et al. 2007, Nobels 2013, Rutten et al. 2014, Auer et al. 2015, Krogull 2018, Rutten 2019).

The notion of *standard language norms*, as the term is used here, must not be confused with the more general term *norms of usage*—even though the two may certainly be “intrinsically connected” as argued, for instance, by Árnason (2003a:245) in an Icelandic context. In the words of Elspaß (2014:303), “[n]orms of usage are inherent to all natural languages.” The culture-cultivational aspect and subsequent symbolic value attached to standard language norms and not (or not necessarily) to norms of usage more generally is what quite sharply distinguishes the two processes. In other words, standard norms are really seen here as a product of nationalism and a part of the ideology of national thought (cf. Leerssen 2006a), in contrast to norms of usage (see also Rutten 2016). It is very doubtful that a strong emphasis on minimal variation in form and maximal variation in function (cf. Haugen 1966) is ever a part of norms of usage in the broader sense. This very aspect is stressed for instance by Haugen (1966) through the notion of the symbolic value of standard norms with material rewards, further elaborated by Bourdieu (1973, 1977, 1991) in his ‘linguistic marketplace’ metaphor, which rests upon the existence of codified prescriptive norms as implemented by norm authorities (see also Angantýsson 2017b and Viðarsson 2017b for discussion in an Icelandic context).

Traditional language histories tend to focus on the adoption of standard norms in public, printed texts, typically of a relatively formal nature and written, moreover, by professional writers. Basing one’s research purely on this sort of material—language history ‘from above’—is disregarding the potentially much more varied nature of textual sources such as private ego-documents written by non-professional writers of the middle and lower ranks of society, the focus of the alternate view ‘from below’ (Elspaß 2012:160f.). It should be stressed, however, that it would be too simplistic to suggest that the lower ranks are somehow categorically different from the higher ranks in terms of the adoption of linguistic norms simply due to the lower ranks’ *unfamiliarity*

with school grammar. For the working class, non-standard language use may signal working-class identity or even an oppositional stance to institutional power, including schools (see Eckert and McConnell-Ginet 2013:255).

The language history ‘from below’ approach has led to the problematisation of the basis on which traditional narratives of standardisation and standardisation effects rests (cf. e.g. Ottósson 1990), calling for a re-evaluation of previously made claims. From the present perspective, to adequately address the question of whether it is possible to reverse linguistic changes (cf. above), public/formal as well as private/informal spheres have to be taken into account. Language historiographies that mainly focus on the propagation of standard norms in printed texts simply fail to consider language use beyond that of a small group of elite speakers. At worst, these traditional narratives may be regarded as **histories of prescription rather than praxis** or even “hallucinations” (see Elspaß 2014:310, with references).

The emphasis on private letters is related directly to the spoken-written continuum and differences in communicative distance as outlined in the work of Koch and Oesterreicher (1985, 2007). Figure 1.1 depicts a variety of different means of communication ranging from (a) spontaneous speech among friends to (i) a legal text. The horizontal axis situates these different modes as belonging to the spoken modality (left) or the written modality (right). The vertical split further categorises these modes in terms of whether the basic units can be conceived of as consisting primarily of graphemes or phonemes, the former aligned with the visual/written modality, the latter with the phonological/spoken modality. From their perspective, colloquial private letters (cf. (c) on the continuum) come closest to spontaneous speech among the grapheme-oriented, written text types. A journal editorial, in contrast, is considered to be primarily written, second only to legal texts on the horizontal axis. Different types of communication may also vary in their degree of communicative distance depending on properties such as private vs. public spheres, referential closeness vs. distance, a high degree of spontaneity vs. a high level of reflection (cf. Koch and Oesterreicher 2007:351).

Against this conceptual background, we can broadly split potential measures of success of language planning efforts as follows (cf. Elspaß 2016:[19]):

- (17) a. **SUCCESSFUL**
Change in formal registers of printed text sources from the elite (‘texts from above’) + in informal, private, registers of non-professional writers (‘texts from below’)
- b. **PARTLY SUCCESSFUL**
Change in formal registers of printed text sources from the elite (‘texts from above’) only
- c. **NOT SUCCESSFUL**
Only “temporary slow down” or no change at all

Traditional reference works tend to note that it is by no means an easy task to measure the effects of language planning (see e.g. Cooper 1989:185, Kaplan and Baldauf 1997:57, Kristinsson 2007). Ambitious attempts to do so reveal mixed results (cf. Auer 2006, 2009, 2014, Anderwald 2014a, Hinrichs et al. 2015). Moreover, it has been suggested that standardisation is more likely to successfully affect attitudes towards language than

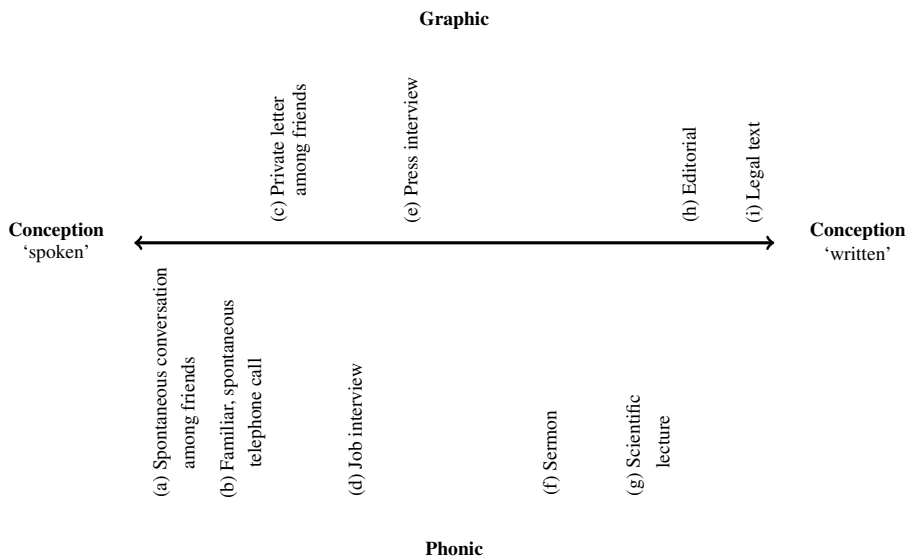


Figure 1.1. Forms of communication on the conceptual continuum (cf. Koch and Oesterreicher 2007:349)

actual behaviour (Cooper 1989:134f., 184). Based on extensive corpus-based research into the effects of prescriptivism in the history of English, Anderwald’s (2014a, 2014b, 2016) studies have revealed little if any lasting effect once a range of different types of data are considered, with larger effects confined to formal, highly edited material such as newspapers. Changes in formal texts ‘from above’ is presumably also what Haugen (1987) is mainly referring to when speaking of the successful recodification of Icelandic, giving the impression that it was a relatively simple case (see below). The problem remains that it is typically near impossible to determine what had happened in the absence of language planning efforts.

1.3.2 Symbolic value and power

Haugen’s (1966) classic model of standardisation identifies four components, the last of which being the one most immediately relevant here: (1) selection of norm, (2) codification of form, (3) elaboration of function and (4) acceptance by society, forming the following matrix (cf. Haugen 1966:933):

		FORM	FUNCTION
(18)	SOCIETY	Selection	Acceptance
	LANGUAGE	Codification	Elaboration

The processes of codification and elaboration are seen as ideal goals of standardisation, striving towards **minimal variation in form** and **maximal variation in function**, respectively (cf. Haugen 1966:931). It has been suggested that linguistic prescription be

regarded as a final, separate stage in this process, as in Milroy and Milroy's ([1985] 2012) revised Haugen-based model (cf. also Percy and van Ostade 2016:3), although these stages or paths need not necessarily be chronologically ordered per se and may work simultaneously. However, in the sense that codification is based on the premise of minimal variation in form and also that it need not be purely *descriptive* in the first place, a separate prescription stage is not always obviously called for. In fact, whether or not standardisation is expected to lead to prescription is arguably an empirical matter that should be assessed for each linguistic feature (see also Auer 2009:189, n. 14). From Haugen's perspective, the contrast between (17a) and (17b) above is to a large extent due to the higher or lower symbolic value of the codified standard language for different groups of speakers based on their social status.¹⁸ In his later work, Haugen (1987) presents a revised model where acceptance of norms at the societal level is subsumed under *implementation*, which is chiefly concerned with educational spread and covers a) correction procedures and b) evaluation (1987:64).¹⁹

Bourdieu's (1977, 1991) more elaborate notions of *symbolic power* and *symbolic capital* associated with employing the standard norms, or 'choices' of the linguistic habitus (cf. Bourdieu 1991:51), are also immediately relevant in this context. In the modern era, this symbolic value of the normalised linguistic product is shaped (or symbolised) by the structures of the linguistic market (cf. Bourdieu 1991:37ff., 46ff.) and the educational system significantly contributes to the ratification of the standard language in presenting it "as the only legitimate one, by the mere fact of inculcating it" (Bourdieu 1991:60). This approach thus neatly captures the change where the linguistic norms themselves have become status symbols, carrying symbolic capital, whereas prior writing conventions were arguably more fluid, allowing for variants to coexist (see also Vandenbussche 2007:33, Auer and Voeste 2012:259, with references). From our 19th-century Icelandic perspective, this value system would be an emerging property.

Moreover, Bourdieu (1991) claims that the linguistic habitus of the dominant class, in particular of those born into that class, is actually the *realisation of the norm* (1991:83). As regards the situation in 19th-century Iceland, again, it could be argued that the linguistic habitus or community patterns in learned circles during the early phases of the implementation stage would clash with the emerging standard due to the selection process. According to the traditional Icelandic narrative, at least, the selection of norms was based on rural varieties associated mainly with the working/peasants class and further legitimised with reference to medieval, Old Icelandic manuscripts. If taken at face value, this leads to an anti-Bourdiesian expectation that it is decidedly not the individuals belonging to the dominant class who could "express all the self-confidence that is associated with a situation where the principles of evaluation and the principles of production coincide perfectly" (Bourdieu 1991:83f.), but it would actually have been those individuals who had to make most effort to adapt.

Returning to the choice of linguistic data in the current study, the three main types of corpora used differ enormously in their level of normativity. Ego-documents such

¹⁸More specifically, the adoption of the standard carries with it a higher value that is "primarily symbolic" for the group of speakers it may admit to positions of power ("material rewards"), or else "the inducement to learn it, except perhaps passively, may be very low" (Haugen 1966:933).

¹⁹Furthermore, the societal tier in the revised model represents *status planning*, the language tier represents *corpus planning* and the form and function tiers represent *policy planning* and *cultivation*, respectively.

as private letters are unedited and in terms of language use represent closeness rather than distance, being “texts which are as close to actual speech as possible” (Elspeß 2007:153). Newspapers, in contrast, are edited, public texts which are further removed from the spoken language and tend to conform to prescribed norms. Student essays are expected to represent an even more strictly normative type of texts due to the fact that they were graded, forming a part of the students’ graduation score. The incentive for students to adopt the prescribed norms is thus particularly great, resulting in other words in an “inflation” of the symbolic value associated with each variant (see also Viðarsson 2017b). Arguably, these assignments, therefore, constitute quite a separate linguistic market from Bourdieuan perspective, which need not go entirely hand in hand with what might be referred to as the “common market” of linguistic variants that we are more likely to find in traditional published works, newspapers and the like.

But what is the scope and limits of these prescriptive norms in praxis? Is it at all possible to manipulate the linguistic habitus from above?

1.3.3 Deliberate change

[L]anguage cannot be changed or moulded by the taste, the fancy, or genius of man.
(M. Müller 1861:39)

It would be childish to try to phrase everything the way Snorri [Sturluson] or the author of *Njála* would have put it, or the best of the other writers of ours during the Icelandic antiquity; we would not succeed even if we tried[.]

(H.K. Friðriksson 1871:107, my transl.)

As Thomason (2007) observes, there is a long tradition in historical linguistics, going back at least to M. Müller (1861), that it is impossible for speakers to deliberately alter structural aspects of language or the direction of language change. Thus, M. Müller (1861:36) states that “although there is a continuous change in language, it is not in the power of man either to produce or to prevent it.” He goes even further, writing: “Try to alter the smallest rule of English, and you will find that it is physically impossible.” (M. Müller 1861:37f.).²⁰ As Thomason (2007) points out, Chomsky and Halle (1968) also assumed limited user agency in terms of change, albeit more permissive than Müller’s view, when they refer to “the adult’s inability to modify his grammar except by the addition or elimination of a few rules” (Chomsky and Halle 1968:251, fn. 3). The same view is expressed by Bennis et al. (2004:48), stating that language intervention seems to have little or no effect, regardless of whether it aims towards consciously halting change or consciously propagating change. An opposing view is presented by various scholars, suggesting that deliberate change is certainly possible, at least given the right circumstances (cf. e.g. Ottósson 1987, Jahr 1989, 2014, Árnason 2003b, van der Sijs 2004, Thomason 2007, Hinrichs et al. 2015).

In the present context, great reservations are made in the work of Labov as regards social evaluation and stigmatisation of features other than rather superficial phonological

²⁰M. Müller (1861:38) further suggested that if some change were to take place, “it will not be by the will of any individual, nor by the mutual agreement of any large number of men, but rather in spite of exertions of grammarians and academies.” The second, revised edition of M. Müller (1862:49f.) makes this even more explicit, referring to the individual, as such, as powerless.

and lexical ones (cf. discussion above). These reservations are not shared by Ottósson (1987) and Árnason (2003b), for instance, although both hedge the significance of their findings, at least provisionally. Ottósson (1987) concludes that deliberate change may be possible so long system-defining properties are not affected, although he voices doubt if there exist any clear-cut boundaries to what sorts of effects can be achieved. Having reviewed Icelandic language planning in terms of syntax, morphology and phonology, Árnason (2003b) shares Ottósson's doubts and points out that the successful eradication of the highly stigmatised innovation of *flámaeli* 'slack jawed speech', involving a merger of /i/ and /u/ with /e/ and /ö/, respectively, actually targeted system-defining properties, such that the phonemic merger was reversed.

Interestingly, from this perspective, Árnason (2003b:202f.) regards syntax "in general to be rather "manageable" and easily influenced by ideas of proper usage", an assertion that is made mainly on the basis of a morphosyntactic variation claimed to be insignificant enough structurally to make it well suited as a shibboleth—the so-called "Dative Sickness".²¹ Árnason's (2003b) overall conclusion is that, provided the fact that each individual plans his/her linguistic behaviour (perhaps unconsciously) in one way or another, and "given the right social circumstances, [...] it seems that almost anything can happen." (Árnason 2003b:214). Essentially the same view is shared by Jahr (1989, 2014), emphasising the role of the dominant contemporary ideology and powerful political backing (see also Wright 2004:48).

Thomason (2007:41) explicitly defends the view that "adult speakers can and do make deliberate choices that bring about nontrivial lexical and structural linguistic change". The main evidence for this comes from language contact (cf. also Thomason 2001:84f.), but also the phenomenon of deliberate non-change (resistance or refusal to change), groups seeking a language to symbolise a new identity, as well as the vast effect or "the zeal" of language standardisers in terms of language planning (cf. Thomason 2007:53; 2001:149f.). Thus, Thomason (2001:149) claims that while the effect of relatively isolated prescriptive remarks and invented grammarians' rules has turned out to be limited, "the same cannot be said of everyone who tinkers with a standard language." The same point is emphasised by van der Sijs (2004), who also defends the view that the standardisation of Dutch led to various deliberate changes at the phonological and grammatical levels, categorically rejecting the view advocated by Bennis et al. (2004), mentioned above. However, van der Sijs's results also suggest that the effects on grammatical aspects were either temporary and limited to the higher echelons or mainly targeted fixed expressions (cf. e.g. van der Sijs 2004:544).

²¹The innovation, also referred to as Dative Substitution, typically involves the spread of dative case to former accusative oblique subjects of psych verbs such as *langa* 'to want' and *vanta* 'to need, to lack'. This innovation is considered a 'language blemish' already in Jónsson (1900). Research conducted by Svavarsdóttir (1982), Svavarsdóttir et al. (1984) and Jónsson and Eyþórsson (2003) on students reveals that the use of the dative is not in remission, but that there is a correlation with school performance and the educational background of parents with regard to the internalisation of the prescriptive rules for assigning accusative case to these subjects.

Similarly, the use of another highly stigmatised syntactic innovation involving the passive, initially referred to as the "Castrated Passive" (*geld þolmynd*), better known as the New Construction or the New Passive (*nýja setningagerðin/nýja þolmyndin*), has been shown to exhibit geographical patterns suggestive of differences in social rank and/or socioeconomic status (cf. Sigurjónsdóttir and Maling 2001). Both these innovations clearly attest to the fact that syntactic variation may be socially evaluated and may participate in the social embedding of language variation and change.

Recent studies have tried to assess the scope and limits of Dutch top-down standardisation in a range of corpora, from private letters, diaries and travelogues to newspaper texts, incorporating a view ‘from above’ and ‘from below’ (see e.g. Rutten et al. 2014, Krogull et al. 2017, Krogull 2018, Rutten 2019). Krogull (2018) studied the influence of a state-commissioned Dutch orthography and Dutch grammar, published in the early 19th-century, focussing on five orthographic and three morphosyntactic features in three distinct types of corpora. The grammatical variables were shown to exhibit much more variation across all three corpora than the orthographic ones, suggesting that morphosyntactic features were harder to implement. The features involved neuter relativiser strategies, masculine and feminine sg./pl. relativiser strategies and the genitive case, normative effects being more successful in the realm of relativisers than with regard to the genitive. There was moreover a cline from a relatively weak effect in the private letters to a stronger effect in the newspaper texts, where diaries and travelogues were shown to occupy an intermediate position, being more uniform than the private letters.

Estonian is also commonly cited with regard to successful language planning but at a relatively abstract level. According to Ehala’s (1998) study, Estonian syntax was successfully changed from the basic order SOV, that was considered to be German, to SVO. The effects of this choice can be observed by a sudden change in the basic embedded word order in newspapers in a couple of decades in the early 20th century until 1940. It may be relevant that the SVO pattern had already existed alongside the SOV pattern, albeit to a limited extent. However, compare this piece of evidence, for instance, to van der Sijs (2004:442-445) on the failed re-implementation of a case system in the history of Standard Dutch and Poplack and Dion (2009) on prescribed future temporal reference in French.

In the case of Dutch, the proposed system, according to most accounts, was invented rather than based on existing dialect forms, although more recently this claim has been questioned on the basis of late 16th-century ego-documents from the northern province of Holland (cf. Hendriks 2012); on the contrary, these documents reveal that 16th-century grammar codifiers were actually *describing* existing case systems, as opposed to prescribing them.²² However, clearly by the 19th century, at least, the sort of case-marking phenomena attempted as a part of the written language was merely artificial and cost a great deal of effort, giving rise to opposition during the mid-19th century (cf. van der Sijs 2004:444). On the basis of (spoken) French, Poplack and Dion (2009) and Poplack et al. (2015) suggest that in cases where community patterns are far removed from the standard, the cost for the speaker to align with the standard is too great. (Elspaß 2014:317) also suggests that although it may appear as if norms of usage were established through normative prescription, in many cases grammarians presumably formulated a *descriptive* norm based on an observed tendency at their time.

In an English context, Anderwald (2014b) specifically warns against drawing conclusions based on seemingly ‘simple’ cases and calls traditional accounts into question that fail to engage with the topic empirically: “simple stories may turn out to be quite complicated, once we try to empirically substantiate (or, indeed, deflate) them.” (Anderwald 2014b:436). Haugen (1966:932), himself, already warns that “[w]here a new norm

²²Hendriks (2012) furthermore argues that there is a correlation between rapid growth (doubling, tripling) in the population size of Dutch cities, including dialects in the southern provinces of the Low Countries (now Belgium), and subsequent morphological case loss.

is to be established, the problem will be as complex as the sociolinguistic structure of the people involved.” (cf. also Auer 2009:1). Indeed, this is what actual empirical research into the implementation of the norms has revealed, as shown e.g. by Heimisdóttir’s (2008) study (see Section 1.3.4). An attempt at a problematisation of these aspects is thus a major contribution of the present work.

1.3.4 The implementation of the standard

Standardisation is in part a fiction. We have imagined languages in the same way that we have imagined communities. (Wright 2004:53)

The emergence of a national standard language in 19th-century Iceland, and elsewhere in Europe, was part and parcel of the advent of nationalistic thought and the cultivation of culture (see Leerssen 1999, 2006a,b, 2008). In the context of language and (standard) language ideology, these initiatives ranged from grammar-writing to language purism and language revivalism to language planning (cf. Leerssen 2006b:569). All of these elements are present in the well-known narrative of the standardisation of Icelandic, which up until quite recently has gone relatively unquestioned—at least in comparison to corresponding historical narratives in fields such as social and political historiography and history writing in general (see e.g. Hálfðanarson 1993, 1995, 2001, 2005).

According to a widely cited narrative, the emergence of a national standard language in 19th-century Iceland resulted in standard norms that sought inspiration, not to mention justification, in Old Norse-Icelandic. These linguistic norms were closely associated with Icelandic medieval literature, in particular the Icelandic sagas, where standardisation effectively led to “halting and reversing the changes in Icelandic” (Kusters 2003:184). The standard norms were selected and implemented already in the 19th century through the education system and it is the simplicity of the social structure and educational institutions at the time, with only one secondary school, that tends to be emphasised in this context (cf. e.g. Haugen 1987, Kusters 2003). What is often overlooked, however, is the fact that establishing a standard norm, even if successfully implemented in particular circles and/or carefully edited texts, does not automatically entail the adoption of said norms across all layers of the society. As argued by Leonard and Árnason (2011):

The form of the modern Icelandic ideal standard has been clearly defined: it is “pure Icelandic” which is effectively the language of the sagas. When it comes to defining “non-standard usage”, the myth has been that there is no such thing. (Leonard and Árnason 2011:94)

The standardisation of Icelandic in the 19th century onwards is widely regarded to have been far-reaching both in scope and effect, targeting most (if not all) linguistic levels—from vocabulary to phonology, morphology and even syntax (for discussion and overview, see Ottósson 1987, 1990, 2003, Sigmundsson 1990-1991, Sigtryggsson 2003, Árnason 2003a,b, Hilmarrson-Dunn and Kristinsson 2010, Viðarsson 2014, 2016, 2017b, Bernharðsson 2017, 2018). The general view on the standardisation of Icelandic thus implies extensive conscious tampering with the language even at relatively abstract levels. While the received wisdom among linguists has been “that prescriptive dicta generally have no effect” (Anderwald 2014b:435), the line of reasoning on traditional accounts of Icelandic language purism/standardisation suggests otherwise. Thomason

(2001:9), for instance, writes that “the creators of Standard Icelandic deliberately archaized the language’s structure, making it look older so as to bring it closer to the language of the Eddas” (cf. also Thomason 1999, 2007 for analogous remarks). A similarly bold statement is found in Haugen (1987:74), who claims that Icelandic was “recodified from the half-Danicised language of the 1584 Bible by reference to the classic models from Old Icelandic” (cf. also Sigmundsson 1990-1991).

It is beyond the scope of this study to detail the potential aspects of grammar that were allegedly ‘tampered’ with in this way, i.e. where there was arguably a mismatch between vernacular and standard norms, as a part of the implementation of the national standard language. Common examples include the supposed revival of a variety of morphological declension classes (e.g. *ia*-stems discussed below, but also kinship nouns), the formation and structure of the *st*-middle, reinstating a plural:dual distinction in *wh*-pronouns (*hver:hvor*, as opposed to underspecified *hvur*, *hvör*, *hver*), the adoption of *hin* as opposed to *sá* as a free-standing definite article, and particularly in the 20th century, the suppression of a phonological innovation involving the simplification or merger of /i,e/ on the one hand and /u,ö/ on the other, referred to as ‘skewed speech’ (*flámaeli*), to name but a few (see especially Ottósson 1990).

When measuring the uptake of prescriptive dicta or standard norms, a difference should arguably be made between relatively isolated prescriptive remarks in metalinguistic discourse as opposed to the prescription that is a part of carefully planned and orchestrated standardisation efforts. The latter are typically backed up by various state-funded institutions, taught in schools, adopted to a greater or lesser extent by the media and so on, and typically rest on nationalistic sentiments and ideologies. Indeed, as scholars have tried to argue, language planning may have little to no limits under such circumstances (see Jahr 1989, 2014, Árnason 2003b). This is not necessarily the case with prescriptive dicta more generally speaking.

As regards standardisation, previous scholarship points to two aspects in particular that are considered to have been of importance in the case of Icelandic: (a) **small-scale** norm implementation initially, reaching a limited, privileged group from where the norms could spread to the rest of the population; (b) the **rural basis** of the norms, which meant that the standard language was already closely aligned with the vernacular of the vast majority of speakers: “In essence, what the majority of Icelanders did was accepting their own linguistic standard” (Friðriksson 2008:99, see also Sigmundsson 1990-1991). Both these points are problematic for a number of reasons, as we will see.

Let us begin with the former point. A striking aspect in this regard is the implicit assumption of uniformity in accounts of the role of the educational system. So Haugen (1987), for example, emphasises the simplicity of implementing norms in a relatively homogeneous, small-scale society like (19th-century) Iceland:

As long as a small, elite group has a monopoly on education, *it is relatively simple to implement a given norm*. But the spread of schooling to entire populations in modern times has made the implementation of norms a major educational issue. [...] The range of heterogeneity from a simple Iceland to a complex Nigeria is vast and disturbing. (Haugen 1987:61, emphasis added)

The fact that a small number of people were engaged in developing standard norms does not entail that these individuals all spoke in one voice. On the contrary, while an

emerging puristic but rather tempered standard was already forming in the 1830s to 1840s, ascribed mainly to the intellectual society and the journal *Fjölur*, who were “the childhood disease of excessive archaism” (Ottósson 1990:75, my transl.) that had dominated the earlier but rather unorchestrated attempts, there were still loud voices of criticism deep into the 19th century (see Ottósson 1990:77f. for some examples). Icelandic orthography is one such aspect, with vastly different spelling adopted and implemented by the principal Icelandic teacher at the Reykjavík Grammar School, Halldór Kr. Friðriksson, as opposed to the proposed spelling of his colleague, Jón Þorkelsson, based on Icelandic medieval manuscripts (see Sigtryggsson 2017)—the latter eventually becoming school principal to the great frustration of the former. Their disputes went far beyond orthography, carrying over to lexical and grammatical features as well. Such striking differences of language ideology and policy at this point in history clearly merit further study.

In an article on a number of “wrong word forms and word orders”, Þorkelsson (1870) describes standardising and puristic efforts of the period 1830-1870 as a great achievement, but he feels that much more can be done to this end. Unsurprisingly, Þorkelsson’s measure of correctness throughout his article is Old Norse as attested in medieval manuscripts. These prospects for a more archaic standard were met with great and fierce resistance, as evidenced by Friðriksson’s (1871) response to Þorkelsson’s article. While Friðriksson (1871:106) agrees that “foreign and unnecessary un-Icelandic words or un-Icelandic word order” (my transl.) should be counteracted and one should never lose sight of Old Norse, he warns about going to the sorts of extremes suggested by Þorkelsson (1870), pleading for a happy medium (*meðalhóf*). In this context, Friðriksson (1871:107) remarks that it would be childish (*barnaleikur einn*)²³ to phrase everything like the medieval authors of the Icelandic sagas—in fact, we would not succeed even if we tried (*op. cit.*). Friðriksson (1871:106) voices harsh criticism over such dogmatism (*eintrjáníngsskapur*) and miswander (*afvegur*), which he reckons would enchain our language and thoughts and be just as bad as being completely heedless, if not more dangerous still to the “rightful and natural advances of the language” (*rjettum og eðlilegum framförum tungunnar*). There are even some examples in Friðriksson’s (1861) Icelandic grammar where he departs from later ‘puristic’ ideals and those of some of his contemporaries, instead preferring certain attested spoken variants and permitting two alternative forms (for examples, see Ottósson 1990:97)

Still, the rather permissive tone in some of Friðriksson’s remarks on language change and words of warning concerning linguistic revivalism and archaisms is unexpected for a number of reasons. The sources clearly state that Friðriksson was relentless when teaching his norms, regardless of whether these concerned spelling or grammar (see e.g. Ottósson 1990:95f., with references). However, it is much less obvious that the written norms he taught to his pupils were necessarily much more to him than just that—standard norms. As Friðriksson’s correspondence with the Copenhagen-based Icelandic politician Jón Sigurðsson (1811-1879) reveals,²⁴ Friðriksson was not

²³The phrase *barnaleikur einn* can in principle also be rendered as ‘pure child’s play’, suggesting perhaps that it would in fact be *easy* to copy medieval writers. However, it is clear from the context that this is not the intended reading.

²⁴Jón Sigurðsson was a key figure in matters relating to Iceland’s independence from Denmark, often referred to as ‘president’ (*forseti*). The title refers to his function as the president of the Copenhagen department

on good terms with Þorkelsson at the time, to say the least (see also Sigtryggsson 2017:165ff. regarding spelling). This should of course be kept in mind when interpreting Friðriksson's criticism.

In Friðriksson's private letters to Jón Sigurðsson, one can find various sub-standard features and these include the linguistic variables under study here, which were all corrected in student assignments, by Friðriksson himself and by at least some of his colleagues. Despite correcting verb-adverb placement in student essays (see further below), the phenomenon appears to have been no particular thorn in his side outside the classroom and the public sphere. The same applies to Danish lexical features, even the generic pronoun *maður*, as well as the definite article *sá*, which all occur in his letters. Rather, this suggests that Friðriksson had a clear sense of a written public standard but at the same time allowed for a certain leeway—a distinction between the public and the private sphere (see also Rutten 2016:198f., with references, on “dual standards” with regard to spelling in public vs. private settings in an English context). In other words, Friðriksson's objection to Adv-Vfin, *maður* and *sá*, as attested in his corrections of student assignments, are not obviously *personal objections* and it would be misleading to suggest otherwise.

The point above about the rural basis of the norms is obviously a complex and thorny issue. To put things in perspective, let us start off with a well-known example from the literature. The alleged revival of the Old Norse inflection of *ia*-stems in Icelandic, frequently mentioned in this context, is widely regarded as a conscious change in morphological patterns that was (more or less) successfully implemented through the educational system (see e.g. Ottósson 1990, Hilmarsson-Dunn and Kristinsson 2010). The change affected a range of nouns where the *-r* of the nominative (masc.) in final position was reinterpreted as belonging to the stem, affecting nouns such as *hellir* ‘cave’, *hirðir* ‘shepard’ and *læknir* ‘doctor’, but also given names such as *Heimir* and *Freyr*, for instance (see Benediktsson 1969, Ottósson 1990, Sigmundsson 2002, Hilmarsson-Dunn and Kristinsson 2010). The previous *r*-less oblique forms (e.g. masc.sg.acc./dat. *helli*, gen. *hellis*) were thus superseded by generalised *r*-forms (acc./dat. *hellir*, gen. *hellirs*). According to traditional language descriptions, the newer pattern had taken over and the older *r*-less patterns (nominative aside) were reinstated, all on the model of Old Norse (see Ottósson 1990:71, pointing to Konráð Gíslason as the instigator of the norm). An overview of these patterns is given in Table 1.1.

Indeed, early 19th-century grammars such as the grammar manuscript of Guttormur Pálsson (c. 1805, p. 27)²⁵ and Rask's (1811:258) grammar of Old Norse/Icelandic show the younger declination patterns for contemporary Icelandic. It is striking that the paradigm provided by Pálsson is a mixed one, the dative singular optionally allowing for the older *r*-less form. The plural, similarly, features the older form in the nominative and the dative. The *-ara* forms in the accusative and the genitive furthermore suggest a competition of two distinct forms: *læknir* vs. *læknari*, *-ari* being a typical agentive suffix. Rask (1811:258) also notes forms with *-ari* in a remark, saying these are frequent.

of the Icelandic literary society.

²⁵The original is thought to be written ca. 1805 (cf. Sigmundsson 1981:286), whereas the preserved copy, Lbs. 1238 8vo, dates from ca. 1815 by the hand of Sigfús Árnason (see Bjarnadóttir et al. 1988-1989:201). This means that the manuscript presumably originates from around the time Guttormur taught Icelandic at the grammar school in Bessastaðir (1806-1807), although there is no official record of it having been used in teaching (cf. Möller 2017:8f.).

	Pálsson (1805?)	Rask (1811)	Friðriksson (1861)	Current standard
SINGULAR	(Contemporary)	(Contemporary)	(Rask's Old Norse)	(Old Norse-based)
nom.	læknir	læknir	læknir	læknir
gen.	læknirs	læknirs	læknis	læknis
dat.	lækni(r)	læknir	lækni	lækni
acc.	læknir	læknir	lækni	lækni
PLURAL				
nom.	læknar	læknirar	læknar	læknar
gen.	læknara	læknira	lækna	lækna
dat.	læknum	læknirum	læknum	læknum
acc.	læknara	læknira	lækna	lækna

Table 1.1. The declension of *ia*-stems as codified in 19th-century grammars, exemplified on the basis of the noun *læknir* ‘doctor’.

He appears to suggest that these are mostly found in the plural, seldom *læknari* in the nominative (singular). Already at the level of codification, there is thus contemporary evidence of variation at the beginning of the 19th century (see also Kvaran 2005:67 and Heimisdóttir 2008:20f. regarding variation in later 19th and 20th-century grammars). In contrast, the Icelandic grammar of Friðriksson (1861:20) simply presents us with the older declension of Old Norse, with no reservations—now being regarded as the standard norm.

In the study of Heimisdóttir (2008), the supposed revival of the Old Norse *ia*-stem inflection was re-evaluated on the basis of empirical evidence. Her study made use of 19th- and early 20th-century private letters, written mostly by scribes with little or no formal education, in addition to a questionnaire administered to 21st-century elementary school and high school students. Her results showed that, contrary to received wisdom, the older inflection pattern was apparently neither fully lost in the 19th century nor was it as successfully implemented as often suggested. This led her to conclude that the older inflection was reinforced rather than revived in the 19th century and that the sub-standard inflection is still attested to some degree in Modern Icelandic.

Similar evidence is provided by a study of the irregular noun *hönd* ‘hand’ carried out by Árnadóttir and Einaradóttir (2007). These authors concluded based on a study of 171 elementary school students (a fill-in task) that no less than *twenty* different declension paradigms were attested *alongside* the standard declension of this obviously not uncommon lexical item. The students were all 6th graders, between 11-12 years of age, from five different schools and all native speakers of Icelandic. In fact, the standard pattern ranked fourth among the most common declensions, used by mere 8.8% of the participants. That is quite a remarkable result given that the four possibilities in the singular (nominative, accusative, dative, genitive) maximally add up to 24 possible combinations. Perhaps surprisingly, a group of university students used for comparison did not perform that much different; while producing “only” 9 different paradigms, not 21 like the 6th graders, the standard declension was still used only 12% of the time.

Contrasting the above with 19th-century data, Friðriksson’s (1861:27) grammar simply presents us with the statement that *hönd* has an irregular dat.sing. form *hendi*, with no further remark on its use. The paradigm tacitly assumed is thus the same as

the standard declension reported on above (*hönd-hönd-hendi-handar* for nom., acc., dat., gen., respectively). While I have not carried out a systematic study of this noun in the 19th-century private letter corpus, it is abundantly clear that the codification of this particular norm was by no means a faithful representation of the attested paradigms; it was a prescriptive norm. Thus, we also find *hendi* forms in the nominative and accusative (nom./acc.sg.indef. *hendi*, nom.sg.indef. *hendin*, acc.sg.def. *hendina* for std. *hönd*, *höndin*, *höndina*), as well as a definite *hönd* form alongside *hendi* in the dative (dat.sg.def. *höndinni* for std. *hendinni*). There are not many examples of the genitival form but this appears to be *handar* in all cases, as the standard would prescribe as well. The plural also exhibits variation with regard to nom./acc. *höndur* alongside the standard (and more common) form *hendur*. In other words, there is a lot of variation here that Friðriksson's (1861) grammar plainly ignores, or better, erases, and little to suggest that the situation is much different now, over 150 years later.

1.3.5 The importance of the grammar school(s)

[S]tandardization of language typically radiates outward from metropolitan centers of power. (Cooper 1989:133)

Given the importance of the Reykjavík Grammar School in the implementation of standard norms, some remarks on the education system are in order. Advances in the establishment of primary and secondary schools in 19th-century Iceland are a relatively well-documented topic (see e.g. Magnúss 1939, Ármannsson et al. 1975, Þorsteinsson 2012, Möller 2017). The foundation was mainly in the form of homeschooling as there were very few primary schools until the 1870s (e.g. Ottósson 1990:104), a notable early exception being Hausastaðaskóli (1791-1812) in the Southwest of Iceland. The primary aim of the primary school at Hausastaðir was instruction to children of poor commoners, teaching basic children's learning, reading, writing, math, religious doctrine, among other things. Other early establishments include the primary schools in Reykjavík (1830-1848, 1862-) in the Southwest, Eyrarbakki (1852) in the South, Akureyri (1871) in the North, Ísafjörður (1874) in the Westfjords and, finally, Seltjarnarnes (1875) and Hafnarfjörður (1877) neighbouring Reykjavík in the Southwest. Later followed the establishment of 30 district elementary schools in 1887, introducing subjects like Icelandic, Danish, (English), geography, natural history and singing.

Despite the inclusion of instruction in the use of the mother tongue, there was actually very little grammar teaching involved (cf. Ottósson 1990). According to Ottósson (1990:104), the major change with the establishment of compulsory education for 10-to 14-year-olds in 1907 implemented the official policy of 'right' and 'wrong' in linguistic matters for all children during a formative period. The establishment of compulsory education for the masses may be regarded as acquisition planning to plan the spread of the standard (see Cooper 1989:33f., Wright 2004:61-64), partly in place already through compulsory literacy skills (see Section 2.3). School regulations up until 1907 were vague about the nature of Icelandic teaching, with a noticeable lack of consistency and centralisation (cf. Þorsteinsson 2012:111). Furthermore, it has been claimed that the language policy after 1907 took on a simpler and stricter, more direct guise to match the younger age groups that were being targeted (cf. Ottósson 1990:105). Note in comparison that the lower bounds for grammar school entry students had been the age

of confirmation (around the age of 13 to 14), but was brought down to 12 years of age in 1877, with the upper bounds for entry set to 18 years of age, and was intended as a 6-year study (see Ármannsson et al. 1975:38, with references).

During most of the 19th century, secondary education was confined to the grammar school first situated in Hólavellir in Reykjavík (SW, 1786-1804), then in Bessastaðir (SW, 1805-1846), before it finally moved back to Reykjavík in 1846, where it was known either as *Reykjavíkur lærði skóli* or *Hinn lærði skóli í Reykjavík*. This school, or succession of schools, was intended as preparation for higher education, e.g. at the priest school (1847-1911) or doctor school in Reykjavík (1876-1911), or at the University of Copenhagen. Later secondary schools included the Women's School (*Kvenmaskólinn*) in Reykjavík (1874-) and Möðruvallaskóli in Möðruvellir (1880-1902) in the north.²⁶ A king's directive from 4th December 1886 granted females the same right as male students to take the 4th-year exam as well as the graduation exam (see Ármannsson et al. 1975:54, with references). However, despite the directive, just three females enlisted at the school during the period 1886-1904 and only one of those graduated, in 1897 (cf. Ármannsson et al. 1975:56). In terms of social background, the students mainly came from the higher echelons and predominantly entered into positions as officials, doctors, priests, teachers, etc. (see Ólafsson 2004:40,44). As such, the main purpose of the Reykjavík Grammar School could be said to have been to populate the officials class (cf. Ólafsson 2004:40, with references). Highly illustrative of this state of affairs is the fact that when Halldór Kr. Friðriksson retired from the grammar school after 47 years of service, each and every official of the country, except one (magistrate or *landfógeti* Árni Thorsteinsson), had been a pupil of his (see Þorkelsson 1903:5).

During the time the grammar school operated in Bessastaðir, Icelandic could barely be regarded as an independent subject (see Möller 2017). According to legislation, Icelandic was to be on par with Danish in Danish schools, 3 lessons per week practising writing Icelandic skills, grammar and literature. However, in practice it was much less, mainly featuring translations from Greek, Latin, Danish into Icelandic with correction on language use, style and spelling. Interestingly, moreover, the 'famous figures' as regards Icelandic teaching were not the Icelandic teachers but the teachers of Latin and Greek, Hallgrímur Scheving and Sveinbjörn Egilsson, respectively (see e.g. Egilsson 1999:32f., with references).

Sveinbjörn Egilsson's translations of Homer's *Iliad* and *Odyssey* are typically seen as a major milestone in the emerging standard language, although according to Hermannsson (1919), Steingrímur Thorsteinsson's Icelandic rendering of the Arabian *One Thousand and One Nights*, published between 1857-1864, was more influential than Egilsson's prose translation of Homer, as it was read more widely than Egilsson's work, while still 'living up to the standard' set by Egilsson (see Hermannsson 1919:51f.). Another important milestone was the 'first' Icelandic novel *Piltur og stúlka* (see Bernharðsson 2017, 2018). As Bernharðsson (2018) emphasises, setting the standard by

²⁶The narrative on the standardisation of Icelandic focuses on the grammar school at Bessastaðir and later in Reykjavík, but the other schools are also of great interest. For example, the extent to which the education offered at Möðruvellir was to be regarded on par with that of the Reykjavík Grammar School was hotly debated at the parliament, with proposals suggesting that the Möðruvellir students should be exempt from entry exams into the first grade in Reykjavík or they should be able to enter the third or even the fourth grade (see Ármannsson et al. 1975:56, with references). To retain a reasonably narrow focus, I must leave this interesting topic for future research.

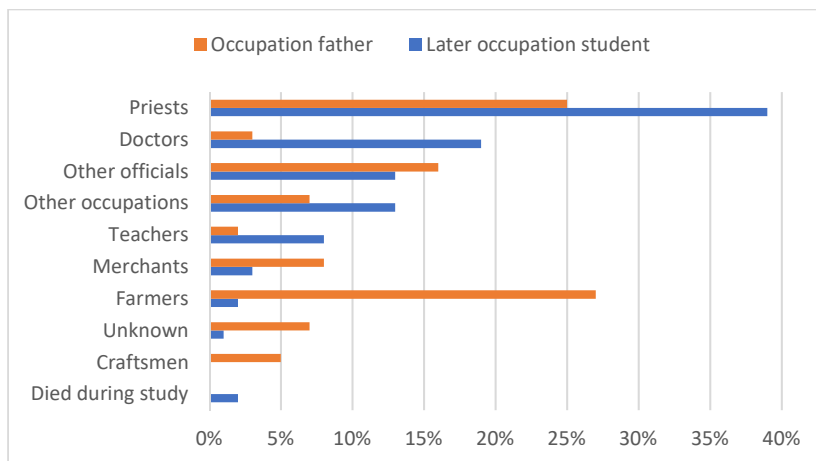


Figure 1.2. Social background of Reykjavík Grammar School students (1846-1904) based on the occupation of the father and the occupation of the students later in life (cf. Ólafsson 2004:40,44).

example in published works was of utmost importance in a society like 19th-century Iceland, in the absence of centralised language authorities.

After the grammar school moved (back) to Reykjavík, Icelandic finally become a subject in its own right (cf. Ármannsson et al. 1975, Möller 2017). According to a 1846 directive, the main purposes of Icelandic teaching were to: 1) introduce the general foundations of linguistics, 2) make students able to write Icelandic in accordance with norms, unmixed and in good taste (*samkvæmt réttum reglum, óblandað og með góðum smekk*), and 3) acquaint students with Icelandic literary history (see Ármannsson et al. 1975:110-112, with references). As discussed by Möller (2017:22f.), the directive was identical to a provisional plan from 1844 for Danish teaching at grammar schools in Denmark, the only difference being that the word “Danish” was substituted for “Icelandic”. Few lessons were allotted to Icelandic teaching in the beginning: in 1846-1847 this was only once per week. Icelandic lessons were increased to 12 in 1847-1848, 15 in 1877, ultimately 17 until 1904 (*op. cit.*, p. 112). Möller (2017:25-34) provides a thorough overview of the first few years of Icelandic teaching at the grammar school. Exams were written, consisting of an essay (*íslensk ritgjörð*), and three examiners, two besides the teacher, would then grade the assignments (Möller 2017:33).

Ármannsson et al. (1975:112) provide a concise overview of Icelandic grammar teaching at the Reykjavík Grammar School, including the later periods not covered by Möller (2017). Between 1850-1877, reading mainly consisted of Old Norse texts, in addition to chapters from the periodical *Fjölnir* (including book reviews on grammar), Friðriksson’s (1861) textbook on Icelandic morphology and Friðriksson’s (1846) reader. Friðriksson’s textbooks, including his book on spelling (Friðriksson 1859), were taught in the first grade until 1885, when Friðriksson’s grammars were replaced by the Wimmer’s Old Norse morphology and his Old Norse reader in 1887 (see Bjarnadóttir et al. 1988-1989 and Ottosson 2002 for comprehensive overviews of Icelandic grammars and language history writing).

In the regulations implemented in 1877, it is specified for the graduation exam (*burtfararpróf*) that the students write two Icelandic essays, one concerning a general topic, another on a topic related to a subject that had been taught at the school, with separate scores given for each. This arrangement was based on the corresponding Danish regulation (see Ármannsson et al. 1975:46). According to the 1877 regulation, the students should learn “to speak and write their mother tongue in a pure, correct and agile manner (*“hreint, rjett og lipurt”*, *ibid.*, p. 38). A further change involves the number of years (grades) when Icelandic should be taught, which the 1877 regulation states should be five instead of four, and finally increased to six in 1879 (cf. Ármannsson et al. 1975:38; 54).

As we will see later on, the data analysed in this thesis shows a statistically significant correlation between (more) standard-like language use and high graduation scores and/or longer exposure to linguistic norms through the progression of study (cf. also Viðarsson 2017b). Möller (2017) independently observes on the basis of a study of essays from spring 1852 that there are many corrections of spelling and punctuation, as well as linguistic traits in the material from the first-year students, “but all of this changed under the “army discipline” of Halldór [Kr. Friðriksson], as the students climbed up the grade ladder.” (Möller 2017:34, my transl.). As she observes for the spring 1852 essays, no non-standard morphological features were used by the fourth-year students and hardly any spelling errors. To get some idea of the extent and nature of linguistic corrections in the Reykjavík Grammar School student assignments, I provide a tentative overview in (19) based on my samples from 1847-1848, 1852, 1860-1861, 1870, 1875, 1882, 1890.²⁷ Note that these samples are not the same as and should not be confused with the sample corpus of student essays published by Ólafsson (2004), which is the one used and reported on in Section 3.4.3 on verb-adverb placement, Section 4.3.4 on the definite article *sá/hinn*, and Section 5.4 on the generic pronoun *maður*.

- (19) A breakdown of teachers’ linguistic corrections in Reykjavík Grammar School assignments (excl. spelling and punctuation):

²⁷The overview covers all grammatical (including lexical) corrections in the student assignments (mainly Icelandic essays) from each year listed, exactly one archival storage box per year/year range except for 1890. In 1890, the overview as yet only covers essays from second- and fourth-year students. Some phenomena that are listed under the heading ‘morphology’ in the overview may be regarded as being morphosyntactic and some listed under ‘syntax’ are also rather lexical in nature (e.g. the generic pronoun). The categorisation of the vocabulary is a very rough attempt at indicating the extreme emphasis on avoiding (potential) loanwords that ought to be acceptable to most writers nowadays in standard writing. For a range of examples, see Viðarsson (2017b) and discussion below.

Furthermore, note that the database *Corpus of Corrections* at this time of writing contains 904 entries, with a total of 476 potential grammatical corrections by teachers (with image links)—over one hundred more than shown in the overview, awaiting further classification. See: <https://github.com/heimirfreyr/RLSS>.

DOMAIN (aspect)	CORRECTIONS
vocabulary	120
acceptable currently	88
unacceptable currently	32
morphology	116
case morphology	57
verbal morphology	27
other	32
syntax	106
superfluous complementisers	28
lacking complementisers	25
foreign complementisers	8
definite marker	20
word order	15
generic pronoun	9
reflexivisation	1
other	27
	369

In this context, Vandenbussche (2007:29) emphasises “the ‘black box’ of historical pedagogy” as “the crucial points for future advances in the study of the spread of literacy and standardized writing behaviour”. The preservation of student assignments from the Reykjavík Grammar School (1846-1904) at the Icelandic National Archives, in fact extending further into the 20th century, is truly remarkable and a ‘game changer’ as far as future research into norm implementation and language ideological factors is concerned.²⁸

²⁸I have recently learned that similar 19th-century Swedish material has also been preserved in Swedish archives (cf. Kalm and Sahlée 2018) and it would be truly fascinating if more were found to exist across Europe. Some of the Icelandic essay topics are very similar to the Swedish topics exemplified by Kalm and Sahlée (2018), some even identical, though this could be a coincidence. The sheer quantity of the Icelandic material is truly massive, preserved in over a hundred large archival storage boxes.

Jóhannes B. Sigtryggsson (p.c.) has carried out a preliminary inventory of these assignments. According to an acquisition receipt, there is typically one storage box per year from 1847 until 1880, but two or more often 3-4 boxes yearly after that, in total 123 boxes from 1847-1909. These include questions for oral exams (1847-1850, 1862-1913), in addition to general entry exams, mid-term, end-term exams and final exams, presumably covering the whole period until 1909. Sigtryggsson’s examination of a box dating from 1870 revealed 378 pages of Icelandic translations (*ísl. stíll*) by first- to fourth-year students and 89 pages of fourth-year spring exam essays and graduation essays (*ísl. ritgjörð*), totalling 467 pages. If we simply multiply this number by 123 as a very rough estimate, we thus arrive at a figure of nearly 60,000 pages of Icelandic assignments.

2 Methodology

2.1 Corpora

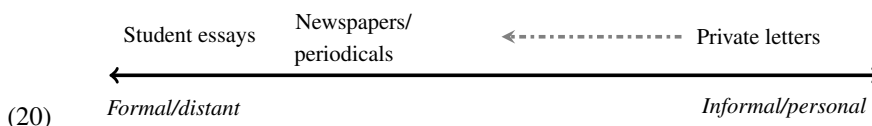
In historical linguistics, the researcher seldom has access to more than a small fraction of language data and crucially no negative evidence. Rather, the material available is usually produced by scribes belonging to the cultural elite who would by most modern (socio)linguists not be considered to be the most typical language users and often these data are not particularly well suited for linguistic research, frequently referred to as the “bad data” problem (see Labov 1972a:100). Corpora on Early Modern Icelandic have long been limited, especially as far as linguistically annotated corpora are concerned, such as tagged corpora that may be used to search for parts of speech and morphological categories, or parsed corpora and treebanks where the data have been annotated syntactically (see Svavarsdóttir et al. 2014).

The first and only Icelandic corpus to combine these two aspects is the Icelandic Parsed Historical Corpus (IcePaHC, Wallenberg et al. 2011). While IcePaHC was used extensively in my search for particular patterns and to get an idea about the major diachronic trends over the centuries (see Viðarsson 2017a and Chapter 3), the corpus design was not optimal for the purposes of this dissertation. The reasons for this are two-fold. IcePaHC is not only quite limited in size, containing only about 121,000 words from the 19th century, the text types included are also quite varied but limited in scope. Since the present research is couched in a particular framework with very specific assumptions about individual text types, and because my project was a part of a larger team involved in the production of the 19th-century corpora, the visible role of IcePaHC in the present work is admittedly rather limited.

The three major corpora used in this dissertation, outlined in the following subsections below, were all limited to a text-based string search using *AntConc* (Anthony 2012), a freely available concordance tool. To facilitate manual annotation of the search results, sorting of lexically or structurally similar hits and to further explore certain patterns in the corpora, the natural language processing tools that are a part of the IceNLP package (Loftsson and Ingason 2011) were used in order to automatically tag parts of the hits. However, automatic tagging was not used as a part of the search process itself so as not to skew the results. For each variable, a basic string search was carried out with the relevant adverb (in the case of the verb-adverb variable), any inflected form of *sá* or *hinn* (for the definite determiner) or *maður* (for the generic pronoun). At no point could potential tagging errors, therefore, influence the collecting of data or the annotation process.

While the focus of the present work is mainly syntactic, there is an additional (but integral) sociolinguistic focus. Here the aim is to study and describe the observed

socio-syntactic variation from the perspective of the community over and above the variation found in the individual, focusing on traditional macro-level contrasts such as social status/education level, sex/gender and geographical origin, in addition to text type, especially in terms of different positions on the formal/informal, distant/personal spectra. As for text type, there is a major contrast between private letters on the one hand, and newspapers and periodicals on the other, with the former (generally speaking) representing language use that is informal/personal and the latter representing formal/distant language (see Finegan and Biber 2001, Elspaß 2005a, 2012). Student essays, in addition, constitute a third genre, expected to be at the top of the formality/distance spectrum in that these were graded by examiners; as such, whatever symbolic value linguistic signs may have on the ‘linguistic marketplace’ (Bourdieu 1977), in terms of their economics metaphorically speaking, that value is arguably subject to ‘inflation’ (cf. Viðarsson 2017b). (Due partly to the more varied speakers’ social background and complex writer-addressee relations, the private letters may be expected to occupy a broader range than depicted in (20) below, as indicated by the dotted arrow.)



A brief overview of these three corpora will be given in the following subsections, beginning with the private letters.

2.1.1 Private letter corpus

Two collections of private letters will be used in the present study. The main corpus consists of an electronic diplomatic/facsimile edition of 19th- and early 20th-century private letters, written mainly to friends and family. This corpus was used jointly in the 19LCLV research project, the size of which is approximately 1 million words.²⁹ The letters were transcribed by members of the project, among others, and the corpus currently contains 1,928 letters written by 343 scribes (222 male, 121 female), of which 860 letters were written by male writers and 1,068 by female writers. A secondary corpus was compiled as a part of the present PhD project as an extension to the 19LCLV letter corpus. The reason for this was mainly that the latter contained only very limited data from up until the middle of the 19th century. These additions consist of late 18th- and early 19th-century private letters, being an electronic rendition of a selection of published diplomatic and semi-normalised editions of private letters, scanned and post-processed using *Google Tesseract-OCR* and *Skrambi* for post-correction.³⁰ The

²⁹This is an estimate based on an automatic extraction of only the <body> portion of the XML version of the letters, excluding all XML/TEI tags. That word count amounts to 1,006,159 words. A more precise estimate of word count per period/time frame is not available at the moment. The 19LCLV private letter corpus is freely accessible online: <http://brefasafn.arnastofnun.is>.

³⁰While the semi-normalised editions have been used in previous historical research (see in particular Hróarsdóttir 1998, 2000, 2009), it should be mentioned that due to the normalisation, these texts are not always suitable for linguistic research. However, based on a cursory comparison against some of the original manuscripts, they appear to be safe to use for most syntactic and lexical purposes, much less so for morphology, and unsuitable, of course, for studying orthographic features.

secondary corpus currently consists of 670 letters written by 26 scribes, and contains approximately 425,000 words. Unless otherwise noted, reference to the private letter corpus includes both of these collections. Where needed, the primary letter corpus will be referred to as the 19LCLV corpus and my extension corpus as the Icelandic Corpus of Early Nineteenth-Century Correspondence (ICENCC).³¹

The data imbalance introduced by well-represented writers, whose letter-writing spans years or even decades, producing tens and in some cases even hundreds of letters, and writers who only wrote few or a handful of letters inevitably leads to a bias which needs to be controlled for in any statistical analysis of the data. A rather cumbersome solution would be to set a maximum per scribe and select the data by random, producing a (more) balanced sample, such that speakers can more easily be grouped together without certain individuals being overrepresented. However, at the same time, one would like to have access to the maximum amount of data from each speaker, especially for qualitative purposes when analysing individual grammatical systems. A number of more elegant techniques are available to circumvent this problem, the most powerful one being the use of statistical mixed-effects models, which can incorporate the individual speaker into the analysis as well as being particularly well-suited to dealing with unbalanced data (see further Johnson 2009, Gries 2015b and Section 2.4).

2.1.2 Newspaper corpus

The newspaper corpus consists of a collection of 46 newspapers and periodicals, covering 250 issues, for a total of 2.017.251 words. The corpus is based on a selection of titles from the larger *Tímarit.is* corpus of the National and University Library of Iceland, featuring a corrected version of the OCR text of each of the 250 issues by members and employees of the 19th-century project.

- (21) Overview of word counts in the newspaper corpus per time period:

PERIOD	TIME FRAME	WORD COUNT
Period 1	1803-1825	97,573
Period 2	1825-1849	351,739
Period 3	1850-1874	187,055
Period 4	1875-1899	508,785
Period 5	1900-1924	872,099

Newspapers are widely regarded as closely representing the national identity and ideology and it is here, of course, that we expect to find the emerging national standard language. As Anderson (2006:25) states, the introduction of both the newspaper and the novel “provided the technical means for ‘re-presenting’ the *kind* of imagined community

³¹The ICENCC is freely accessible at <https://github.com/heimirfreyr/ICENCC>, along with a letter inventory, and more detailed bibliographical information. The corpus has been automatically tagged using the IceNLP tools and parsed using Berkeley Parser (BerkeleyParser-1.7.jar) in combination with the IcePaHC grammar set (ice.gr). For further information, see <https://github.com/antonkarl/icecorpus/tree/master/parsald>. Note that references to the ICENCC and 19LCLV corpora can be distinguished by the .txt file extension in the former vs. .xml in the latter.

that is the nation”, the newspaper being an ‘extreme form’ of the book, mass-produced and sold on a colossal scale (*op. cit.*, p. 33f.).

Even more importantly from our point of view, some of the 19th-century newspaper editors have been described as ‘fervent language purists’ (Gíslason 1972:17). These include editors such as Baldvin Einarsson (*Ármann á Alþingi*), Konráð Gíslason (*Fjöl-nir*), Björn Jónsson (e.g. *Þjóðólfur, Ísafold*) and Jón Ólafsson (e.g. *Baldur, Þjóðólfur, Iðunn, Skuld, Reykjavík, Dagblaðið*), among others. Björn Jónsson, whose language was “pure and well-crafted” (*hreint og vandað*, cf. Gíslason 1972:141), is reported to have been extremely particular about not letting non-standard linguistic features slip into his newspapers (see e.g. Ottósson 1990). He also produced dictionaries and textbooks, e.g. an Icelandic spelling dictionary in 1900, which was long popular (cf. Gíslason 1972:136) and included a rather comprehensive section on common language blemishes (*mállýti*). Jón Ólafsson, too, was greatly influenced by the journal *Fjöl-nir*, the language commentaries of which (in the form of book reviews) were widely read, as well as taking great interest in Icelandic linguistics (Gíslason 1972:146).

Besides academies and educational institutions, book reviews are an exuberant source of prescriptive remarks on language. From an English perspective, Percy (2010:57) argues that “reviewers exploited and furthered the standardization of English in their own quest for professionalization” and “contributed to the social prestige and commercial value of education generally and of good English specifically.” The same is true of Icelandic book reviews in newspapers and periodicals, which in many cases would have reached a much wider audience than the work reviewed ever did. There is reason to suspect that the reviews published in *Fjöl-nir* were particularly influential given that *Fjöl-nir* actually became a part of the school curriculum (Möller 2017:31f.).

As Gíslason (1972:11) observes, on the one hand, late 18th and early 19th-century Icelandic newspapers often tended to be closer to historiographies and leisure stories than they are to modern newspapers, while on the other hand, the newspaper as a phenomenon can be traced back to medieval times in skilful storytellers and annal-writing; the short or dense sentences of the annals may even resemble the modern headline (Gíslason 1972:11f.). In the first issue of the paper *Ísafold* in 1874, editor Björn Jónsson also situated the newspaper genre among (private) letters, referring to the former as a sort of ‘open letter’, not originating from the chancery but *from the nation to the nation*, “*from everybody to everybody*” (*Ísafold* 1:1, 19 Sept. 1874, p. 1, my translation and emphasis; cf. also Gíslason 1972:10). Newspapers in the more traditional, modern sense of the word emerged with publications such as *Klausturpósturinn* from 1818 and *Þjóðólfur* from 1848 (*ibid.*, p. 13).

2.1.3 Student essay corpus

The student essay text corpus consists of 189 essays written by 170 students, being a selection of student assignments published by Ólafsson (2004). The essays were provided to the 19th-century project in electronic form, making it possible to use similar search and analysis techniques as for the two other corpora. The time period covered by this selection is 1852-1906, and comprises in total about 83,000 words. A rough overview of the general make-up of the corpus per period is provided below.

(22) Overview of word counts in student essays (ed. Ólafsson 2004):

YEAR	WORDS	PERIOD	WORDS	TEACHER(S)	WORDS
First	6,421	1852-1874	16,408	HKF (1848-'80)	24,465
Second	13,175	1875-1884	16,798	HKF a.o. (1881-'95)	27,074
Third	7,614	1885-1894	14,922	Post-HKF (1896-)	31,714
Fourth	15,534	1895-1900	17,199		
Fifth	15,809	1901-1906	17,926		
Sixth	23,859				

The variables YEAR (or grade) and TIME PERIOD should be self-explanatory. The TEACHER(S) variable refers to three (time) periods during which 1) Halldór Kr. Friðriksson (=HKF) taught Icelandic alone at the Reykjavík Grammar School, 2) HKF taught with others, and 3) after HKF retires as Icelandic teacher. Halldór Kr. Friðriksson (1819-1902) taught Icelandic at Lærði skólinn between 1848-1895 and is credited with having been “more influential than most” in implementing the puristic norms of *Fjölnir* during this period (cf. Ottósson 1990:95f., my translation). The student essays were produced by the students as a part of grading, thus making it clear that the language use appearing in this corpus can be expected to be an honest attempt at producing the standard norms supposedly so successfully implemented by the Reykjavík Grammar School (see also Viðarsson 2017b).

Briefly mentioned above with regard to Icelandic standardisation more generally, the student essays are a remarkable source of not only students’ language use in a formal setting, but they also provide insight into both the emerging standard and the standard language ideology, as these appear through the linguistic corrections of the teachers. The student assignments are an important, yet neglected, perhaps hitherto even altogether forgotten, source, at least as regards the present context of language standardisation. Due to the fact that the assignments were corrected by teachers (typically by means of underlining), they attest not only to language variation and change, but in my view constitute a metalinguistic gold mine meriting much further study than I could permit myself in the present work. When considering the fact that 19th-century Icelandic grammars and (meta)linguistic discussions were limited not only in number, being very scarce, but also in their coverage, in particular with regard to syntax, we simply cannot afford to ignore such a source. Importantly, the assignments are not just a straw in the wind—the writers of the future in training—but the smoking gun of norm implementation and norm evaluation, no less. We arguably cannot get much closer to the actual execution of language standardisation than this.

Ólafsson’s edition only reproduces the student texts, but lacks the teachers’ corrections thereof. I have therefore complemented Ólafsson’s selection of texts with my own selection of two kinds of materials that I have photographed and annotated for corrections of linguistic features—well over 200 assignments dating from the years 1847-48, 1852, 1860, 1870, 1875, 1882 and 1890.³² Most of these assignments consist of the same sort of Icelandic essays (so-called *íslensk ritgjörð*) used by Ólafsson. However,

³²My database comprises about 500 corrections, most of which may be traced to specific teachers or instructors based on the bundle wrapping paper, listing examiners, and based on the colour and texture of the ink (see also Viðarsson 2016). It is my hope that this corpus of corrections will eventually be published and even extended further in the future (for examples and overview, see Viðarsson 2014, 2016, 2017b).

my selection also contains Icelandic translations from Danish (so-called *íslenzkur stíll*) and, similarly, Danish-Icelandic translations for entry exams. The student texts from Ólafsson's edition will be used during the analysis to measure the uptake of standard norms, whereas my corpus of corrections will be used to complement the often fragmentary descriptions and codification of the standard norms in grammars and common reference works—many of which are particularly unreliable as regards standard norms relating to syntactic phenomena.

The printed edition of the student essays already lists a number of features that may be used as a part of the statistical analysis. These include the year an essay was written, the student's name and the grade (year) level at the time of writing. These aspects could be turned into variables: TIME PERIOD, the INDIVIDUAL and the educational variable GRADE. To these variables, I manually added further linguistic variables corresponding to those of the other case studies (see below) and, last but not least, the educational variable GRADUATION SCORE. Unlike student scores today which are considered private, student graduation scores in the 19th century were matters of public record, and were simply published in the newspaper. On the occasion of the 50th anniversary of the Reykjavík Grammar School, all scores were also conveniently published in a small bundle featuring information about the teachers and the students' occupation later in life (see Helgason 1896). These scores will be used as an educational variable below, as an approximation of the skill or competence of the student. Student grade is another such educational variable, based on the grade in which an essay was written, according to the information provided by Ólafsson (2004). SCORE is obviously not a direct grading of the essay but a final score upon graduation. However, this arguably might be used as an index to or approximation of highly skilled vs. less proficient students, which is likely to correlate in at least partly predictable ways with the uptake of the standard norm, as indeed revealed by the case studies below (cf. also Möller 2017, Viðarsson 2017b).

2.2 Social structure

A fundamental principle in historical linguistics, whether sociolinguistic in nature or not, maintains that the forces underlying language variation and language change are the same in present-day communities as they were in the past, known as the UNIFORMITARIAN PRINCIPLE (see Labov 1972b:101). As a consequence of this, we expect our (socio)linguistic toolkit to be applicable, *mutatis mutandis*, to analyse both past and present situations. It is important to bear in mind that while the principal components of analysis arguably will be the same, the features and social categories shown to be relevant in modern societies may just as well be completely different historically. One must be particularly careful about 'constructional' anachronisms when referring to social groups in the context of language variation and change (see Bergs 2012). This is most obviously the case with social categories that arguably existed but were fundamentally different, e.g. the social status of men vs. women, children and adolescents, and social class division.

The status of women in pre-modern periods was not at all comparable to the way women have been conceived of as a group in most sociolinguistic studies on western

societies in the 20th and 21st centuries. This naturally raises the question of whether common sociolinguistic findings such as sexual differentiation in language change (see e.g. Labov 1990, 2001) actually apply to historical periods. If the special status of women depends on their access to prestige/standard forms (Labov 1990:213), there is every reason to expect that their lack of formal education in 19th-century Iceland will affect the linguistic patterns we find. Moreover, if conformity to prestige patterns is also strongest among groups with high linguistic insecurity (Labov 1990:213), it is not obvious that the same social groups would be affected by such insecurities.

The age factor also presents numerous issues and cannot be directly compared to present-day categorisation by any means. The studies of Guttormsson (1983) and Magnússon (1995) are particularly revealing in this regard. To the extent that we can speak of adolescence in 19th-century Iceland at all, the significance of these terms from a sociolinguistic perspective is rather unclear. In his study of childhood in Iceland during this period, Magnússon (1995:314) concludes: “The absence of adolescence in nineteenth-century Iceland peasant society was simply a fact of life.” By and large, children would start to work between the age of five and seven, entering adulthood around the age of fourteen (Magnússon 1995:300). In fact, as labour, children aged 8-10 years were simply considered miniature adults, and during the most hectic periods of the year, they were not allotted much more time to play than grown-ups got to rest (Guttormsson 1983:198f.).

Magnússon (1995) identifies three life stages: (i) *infancy* from birth to about five to seven years of age; (ii) *childhood*, delimited by the ‘first work assignment’ and spanning the period from the end of infancy until the age of fourteen; (iii) *adulthood* from the age of fourteen onwards, delimited both by their ‘first work task’ assigned and the act of confirmation. Within the second life stage, Magnússon (1995:307ff.) also emphasises changes around the age of ten, when children were assigned more substantial tasks and entered into a closer relationship with adults. Based on evidence from the early 18th century, children who had reached 15 years of age would typically permanently relocate to other farms in order to join the general workforce (Guttormsson 1983:84).

In terms of social economic status, this system gave rise to two distinct class structures within the peasant society, where people would either become independent farmers or belong to the underclass, forming “a permanent corps of cheap labour” (Magnússon 1995:316). Before turning to class division in more detail, it is worth mentioning that pre-modern Icelandic community structure is widely argued to have given rise to the relatively conservative character of Icelandic. More specifically, the lack of peer groups for both children and adolescents, in addition to limited social stratification, is thought to have retarded linguistic differentiation (see Kroch 2005). While the lack of peer groups (in a modern sense) may have been a more frequent feature of earlier times even outside of Iceland, the limited number of people, isolation and often harsh conditions hindering mobilisation could be regarded as different from, say, the European mainland. Similarly, Kusters (2003:185) emphasises the lack of urban centres such as cities or even villages until the end of the 19th century, as a result of which, there was no ‘street life’, inhibiting the formation of language varieties associated with certain age groups (‘chronolects’). With the rise of a more urbanised society dependent on fishing during the period 1880-1930, these traditional characteristics of Iceland started changing (Gunnlaugsson and Guttormsson 1993).

The main social categories that have figured in discussions of variation and change in Icelandic in the traditional literature concern the opposition between EDUCATED INTELLECTUALS (*menntamenn*) and COMMON PEOPLE (*alþýða*), the latter of which mainly refers to the lower echelons of society, especially the (highly diverse) class of peasants, workers and day-labourers. Contemporary 19th-century writing assumed three classes: the office-holder-, clerical-, and peasant-classes (see Hálfðanarson 1995:768f.). Subsumed under peasants according to this classification are merchants, craftsmen and civilians, and the clerical class includes teachers.

The traditional *learned vs. common* opposition is often taken to represent a sharp divide in language use. Supposedly, the common people exhibited a ‘pure’ vernacular, close to medieval Old Norse, whereas the language of the educated intellectuals had been ‘contaminated’ with foreign features, either through their education abroad and/or through interaction with Danish crown officials, merchants etc. Language traits of the former group tend to be depicted as contrived in the sense of belonging only to a particular intellectual style (or *lærður stíll* ‘learned style’) and these speakers regarded as being under considerable foreign influence from the Danish colonial aristocracy, through their education at the University of Copenhagen, foreign authors they were exposed to and so on (for a recent evaluation along these lines, see Heycock and Wallenberg 2013:152f.). Female speakers are of a particular interest in this context in that they did not have the same opportunities as males to receive formal education. At the same time, however, they could still belong to the higher classes in a social and economic sense, arguably having a similar sort of prestige status, by affiliation, as their male counterparts who graduated with college and university degrees.

Again from the traditional perspective, the implementation of standard norms is considered to have been relatively straightforward because the idiom of the privileged few had insufficient representatives to survive and was antithetical to nationalistic ideals, while the standard at the same time had a precedent in “popular/folksy historians” and on the “lips of rural people” (cf. Sigmundsson 1990-1991:139, my translation; see also Smári 1920:13f., Ottósson 1990:75 for similar remarks). The problem with such claims is that they are mainly based on assumptions which are ultimately rooted in the nationalistic discourse of the 19th century (see Leerssen 1999, 2006a). While this does not automatically entail, of course, that these assumptions are completely unfounded, equating the language of commoners with that of the standard or, for that matter, with a long-standing, unbroken tradition going back to medieval times, are each a set of claims that cannot be considered in isolation from their historical context.

Modern scholarship has engaged with these distinctions critically and revealed more intricate hierarchies (see Gunnlaugsson 1988, Hálfðanarson 1995, Bjarnason 2006, Magnússon 2010). The most elaborate of these is arguably that of Gunnlaugsson (1988), who argues for at least a basic ten-partite classification (cf. also Magnússon 2010:28f.):

- (23) 1) Crown officials (*embættismenn*) (who were often farmers as well).
- 2) Landowning farmers (*sjálfseignarbændur*) who were not crown officials.
- 3) Merchants (*kaupmenn*) and artisans (*handverksmenn*).
- 4) Tenant farmers (*leiguliðar*).
- 5-6) Sub-tenants (*hjáleigumenn*) and cottars (*búðsetumenn*).
- 7-8) Lodgers (*húsmenn*) and boarders (*lausamenn*).

- 9) Servants (*vinnuhjú*).
- 10) Paupers (*þurfamenn*).

At this time of writing, as fine-grained an approach to the class structure of 19th-century Iceland as that of Gunnlaugsson (1988) is far beyond the background information available to the 19th-century project of which my study is part. My hand is therefore forced to simplify these distinctions grossly, while at the same time trying to avoid oversimplifying so much that my data become interpretable. As a very basic scheme, I adopt a tripartite division, different from the traditional contemporary 19th-century one, tailored to better reflect the social and educational background of the speakers:

- (24) *(Over)simplified working definition of the class-status variable*
- 1) Officials/educated intellectuals and their partners by association.
 - 2) Other professions typically requiring formal training other than farming.
 - 3) Peasants/labourers.

By far the most controversial group in this operationalisation of class-status categories is the peasants/labourers group. Peasants/farmers, for one, is a highly diverse group where socio-economic status arguably plays a very large role (cf. groups 3), 4) and 5) in (23) above). That dimension is, at the moment at least, completely hidden in my data. However, I feel that I have engaged in such a way with the data that aggregating these individuals, in addition to labourers (i.e. 9) above, perhaps 7)-8) as well), appears to be defensible. I will provide remarks below with regard to the distribution of linguistic variants for further justification.

Another aspect where the uniformitarian principle would often have to give rise to anachronisms when regarded from a modern perspective directly concerns literacy and writing practices. While literacy nowadays is usually taken to comprise both the ability to read and to write, it is frequently pointed out that these must be treated as two separate skills in early modernity. Since the ability to read was closely tied to the religious sphere, as a part of people's religious upbringing, this level of literacy has been referred to as 'religious literacy' (see next section).

2.3 Compulsory learning and literacy

In Iceland, the so-called household discipline act of 1746 and later a king's directive issued out in 1790 made reading skills mandatory for all children before their confirmation around the age of thirteen, resulting in near-universal (religious) literacy; mandatory writing skills, in contrast, made it into legislation only as late as in 1880 (see e.g. Guttormsson 1990, 2012, Rastrick 2003, Ólafsson 2012, Magnússon and Ólafsson 2012). Notwithstanding the limited effort made to implement 'full' literacy from above by law, writing skills were transmitted informally by those in possession of this ability as shown by the vast amount of 19th-century private letters as well as "the abundance of preserved manuscript material in ordinary people's hand-writing" (Ólafsson 2012:67f.)—effectively emerging *from below* (see e.g. Guttormsson 1990, cf. also Halldórsdóttir 2003:249, Magnússon and Ólafsson 2012:114, with references). Nonetheless, the ability

to write in the early 19th century was a skill of the privileged few (see e.g. Halldórsdóttir 2014:185).

Based on questionnaires sent out to priests by the Icelandic Literary Society in 1839 regarding the life situation of people in their parish, including people's writing abilities, it has been possible to roughly estimate the spread of writing skills among the general public in the early 19th century. Responses indicate that around 1840, approximately 20-50% of adult males and 10-30% of adult females could write (cf. Halldórsdóttir 2003:250, 2014:185, with references). As has been pointed out, the majority of those who could not write, women aside, were elderly people and paupers; in a corresponding questionnaire from 1873, the rate of literacy had increased considerably, especially among younger people Halldórsdóttir see 2003:250, with references). Guttormsson (2003:59) also suggests that the growth in the production of handwritten material in the period 1820-1870 and the increasing rate of writing abilities during that period was tied to the rise in population size in the countryside which led to disguised unemployment, providing many with more leeway than before.

Legislative changes show that being able to write was not considered desirable for the general public until the latter half of the 19th century. The same case can be made for women more generally. However, women of better social standing, whose fathers were e.g. officials or wealthy farmers, arguably had more chances of acquiring this craft than women who were lower-placed (see e.g. Halldórsdóttir 2003:251). This is also what the historical record suggests, although here we can only judge, of course, by what has been preserved. Most female letter-writers from the late 18th and early 19th centuries have ties to the higher echelons of society, cf. e.g. the collection of women's letters in Sigmundsson (1952, 1961) and the female writers discussed by Halldórsdóttir (2003, 2014). Similarly with regard to autobiographical writings, it was first in the latter half of the 19th century that individuals belonging to the lower classes, who lacked a household status of their own, as well as the female population, began writing, writing abilities having been rare among these groups up until then (Guttormsson 1983:164).

Unlike elsewhere in Europe, no letter-writing manuals were published for Icelandic letter-writers. However, that does not rule out the possibility that manuals produced for other markets were not used, e.g. Danish guides (see Sandersen 2003:375), although no such cases have been documented to the best of my knowledge (see also Halldórsdóttir 2014:186). Indeed, many probably learned from family members and friends, and by practice, how to pen a letter. As Halldórsdóttir (2014:186) states for two female writers she studied: "The actual letter writing was their manual". Although the letter-writers in the private letter corpus clearly are not all experienced writers, as revealed for instance by their unconventional punctuation and various orthographic features, they do exhibit familiarity with the basic components of (familial) letter-writing: greeting the addressee, date and place written, thanking for a previous letter, wishing well, followed by important news and the main body of the letter, closing with giving regards, a farewell and signature. Letter-writers even occasionally remark upon the structure of their letter, bringing attention to the norms and conventions of the private letter genre.

A case in point is Friðbjörn Bjarnarson (1860-1943) in his early adulthood writing to the priest Einar Friðgeirsson, whom he addresses as his relative (*kæri frændi*) and Soffía Daníelsdóttir (1858-1907), a priest's wife in the north-east of Iceland, writing to her parents (my translation):

- (25) I just have to mention that I have not so much as wished that these lines would visit you in good health, and yet I believe it is **the custom** (*siður*) of various good letter-writers to begin their letters with that (FriBja-1879-05-12.xml)
- (26) I think I have now gone **against the customs** (*farið aptanað siðunum* [idiom. ‘broken the rules’]) by not beginning the letter by wishing you good fortune and blessing (FriBja-1881-01-08.xml)
- (27) It should be fine for me to **skip the introduction** (*formálanum*), that letters so often begin with, namely to thank you for the letter with the last mail (SofDan-1888-02-29.xml)

Letter-writers may also bring more subtle attention to the structure using phrases announcing the order or arrangement of the contents, for instance: “... I commence it [i.e. the letter] and the first then is to thank for your good and much appreciated letter ...” (SigPal-1868-08-04.xml) or “... then it is [time] to begin with the news” (SigEir-1867-09-12.xml). These remarks and often systematic organisation of structure in the letters attests to the fact that even colloquial letter-writing adheres to strict rules and should not be equated with fully spontaneous, unattended speech.

2.4 Statistical analysis

A fundamental claim that Weinreich et al. (1968) make is that there is underlying structure to the apparent chaos of linguistic variation in its distribution among speakers—*structured heterogeneity* in terms of the social and stylistic function(s) a variant carries. Statistical analysis is an important way of revealing such distributional properties in linguistic data by establishing statistically significant correlations between an independent linguistic variable and dependent grammatical and social variables.

The present study makes extensive use of a multifactorial statistical methodology which has been gaining currency over the past decade or so, known as mixed-effects modeling (for recent applications in linguistics, see e.g. Johnson 2009, 2014, Tagliamonte and Baayen 2012, Gries 2015b,a). Mixed-effects models are particularly useful for historical sociolinguistics in that they deal more easily than other models with unbalanced data sets, e.g. where certain social background data are missing for some speakers or where a particular individual contributes relatively many tokens to the corpus.

Unlike traditional variable rule analysis using logistic regression, popular in sociolinguistic research since the 1970s as a part of the GoldVarb and VARBRUL software packages (see e.g. Johnson 2009, Tagliamonte 2012),³³ and traditional descriptive statistics such as cross-tabulation summing over individuals with chi-square testing for statistical significance,³⁴ mixed-effects models allow the researcher to introduce fixed effects like *biological sex*, *social class* or whether grammatical subjects are full

³³Johnson (2009) implements a mixed-effects variable rule model (Rbrul) in combination with R’s *glm* and *glmer* functions.

³⁴Cross-tabulation can, of course, be used to test the effects of each individual separately and thus to measure the effect of individual variation, referred to as “[b]est practice in variationist methodology” (Tagliamonte and Baayen 2012:144), in cases where individual speakers’ responses have been treated in aggregate.

noun phrases or *pronouns*, and random effects like *individual speaker* into the model. This is important because if a multivariate analysis fails to take the individual level into account, we run the risk of finding a statistically significant correlation between a linguistic feature and a social group where the correlation may actually not hold at the level of the group. The reason for this is that most other statistical methods will assume that each data point in the data set is independent of the other (cf. e.g. Gries 2015b:99), potentially overestimating group behaviour and underestimating individual speaker variation.

Treating every observation in a data set as independent may be perfectly fine for studies where each subject contributes exactly one data point. However, this design is highly unusual within corpus linguistics where dozens, hundreds or even thousands of observations are usually collected from a single author and/or text such that the independence assumption is violated (for discussion, see Gries 2015a:103). Corpus linguistics and quantitative studies in linguistics in general thus resemble repeated measure studies where measurements for each subject in a study are carried out more than once, e.g. longitudinal studies where a participant in an experiment is later subjected to re-testing or where a subject is tested before and after receiving a particular treatment.

By not including the individual in the model, a corpus linguist may miss out on potentially valuable data encoded in the corpus. As Gries (2015a:101) points out, corpora are often hierarchically nested such that speakers are nested into files, files into sub-registers, which in turn belong to registers, which may consist of different modes such as spoken versus written language and so on. Since effects may be located at each of these levels, we make a stronger case for each correlation that can be shown to be significant, if such factors are included in the model. The same is true for individuals versus group aggregates. In addition, factoring in the individual makes it possible to assess the extent to which “individuals differ systematically over and above the differences between the groups to which they belong” (Tagliamonte and Baayen 2012:158).

Mixed-effects models are not as sensitive as classical models to biases introduced into a data set when a large number of tokens are sampled from a particular individual: “An advantage of the mixed-effects modeling framework is that it allows the researcher to sample as many tokens from a given individual as is feasible, thereby increasing statistical power.” (Tagliamonte and Baayen 2012:158). An unbalanced data set (or subset of data) consisting of five data points from speaker A, 50 data points from speaker B and 500 data points from speaker C thus need not interfere with our hypothesis-testing in that the model will produce a result taking these effects into consideration, as opposed to treating the data set merely as 555 (independent) data points.³⁵

While quantitative research in historical sociolinguistics is increasingly being combined with statistical models (see Auer et al. 2015), issues of the kind just described have led researchers in this field *not* to attempt a full-blown statistical analysis. Instead, some may for instance rely mainly on less elaborate descriptive statistics such as aggregated cross-tabulation paired with chi-square tests of statistical significance, the main reason

³⁵Note that I am not suggesting here that five data points will be sufficient to support a full variable analysis, but rather that models which are aware of (individually) correlated observations in their calculation procedures will be in a better position to evaluate such data in a meaningful way within a larger data set. Paolillo (2013:96) maintains that for a full variable linguistic analysis, “generally, 100 tokens per speaker is a minimum to ensure reliable estimates.”

being that the newer mixed-effects techniques have not been considered feasible yet (see e.g. Nobels 2013:52). Indeed, scholars have warned against over-reliance on powerful tools like statistical models where data are scarce, as is often the case in sociolinguistic research: “No statistical model can correct for an insufficiency of data, nor for failure to meet distributional assumptions.” (Paolillo 2013:114).

Resolving sampling problems is also of particular relevance for linguists interested both in the inter- and intra-speaker dimension. Arguably, limiting the sampled individual tokens in the statistical analysis of the aggregate group values, i.e. restricting the number of tokens collected for each speaker, as advised for the traditional models (see e.g. Tagliamonte and Baayen 2012:158, Paolillo 2013:115), might help resolve biases present in the data with traditional models. For an individual linguistic analysis, one could then proceed to embrace the full range of data available from a given individual, which (inconveniently) had to be left out in the statistical analysis. However, that solution is not optimal in that the sociolinguistic analysis and, say, syntactic analysis of individual grammars would then not be based on the same set of data. Additionally, this obviously also leaves us with less data to interpret, which translates into less evidence for distinguishing the linguistic properties of an individual, leading in turn to a less clear picture of that speaker’s grammatical constraints. Moreover, artificially limiting the set of data for one application and not the other also seems to go against the ‘unified’ approach to linguistic variation envisioned by Weinreich et al. (1968). In this respect, the mixed-effects approach is definitely an improvement over previous techniques, offering great potential despite being in its infancy in being applied to historical sociolinguistics.

Mixed-effects modelling offers a powerful new toolkit for statistical analysis (see e.g. Tagliamonte 2012:147ff.). However, here as elsewhere it is important to apply Occam’s razor, for in a sense, less is more as far as the model design is concerned. In order to include an effect it is important to make sure it offers a significant contribution to the model, in our case as measured by the *glmer* function of the *lme4* package in R. Moreover, a model with the effect in question must be a statistically significant improvement over an identical model lacking that very effect. One recommended way of achieving this is by model comparison as measured by the *anova* function of R (cf. Gries 2015b:115); this method is adopted and applied here (see also Viðarsson 2017b). Statistical modelling implies an abstraction over the data based on the effect interaction and this abstraction may be more or less faithful to the actually observed data; the predictions of a good model thus match relatively well to the observed data. To test the goodness of fit, the *C* index of concordance and Somers’ D_{xy} value were obtained using the *Hmisc* package, as described by Baayen (cf. 2008:281) (see also Tagliamonte 2012:149f.). Another measure is the classification accuracy of the linguistic variants of a model, calculated here on the basis of the *fitted* function (following Gries 2015b and accompanying data).³⁶ Finally, the fit has been plotted using the *plot.logistic.fit* function of the *languageR* package (cf. Baayen 2008) by fitting them on a line as a way to visualise how well the observed proportions match the mean expected probabilities.

³⁶The .r data file accompanying the article can be obtained at: http://www.linguistics.ucsb.edu/faculty/stgries/research/2015_STG_MultilevelModelingInCorpLing_Corpora_vpcs.r.

3 Verb-adverb placement

Introduction

A well-known feature of Icelandic syntax is the ability for the finite verb to occur in second position of both main and embedded clauses, immediately following an initial constituent, be that a subject, a topicalised (fronted) object or adverbial phrase. This is in stark contrast to most of the other Germanic languages, where verb second (V2) is typically a main clause phenomenon; in subject-initial embedded clauses, the finite verb (Vfin) canonically occurs later in these languages, e.g. clause-finally as in Dutch and German or in third position (V3) as in most of the other Scandinavian languages, thus following adverbs and the negation (see e.g. den Besten 1977, 1989, Eythórsson 1995, Eyþórsson 1997-1998, Vikner 1995, Zwart 1997, 2008, Þráinsson 2001, Þráinsson 2010, Holmberg 2015, Angantýsson 2011, 2017d, 2018). In the following, embedded subject-initial V2 and V3 clauses containing sentence-medial adverbs will be referred to throughout as **Vfin-Adv** and **Adv-Vfin**, respectively. Both configurations are exemplified immediately below based on 19th-century Icelandic data:

(28) *Verb-adverb placement: finite verb – adverb vs. adverb – finite verb*

- a. *eg finn eg get ei talað eða skrifað mikið um það.* (VFIN-ADV / V2)
I feel I can not speak or write much about that
'I feel that I cannot speak or write much about that.'
(GudMag-1819-00-00.txt)
- b. *og gefi það nú guð, að þetta bréf ei angri þig.* (ADV-VFIN / V3)
and give that now god that this letter not bothers you
'And I hope to God that this letter does not bother you.'
(GudMag-1819-00-00.txt)

As has frequently been noted in the literature, Adv-Vfin is indeed attested to some degree in subject-initial embedded clauses in Icelandic (Maling 1980, Sigurðsson 1986, 1989, Bobaljik and Þráinsson 1998, Angantýsson 2011, to appear, Angantýsson and Jonas 2016), although corpus-based evidence has suggested that it is a rare phenomenon in the modern variety (see Bobaljik and Þráinsson 1998, Angantýsson 2001, 2011, Heycock and Wallenberg 2013). Despite the paucity of Adv-Vfin in modern (post-19th century) texts, it is still judged acceptable by up to a third of the population on average, depending on the linguistic context—receiving the highest score (up to over 50%) in relative clauses (cf. Angantýsson 2011:74ff.). In contrast, traditional works such as Smári (1920), as well as more recent corpus-based research by Heycock and Wallenberg (2013), both indicate that Adv-Vfin was gaining ground in the period after

1600 up until the mid-19th century, whereafter there is a dramatic decrease in its use in published works.

More importantly in the present context, Pettersson (1988) suggests that social language ideological and normative factors prevented the Adv-Vfin order from becoming the canonical embedded word order in Icelandic (unlike e.g. Danish, Norwegian and Swedish) in that Vfin-Adv was selected as the written norm in Icelandic as opposed to Adv-Vfin e.g. in Swedish (cf. also Håkansson 2011; Gregersen and Pedersen 2000, Sandersen 2007 for similar claims regarding Danish).

In this chapter on verb-adverb placement, the social embedding of Adv-Vfin/Vfin-Adv will be the major point of attention, taking into account the negative social evaluation of the Adv-Vfin variant measured against actual language use as attested in three types of sources: newspapers, student essays and private letters. It will be shown that there is a downward trend in the frequency of Adv-Vfin over time. While Adv-Vfin is clearly more deeply entrenched in 19th-century Icelandic than the literature suggests, scribes in the higher echelons do tend to make more extensive use of the feature than, for instance, peasants and labourers. Interestingly, this is true of both male and female scribes, suggesting that the variation is more about social networks and engagement in similar practices in general than strictly about educational background.

Overall, the rate of Adv-Vfin decreases over time but in the case of the females, this trend is visible across social categories such as social rank, carrying over to the class of peasants and workers. While this is in apparent compliance with the frequently made claim that women tend to deviate less than men from prescribed linguistic norms (cf. Labov 2001:366f.), it is more likely to be due to a greater sensitivity among females than among males in the lower ranks to the lessened prestige of Danish in the latter half of the 19th century onwards than strictly about deviations from prescribed norms.

The outline of the chapter is as follows: Section 3.1 outlines the negative social evaluation of the Adv-Vfin variant, building on my own earlier work in Viðarsson (2014, 2016). Section 3.2 mainly deals with the historical background of verb-adverb placement in Icelandic prior to the 19th century (Section 3.2.1). Section 3.2.2 briefly presents a comparative perspective, focusing on Danish. Section 3.3 introduces the potentially relevant grammatical factors that may influence the choice between the two variants in 19th-century Icelandic and it also serves the function of circumscribing the variable context as defined here. Section 3.4 presents the results of three corpus studies, beginning with newspapers and periodicals, followed by results from the private letters and, finally, the student essays, with subsequent statistical analysis. Section 3.5 offers some discussion about the main patterns observed in the previous section and consequences for linguistic theory. Section 3.7 is an attempt at analysing the data syntactically, making use of a grammar competition model to account for the embedding of variation within the speakers' linguistic systems. The chapter concludes with a brief summary of the main findings.

3.1 Codification and implementation

19th-century Icelandic grammars converge insofar as the position of the finite verb is described using statements of a basic word order along the lines of (29):

(29) SUBJECT - FINITE VERB - ADVERBS - OBJECTS

This arrangement is stated explicitly in Rask's (1818:§424) grammar and Jónsson's (1893:38ff.) syntax of Icelandic. The same is at least implicit in Friðriksson's (1859, 1861) description of adverbs accompanying verbs, which is then exemplified with linguistic data adhering to the pattern in (29) above (see Friðriksson 1859:29). Interestingly, there is no mention at all of a different word order scheme for main and embedded clauses in the grammars. Given its prevalence in contemporary published texts, the absence of any explicit mention of the Adv-V_{fin} variant in 19th-century Icelandic grammar books is striking. Rather than acknowledging the existence of Adv-V_{fin} and denouncing it explicitly, the grammarians instead have given the false impression that the variation did not exist to begin with—effectively “erasing” the variability, to use Irvine and Gal's (1995, 2000) term (see Viðarsson 2016). As we will see, the absence of overt rejection in grammars should not be confused with tolerance towards variation.

It is not until very late in the period under study, in the early 20th century, that we find more explicit references to Adv-V_{fin} in grammars. The first is Jónsson's textbook grammar, published in three revised editions from 1909-1920. Following a statement about the canonical Icelandic clause structure, corresponding to the arrangement in (29) above, the reader is presented with sample clauses featuring the Adv-V_{fin} order and is given the task to correct the sentences (see e.g. Jónsson 1920:59, the first and second edition being identical in this regard). The reader is thus given an impression of Adv-V_{fin} as a linguistic error, no less. A similarly harsh negative evaluation is provided by Smári (1920:258) in a more widely known remark, describing Adv-V_{fin} as a “(foreign) practice” which “ought to be avoided” (*Varast skal þá (útlendu) málvenju*).

Widely argued to have been very influential with regard to the standardisation of Icelandic, a series of book reviews appeared in the journal *Fjölnir*, published between 1835 and 1847 (see e.g. Ottósson 1990:70ff.). In what appears to be the first published indication of the stigma attached to Adv-V_{fin}, the linguist Konráð Gíslason (1844:85) addresses the Adv-V_{fin} variant in a recent publication, stating in his review that the word order is “Danish-like” (*dönskuleg*) and ought instead to have been V_{fin}-Adv. However, despite the widely assumed importance of *Fjölnir*, instances of Adv-V_{fin} managed to slip into high-profile grammars including Halldór Kr. Friðriksson's (1861:vi, viii) morphology textbook, taught at Lærði skólinn, as well as Bjarni Jónsson's (1893:19) textbook on the syntax of Icelandic.

Further evidence of the social stigma concerning Adv-V_{fin} is Sigurður Gunnarsson's (1878) counter-review, a newspaper article written in reaction to a harsh, anonymous book review. The original reviewer had voiced complaints concerning supposed Danicisms (pertaining to word order and lexis) in a recent publication and Gunnarsson lashes out at the reviewer's own language use. Gunnarsson (1878:10) agrees with the reviewer that “clauses too often have un-Icelandic word order” (*of víða óíslenzk orðaskipun í málsgreinum*) in the reviewed work. His criticism of the reviewer mainly targets Adv-

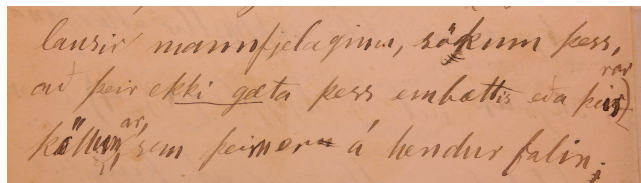


Figure 3.3. Teacher correcting Adv-Vfin in a student essay (1861, 4th grade).

Vfin constructions, which Gunnarsson iconically claims to have “some rotten³⁷ taste of Danish” (*eitthvert danskt óbragð*). He further remarks that “words are seldom so arranged in daily speech” (*Sjaldan er orðum svo skipað í daglegu tali*), giving the impression that Adv-Vfin was marginal. While such anecdotal evidence is arguably a poor indicator of the actual scope of variation during the last quarter of the century, it speaks volumes for its perceived extent. In a later, 20th-century book review by the linguist Björn Guðfinnsson (1940), entitled “An assault on the Icelandic language” (*Tilræði við íslenskt mál*), he refers to Adv-Vfin as *röng staða neitunar* or the ‘wrong position of negation’ (? :74), shown in italics to denote *helztu málleysurnar* ‘the main linguistic aberrations’ (1940:74, fn. 1).

Important evidence of a negative evaluation of Adv-Vfin already in the 19th century is provided by teachers’ corrections (usually by underlining) in the Reykjavík Grammar School student essays. My preliminary survey of this material reveals that Adv-Vfin was indeed considered sub-standard at the Grammar School, although it was not systematically corrected in the essays, a substantial number of attestations slipping through. The oldest corrections of Adv-Vfin that I have found so far date from 1861, exemplified in Figure 3.3,³⁸ but the corrections likely go further back (for more examples, see Viðarsson 2014, 2016).³⁹

Nothing suggests that consistent verb-adverb placement had any symbolic value during the early 19th century prior to Konráð Gíslason’s (1844) raising of awareness and explicit rejection of the Adv-Vfin variant. For one, some of Sveinbjörn Egilsson’s ceremonious opening/closing speeches at the Grammar School covering a period of roughly 30 years, first at Bessastaðir and later in Reykjavík, have been published by

³⁷Or: ‘bad taste’. However, while perhaps more canonical, *bad* considerably milder in tone than in the original *óbragð*, lit. ‘un-taste’. The fact that the word usually refers to a foul or rotten taste when used in corpora that I have consulted, I have decided to render *óbragð* here as *rotten flavour*. The reference to Shakespeare does not hurt either, although no such connotation is suggested in the original.

³⁸Viðarsson (2016:157, Fig. 10.1) erroneously dated this correction to 1860.

³⁹To give some idea about the correction ratio, I surveyed Adv-Vfin in 31 essays from the same 2nd grade in two separate examination rounds, both dating from 1890:

(1) 14 essays (19.2.1890)	(2) 17 essays (20.6.1890)
Adv-Vfin: 16 (8 of which Neg-Vfin)	Adv-Vfin: 7 (2 of which Neg-Vfin)
Corrected: 6 (all Neg-Vfin)	Corrected: 0

The teachers/examiners in the first round were Halldór Kr. Friðriksson and Pálmi Jónsson (both corrected the variant) as opposed to Jón Þorláksson, Pálmi Jónsson and Hannes Þorsteinsson in the second round. All in all, roughly 25% of the attestations thus appear to have been corrected, whereas considering only Neg-Vfin, the correction rate raises to at least 60%. Interestingly, the corrections that I have found in previous periods have also only targeted Neg-Vfin. The special status of negation is an aspect that merits further study.

Haraldur Sigurðsson (1968) and these quite regularly feature the Adv-Vfin order (but always alongside Vfin-Adv). Examples from various periods are shown in (30) (date of speech is given in brackets, page numbers refer to H. Sigurðsson's edition):⁴⁰

- (30) a. *þeirri ákjæru, að þecking og dygð ei séu ætíð samfara* (1819, 10)
the charge that knowledge and virtue not are always associated
- b. *samveran með honum varð yður ónýt, ef þér ekki brúkuðu hana*
the communion with him became you useless if you not used her
(1828, 47)
- c. *setji sig ekki upp yfir aðra, þar sem bein skylda ekki býður það*
puts REFL not up over others, where direct duty not demands it
(1834, 52)
- d. *vér verðum [...] að játa, að vér enn þekkjum ekkert, eins og það er.*
we have to admit that we still know nothing as it is
(1846, 78)
- e. *sem þér svo leingi hafið hreykt af* (1850, 88)
which you so long have boasted of

Similar evidence is provided by the periodical *Ármann á Alþingi* (1829-1832), featuring fictitious characters that represented vastly different political attitudes and value orientations, partly reflected also in their (non-)use of foreign language features; verb-adverb placement is clearly not among those either (see Viðarsson 2014:16f.).

The stigma attached to the Adv-Vfin variant as a part of (the implementation of) the standard language has received scant attention in previous studies of the phenomenon and it appears that its low occurrence rate in texts from the 20th-century onwards has not been associated with its non-standard status. Angantýsson (2001:96, fn. 1, 2011:62f., fn. 41) briefly mentions the negative evaluation of Adv-Vfin by Smári (1920:258) and Böðvarsson (1992:263). The latter of these two authors does not reject Adv-Vfin but refers to it as a “bad habit” when used “unnecessarily” (*ósið ... að þarflausu*). Maling (1980:176f., fn. 4) claims that Modern Icelandic permits Adv-Vfin for emphasis with certain adverbs (such as *fyrir löngu* ‘long ago’ and *loksins* ‘finally’) but not with negation, the latter she states “is felt to be a Danicism.” With negation, Maling (1980) judges Adv-Vfin as ungrammatical (a full star notation) and the same is stated or implied e.g. by Platzack (1986:177), Sigmundsson (1987:102f.) and Thráinsson (1994:184). Sigurðsson (1986) appears to be first in more generally acknowledging the grammaticality of the Adv-Vfin pattern in Icelandic, even with negation, stating that “its use seems to be on the increase” (1986:146). In my view, this reluctance among most scholars to fully acknowledge the presence (and perhaps rise) of Adv-Vfin in Icelandic up until quite recently is clearly rooted in standard language ideology and is appropriately analysed in terms of erasure.

⁴⁰Interestingly, the same piece of text fragment in (30b) from 1828 is recycled in 1851, where the word order mysteriously changes from Adv-Vfin to Vfin-Adv, with subsequent Object Shift (see e.g. Section 3.7):

- (1) *samveran með honum varð yður ónýt, ef þér brúkuðu hana ekki* (1851, p. 93)
togetherness-the with him became you ruined if you used her not

Was substituting Adv-Vfin for Vfin-Adv an act of will in order to conform to the emerging standard language or was it purely whimsical?

3.2 A diachronic sketch of Adv-Vfin

3.2.1 The Adv-Vfin order in older Icelandic

Previous studies have revealed quite extensive variation with regard to the position of the finite verb in Icelandic subject-initial embedded clauses, especially in the period 1600-1850 (cf. Heycock and Wallenberg 2013). Incidentally, this period in the history of Icelandic is often regarded as an era of ‘degradation’ in the nationalistic discourse of the 19th and 20th centuries. Thus, Adv-Vfin is usually assumed to be closely associated with supposedly impure, un-Icelandic, foreign (i.e. Danish) features, taken to be a characteristic of the language of the educated elite, as opposed to the supposedly pure language of the common people (cf. e.g. a general note in Þorkelsson 1870:82, and Smári 1920:13f. on Adv-Vfin specifically). Some of these traditional assumptions are taken over by Heycock and Wallenberg (2013), who suggest that the rise of Adv-Vfin in 1600-1850 was a contrived trait of the very formal written language only. Their hypothesis is that certain scribes attempted to calque the syntax of Danish, although they never fully succeeded due to interference from their native Icelandic grammar. More specifically, they hypothesise that the rise in Adv-Vfin order in Icelandic was due to a conscious imitation of Danish authors and the language of the local Danish colonial aristocracy, also pointing out that many educated writers received their formal education at the University of Copenhagen.

While this is not necessarily obvious from the historical record, the previous literature has apparently never entertained the hypothesis that Adv-Vfin may have Icelandic roots. In fact, Gunnarsson (1878:10) explicitly mentions an Old Norse-Icelandic example known to him from the 13th-century Egils saga, only to discount it. Clearly wishing to undermine this piece of data, he suggests it must be due to a scribal error. Corroborating Håkansson’s (2011) observations regarding the presence of Adv-Vfin in Old Swedish, Heycock and Wallenberg (2013) also indicate that Adv-Vfin was an option during the Old Icelandic period, albeit a rarely used one. They do not elaborate on this point but from their graph (2013:153, Fig. 3) can be gleaned that sporadic examples of Adv-Vfin are attested as early as the 12th century. As far as I can see, examples from this early period always feature a subject pronoun, just as observed for Old Swedish by Håkansson (2011), and Adv-Vfin with nominal subjects does not start to appear until Heycock and Wallenberg (2013) observe a rise in the frequency of Adv-Vfin, as late as the 17th and 18th centuries. As their study shows, the use of Adv-Vfin peaks around the 18th/19th century, mainly in the same environments where the V2 word order (=Vfin-Adv) is not found in Mainland Scandinavian, suggesting that it partly had Danish characteristics (for a similar finding based on a sample of 19th-century newspaper texts, see Viðarsson 2014 and discussion below).

A potentially confounding factor concerns the fact that Heycock and Wallenberg (2013) did not study the use of Adv-Vfin in less literary contexts for comparison but confined themselves to the Icelandic Parsed Historical Corpus (Wallenberg et al. 2011) which mostly contains rather formal/distant narrative texts. Heycock and Wallenberg’s (2013) suggestion that Adv-Vfin was a contrived trait restricted to a subset of native Icelandic speakers consciously trying to imitate Danish, yet failing to do so completely, thus rests upon a number of traditional assumptions they make about Icelandic language

history. This is not to say that they cannot be right, but the calquing hypothesis clearly remains to be tested.

Hróarsdóttir (1998) has already voiced skepticism over treating Adv-Vfin as a Danish feature ignited by studies abroad or otherwise confined to lettered males. The main reason for her skepticism is that Adv-Vfin is also found in the letters of 19th-century women lacking the educational background in question. This state of affairs suggests that Adv-Vfin probably constituted a part of (some) speakers' casual code rather than being something quite as artificial as suggested by Heycock and Wallenberg (2013) and traditional scholarship such as Smári (1920). Although Hróarsdóttir's study was confined to a "numerical study" of the Adv-Vfin phenomenon in the sense of Labov (2016:582), i.e. one which does not report "the *absence* of variants" (here: Vfin-Adv), and, furthermore, only considers sentences collected independently for the study of a change from OV/VO to VO in the Icelandic verb phrase, it still revealed more variation than expected on this traditional view.

Moreover, Heycock and Wallenberg's (2013) corpus study indicates that the Adv-Vfin order is found sporadically in the Old Icelandic period, reporting some hits around 1150 and again in 1300 and 1400. In the early modern period, Adv-Vfin shows up again in the texts around 1600 and increases gradually until around 1850, where the innovation apparently fades out again, eventually reaching 0% (see Heycock and Wallenberg 2013:153, Figure 3).⁴¹

A closer look at the Old Norse part of the Tagged Icelandic Corpus (MÍM), searching for *conjunction-pronoun/nominal-adverb-V_{fin}*, reveals a number examples of what appears to be the Adv-Vfin variant (see (183), Appendix A, page 249). Being historical data, philological factors are a potential complication as not all of these texts are actually preserved in medieval manuscripts, surviving in later copies. At least four of these examples unquestionably appear in manuscripts from the Old Norse-Icelandic period based on information obtained from *Ordbog over det norrøne prosasprog*.⁴² Among the philologically solid ones, only (31) (=183a) features negation:

- (31) *Kom hann eitt sinn á minn fund og sagði eg honum að eg ekki vildi hér*
 came he one time on my meeting and told I him that I not wanted here

⁴¹These results appear to contradict Rögnvaldsson (1995), who states that in a corpus containing the vast majority of the Old Icelandic canon, he found only one instance of the Adv-Vfin variant. There are two likely reasons for this discrepancy. First, it could simply be due to the selection of texts, although there is a considerable overlap between the texts used by Heycock and Wallenberg (2013) and Rögnvaldsson (1995). Second, Heycock and Wallenberg (2013) defined the Vfin-Adv/Adv-Vfin variable quite broadly, which means that some of their hits may actually have been disregarded by Rögnvaldsson (1995) as not belonging to the linguistic variable in question (see Section 3.3).

⁴²These are (183a), (183g) and (183h), all dating from c. 1330-1370, and (183f) from c. 1500. Examples (183a-d) are most reminiscent of the data discussed in this chapter featuring the negation *ekki/eigi* and the temporal adverbs *senn* 'soon' and *þá* 'then', which at least from a linguistic perspective seem rather solid. Then there are adverbs which are lower on Cinque's hierarchy: *vel/betur/best* 'well/better/best', and finally *svo* 'so, thus'. Not all of these are reliable medial adverbs in that most, except negation, may probably occur in post-VP position. However, they arguably all have the one property in common of not being able to occur in the verb-third configuration in main clauses. Note also that these examples involve pronominal subjects, which is the only type of subject possible in this construction in Old Swedish (cf. Håkansson 2011)—nominal subjects arguably being an extension of an earlier pattern which only accommodated (light) pronouns.

TEXT	AUTHOR/*TRANSLATOR	ADV-VFIN	<i>n</i>
<i>Píslarsaga</i> (1658-59)	Jón Magnússon (1610-1696)	3.0%	164
<i>Vídalínspostilla</i> (1718-1720)	(*)Jón Vídalín (1666-1720)	32.5%	83
<i>Nikulás Klím</i> (1745)	*Jón Ólafsson (1705-1779)	8.4%	179
<i>Uppkast ... um brúðkaupssiðu</i> (1757-1767)	Eggert Ólafsson (1726-68)	10.3%	58
<i>Ævisaga</i> (1784-1791)	Jón Steingrímsson (1728-1791)	6.5%	340
<i>Stuttur siðalærðómur ...</i> (1799)	*Guðlaugur Sveinsson (1731-1807)	58.6%	70

Table 3.2. Verb-adverb placement in texts from 1600-1800 (*Íslenskt textasafn*).

í landi vistir hans af þeim sökum sem yður er áður kunnigt.
 in land supplies his of the reasons which you are before known
 ‘... I told him that I did not want his supplies here in this country for reasons
 already known to you.’ (MÍM: Egils saga)

While this does not necessarily mean the rest of the examples must be disregarded as not representing Old Icelandic, it does make it more difficult to maintain that Adv-Vfin occurred in the medieval period, although more research is clearly needed. I take the lack of clear-cut evidence that the Adv-Vfin variant existed in Old Icelandic, beyond a handful of cases (all of which featured pronominal subjects as in Old Swedish), to suggest that ‘real’ violations of the V2 constraint are a more recent phenomenon.

Turning to the early modern period, I have also studied the use of the Vfin-Adv/Adv-Vfin variable in texts from 1600-1800 in the Icelandic text collection of the Árni Magnússon Institute for Icelandic Studies (*Íslenskt textasafn*), confined to subject-initial embedded clauses with negation.⁴³ Studying the phenomenon is complicated by the fact that V_{final} is also an option for some of the writers during the early modern period and it is often the case that such an analysis cannot be ruled out completely. Adv-Vfin should therefore be interpreted here as *third position or later*. Table 3.2 lists the texts, along with year written, the author/translator, percentage of Adv-Vfin vis-à-vis Vfin-Adv and the number of attestations.⁴⁴

There is clearly a lot of variation in this period, but strikingly, every text had at least some instances of Adv-Vfin. The highest frequency is found in the translated texts. However, from this perspective, it is surprising that Adv-Vfin even figures in the writing of an author like Eggert Ólafsson, a fervent language purist who criticised his contemporaries for the use of foreign linguistic features.

3.2.2 Historical variation in Danish

As Icelandic was in prolonged contact with Danish, it is important that we begin with a picture of the attested variation in verb-adverb placement in Danish during the early modern period and in the 19th century in particular. To what degree was the older Vfin-Adv word order still attested in Danish subordinate clauses, spoken and written? Just how unequivocal was the Danish input, which the Icelandic use may (or may not) have

⁴³The text *Vídalínspostilla* is in six parts but I only studied two (*vidpost* and *vidalin1*). In addition, I used the more recent edition of *Ævisaga síra Jóns Steingrímssonar* (ed. Sigtryggsson 2011).

⁴⁴In Table 3.2 I have left out the cookbook *Einfalt matreiðsluasakver fyrir heldri manna húsfreyjur* (1800) due to a very small number of examples, seven in total, two of which featured the Adv-Vfin order.

been modelled after, whether spoken/written in Reykjavík, Copenhagen or elsewhere, in terms of the position of the finite verb?

Late medieval and early modern Danish sources show a clear trend towards replacing the older embedded Vfin-Adv with Adv-Vfin word order (see e.g. Gregersen and Pedersen 2000, Sundquist 2003). It appears that the Adv-Vfin order was innovated in negated relative clauses in the early 14th century, reaching levels as high as 41% in diplomas from 1403-1482 (see Gregersen and Pedersen 2000:407, with references). Based on a study of texts from 1500-1700, Sundquist (2003) reports a drop in the frequency of Vfin-Adv from approximately 45% in the first half of the 16th century to 12% in the 17th century, with considerable intra-speaker variation (2003:241).

Areas that were less affected by language standardisation, especially in the Jutland dialects, appear to exhibit a greater degree of variation, with more embedded Vfin-Adv in the western than in the eastern dialects (cf. Pedersen 1996:244).⁴⁵ A recent study of spoken Danish data at least partly corroborates this, indicating that embedded Vfin-Adv is vastly more frequent than the literature suggests and may in some cases even outnumber the so-called canonical embedded Adv-Vfin order (see Jensen 2011, Jensen and Christensen 2013).⁴⁶

For 19th-century Danish, Sandersen (2003, 2007) observes on the basis of letters of private soldiers during the Three Year's War (1848-50) that embedded Adv-Vfin order was firmly in place around 1850 in the written language. However, she maintains that there was variation in the spoken language coming through, appearing mainly in the writing of those with the highest number of deviant/non-standard features in their letters (Sandersen 2007:273). Sandersen does not elaborate further on this point, but her (made-up) linguistic examples demonstrating the linguistic variable suggest that declaratives embedded under verbs of saying may be typical cases.

More detail is provided by Sandersen (2003:373f.) on the distribution across different authors and on types of adverbs, but not regarding clause types. The overall frequency of embedded Vfin-Adv in the most advanced group of letter-writers is in the range 5-6.5%, depending on the group of adverbs included, whereas in the least advanced group, the frequency of embedded Vfin-Adv is much higher, between 31-36%. Furthermore, as Sandersen (2003:371) points out, the soldiers of the Three Year's War belong to the generation in which language standardisation came to encompass the whole population through the implementation of compulsory education in 1814. The apparent lack of embedded Vfin-Adv examples in the most advanced letter-writers might therefore be interpreted as an avoidance strategy of a spoken feature on their part (see also Sandersen 2003:372). The number of examples is admittedly not very high, especially within the least advanced group of letter-writers, but still suggestive of differences between the two groups.

⁴⁵ In Jutland recorded 20th-century dialect data, for instance, 55% of subordinate clauses were shown to have the embedded Vfin-Adv order, most of which were final/consecutive clauses, but the order was also frequently found in indirect questions and quite frequent in relative clauses (cf. Pedersen 1996, see also Bentzen 2014). Embedded Vfin-Adv is also reported to be frequent in present-day spoken language in Copenhagen, especially in clauses introduced by *at* 'that', *fordi* 'because' and *hv*-questions (Pedersen 1996:244).

⁴⁶ Bentzen (2014) points out that Jensen's (2011) study reveals a much lower frequency of Vfin-Adv in relative clauses than Pedersen's earlier study of the Jutland area, stating that this is likely due to the informants in Pedersen's study being older than the ones included in Jensen's study (cf. Bentzen 2014:245).

Vikner (1994:146, note 6) also mentions the use of embedded Vfin-Adv in 19th-century Danish, drawing a parallel between the situation then and the Vfin-Adv/Adv-Vfin variation found in Modern Faroese, where Vfin-Adv is still acceptable to many speakers (see e.g. Heycock et al. 2012, Angantýsson 2018). However, Vikner (1994) seems to suggest that Vfin-Adv may have been associated with a particular register, pointing out that the example used to demonstrate this property is found in “the written down version of a folk tale” (citing Skautrup 1953:233):

- (32) *Prindsessen havde faaet Vulle saa kjær, at hun vilde ikke have Kongen*
 Princess-the had got Vulle so fond that she would not have king-the
 ‘The princess fell so in love with Vulle that she would not have the king.’

Skautrup (1953:233) also appears to be describing stylistic variation when he claims that in the narrative language (*fortællersproget*), adverb placement in subordinate clauses is the old one (i.e. Vfin-Adv) as a rule (*som regel*), explicitly defined in terms of subject-verb adjacency. Later, Skautrup (1953:241) also claims that the Vfin-Adv word order is found in *at*-clauses in the language of H.C. Andersen (1805-1875), giving examples of negation embedded under the matrix verbs *synes* ‘think, feel’ and *sige* ‘say’. The status of consequence of degree clauses as in (32) is perhaps not as clear-cut as complements of verbs of saying in terms of allowing root phenomena, at least not given the traditional conception of ‘bridge verbs’ (cf. Vikner 1995). Given that both of these types may be asserted (cf. Julien 2009, see also Heycock et al. 2012, Heycock and Wallenberg 2013) these distributional facts do give the impression that the so-called old word order was available (for H.C. Andersen) in root-like subordinate clauses (see further discussion below). Interestingly, Andersen’s place of origin, Odense, is among the dialect areas where previous research has shown Vfin-Adv to be possible in typical, non-V2 contexts such as relatives (cf. Pedersen 1996, Bentzen 2014).

In the 19th-century Danish material studied by Gregersen and Pedersen (2000), the frequency of the Vfin-Adv word order is very low in most speakers writing after 1850 but still is quite amply attested e.g. in Grundtvig’s treatises and sermons, written between 1816-1832, which had nearly 40% Vfin-Adv word order with a higher than average percentage reported mainly in *that*-clauses and consecutive clauses. According to Gregersen and Pedersen (2000), H.C. Andersen’s fairy tales (1835-1872) feature the Vfin-Adv order at an average frequency of 20%, mainly in *that*-clauses and consecutives, but this word order is also found in his private letters, occurring at an avg. rate of 14%. While this discussion does not really weigh in on the root versus non-root distinction mentioned above, it does seem that the variation is very rare in clear non-V2 (non-root) environments (see next section).

The variation attested in 19th-century Danish thus shows that verb-adverb placement may have been rather flexible in principle. Importantly, this suggests that if we were treat Danish as the “target grammar” on which certain Icelandic scribes may have based their written norms (cf. e.g. Heycock and Wallenberg 2013), it should be kept in mind that the target grammar itself was also inherently variable.

3.3 Grammatical factors

3.3.1 Clause type

Historical variation in Icelandic verb-adverb placement has been studied in the greatest syntactic detail by Heycock and Wallenberg (2013). On the basis of data gathered from the Icelandic Parsed Historical Corpus (Wallenberg et al. 2011), they show that the syntactic environments argued to be relevant for the distribution of embedded V_{fin}-Adv vs. Adv-V_{fin} in Mainland Scandinavian were also an important conditioning factor in Icelandic historical record during the period 1600-1850. The same effect was reproduced by Viðarsson (2014) in a pilot study of Icelandic newspapers and periodicals from three periods: 1800-1850, 1875 and 1900. The distinctions made by Heycock and Wallenberg (2013), also adopted in Viðarsson (2014), involved a three-way approximation of a much finer distinction between clause types based on semantic and structural factors, dividing the data into non-V2 environments (comparatives, relatives, indirect questions), V2 environments (declaratives, consequence of degree clauses) and mixed environments (other adverbials). This classification refers to the typical distributional properties of V_{fin}-Adv/Adv-V_{fin} in subject-initial embedded clauses, found, for example, in standard varieties of present-day Danish.

In comparing Faroese and Danish, Heycock et al. (2012) distinguish between eight different clause type categories, the last four of which (=33e-h) are least likely to allow embedded V2:

- | | | |
|------|--------------------------|------------------------|
| (33) | a. Declarative clauses | e. Adverbial clauses |
| | b. Result clauses | f. Indirect questions |
| | c. Consequence of degree | g. Relative clauses |
| | d. Cause clauses | h. Conditional clauses |

I follow Heycock et al.'s (2010) coarse split between typical V2 and non-V2 environments, in addition to placing the rest of the adverbials in (33e) in the mixed category, along the lines of Heycock and Wallenberg (2013). As Heycock and Wallenberg (2013:142) remark, adverbial clauses are “something of a mixed bag” when it comes to permitting or excluding embedded V2.⁴⁷ For a more accurate comparison with the rest of the Scandinavian languages, I also attempt to categorise a subset of the data in terms of the more fine-grained distinctions known to be relevant to the distribution of V_{fin}-Adv/Adv-V_{fin} in these languages, in particular with regard to different classes of declarative complement clauses. However, first, let us briefly consider adverbial clauses in order to determine which adverbials belong to the ‘non-V2’ and ‘V2’ environments, respectively.

In recent work, Haegeman (2004, 2012) proposes a dichotomy between *central* vs. *peripheral* adverbial clauses depending on whether or not these allow main clause phenomena such as argument fronting (topicalisation). Peripheral adverbials are argued

⁴⁷Due to slightly different partitioning, adverbial clauses did not include comparative clauses in Heycock and Wallenberg (2013), which they treated as a distinct (non-V2) category. In contrast, Heycock et al. (2012) separated conditional clauses (non-V2) and cause clauses (V2) from the mixed category of adverbial clauses, leaving in comparatives. Here, I follow the partitioning scheme in Heycock et al. (2012).

to be less integrated than central adverbials and may pattern with main clauses with regard to argument fronting. Also, a single conjunction can sometimes introduce a central adverbial clause in one context and a peripheral in another. A case in point is *while*. Depending on whether it has a temporal specification (event reference) or whether it merely has a discourse structuring function (lacking the temporal sense), *while* can carry either function. This is exemplified in (34) (see Haegeman 2012:160, with references):

- (34) a. According to Smith, a group of Arkansas state troopers who worked for Clinton **while** he was governor wanted to go public with tales of Clinton's womanising.
b. **While** his support for women priests and gay partnerships might label him as liberal, this would be a misleading way of depicting him uncompromisingly orthodox espousal of Christian belief.

Only the peripheral (non-temporal) type allows fronting (cf. Haegeman 2012:159, with references):

- (35) His face not many admired, **while** *his character* still fewer felt they could praise.

While this distinction is thus highly sensitive to interpretation, Haegeman argues that central vs. peripheral clauses manifest a difference in their syntactic structure. Peripheral adverbial clauses are assumed to have an additional syntactic position, viz. a Topic Phrase within an articulated CP structure, which can accommodate fronted arguments, adverbs, etc., lacking in central adverbial clauses. The expectation would then be that peripheral clauses ought to (be able to) pattern with main clauses with regard to verb-adverb placement in Mainland Scandinavian, by means of the available Topic position, whereas central adverbial clauses ought to pattern with typical non-V2 environments such as relative clauses and indirect questions.

Angantýsson and Jonas (2016) have recently argued that embedded topicalisation in Icelandic adverbial clauses fits quite nicely with Haegeman's central vs. peripheral dichotomy. Moreover, embedded Adv-Vfin word order and embedded topicalisation in present-day Icelandic are argued to be, in a way, in complementary distribution (cf. Angantýsson 2011, 2017a, Angantýsson and Jonas 2016), the proposal being that peripheral subclauses have a more elaborated left periphery, similar to main clauses, where Adv-Vfin (of this type) is not possible. I come back to this in Section 3.7.

With regard to declarative complement clauses, it is frequently observed that Vfin-Adv/Adv-Vfin patterns in markedly different ways in the Mainland Scandinavian languages depending on the type of *that*-clause. In contrast, the literature on Icelandic Adv-Vfin merely states that the Adv-Vfin pattern is usually very degraded in *that*-clauses, with no reference to certain types of *that*-complement clauses being any better or worse than others. Thus, if the availability of Icelandic Adv-Vfin can be shown to interact with *that*-clause types along the lines of Mainland Scandinavian, that might be taken as evidence in favour of analysing Icelandic on par with, say, Faroese, where there is variation between an earlier V-to-I grammar and a more recent Danish-like V-to-C grammar (cf. Heycock et al. 2010, 2012, Angantýsson 2018). In Section 3.7, I will try to argue for such an analysis for 19th-century Icelandic as well.

The availability of embedded V2 in Mainland Scandinavian is highly sensitive to aspects of meaning, viz. whether or not a sentence is asserted (cf. e.g. Julien 2009, see also Platzack 1986:224). In what follows, I take this to be essentially the same property as referred to by, e.g. Hooper and Thompson (1973) as the ‘core meaning’ or ‘main proposition’, by Erteschik-Shir (1973) as ‘the dominant part of the sentence’ or ‘the central piece of information’, and by Wiklund et al. (2009) as the ‘main point of utterance’. Hooper and Thompson (1973:473f.) make a widely-used five-way distinction between verbs in matrix clauses taking *that* complements which, in their view, are non-factive, viz. Class A, B and C, and factive, viz. Class E, D. These are shown in (36)-(37):

	NON-FACTIVE		
	A	B	C
	say	suppose	be (un)likely
	report	believe	be (im)possible
	exclaim	think	be (im)probable
(36)	assert	expect	doubt
	claim	guess	deny
	vow	imagine	
	be true	it seems	
	be certain	it happens	
	be sure	it appears	
	be obvious		
	FACTIVE		
	D	E	
	resent	realize	
	regret	learn	
(37)	be sorry	find out	
	be surprised	discover	
	bother	know	
	be odd	see	
	be strange	recognize	
	be interesting		

Hooper and Thompson (1973) argue that some of these classes have properties in common with main clauses in that they allow root transformations, e.g. VP-preposing, topicalisation and left/right-dislocation, whereas others disallow such operations. A main clause such as (38a) is an example of a main assertion (or ‘speaker assertion’), whereas (38b) in addition has an ‘indirect assertion’ associated with the matrix verb (Hooper & Thompson 1973:475):

- (38) a. It’s just started to rain.
b. He said it’s just started to rain.

Example (38b) thus has two assertions, namely:

- (39) He said X.
(40) It’s just started to rain.

		Swedish	Norwegian	Faroese	Icelandic
Class A	V>Neg	✓	✓	✓	✓
	Top	✓	✓	✓	✓
Class B	V>Neg	✓	✓	✓	✓
	Top	✓	✓	✓	✓
Class C	V>Neg	*	*	✓	✓
	Top	*	*	*	*
Class D	V>Neg	*	*	✓	✓
	Top	*	*	*	*
Class E	V>Neg	✓	✓	✓	✓
	Top	✓	✓	✓	✓

Table 3.3. The distribution of embedded V2 (Wiklund et al. 2009:1921).

Hooper and Thompson (1973) show that root phenomena are possible with Class A, B and E (much as in (35) above), which are assertive, but ungrammatical for most speakers with classes C and D, which are both non-assertive. Julien (2008, 2009) provides corpus-based evidence for the relevance of this important distinction for Vfin-Adv in the Mainland Scandinavian languages. Class A, B and E correspond roughly to Vikner's (1995) bridge verbs, and Class C and D to their opposites, which have been influential concepts in analyses concerning the availability of embedded V2 in the Scandinavian languages and beyond (see Vikner 1995, Þráinsson 2001, Þráinsson 2003, Heycock et al. 2010, Angantýsson 2011). For Icelandic, Swedish, Norwegian and Faroese, Wiklund et al. (2009) provide the neat overview in Table 3.3, including whether or not each class permits embedded topicalisation. Similar results are reported by Angantýsson (2011:140) based on a large judgment study, although *that*-clauses that were complements to Class C and D predicates were accepted by some speakers (for discussion, see also Hrafnbjargarson and Wiklund 2009, Heycock et al. 2010, Þráinsson 2011, Gärtner 2016).

The results of two judgment studies on Faroese carried out by Petersen (2000) and Þráinsson (2003), respectively, both showed a clear difference between bridge (*sigja* 'say') and non-bridge verbs (*harma* 'regret', *tað er spell* 'it is bad') with regard to Vfin-Adv. Þráinsson's (2003:176) corpus study on Faroese texts, written by 14 different authors born between 1819 and 1950, similarly revealed a very clear contrast between bridge and non-bridge contexts. The proportion of Vfin-Adv was high (80%) in complements of bridge verbs but much lower in non-bridge complements (37%), on par with adverbial clauses (49%), indirect questions (40%) and relative clauses (24%). Note that Petersen (2010:144) suggests that the innovation of Adv-Vfin is due to Faroese-Danish language contact. As we will see below, such a mixed, non-categorical distribution is strikingly similar to some 19th-century speakers as well, suggesting that they have internalised a grammar where the position of the finite verb is partly sensitive to whether or not a sentence is asserted.

In addition to the five types identified by Hooper and Thompson (1973), Julien (2009) discusses Vfin-Adv (=embedded V2) in other contexts, viz. in copula constructions, in complements to (pro)nouns, in causal clauses introduced, e.g., by *fordi at*

‘because (that)’ and in consequence of degree clauses introduced by *så X at* ‘so X that’. All of these are (or can be) asserted, according to Julien (2009:14-20) and thus allow Vfin-Adv in Mainland Scandinavian, as exemplified for Danish below:

(41) *Other contexts: Often assertive*

- a. *Resultatet er, at borgerne får ikke den hjælp, de har brug for*
 the result is that the citizens get not the help they have use for
 ‘The result is that the citizens don’t get the help they need.’ (2009:15)
- b. *Jeg siger det bare fordi at man skal ikke tro at jeg hader alle tyskere*
 I say it just because that one shall not think that I hate all Germans
 ‘I say it only because one mustn’t think that I hate all Germans.’ (2009:17)
- c. *Hun var så optaget at hun havde ikke lagt mærke til at han var kommet hjem*
 she was so busy that she had not noticed that he was come home
 ‘She was so busy that she had not noticed that he had come home.’ (2009:19)

The type in (41a) is assertive, fully on par with Type A above. Furthermore, it is possible to render types (41b,c) pragmatically equivalent to an asserted main clause by paraphrasing it, indicating that such embedded clauses with the Vfin-Adv order are, indeed, asserted as well. Based on the above observations, Julien claims that the availability of embedded V2 is not simply a property of the matrix predicate, but rather whether or not the matrix clause as a whole allows for an asserted embedded clause.

3.3.2 Adverb type

The literature on verb-adverb placement in Icelandic and other closely related languages largely bypasses problems that different classes of adverbs present by focussing on a particular class of so-called central adverbs. These are mainly sentential negation and temporal or frequency adverbs such as *aldrei* ‘never’, *alltaf* ‘always’, *sjaldan* ‘seldom’ as well as *loksins* ‘finally’, that have distinct properties in terms of their structural position (see e.g. Bobaljik and Thráinsson 1998, Angantýsson 2001, 2007, 2011, Thráinsson 2007:38ff., 2010). Unlike focus particles like *bara* ‘merely, just’ and speaker-oriented adverbs like *sennilega* ‘probably’ or *því miður* ‘unfortunately’, which have a relatively flexible position, the central adverbs in question, which I refer to as sentence-medial adverbs, are frequently argued to have a relatively fixed position in these languages, regardless of the analysis assumed. This difference is exemplified in (42) and (43) below (based on Thráinsson 2010:1076f., see also Maling 1980, Thráinsson 1986, 2010, Rögnvaldsson and Thráinsson 1990, Angantýsson 2001, 2007, 2011):⁴⁸

- (42) a. *Nemandinn bara/sennilega náði ekki prófinu*
 the student just/probably passed not the test

⁴⁸To the best of my knowledge, no research has been done on the position of focus particles in Icelandic main clauses from a historical perspective. As far as I have been able to see, these elements follow the finite verb in main clauses during the early modern period and are first found preceding the verb in embedded contexts. More research is needed in order to discover the chronology of these different patterns and whether the variation indeed always starts out in embedded contexts before spreading to main clauses.

- b. *Nemandinn náði bara/sennilega ekki prófinu*
the student passed just/probably not the test
- c. **Nemandinn ekki/aldrei/loksins náði prófinu*
the student not/never/finally passed the test
- d. *Nemandinn náði ekki/aldrei/loksins prófinu*
the student passed not/never/finally the test
- (43) a. *Þetta er prófið sem hann bara/sennilega/ekki/aldrei/loksins náði*
this is the exam that he just/probably/not/never/finally passed
- b. *Þetta er prófið sem hann náði bara/sennilega/ekki/aldrei/loksins*
this is the exam that he passed just/probably/not/never/finally

There are not only upper but also lower bounds on the structural position of these sentence-medial adverbs in that they cannot easily follow the VP either. This is exemplified in (44) below on the basis of a contrast between *aftur* ‘again’ as opposed to *aldrei* ‘never’ (from Angantýsson 2011:64):

- (44) a. *María hafði aftur séð Jón*
Mary had again seen John
‘Mary had seen John again’
- b. *María hafði aldrei séð Jón*
Mary had never seen John
‘Mary had never seen John’
- (45) a. *María hafði séð Jón aftur*
Mary had seen John again
‘Mary had seen John again’
- b. **María hafði séð Jón aldrei*
Mary had seen John never
‘Mary had never seen John’

The contrast between (45a) and (45b) shows that a low adverb like *aftur* ‘again’ can follow the VP whereas pre-VP adverbs like *aldrei* ‘never’ cannot. The same carries over to the rest of the adverbs above belonging to the class of sentence-medial adverbs, both in main and embedded contexts (for further discussion, see Angantýsson 2011:64ff., 2017c). As a result, these adverbs have been regarded as a signpost for determining the structural position of other elements that either occur higher or lower in the clause (precede or follow in linear order), especially in the seminal work of Pollock (1989) and much subsequent work. As pointed out by Angantýsson (2011:66, fn. 45, with references), the natural semantic class of adverbs that may arguably be seen as reliable diagnostics for verb raising into the inflectional domain would be negation, in addition to adverbs of quantification, belonging to Cinque’s (1999) Lower (pre-VP) AdvPs.

To avoid complications, my quantitative study of verb-adverb placement was confined to the handful of sentence-medial adverbs already known from the previous literature on Icelandic adverbial syntax. The following adverbs were selected, including spelling variants and close synonyms exhibiting the same behaviour:

- (46) a. *ekki, eigi, ei* NEGATION
- b. *aldrei, aldregi* NEVER
- c. *alltaf, alltið, ávallt, einatt, ætíð* ALWAYS
- d. *stundum* SOMETIMES

The adverbial phrase *á stundum* ‘at times; sometimes’, although largely synonymous with *stundum* ‘sometimes’, was not included here since it fails to meet the distributional

criteria. It may, for instance, occur after the VP, as the following 19th-century Icelandic attestations show:

- (47) *enda munu þeir stínga stöku við á stundum*
 and.also will they stab occasionally PRT at times
 ‘As they [the angels—HFV] allegedly limp occasionally at times.’
 (GeiVid-1805-09-04.txt)

The upside of this approach is that we end up with a relatively clean data set in which each data point arguably passes as a reliable diagnostic for verb raising. However, the downside is that we miss out on a number of potential but perhaps less reliable data points featuring other adverbs, the status of which is either less clear with regard to traditional V-to-I diagnostics or is yet to be determined.

Interestingly, Heycock and Wallenberg (2013) approached this issue very differently, allowing them to make (better?) use of a more limited amount of data. Heycock and Wallenberg’s (2013) method was to extract all adverbs in IcePaHC (Wallenberg et al. 2011) that occurred at least once between an auxiliary and a non-finite verb and then to filter out adverbs that occurred directly after a non-finite verb (i.e. post-VP adverbs). While I find this approach not only ingenious but also very appealing in that the category of medial adverbs is defined on the basis of usage-based evidence, this can be very risky. Their corpus was limited in size and it is likely that there is insufficient distributional data to define post-VP adverbs reliably in such a way. More importantly, perhaps, this does not filter out main clause Adv-V_{fin} adverbs. Structures that are ambiguous between Scandinavian-like Adv-V_{fin} and German-like V_{final} configurations, may also complicate the picture considerably during the early modern period, whereas V_{final} hardly occurs at all in the 19th-century material.

3.3.3 Subject type

One of the major pieces of potential evidence used by Bobaljik and Thráinsson (1998) to argue that the Adv-V_{fin} word order in Icelandic is not to be analysed in the same way as in Faroese and Mainland Scandinavian (as V-in-situ, see Section 3.7) has to do with subject-type restrictions (cf. also Angantýsson 2001, 2007, 2011, Thráinsson 2003, 2010). Their main argument is as follows: Richly inflected languages have two subject positions in the IP-domain, viz. in the specifier of the tense projection TP and the agreement projection AgrP, respectively, whereas poorly inflected languages only have one subject position. This can be seen by the fact that expletive constructions are compatible both with an expletive and a subject (associate of the expletive) simultaneously, both in a VP-external position, only in the former but not in the latter (Bobaljik and Thráinsson 1998:56):

- (48) a. *Það hefur { einhver köttur } étíð { *einhver köttur } mýsna* (Ic.)
 EXPL has some cat eaten some cat mice-the
 ‘A cat has eaten the mice.’
 b. **Det har { en katt } ete { en katt } mysene* (No.)
 EXPL has a cat eaten a cat mice-the
 Intended: *‘There has a cat eaten the mice.’

In addition, indefinite subjects prefer the lower subject position while definite subjects do not (see Bobaljik and Thráinsson 1998:57). Correspondingly, the Adv-Vfin order is claimed to be incompatible with subjects occurring in the lower subject position (Bobaljik and Thráinsson 1998:63, fn. 31, Angantýsson 2001:115). This is shown in the contrasting judgments of *hann* ‘he’, in the higher subject position, and *einhver* ‘somebody’, in the lower subject position, suggesting that the adverb *ekki* ‘not’ here is adjoined to TP (see Thráinsson 2003:184, his judgments):

- (49) a. *Það var Hrafnkelssaga sem hann ekki hafði lesið*
 it was Hrafnkel’s saga that he not had read
 b. ?* *Það var Hrafnkelssaga sem einhver ekki hafði lesið*
 it was Hrafnkel’s saga that someone not had read

Wiklund et al. (2007:210, fn. 8) mention potential subject-type restrictions but suggest that these hold for certain speakers only. Being a clear prediction of Bobaljik and Thráinsson’s (1998) V-to-T analysis of Adv-Vfin in Modern Icelandic, the distinction between definite and indefinite (and non-specific) subjects must be taken as a potential factor in the analysis of 19th-century Icelandic as well.

Restrictions on Modern Icelandic Adv-Vfin in subject-initial embedded clauses have been documented in most detail by Angantýsson (2001, 2007, 2011, 2017a, to appear). In addition to non-specific indefinites, Angantýsson (2011:72) argues that while Adv-Vfin is in principle possible in any clause type, the Adv-Vfin order is most easily found with indirect questions with a *wh*-pronoun (*hverjum* ‘whom’) and relative clauses modifying noun phrases, while certain other types of clauses are questionable/marginal with non-pronominal subjects, viz. relative clauses modifying adjective phrases, indirect questions with a *wh*-adverb (*hvers vegna* ‘why’), concessive (*þótt* ‘although’), temporal (*þegar* ‘when’) and conditional (*ef* ‘if’) clauses (for examples, see Angantýsson 2011:67-71). Finally, some clauses do not easily occur with the Adv-Vfin order regardless of subject type, viz. *that*-clauses, indirect questions with *hvort* ‘whether’ and adverbial clauses connected with *að* (e.g. *svo að* ‘so that’).

The final factor concerning subject type that needs to be considered here involves morphological case marking. Vfin-Adv/Adv-Vfin variation with sentence-medial adverbs **only occurs with subjects**, which effectively makes Adv-Vfin a reliable test for subjecthood. The fact that Adv-Vfin can also occur with so-called quirky/oblique subjects is a clear indicator that these are, indeed, to be regarded syntactically as subjects fully on par with nominative subjects:

- (50) a. *Þetta er prófið sem honum ekki tókst að ná*
 this is the exam that him_D not managed to pass
 b. *Þetta er prófið sem honum tókst ekki að ná*
 this is the exam that him_D managed not to pass
 ‘This is the exam that he did not manage to pass.’
- (51) a. *Jón spurði hvaða spurningar Páli ekki líkuðu*
 John asked which questions Paul_D not liked
 b. *Jón spurði hvaða spurningar Páli líkuðu ekki*
 John asked which questions Paul_D liked not
 ‘John asked which questions Paul did not like.’

Since there is a broad consensus among linguists working on Icelandic and related languages that oblique subjects go back to the medieval period and beyond (cf. Rögnvaldsson 1996, Eythórsson 2002, Eythórsson and Barðdal 2005 and references therein), reconstructed even for Proto-Indo-European (see Barðdal and Smitherman 2013), I will simply treat all subjects alike in what follows, regardless of case.

3.3.4 Verb type

The effect of verb type (e.g. main, auxiliary or modal) has not been documented with regard to Icelandic verb-adverb placement, as far as I know, and Wiklund et al. (2007:210, fn. 8) explicitly state that variation in subject-initial embedded clauses is not influenced by whether or not the finite verb is a main verb. Interestingly, there does appear to be an effect in Övdalian such that embedded V2 is more acceptable if the finite verb is an auxiliary (cf. Garbacz 2010:126f., Angantýsson 2011:90, 111). Also, Stylistic fronting in Icelandic was more frequently accepted by the oldest age group in Angantýsson (2011:153) in a relative clause with a main verb than in a corresponding relative with an auxiliary. Similarly, Faroese speakers accepted transitive expletive constructions more readily with a finite verb than an auxiliary (Angantýsson 2011:173). Garbacz (2010:122) suggests that finite verbs and modal auxiliaries (perhaps auxiliaries in general) behave differently in terms of their structural position potential and points out that the embedded order *finite verb – sentential adverb* has also been argued to occur earlier with modal auxiliaries in the speech of young children acquiring Swedish than with other types of verbs. These findings clearly suggest that verb type might be a relevant factor.

Truckenbrodt (2006:288f.) also discusses the fact that modal auxiliaries in German may behave differently from main verbs with regard to the possibility of embedded V2:

- (52) a. * Maria bittet Peter, er **geht** nach Hause.
 ‘Maria asks Peter, he is going home’
 b. Maria bittet Peter, er **soll** nach Hause gehen.
 ‘Maria asks Peter, he should go home.’

Instead of relating these differences between main verbs and modals to syntactic properties, he argues that the contrast follows from the interaction between syntax and semantics. The idea is that for embedded V2 to take place in German (taken to involve V-to-C), the meaning must entail an epistemic context, in which these two examples crucially differ (cf. Truckenbrodt 2006:289, his examples):

- (53) a. Maria asks Peter that he goes home.
 b. → Maria says that Peter goes home.
 c. ≈ Maria commits to her believing that Peter goes home.
- (54) a. Maria asks Peter that he **should** go home.
 b. → Maria says that Peter **should** go home.
 c. ≈ Maria commits to her believing that Peter should go home.

Truckenbrodt (2006:289f.) furthermore claims that this entailment fails under likelihood, possibility, causation and purely deontic embedding, resulting in the ungrammaticality of embedded V2 in these contexts as well. In addition, modal particles and use of the subjunctive mood can interact with clause structure and render embedded V2 possible in ‘V2-hostile’ environments (cf. Meinunger 2004, Frey 2012, Gutzmann 2017).

The contrast between modals/auxiliaries and main verbs is also very well-known from the literature on English (see Lightfoot 1979, Kroch 1989, Pollock 1989, 1997, Warner 1993, Bybee and Thompson 2000, Han and Kroch 2000, Ecay 2015). Thus, for Pollock (1989), the contrasts between English and French (as well as diachronically in English), where the adverb *often/souvent* are assumed to mark the boundary between IP and VP, indicates obligatory V-to-I movement in French but V-in-situ in English (cf. Pollock 1989:367):

- (55) a. * John **kisses** often Mary.
 b. Jean **embrasse** souvent Marie.
 c. John often **kisses** Mary.
 d. * Jean souvent **embrasse** Marie.
- (56) a. * John **likes** not Mary.
 b. Jean (n’) **aime** pas Marie.

In contrast, English auxiliaries occur in exactly the configuration that are ungrammatical in (55a)/(56a), namely preceding adverbs/negation:

- (57) a. John **has** often kissed Mary
 b. John **has** not/**hasn’t** kissed Mary

A lot of work on the contrast between English auxiliaries and main verbs in the generative tradition has emphasised the potential relation to deflection in the verbal inflectional paradigms. Thus, Pollock (1989), for instance, argues that the contrast is due to the English inflectional phrase being opaque, resulting in a fused, unsplit IP that is unable to attract lexical verbs (verbs with a thematic grid, see Pollock 1989:386, Chomsky 1991:422-424), whereas IP was transparent in earlier stages of English, with separate functional projections for Agr and T, as in present-day French as well (but see Koenenman and Zeijlstra 2014). Bybee and Thompson (2000) have suggested instead that this split is actually due to frequency effects in syntax (see also Erker and Guy 2012), which they argue can have a conserving effect. Thus, high frequency sequences such as verbal auxiliaries “take on a life of their own, and resist change on the basis of newer productive patterns for juxtaposing words and morphemes” (2000:381). This effect can lead to their retaining old conservative syntactic characteristics that previously, going back to Middle English, held for the verbal system as a whole.

Whatever the reason(s), it might thus be argued that lexical versus auxiliary/modal asymmetries can be expected, or at least that these categories need not necessarily be expected to exhibit exactly the same developmental patterns or rate of use. With the potential grammatical factors in place, let us now turn to the results of the corpus study.

	Vfin-Adv	Adv-Vfin	<i>n</i>
1800-1824	44.5%	55.5%	182
1825-1849	62.6%	37.4%	1167
1850-1874	81.3%	18.7%	562
1875-1899	87.9%	12.1%	1297
1900-1924	92.9%	7.1%	1836

Table 3.4. Verb-adverb placement in newspapers/periodicals, per quarter century.

3.4 Results

This chapter presents the results of this thesis' first case study on verb-adverb placement, reported separately for each corpus, in Sections 3.4.1, 3.4.2 and 3.4.3. For a general overview, simple descriptive statistics will first be shown, with the aggregated data split over periods of approximately 25 years each for the newspaper and private letter corpus, but into shorter periods in the case of the student essays. For each corpus, this overview is followed by a second section incorporating grammatical and social factors as predictors in a mixed-effects statistical model, with a random effect for individual variation (see Section 2.4). For the two larger corpora, newspapers and private letters, we then zoom in on individual variation. In the private letter corpus, individuals are considered with respect to language-internal variables but also with regard to variation within individuals (intra-speaker variation), both over time and across different recipients.

3.4.1 Newspapers and periodicals

3.4.1.1 Overall trends in the data

The newspaper corpus exhibits rather extensive variation in verb-adverb placement, indicating a rise and sharp decline of the Adv-Vfin order very much in agreement with the previous study of Heycock and Wallenberg (2013) on public narrative texts. Table 3.4 shows the overall proportion of the two variants per quarter century. Note that this presentation of the data is prior to any further partitioning in terms of language-internal factors (see below). However, purely from the perspective of the implementation of linguistic norms, where Adv-Vfin is categorically considered non-standard regardless of the environment in which it occurs, the sharp decline of Adv-Vfin arguably attests to increasingly standard language use in relatively formal, printed sources (taking a view 'from above').

Rather surprisingly, perhaps, the Adv-Vfin variant turns out to be in the majority during the first period, occurring at an average rate of over 55%. However, it should be kept in mind that during this period, the data are not only rather scarce but also instantiated by a very limited number of titles. Indeed, as we will see, individual newspapers exhibit different rates of Adv-Vfin across the whole period, suggesting variation beyond the TIME factor. Assuming that those responsible for most of the contents of the newspapers belonged to the higher echelons (the educated class), a high proportion of Adv-Vfin early in the century is not unexpected based on previous

scholarship, albeit considerably higher than what Heycock and Wallenberg (2013) have reported on the basis of the IcePaHC corpus, partly from the same period.

As regards the potential visible effects of language planning, it is telling that we observe an increasing number of newspapers approaching complete or near-complete absence of Adv-Vfin towards and beyond the year 1900 (see also Viðarsson 2014, 2016). While the overall trend towards less use of Adv-Vfin is probably due to a conspiracy of factors, the effects of prescriptivism being only one of them, the marginal presence of Adv-Vfin towards the end of the period does suggest some such intervention effects.

Let us now move on to other dimensions fully obscured in the Table 3.4, viz. the potential linguistic and extra-linguistic factors involved in determining which contexts may be more or less favourable to either of the two variants, based on mixed-effects statistical modelling of the data (cf. Sections 2.4 and 3.3).

3.4.1.2 Modelling the variation

This section presents a statistical analysis of variability in verb-adverb placement in the newspaper corpus on the basis of six fixed effects, introduced with reference to the previous literature in Section 3.3, with a random effect for the individual. The fixed effects were as follows (the first value listed being the reference level for comparison of the categorical variables):⁴⁹

- (58) a. DEFINITENESS STATUS of subject (two-level factor: *definite*, *indefinite*)
- b. NOMINAL STATUS of subject (two-level factor: *nominal*, *pronominal*)
- c. CLAUSE TYPE (three-level factor: *typical V2*, *typical non-V2*, *mixed contexts*)
- d. VERB TYPE (three-level factor: *auxiliary*, *modal*, *lexical*)
- e. ADVERB TYPE (two-level factor: *sentence-medial adverb*, *negation*)
- f. TIME PERIOD (continuous: per decade, 1803-1924)

A forest plot of the estimates is provided as a basic overview of the model in Figure 3.4.⁵⁰ The factors in Figure 3.4 are sorted top to bottom from the factors most strongly favouring Adv-Vfin (in blue) to the factors most strongly favouring Vfin-Adv (in red). The values on the x-axis refer to odds ratios (OR), where ‘1’ may be interpreted as neutral (exposure to this the factor does not affect the odds of outcome), whereas a value such as ‘4’, as in the case of CLAUSE TYPE (*non-V2*) such as relative clauses (OR=4.09), means that this environment has four times the odds of occurring with Adv-Vfin than in typical V2 environments, such as declaratives.

The factors CLAUSE TYPE and VERB TYPE are the most important language-internal factors according to the model and these are both highly statistically significant. Non-V2

⁴⁹One additional factor, PLACE OF PUBLICATION, was originally included in the model but did not improve the model. This was even the case when simplified to provide only a three-way distinction between a) Reykjavík, b) areas outside Reykjavík, and c) abroad. A split with Reykjavík (SW) and Akureyri (NE) serving as centres of potentially elevated levels of Danish influence, as compared to other areas in the country, in addition to Denmark, was also attempted, with no significant result. As the place of publication did not improve the model, that variable was excluded.

⁵⁰Model evaluation: C index of concordance = 0.843272, Somers’ D_{xy} = 0.6865450. Classification accuracy: 84.67% of the choices among variants predicted correctly (number of observations = 5,043).

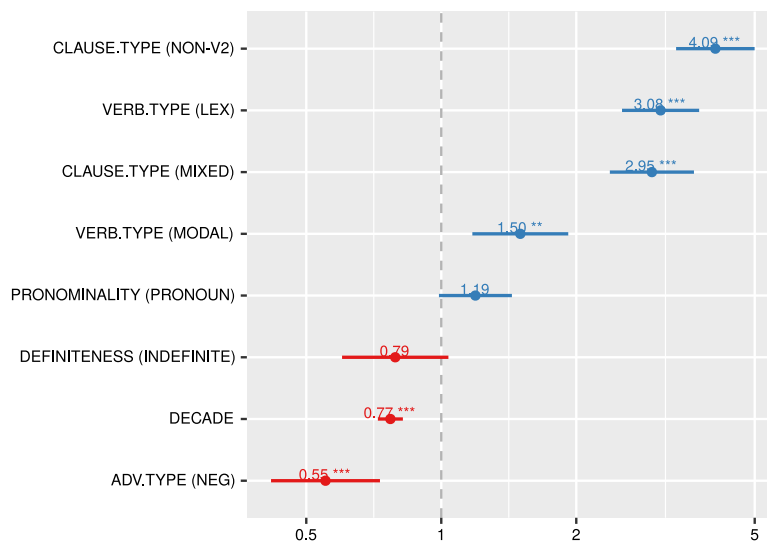


Figure 3.4. A forest plot of estimates with odds ratios of fixed effects: OR>1 (blue) favours Adv-Vfin, OR<1 (red) favours Vfin-Adv. (*' $p<0.05$, ***' $p<0.01$, ****' $p<0.001$.)

environments have 4.09 times the odds of exhibiting Adv-Vfin (as compared to typical V2 environments), mixed environments have 2.95 times the odds of Adv-Vfin (again, as compared to typical V2 environments) and lexical verbs have 3.08 times the odds of Adv-Vfin (as compared to auxiliaries), all of which are significant at the $p<0.001$ level. Modal verbs have 1.5 times the odds of Adv-Vfin (again, as compared to auxiliaries), significant at the $p<0.01$ level. The fact that modals are significantly more and not less likely to occur with the Adv-Vfin than auxiliaries is rather surprising given their potential to facilitate embedded V2 (cf. Section 3.3.4). The same effect is even stronger for lexical verbs and clearly merits further study.

Figure 3.5 provides a visualisation of the predicted probabilities of the fixed effect CLAUSE TYPE over time, conditioned on random effects for individual newspaper titles.⁵¹ The raw data before and after 1850 are also shown in Table 6.29, Appendix A, page 245.⁵² In accordance with previous research (cf. Heycock and Wallenberg

⁵¹ For a visualisation of the remaining fixed effects, see Appendix A, page 246.

⁵² Table 6.29 contrasts private letters and newspapers before and after 1850 across the eight basic categories used by Heycock et al. (2012) when contrasting Danish and Faroese, cf. also (33) above. Although the percentage of Adv-Vfin is slightly higher in the newspapers than in the private letters on average, the two genres are almost in perfect lock-step. The first four types (cf. (33a–d) above) exhibit Adv-Vfin at nearly half the rate found in the last four types (cf. (33e–h)) in the period until 1850, ranging from 25–27% vs. 46–52%. The difference between these two categories is even greater in the period after 1850, although the rate of use is much lower, ranging from 6–7% vs. 18–20%. This difference between the two periods may be due to sensitivity to the ASSERTION factor in declaratives (see Section 3.7.1) in the former period which may not be present (or is greatly reduced) after 1850, when the overall frequency of Adv-Vfin is considerably lower.

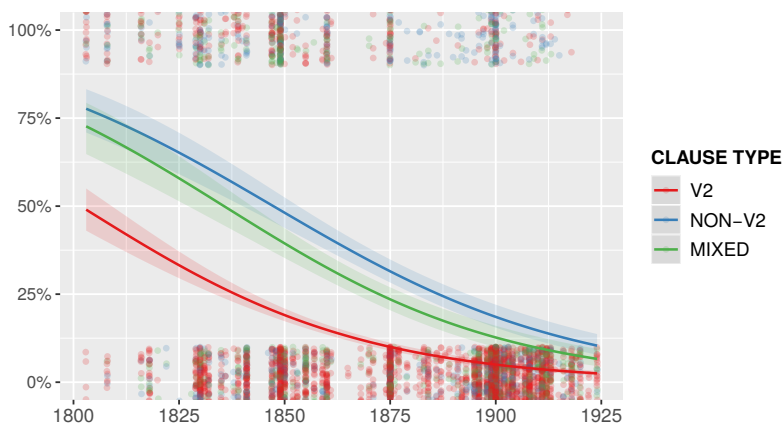


Figure 3.5. Predicted probabilities of the fixed effect CLAUSE TYPE over time, conditioned on random effects (lower values = less Adv-Vfin).

2013, Viðarsson 2014), the results indeed reveal that typical non-V2 contexts are more strongly associated with the Adv-Vfin variant as compared to typical V2 contexts. The same applies to mixed environments, but slightly less so. This is largely as expected given that this category may align with either of the other two environments, although judging by the size of the effect, the mixed environments clearly have more in common with typical non-V2 than with typical V2 environments.

The factors PRONOMINALITY and DEFINITENESS, in contrast, do not turn up as significant at all, although leaving them in still made a statistical improvement to the model based on the *anova* function (see Section 2.4). At the other end of the plot, the temporal factor DECADE and the language-internal factor ADV-TYPE exhibit significantly decreased odds of the outcome Adv-Vfin, i.e. there is statistically less Adv-Vfin over time and the odds of negation occurring with Adv-Vfin are lower than with the other sentential adverbs. ADVERB TYPE is a crucial factor from the perspective of Mainland Scandinavian, negation being expected to act as an intervener for Vfin-Adv, strongly favouring Adv-Vfin. However, these data show the opposite effect: the negation is less – and not more – strongly associated with Adv-Vfin vis-à-vis the other sentence-medial adverbs studied.⁵³

3.4.1.3 Individual variation

Table 3.5 shows the rate of use per individual title with at least 20 examples of the variable. Period I is striking in that the Adv-Vfin variant is more frequent than Vfin-Adv in most of the titles, *Margvíslegt gaman og alvara* (1818) being an exception with the

⁵³Due to data limitations, sentence-medial adverbs ($n=459$) were consolidated and contrasted with negation as a single group, the largest member of which being *aldrei* ‘never’ ($n=273$), followed by *always* ($n=119$). The effect is strongly significant with NEVER ($p<0.001$, ‘***’), but ALWAYS and SOMETIMES point in the same direction. The adverb *sjaldan* ‘seldom’ only occurred twice and was not considered separately.

	Time period	Place of publ.	Vfin-Adv	Adv-Vfin	<i>n</i>
<hr/>					
1800-1824					
Íslenzk sagnablaðið	(1816)	DK	36.4%	63.6%	33
Klausturpósturinn	(1818-1822)	IS (W)	40.0%	60.0%	20
Margvíslegt gaman og alvara	(1818)	IS (W)	80.4%	19.6%	51
Minnisverð tíðindi	(1803-1808)	IS (W)	25.6%	74.4%	78
<hr/>					
1825-1849					
Ármann á Alþingi	(1829-1832)	DK	68.8%	31.2%	279
Búnaðarrit Suðuramtsins ...	(1839)	IS (SW)	55.6%	44.4%	72
Norðurfari	(1849)	DK	55.9%	44.1%	315
Ný félagsrit	(1841)	DK	86.7%	13.3%	143
Reykjavíkurbósturinn	(1846-1849)	IS (SW)	50.3%	49.7%	151
Skírnir	(1828-1840)	DK	67.6%	32.4%	71
Sunnanpósturinn	(1835-1838)	IS (SW)	56.0%	44.0%	84
Þjóðólfur	(1848)	IS (SW)	92.0%	8.0%	25
<hr/>					
1850-1874					
Ársritið Gestur Vestfirðingur	(1850-1855)	IS (SW)/DK	95.5%	4.5%	134
Ingólfur	(1853-1855)	IS (SW)	77.4%	22.6%	31
Íslendingur	(1861)	IS (SW)	66.7%	33.3%	36
Norðanfari	(1865-1871)	IS (NE)	86.3%	13.7%	51
Norðri	(1853-1859)	IS (NE)	85.3%	14.7%	95
Ný sumargjöf	(1860)	DK	63.6%	36.4%	88
Þjóðólfur	(1852-1860)	IS (SW)	73.7%	26.3%	76
<hr/>					
1875-1899					
Austri	(1884)	IS (E)	96.4%	3.6%	28
Bjarki	(1896)	IS (E)	96.8%	3.2%	31
Dagskrá	(1896)	IS (SW)	87.5%	12.5%	24
Eimreiðin	(1895-1899)	DK	92.2%	7.8%	499
Fjallkonan	(1884-1889)	IS (SW)	81.6%	18.4%	147
Fréttir frá Íslandi	(1875)	IS (SW)	68.6%	31.4%	35
Ísafold	(1875-1899)	IS (SW)	92.3%	7.7%	208
Íslendingur	(1875)	IS (SW)	72.1%	27.9%	61
Norðanfari	(1875)	IS (NE)	87.1%	12.9%	70
Norðlingur	(1875-1881)	IS (NE)	87.7%	12.3%	57
Suðri	(1883)	IS (SW)	84.0%	16.0%	25
Þjóðólfur	(1875)	IS (SW)	76.3%	23.7%	76
<hr/>					
1900-1924					
Austri	(1900)	IS (E)	77.8%	22.2%	72
Austurland	(1907)	IS (E)	92.6%	7.4%	27
Bjarki	(1900)	IS (E)	100.0%	0.0%	32
Eimreiðin	(1900-1920)	DK/IS (SW)	92.6%	7.4%	597
Fjallkonan	(1900-1911)	IS (SW)	94.5%	5.5%	273
Framsókn	(1900)	IS (SW)	93.3%	6.7%	75
Ísafold	(1900-1924)	IS (SW)	99.1%	0.9%	221
Norðurland	(1902)	IS (NE)	100.0%	0.0%	30
Reykvíkingur	(1900)	IS (SW)	62.5%	37.5%	24
Stefnir	(1900)	IS (NE)	77.1%	22.9%	48
Þjóðólfur	(1900-1920)	IS (SW/S)	94.7%	5.3%	319
Þjóðviljinn	(1900)	IS (Wf)	89.1%	10.9%	46

Table 3.5. Proportion Vfin-Adv/Adv-Vfin in newspapers and periodicals (1800-1924) with a minimum of 20 examples of either variant.

opposite pattern of over 80% Vfin-Adv. A number of titles in Periods I-III with high proportions of Adv-Vfin—the typical Danish embedded word order—were published in Denmark. However, as corroborated by the statistical analysis, place of publication is not a significant predictor of trends in the data. In the first period, for instance, only one of the titles, *Íslensk sagnablöð* (1816), was published in Denmark, by the Icelandic Literary Society in Copenhagen, whereas *Klausturpósturinn* (1818-1822), *Margvíslegt gaman og alvara* (1818) and *Minnisverð tíðindi* (1803-1808), edited among others by the official Magnús Stephensen (1782-1833), were published in Iceland.

In Period II, most newspapers exhibit the Adv-Vfin order approximately 1/3 of the time or more, notable exceptions being *Ný félagsrit* (1841) and *Þjóðólfur* (1848) with frequencies as low as 13.3% and 8%, respectively. Despite claims to the effect that the puristic norms of Fjölirnir were extremely effective, allegedly reaching a broad consensus in main respects already in the 1840s (cf. above), the first explicit remark in 1844 by Gíslason on the Adv-Vfin variant being undesirable appears not to have had much effect in *Reykjavíkurlpósturinn* (1846-1849) or in *Norðurfari* (1849). However, a potential candidate may be *Þjóðólfur* (1848), a provocative liberal newspaper, edited by Sveinbjörn Hallgrímsson (1814-1863) and unique at the time in its lack of ties to the authorities and rulers of the printing press (see e.g. Jónsson 1998:82). Sveinbjörn Hallgrímsson, moreover, had ties to the Reykjavík Grammar School and was Sveinbjörn Egilsson's nephew (Jónsson 1998:80), the iconic figure of Icelandic language purism. Contrasting this with the largely contemporaneous titles *Norðurfari* and *Reykjavíkurlpósturinn*, the former edited by Gísli Brynjólfsson (1827-1888) and Jón Thoroddsen (1818-1868), the latter by Páll Melsteð (1791-1861), Þórður Jónasson (1800-1880) and Sigurður Melsteð (1819-1895), these were individuals of the more traditional officials class, most of whom had studied law at the University of Copenhagen and entered into existing power structures, positions of crown officials, sheriffs and members of parliament. Although one should view with reservation the small amount of data sampled from *Þjóðólfur* in this period, the fact that it fails to pattern with the other traditional newspapers, breaking with the language norms associated with the ruling class, could possibly be regarded as a consequence of its liberal, anti-establishment stance.⁵⁴

Similarly, perhaps, *Ný félagsrit* was a liberal newspaper and the product of an Icelandic intellectual society based in Copenhagen that grew out of the puristic Fjölirnir movement. Its chief editor and contributor, Jón Sigurðsson (1811-1879), was a student of grammar, literature and history at the University of Copenhagen, later becoming a politician where he came to be regarded as a chief leader in the struggle towards independence from Denmark. In terms of style, Jón Sigurðsson is considered among those who “set a good example” by writing in “clear unaffected language” (Hermannsson 1919:51). While the sampled issue of *Ný félagsrit* pre-dates the overt stigmatisation of the Adv-Vfin order by Gíslason (1844), it is quite telling that *Ný félagsrit*, too, rather follows its own pattern as regards language use, breaking with the tradition.

⁵⁴Note in this regard also that when the local authorities issued a ban on the publication of *Þjóðólfur* in 1850, the editor instead had it printed with S.L. Møller in Copenhagen. However, since the Danish printer did not have the required title sized versions of the Icelandic letters þ and ð, and the editor decidedly did not want to ‘Danicise’ þ as th, as was not only customary but even prestigious in the very circles of power of which *Þjóðólfur* was critical (whereas d for ð was less salient and more commonplace), the editor instead opted to actually substitute þjóð- ‘nation’ for hljóð- ‘sound’, changing the title of the paper temporarily to *Hljóðólfur* (Jónsson 1998:70-74).

In Periods III and IV, there is still quite extensive variation in verb-adverb placement, but a clear downward trend in the frequency of the non-standard Adv-Vfin has appeared. In both periods, *Þjóðólfur* now scores lower in terms of the use of the standard Vfin-Adv as compared to Period II. However, it should be noted that *Þjóðólfur* underwent a change in editors shortly after the beginning of Period III, with Jón Guðmundsson (1807-1875), a member of parliament and a law graduate from the University of Copenhagen taking over between 1852 and 1874.

Whereas Adv-Vfin had been a prominent pattern in the first two periods, this variant is clearly becoming less and less frequent towards the turn of the century. It is during the end of Period IV and, especially, in Period V that we start observing (near) **categorical use** of the standard Vfin-Adv variant in this type of printed material, including *Bjarki* (0.0%), *Ísafold* (0.9%) and *Norðurland* (0.0%). As already mentioned above, many newspaper editors were avid language purists and it is probably not a coincidence that Björn Jónsson (1846-1912), known to have been very particular about avoiding non-standard features in his newspapers, was the chief editor of *Ísafold* most of the time between 1874-1909. Indeed, already in Period IV, *Ísafold* exhibits a higher relative frequency of the standard Vfin-Adv order (92.3%) than most other papers in that period.⁵⁵

Before proceeding further to contrast verb-adverb placement in the newspapers with the private letter corpus, let us first make an attempt at a crude examination of these linguistic variables in newspapers beyond the 19th century, up until present times.

3.4.1.4 Beyond the 19th century and hypercorrection

For a rough approximation of verb-adverb placement in later newspapers, the online *Tímarit.is* corpus was used in its entirety. Figure 3.6 shows the proportion of Neg-Vfin from 1850 onwards, i.e. over a period of roughly 17 decades, based on a simple string search for patterns matching *that/which he/she/it <not> have/had <not>*. The reader is thus warned that these data are unanalysed and may potentially contain noise. However, the search strings were carefully selected so as to minimise the risk of any irrelevant or confounding factors.

The relative frequency of Neg-Vfin before the turn of the 19th century agrees well with the results reported for the smaller, analysed data set of the previous section. As early as 1900-1909, the average rate of occurrence of Neg-Vfin is already below the 5% mark (68/1691, 4.1%), roughly half that in the 1920s (69/3212, 2.1%) and below the 1% mark in the 1940s (65/6875, 0.9%). The frequency drops even further over the following decades, down to 0.5% in the 1950s (47/8828), 0.3% in the 1990s (34/12590) and to a historical low of **near absolute zero** in the 2010s: 0.07% (4/5534). It therefore appears we are witnessing a more or less successful suppression of a non-standard feature from this relatively formal, written register. The Gigaword Corpus (Steingrímsson et al. 2018) reveals similar numbers, the frequency of Neg-Vfin being 0.04% ($n=309,126$) for clauses introduced with *að* and 1.28% ($n=34,564$) for clauses introduced with *sem* (yielding a total rate of 0.16% ($n=343,690$) for these two environments).⁵⁶

⁵⁵Björn Jónsson also edited *Norðanfari* and *Norðri*, which exhibit in Period III an above average frequency of Vfin-Adv (86.3% and 85.3%, respectively), but which still score lower than *Ísafold*.

⁵⁶It appears that texts that extend further back into the 20th century, such as parliamentary speeches and the

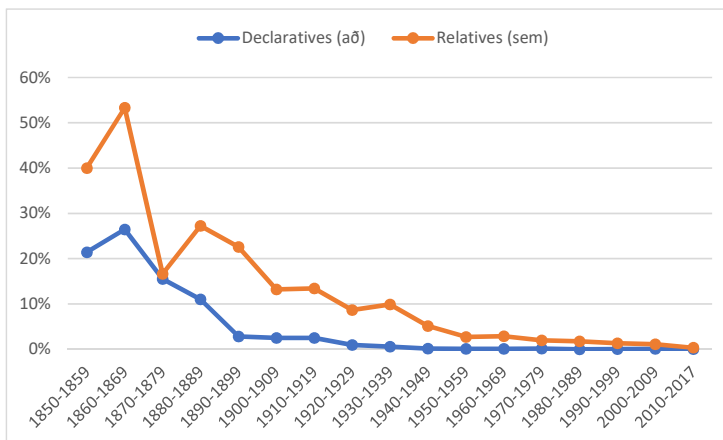


Figure 3.6. Proportion Neg-Vfin in newspapers from 1850 onwards ($n=93,748$). Unanalysed data set containing clauses introduced by *að* ‘that’ (typically DECL) or *sem* ‘which’ (typically REL), immediately followed by 3P.SG. subject (*hann/hún/það*) and the negation (*ekki*) either precedes or follows an inflected form of the verb *hafa* (only 3P.SG.).

For the purposes of identifying potential cases of hypercorrection, it is worth pointing out that the verb-adverb placement variable as defined here (subject-initial embedded clauses) bears a superficial resemblance to a different construction known in the syntactic literature as Stylistic Fronting, featuring a subject gap (see Maling 1980 and Section 3.7). In fact, there is clear evidence of false signals of Neg-Vfin in Stylistic Fronting environments, but there is also potential distributional evidence to this effect. Let us start with the former. As mentioned above, Gunnarsson (1878) criticised the language use of a book review on the basis of its having “some rotten taste of Danish” (see Section 3.1). Interestingly, his examples cover not only the verb-adverb placement variable, but also instances of Stylistic Fronting, such as (59):

- (59) *sem svo víða hefir vel vandað*
 REL so widely has well done.carefully
 ‘... who has done a careful job in so many places.’ (Gunnarsson 1878:10)

Unlike the Adv-Vfin as defined here, there is certainly nothing ‘Danish’ about this arrangement of elements. The reason why this construction appears to be targeted by some commentators presumably lies in the idea that the relativiser *sem* is somehow conceived of as a pronominal element and, thus, a subject-initial clause of sorts (for reasons why this is not so, see Þráinsson 1980). Curiously, Smári (1920:258) makes this (arguably) false assumption in his Icelandic syntax as well (see Viðarsson 2014:18, fn. 18). I would like to emphasise in this context that I have found no evidence that Stylistic Fronting (in relatives or otherwise) was in any way considered non-standard in the Reykjavík Grammar School material. This is unsurprising, given the emphasis on

Morgunblaðið newspaper corpus, have a higher than average rate of Neg-Vfin in relatives; a rate of 3.45% Neg-Vfin in parliamentary speeches ($n=4,581$) and 1.99% ($n=7,972$) in Morgunblaðið. This effect appears to be absent in *að*-environments.

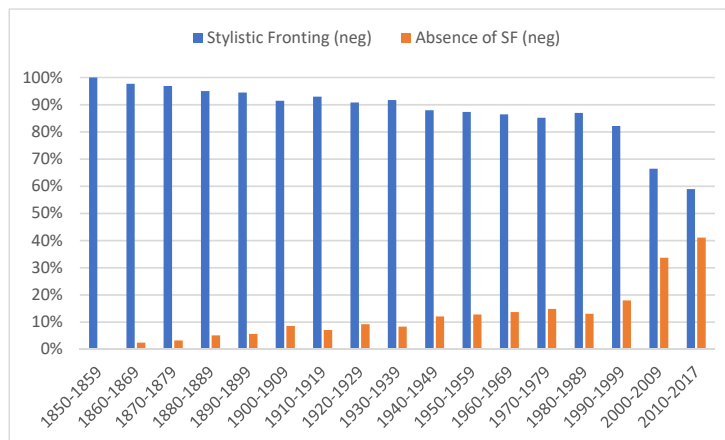


Figure 3.7. Stylistic Fronting: Proportion of fronting versus non-fronting of negation in raw data set ($n=89,686$) compiled from newspapers from 1850 onwards. The data set contains clauses introduced by *sem* ‘which’ (typically REL), followed by the negation (*ekki*) and an inflected form of the verb *hafa* (only 3P.SG.) in either order, and lacking an overt subject.

and vast knowledge at the grammar school of Old Norse grammar, where this order was commonplace (see also Sigurðsson 2017, with references).

However, if even Smári could be so ‘mistaken’ (note that Sigurður Gunnarsson was a pastor and not a linguist), it is certainly not unlikely that other writers could be as well. Indeed, note in this regard that at roughly the same time that Adv-Vfin in subject-initial clauses is receding in the newspapers (cf. Figure 3.6), so is also the “corresponding” Neg-Vfin order in Stylistic Fronting contexts. As Figure 3.7 shows, the order Rel-Vfin-Neg (e.g. *sem hafa ekki verið* ‘which have not been’) vs. Rel-Neg-Vfin (*sem ekki hafa verið*) is not particularly frequent during the 19th century, the order being overwhelmingly Rel-Adv-Vfin. As time progresses, there is a gradual but clear increase in the (non-fronted) Rel-Vfin-Neg order. This might be interpreted as hypercorrection, i.e. avoidance of Stylistic Fronting on identity with the non-standard Adv-Vfin variant. The much higher rate of the non-fronted Rel-Vfin-Neg variant during the period 2000-2017, as compared to 20th century, is very striking. Presumably, the gradual decrease in Stylistic Fronting over time (and in favour of overt expletives) plays a role (see e.g. Angantýsson 2011, Sigurðsson 2017).

3.4.2 Private letters

3.4.2.1 Overall trends in the data

Due to an unbalanced corpus design and an arguably less than representative selection of private letters from the period up until 1850 compared with after 1850, there is a bias in the former period in the private letter corpus towards letter-writers from the higher social ranks. Conversely, there is also a bias towards peasants and labourers in the period after 1850. For this reason, there is simply no way of taking the distribution in the

	Vfin-Adv	Adv-Vfin	<i>n</i>
1784-1799	73.8%	26.2%	42
1800-1824	62.9%	37.1%	833
1825-1849	69.1%	30.9%	1593
1850-1874	90.5%	9.5%	1515
1875-1899	87.8%	12.2%	1995
1900-1918	91.2%	8.8%	476

Table 3.6. Verb-adverb placement in the private letter corpus, per quarter century.

private letters in Table 3.6 at face value and concluding that we are directly witnessing the successful reversal of an ongoing linguistic change.

Nonetheless, there is most definitely a trend from higher to lower rates of Adv-Vfin as time progresses that appears to be real. Interestingly, the trend is visible among the highly educated as well as among the farmers/labourers. Furthermore, in both groups, female writers pattern rather differently from male writers (see next section). The overall pattern from 1850 onwards is quite stable, the percentage of Adv-Vfin being slightly above or below 10%. However, even within this period, a number of speakers exhibit higher occurrence rates of Adv-Vfin, suggesting considerable inter-speaker variation. Individual variation will be discussed in more detail in section 3.4.2.3, including changes over time for individuals with a sufficient amount of data.

3.4.2.2 Modelling the variation

This section presents a statistical analysis of variability in verb-adverb placement in the private letter corpus. The fixed effects are the same as in the newspapers, cf. (58) above, in addition to two effects relating to the SOCIAL STATUS and SEX of the speaker, as shown in (60).⁵⁷

- (60)
- a. DEFINITENESS STATUS of subject (two-level factor: *definite, indefinite*)
 - b. NOMINAL STATUS of subject (two-level factor: *nominal, pronominal*)
 - c. CLAUSE TYPE (three-level factor: *typical V2, typical non-V2, mixed contexts*)
 - d. VERB TYPE (three-level factor: *auxiliary, modal, lexical*)
 - e. ADVERB TYPE (two-level factor: *sentence-medial adverb, negation*)
 - f. TIME PERIOD (continuous: 1-6, per quarter century)
 - g. SOCIAL STATUS (three-level factor: *officials/lettered, peasants/labourers, other professions*)
 - h. SPEAKER SEX/GENDER (two-level factor: *male, female*)

⁵⁷Neither DEFINITENESS nor ADV TYPE showed statistically significant effects but also did not make the model significantly worse so these were left in for comparison with the previous sections. All other effects had significant contributions over a model without the effect in question.

A generational effect (per quarter of a century) was also tested as a part of the model but neither served to improve the model nor did it yield any significant effects. See Figure 3.11.



Figure 3.8. A forest plot of estimates with odds ratios of fixed effects: OR>1 (blue) favours Adv-Vfin, OR<1 (red) favours Vfin-Adv. (*' $p<0.05$, ***' $p<0.01$, ****' $p<0.001$.)

A forest plot of the estimates is provided as a basic overview of the model in Figure 3.8.⁵⁸ As in the previous section, the most important effect is CLAUSE TYPE, such that typical non-V2 environments strongly favour the Adv-Vfin order over V2 environments. As expected, this effect is less pronounced for the mixed environments. Figure 3.9 provides a visualisation of the predicted probabilities of the fixed effect CLAUSE TYPE over time, conditioned on random effects. The raw data before and after 1850 are also shown in Table 6.29, Appendix A, page 245.⁵⁹ For a visualisation of the remaining fixed effects, see Figure 6.36, Appendix A, page 246.

VERB TYPE is a highly significant factor in the private letters, with lexical verbs showing a far higher rate of Adv-Vfin than auxiliaries do. Unlike the newspapers, this effect is nearly absent in the category of modals. Also different from the newspaper results in the previous section is that PRONOMINALITY is a significant factor in the model for private letters. Unlike what would be expected based on Modern Icelandic (see Section 3.3.3), the *indefinites* actually have higher odds of occurring with Adv-Vfin and not the other way round. Note that the direction of this effect in the newspaper results is the same as reported here for the private letters.

ADVERB TYPE is not a significant factor in the private letter corpus and the effect of

⁵⁸Model evaluation: C index of concordance = 0.8710082. Somers' D_{xy} = 0.7420165 (observations = 6,108, individuals = 201). Classification accuracy: 83.79% of the choices among variants predicted correctly. Note that the number of observations in the model is lower than the number given in Table 6.29. This due to cells with missing data which are ignored by the model, typically unknown social ranks or a temporal dimension that is unknown at the level year/decade but known at a bipartite (before/after 1850) level.

⁵⁹For discussion, see also fn. 52, Section 3.4.1.2.

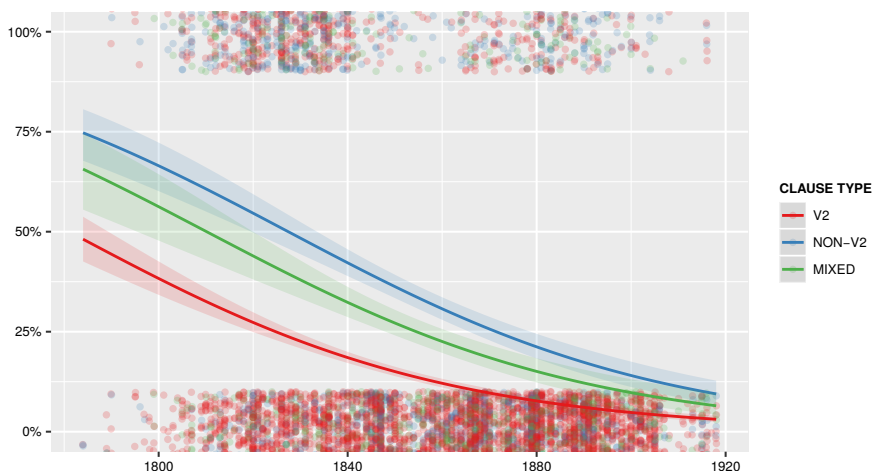


Figure 3.9. Predicted probabilities of **CLAUSE TYPE** over time, conditioned on random effects (lower values = less Adv-Vfin).

negation disfavouring Adv-Vfin is not even observable (slightly *favouring* Adv-Vfin), unlike the newspaper corpus. If the above suggestion is on the right track that the retarding effect of negation with regard to Adv-Vfin as compared to other sentence-medial adverbs in the newspapers is due to negation being prototypical somehow (iconic in terms of prescriptive dicta), it need not be surprising that this effect is not present in the private letter corpus, being less normative by hypothesis.

Turning now to **social factors**, it is clear that the variation attested in the private letter corpus is at least partly socially conditioned. Figure 3.10 plots the effect of social factors over time, conditioned on random effects for the individual. The broad scope of the variability up until the latter half of the century, visible in the shading of each line, is due to the scarcity of data in all categories except for the group of lettered individuals (see raw numbers in Table 3.8 below). Unlike peasants/labourers, the somewhat vague group of “other professions” is not significantly different from officials/lettered, but still appears to disfavour Adv-Vfin to a much lesser extent than peasants/labourers, suggesting that it is intermediate between two extremes.⁶⁰

Interestingly, **SPEAKER SEX** is a significant factor.⁶¹ Overall, females use Adv-Vfin to a lesser extent than males. It goes without saying that the sex/gender category as operationalised here, viz. as a large-scale, binary opposition, is overly simplistic and great care must be taken not to overstate the significance of this finding; after all, this use of gender is for exploratory purposes only. The size of this effect is great and strongly

⁶⁰Furthermore, if we separate unknowns from peasants/labourers and change the base comparison to peasants/labourers, other professions are still statistically distinct from that peasants/labourers.

⁶¹For the sake of completeness, it should be mentioned that the sex/gender effect is more pronounced in alternative models (not shown) that only included speakers the social status of which is known. I must leave this aspect to future research.

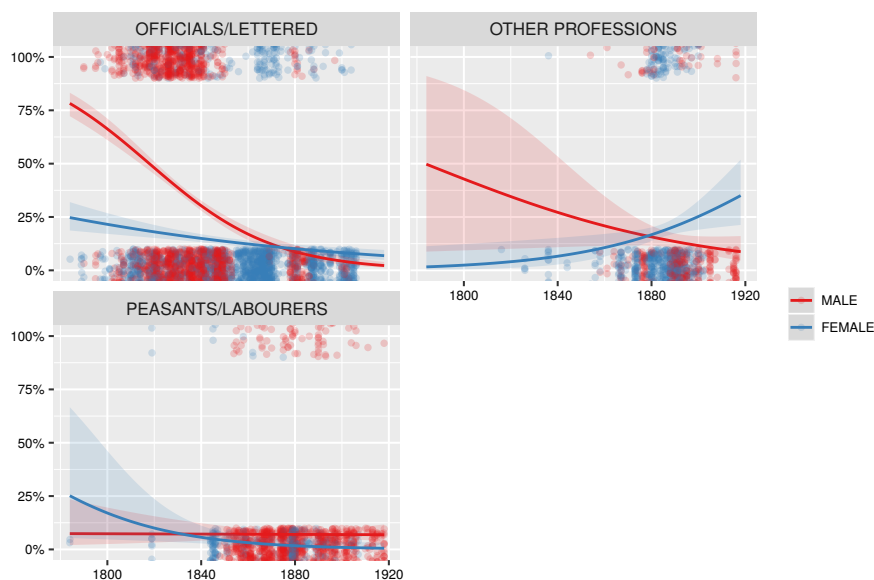


Figure 3.10. Predicted probabilities of social status across time and sex/gender, conditioned on random effects (lower values = less Adv-Vfin).

significant within the model ($p=0.00000533$). The AIC and BIC values are provided below for two equivalent models with and without sex/gender to further estimate the size of the effect based on the *anova* function (cf. Table 3.7, lower AIC/BIC = better fit to the data). Both AIC and BIC values are lower for the main model that includes sex/gender, suggesting that the inclusion of this factor is to be preferred over a model without it, despite the resulting increase in complexity. The difference between the models is strongly significant ($p\text{-value}=0.000004535$).

Df	AIC	BIC	logLik	deviance	Chisq	Chi Df	Pr(>Chisq)
Model ^B (♀)	4415.1	4495.7	-2195.5	4391.1			
Model ^A (♂)	4400.4	4487.8	-2187.2	4374.4	16.633	1	4.535e-05

Table 3.7. Model comparison with and without SEX/GENDER (*anova*).

What is particularly striking in the private letter corpus is that if we take the social status categories based on occupation into account, focusing on females in the officials/lettered intellectuals class (by association) versus females in the peasants/labourers group, the relative proportion of Adv-Vfin is considerably higher in the former group than in the latter (see also Table 3.8). As can be seen in Figure 3.10, elevated uses of Adv-Vfin are clearly observable among the higher echelons, although the effect is generally

	1784-1850		1851-1885		1886-1918	
MALES						
Officials/lettered	39.5%	(714/1809)	10.7%	(16/149)	12.1%	(4/33)
Other professions	–	(0/0)	22.5%	(25/111)	10.0%	(29/289)
Peasants/labourers	0.0%	(0/8)	6.7%	(48/721)	9.1%	(34/373)
Unknown	–	(0/0)	10.3%	(12/116)	10.8%	(8/74)
FEMALES (ranks mainly by association)						
Officials/lettered	17.4%	(112/643)	12.6%	(99/787)	7.5%	(28/375)
Other professions	5.0%	(1/20)	18.9%	(53/280)	19.0%	(62/327)
Peasants/labourers	9.1%	(7/77)	3.8%	(6/158)	1.9%	(1/52)
Unknown	0.0%	(0/6)	11.3%	(11/97)	10.9%	(6/55)

Table 3.8. Percentage Adv-Vfin across social status and gender in private letters over time.

discernible in other social groups as well but more subtle.⁶² Moreover, in the two major social groups, we observe a downward trend for both males and females in the officials/lettered group (by association in the case of women), but only for females in the peasants/labourers group, albeit only a very slight one. Social status in and of itself is also an important factor and evidence of a clear contrast between officials/lettered individuals on the one hand and peasants/labourers on the other is easily discernible in Figure 3.10, especially by abstracting away from the sex/gender differences: peasants/labourers are much less prone to using Adv-Vfin than officials/lettered writers.

A traditional crosstabulation of the raw data is provided in Table 3.8, showing SEX/GENDER, SOCIAL STATUS (*unknowns* as a separate category) and TIME PERIOD, split into three consecutive periods: 1784-1850, 1851-1885, 1886-1918. During the period in which Adv-Vfin occurs at its highest rate (the first period), the overall picture is unfortunately very sketchy. This is especially the case for male writers in the positions of peasants/labourers and other professions, for which there is simply no (or almost no) data from the first period. Evidence from the letters of the women is also rather thin for categories other than officials/lettered, but there is at least some data available from all three groups.

Within the high ranks, there is quite a dramatic decrease in Adv-Vfin over time and the women lag somewhat behind the men, until the last period. However, it must also be borne in mind that there are extensive individual differences within these groups. Interestingly, male writers of all social backgrounds eventually converge around the 10% mark in period III for using the Adv-Vfin order in private letters. There appears to be considerable variation in period II and III for female writers of other professions, with similar figures in period II for men and women. However, most of the cases among the female writers are due to a single speaker, Guðrún Jónsdóttir (GdrJon), responsible for 51 instances of Adv-Vfin (out of 175) in period II and 51 Adv-Vfin instances (out

⁶²Note also that the rather marginal group of other professions patterns partly the same way, except mainly for the last two periods where there is a steep increase in Adv-Vfin among the women. This increase is mainly due to a single individual (GdrJon, on which see below). This does not significantly affect/skew the overall patterns of males versus females in the model, with the effects being measured by the mixed-effects model over and above the level of the individual.

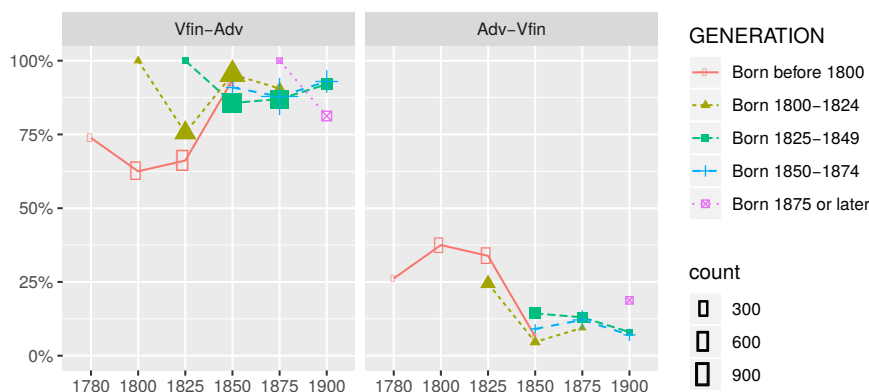


Figure 3.11. Verb-adverb placement across five generations of speakers (1784-1918).

of 242) in period III. Excluding GdrJon, other professions in period II thus amounts to merely 1% Adv-Vfin (2/195) versus 12.9% (11/85) in period III.⁶³

Among peasants/labourers, Adv-Vfin hovers around the 0-10% mark, showing a gradual decrease across the three periods only among the female writers. In a way, these women exhibit a similar trend as female writers in the officials/lettered ranks, although the end result in the officials/lettered ranks is the beginning point in the peasants/labourers ranks in terms of the relative frequency of Adv-Vfin. For the men in the peasants/labourers ranks, it is not at all obvious that any changes are taking place over time, in stark contrast to the male writers in the officials/lettered ranks, with other professions somewhere in between.

Finally, Figure 3.11⁶⁴ shows variation in verb-adverb placement through the lens of generations, divided into five generations from before 1800 until after 1875. There is an obvious temporal dimension in the data, with considerable variation in generations born until 1850, at which point there is a sharp decline in the frequency of Adv-Vfin. Social rank is clearly a part of the explanation for the high frequency of Adv-Vfin during the period before 1850, with a large number of observations for the higher echelons. This bias obviously affects and distorts the results, especially for the first two generations.

What about **regional variation**? In addition to the fixed effects in (60) above, the geographical factor **ORIGIN REGION**, referring to the speaker's place of origin, was originally tested with the following distinctions (the first item being the basis of the

⁶³As a speaker, GdrJon is difficult to categorise in that she is the descendant of commoners who had ties into networks of learned individuals, all her brothers became well-educated and her sister married into the officials class. Guðrún herself never married, but given the upward mobility of her other siblings, it seems disingenuous to place her in the commoners group. This partly applies to her brothers as well, who were placed among the learned individuals, for obvious reasons, but they, too, share the same common ancestry with their sister.

⁶⁴Here and elsewhere, the two facets of Figure 3.11 (left for the standard norm Vfin-Adv, right for the non-standard variant Adv-Vfin) are provided for clarity and in order to give an idea about the amount of data (token count) behind each bullet in the graph for either variant. The two facets represent only one degree of freedom as these are proportions of one and the same bivariate variable: 75% Vfin-Adv (left) thus simply translates into 25% Adv-Vfin (right).

comparison): the Southwest, West, Westfjords, North, East and South. Although all of these regions appeared to use Adv-Vfin to a slightly lesser extent than the Southwest, the effect was far from significant.⁶⁵ In contrast, a model containing an effect for REGION WRITTEN, referring to the region of the location at which the letter was written, did yield a significant effect for Southwest vs. North.⁶⁶ Location *written* is arguably not a very reliable indicator of regional variation, for obvious reasons. And while it certainly makes sense socio-historically that the Adv-Vfin variant could have had a stronger foothold in the Southwest (typically Reykjavík and its immediate surrounds), one would clearly have liked to see this effect corroborated by statistically significant evidence from the ORIGIN REGION variable as well, contrary to fact.

Up until now, we have mainly considered aggregated data, factoring in the individual level as random noise, as it were, by focusing instead on the role of language-internal and external factors. In the next two sections, we will consider the individual level in more detail, focusing first on inter-speaker variation and then on intra-speaker variation.

3.4.2.3 Inter-speaker variation

Let us begin by analysing speakers with at least 20 tokens of either variant, ordered by percentage of Adv-Vfin in descending order. The result is a selection of 52 scribes as shown in Table 3.9, which also shows birth year, the total number of tokens (both variants), the rate of Adv-Vfin, the origin of the speaker at the regional level, an approximation of the scribe's social status and, finally, the sex of the scribe.⁶⁷

Concerning the status of Adv-Vfin as a feature of educated intellectuals, as is often suggested, we indeed see in the “Top 10” a fair number of scribes likely to fit that profile. These range from Icelandic students at the University of Copenhagen (TorEgg, ThoHel), to officials such as bishop Geir Vídalín (GeiVid) and Bjarni Thorarensen (BjaTho), deputy governor, judge and poet. For these particular individuals, Adv-Vfin even appears to be the predominant variant. However, a rather striking fact about the distribution in Table 3.9 is that among the Top 10, four scribes are actually women. What these four scribes appear to have in common is that they have personal ties in the middle and upper classes. A priest's wife from the east of Iceland, Stefanía Siggeirsdóttir (=SteSig), tabulated in row six, interestingly uses Adv-Vfin over half the time (52%). As is the case with all letters in the private letter corpus, the letters were written to family and friends. Stefanía wrote 51 letters in the corpus, all addressed to her uncle, Páll Pálsson, who went by the title “student” (*stúdent*), an educated man and professional scribe. This also applies to the Copenhagen University students, who were writing home; in Torfi Eggerz's case, with a staggering proportion of Adv-Vfin of over 70%, writing to his parents and his brother.

⁶⁵Recoding into the Southwest vs. the rest of the country is also not statistically significant ($p=0.645329$).

⁶⁶The contribution of REGION WRITTEN in the alternative model does not undermine the significance of the other fixed effects discussed above. The North has an OR of 0.63 ($p=0.002503$), which indicates reduced odds for Adv-Vfin. Other areas in Iceland are not statistically significant but the OR is similar to the North: East = 0.60, South = 0.45, West = 0.89, Westfjords = 0.83. Finally, Copenhagen has an OR of 1.65, which indicates higher odds of Adv-Vfin, but the effect is not statistically significant ($p=0.224525$). I leave the further study of potential regional variation to future research.

⁶⁷This overview is complicated by various language-internal factors, most notably clause type and other structural factors as discussed above. See Section 3.4.2.4 below for a more sophisticated analysis of a smaller subset of speakers based on clause type.

INDIVIDUAL	ADV-VFIN	<i>n</i>	ORIGIN	RANK	SEX	BORN
TorEgg	71.4%	77	W	University student	m	1809
FriOlg	59.1%	22	NE	Peasant/Artisan	m	1834
GeiVid	56.0%	234	NE	Official	m	1761
ThoHel	54.9%	51	S	University student	m	1806
BjaTho	52.9%	806	S	Official	m	1786
SteSig	51.7%	143	E	Clergy	f	1842
AndFje	50.0%	22	W	Peasant	m	1835
SteSal	40.6%	32	NE	Sailor/Peasant	f	1834
IngJon	28.6%	290	W	Artisan/Clergy	f	1784
GdrJon	24.3%	415	SW	Other (?)	f	1856
JonTho	21.7%	23	NE	Peasant	m	1825
RagMag	21.7%	23	SW	Clergy	f	1786
MarDan	20.0%	25	E	Clergy	f	1853
SteTho	15.4%	39	NE	Artisan	m	1890
FinJon	11.0%	109	SW	Scholar	m	1858
GriTho	11.1%	90	SW	Scholar	m	1820
BalEgg	10.5%	57	NW	Teacher	m	1857
JohGud	10.0%	30	NW	Peasant	m	1823
KlmJon	9.7%	31	SW	Official	m	1862
JohHal	9.5%	74	NE	Sales clerk	m	1851
RagTho	9.5%	21	NE	Official	f	1738
GudMag	9.2%	65	SW	Peasant	f	1782
BenHal	9.1%	209	E	Servant	m	1845
VilJon	9.1%	33	SW	Postman	m	1870
RagDan	8.9%	90	E	Clergy	f	1859
BryPet	8.4%	83	NW	Lawyer/Official	m	1810
EriOla	8.0%	25	NE	Peasant/Bookbinder	m	1817
ArnHel	7.9%	453	Wf.	Clergy	m	1777
AsgFri	7.7%	52	NE	Servant/Peasant	m	1860
BjoOla	7.7%	26	E	Peasant	m	1801
SofDan	6.9%	233	E	Clergy	f	1858
AlfJon	6.7%	30	NE	Clergy	f	1794
MalJen	6.7%	30	SW	Official	f	1786
ThuHal	6.2%	48	NE	Clergy	f	1789
GunOdd	6.0%	150	NE	Artisan	m	1850
SigEin	4.8%	21	NE	Peasant	f	1819
SigPal	4.5%	449	E	Clergy	f	1809
HilJon	3.8%	26	NE	Merchant	f	1807
ThoSte	3.6%	28	NE	Peasant	m	1861
JakJon	3.1%	223	NE	Clergy	f	1835
ThoPal	2.8%	71	E	Official	f	1811
GudJon	2.7%	37	SW	Official	f	1865
LarBja	2.0%	98	W	Servant/Peasant	m	1849
KleBjo	1.7%	59	W	Freelancer/Carpenter	m	1829
EirJoh	1.4%	71	NW	Peasant	m	1862
AdaBja	0.0%	60	W	Servant	m	1860
AnnGud	0.0%	94	NE	Servant	f	1828
AslTor	0.0%	48	W	Teacher	f	1869
GudSku	0.0%	20	NW	Official	f	1740
GudZak	0.0%	93	Wf.	Housewife/Mistress	f	1845
SgjJon	0.0%	35	NE	Servant	f	1865
SigJon	0.0%	25	E	Peasant	m	1814

Table 3.9. Individuals with at least 20 examples of either variant, ordered by percentage of Adv-Vfin in descending order (n = total number of both variants).

We also find a high proportion of Adv-Vfin among some individuals from the middle to lower classes, although it is difficult to assess their social and economic status. Friðgeir Olgeirsson (FriOlG), a farmer's son and himself a farmer and saddler, uses Adv-Vfin in nearly 60% of the cases. Note, however, that Friðgeir was writing to his son, Einar Friðgeirsson, who is a priest, suggesting that Friðgeir himself was probably reasonably well-off. Similarly, perhaps, Andrés Fjeldsted (AndFje) went abroad to Scotland to study ship-building, tinsmithing and canning, later to become not only a farmer but the head of his parish (*hreppstjóri*). He wrote his letters to Friðgeir Olgeirsson's son, addressing him as *kæri vin!* 'dear friend' in one of his letters. The fact that speakers like Friðgeir and Andrés also used Adv-Vfin extensively shows that high Adv-Vfin rates should not be regarded purely as a phenomenon of the educated elite.

The language use of the other three female speakers in Table 3.9, who produced Adv-Vfin in the 20-40% range, also points in a similar direction. Stefanía Salómonsén (SteSal), daughter of a merchant and farmer, was the wife of Benedikt Jónsson Salómonsén (1827-1862), a farmer and sailor (captain), although widowed by the time she wrote the letters.⁶⁸ Ingibjörg Jónsdóttir, the daughter of a priest, was clearly higher placed, being the wife of a goldsmith and mother of Grímur Thomsen (GriTho), a scholar, Member of Parliament, poet and also farmer. In contrast, Guðrún Jónsdóttir (GdrJon) came from a poor family of common origin, although she may have received some form of education. Her mother was a labourer and her father worked among other things as a bookbinder, a travelling salesman and police officer. He was noted for his eagerness for books and reading and sometimes referred to as a 'folk scholar' (*alþýðufræðimaður*). Guðrún Jónsdóttir never married but her other siblings, all present in Table 3.9, were upwardly mobile. Her sister Guðný married a sheriff and all their brothers received formal education. Interestingly, her family uses Adv-Vfin to a much lesser extent than Guðrún (24%) herself does: FinJon (Finnur) 13%, KlmJn (Klemens) 10%, VilJon (Vilhjálmur) 9%, GudJon (Guðný) 3%, and their mother, AnnGud (Anna Guðrún), as low as 1%.

Perhaps the most unexpected finding in Table 3.9, however, is how relatively few speakers *never* used Adv-Vfin in their letters, merely five out of the 52 scribes selected above. Note also that we have only been referring to the use of Adv-Vfin with sentence-medial adverbs. These speakers may still use Adv-Vfin with other adverbs, e.g. GudZak in the shared bottom place in Table 3.9, who produced the example in (61):

- (61) *annars held jeg þaúg hefðu farið þó þeim kannskje líki ekki*
 by.the.way think I they had_{SUBJ.} gone although them maybe like_{SUBJ.} not
giftingin að ölluleit
 marriage_{DEF} at all.way
 'By the way, I think they would have gone even though they maybe do not like
 the marriage in every way.' (GudZak-1866-06-22.xml)

The adverb *kannskje* 'maybe' belongs to the set of Adv-Vfin adverbs that may precede the finite verb in subject-initial main clauses and therefore 'does not count' the way the variable has been defined—GudZak's example in (61) is Adv-Vfin because the negation follows the finite verb.

⁶⁸The 1880 consensus, when Stefanía is 45 years of age, states that her financial means are provided by her children. In her letters, she also indicates that she lives in poverty.

	Typical V2 env.			Mixed env.			Typical Non-V2 env.		
	Vf.-Adv	Adv-Vf.	<i>n</i>	Vf.-Adv	Adv-Vf.	<i>n</i>	Vf.-Adv	Adv-Vf.	<i>n</i>
<i>BjaTho</i>									
1811-1819	51.6%	48.4%	31	12.5%	87.5%	24	7.4%	92.6%	27
1820-1829	62.2%	37.8%	148	40.0%	60.0%	70	28.0%	72.0%	107
1830-1841	73.1%	26.9%	175	44.7%	55.3%	94	30.0%	70.0%	130
<i>GeiVid</i>									
1790-1805	73.1%	26.9%	26	57.1%	42.9%	14	13.0%	87.0%	23
1806-1815	62.9%	37.1%	35	61.1%	38.9%	18	13.0%	87.0%	23
1816-1823	57.1%	42.9%	35	29.6%	70.4%	27	24.0%	76.0%	25
<i>GdrJon</i>									
1878-1882	81.2%	18.8%	69	78.9%	21.1%	19	65.8%	34.2%	38
1883-1887	84.6%	15.4%	52	61.1%	38.9%	36	36.4%	63.6%	22
1888-1892	95.7%	4.3%	47	78.6%	21.4%	28	60.0%	40.0%	20
1893-1902	90.2%	9.8%	41	72.7%	27.3%	22	57.1%	42.9%	21
<i>IngJon</i>									
1807-1819	88.6%	11.4%	70	67.9%	32.1%	28	58.8%	41.2%	34
1820-1829	88.9%	11.1%	27	50.0%	50.0%	6	42.9%	57.1%	14
1830-1839	82.4%	17.6%	34	66.7%	33.3%	21	53.3%	46.7%	15
1840-1852	81.0%	19.0%	21	16.7%	83.3%	12	50.0%	50.0%	8
<i>SigPal</i>									
1819-1829	100.0%	0.0%	7	40.0%	60.0%	5	66.7%	33.3%	3
1830-1839	70.0%	30.0%	10	75.0%	25.0%	4	83.3%	16.7%	6
1840-1849	97.9%	2.1%	47	91.3%	8.7%	23	86.7%	13.3%	15
1850-1859	100.0%	0.0%	36	92.9%	7.1%	14	100.0%	0.0%	16
1860-1871	99.2%	0.8%	124	98.4%	1.6%	64	98.5%	1.5%	67
<i>SteSig</i>									
1863-1869	60.9%	39.1%	46	47.1%	52.9%	17	43.8%	56.3%	16
1870-1877	54.5%	45.5%	33	40.0%	60.0%	15	12.5%	87.5%	16

Table 3.10. Verb-adverb placement for data-rich letter-writers in V2 and non-V2 environments over time (highest and lowest values boldfaced).

3.4.2.4 Intra-speaker variation

In the previous section, we saw extensive differences in verb-adverb placement at the individual level (inter-speaker variation) as well as variation that is socially conditioned, in terms of social status and/or gender, but we did not consider variation within the individual (intra-speaker variation). Table 3.10 singles out six speakers for whom there is enough data to zoom in further, showing the three-way approximation of typical V2 and non-V2 environments familiar from the previous sections. Observe that the highest and lowest values (in boldface) are nearly always found in typical non-V2 and typical V2 environments, respectively, with mixed environments somewhere in between.⁶⁹ Fluctuations in the frequencies across time periods are usually relatively minor and suggest that only the frequency with which the Adv-Vfin variant is selected undergoes slight changes, but arguably not the underlying syntax of verb-adverb placement—the distribution thus appears to retain its core properties related to clause structure.

Diachronic developments in the language use of these speakers do *not* follow the same downward trend observed in the newspaper corpus, except perhaps in the case of

⁶⁹There are but four minor exceptions: IngJon in the last period and SigPal in the second and the two last periods. In all these cases, the raw numbers are very low.

Sigríður Pálsdóttir (see below). Obviously, both Bjarni Thorarensen and Geir Vídalín were writing in a period prior to the overt stigmatisation of the Adv-Vfin variant by Gíslason (1844), in fact wholly or mostly prior to any influence that the journal *Fjölnir* could have had on their language use more generally with regard to language purism. Geir Vídalín wrote 64 letters in the corpus, dated between 1790–1823. The overwhelming majority of the letters were written to his friend Bjarni Þorsteinsson and the data are too limited to study individually even when all remaining recipients are taken together as a group.

The letters of Bjarni Thorarensen, however, allow for a more detailed breakdown across a number of recipients. Bjarni wrote 203 private letters in the corpus over a thirty-year period between 1811–1841. Table 3.11 shows the proportion of Adv-Vfin in V2 and non-V2 environments across the four most frequent recipients, in addition to a fifth group where the remaining recipients have been combined. Notice that the frequency is not much different when writing to his friends and colleagues Bjarni Þorsteinsson (an ex-schoolmate and official), Finnur Magnússon (a barrister, scholar and secret archivist), Grímur Jónsson (an official), nor his father-in-law Bogi Benediktsson, who was also a graduated student, scholar and wealthy store manager (cf. ÍÆ I:262f.) nor, indeed, any other recipient, including his mother and mother-in-law, brother and sister. There are certainly some fluctuations but the overall pattern is very similar across recipients.

<i>BjaTho</i>	BjaÞor	<i>n</i>	BogBen	<i>n</i>	FinMag	<i>n</i>	GriJon	<i>n</i>	Other	<i>n</i>
Typical V2	42.1%	121	38.5%	39	19.0%	63	35.8%	67	25.0%	64
Mixed	72.4%	58	56.3%	16	48.6%	37	65.9%	41	52.8%	36
Non-V2	74.6%	71	76.9%	26	68.6%	70	72.9%	48	75.5%	49

Table 3.11. Proportion of Adv-Vfin in Bjarni Thorarensen’s letters in V2 and non-V2 environments, per recipient (*n* = total number of both variants).

In contrast to Bjarni and Geir, the women Ingibjörg Jónsdóttir and Sigríður Pálsdóttir wrote letters both prior to and following the prescriptive remarks of Gíslason (1844). Ingibjörg wrote a total of 123 letters between 1807–1852, the overwhelming majority of which were addressed to her brother (*n*=116), Grímur Jónsson, the rest were addressed to her uncle (*n*=2) and niece (*n*=5). Ingibjörg’s overall use of Adv-Vfin is nearly 30% and the typical V2 and non-V2 contexts appear fairly stable over a period of more than four decades, with no discernible decrease towards the end of this period.⁷⁰ Sigríður Pálsdóttir wrote 232 letters in the corpus, dated between 1819–1871 and all addressed to her brother. SigPal does not use the Adv-Vfin variant frequently but most of her examples of Adv-Vfin occur in the period up until 1845, fourteen in total, in addition to two examples in an undated letter. These numbers may be much too small to interpret but it is striking that the number of letters she wrote is much greater between 1846–1871 than between 1826–1845 (172 letters vs. 60 letters), yet there are only *four* examples of Adv-Vfin in the period after 1845. From this, one might conclude that a change took place in her language use over time or else we would have expected dozens more examples.

⁷⁰Ingibjörg’s son, Grímur Thomsen (GriTho), uses Adv-Vfin to a much lesser degree than his mother (see Table 3.9). This may very well have something to do with closeness to standard norms with him being a writer, in addition, perhaps, to his ties to language puritans of the *Fjölnir* society.

In one of her letters, there is also an example of a correction of double negation, where she interestingly strikes out the latter instance of the negation, thereby correcting it to Adv-Vfin rather than Vfin-Adv:

- (62) *mér hefði þókt nógur skadi ad verda fyrir því ad fá ónytt skrifli*
 me had_{subj.} felt enough damage to become for it to get useless piece.of.junk
eins og þú getur til ad hefði kuñad ad verda, hefði verid bedid um
 as and you guess to that had_{subj.} can_{pa.pple.} to become had_{subj.} been asked about
gamla jafnvel þó mér ekki hugkvæmdist ekki añad i fyrstuñi
 old even though me not came.to.mind not other in first_{DEF.}
 ‘I think it would have been enough damage to get a useless piece of junk
 [furniture—HFV] as you can imagine would have been, had the old one been
 asked for, even though I didn’t think of it at first.’ (SigPal-1863-05-23.xml)

If the decline in the use of Adv-Vfin observed in SigPal’s letters over time were due to her consciously trying to avoid the Adv-Vfin variant, this correction is clearly not what you would expect. Aspects relating to changes in her social network and relocation patterns may arguably also be a factor. In the late 1820s and early 1830s, she wrote from Reykjavík in the Southwest and between 1833–1840 she wrote from Reykholt in the West, where she lived with her husband, the priest Þorsteinn Helgason. The reader may recall Þorsteinn from Table 3.9, who had been a student at the University of Copenhagen and apparently a high-frequency Adv-Vfin speaker. It is, therefore, during her stay in Reykjavík and during her later marriage to Þorsteinn that Sigríður produces the most instances of Adv-Vfin.⁷¹ After her husband’s death in 1839, Sigríður married another priest, Sigurður G. Thorarensen, relocating to the east of Iceland (Breiðabólstaður). The fact that she lived in the countryside during most of the later period, far removed from Reykjavík and urban centres in general, may very well have been of some influence as well, though all of this remains guesswork.

Another frequent Adv-Vfin speaker singled out in Table 3.10, and already mentioned above, is Stefanía Siggeirsdóttir, who wrote in total 53 letters (one of which is undated). The letters are addressed to her uncle, Páll Pálsson, over a period of 15 years between 1863-1877. Interestingly, given the fact that the overall frequency of Adv-Vfin decreased sharply in the public domain (cf. newspapers and periodicals above) throughout the latter half of the century, there is no sign of the feature receding in Stefanía’s letters over these 15 years. In fact, the frequency of Adv-Vfin in her letters actually increases in all three environments during this time.

The final letter-writer we zoom in on in this section is Guðrún Jónsdóttir, who wrote 153 letters in the period 1878-1902 to two recipients, her brother Finnur Jónsson and her father, Jón “Borgfirðingur” Jónsson. The letters to her father are fewer in number and are mainly from the two last periods, whereas the letters to her brother cover the whole period. There is a peak in the use of Adv-Vfin in non-V2 and mixed environments in the second period (1883-1887) but the pattern is otherwise quite stable, with no clear decrease in use. Also when comparing letters written to the two different recipients, the pattern is rather similar, cf. Table 3.12.

⁷¹Note that it is not the case that Þorsteinn wrote Sigríður’s letters for her because they are in her handwriting. The only exception to her writing her own letters, as far as I can tell, is a letter from her childhood, written in 1819 when Sigríður is eight years old. In the letters following, written in the early 1820s, Sigríður is already practicing (and apologising for) her handwriting.

<i>GdrJon</i>	FinJon (brother)	<i>n</i>	JonBor (father)	<i>n</i>	Σ
Typical V2	15.6%	154	5.5%	55	12.9%
Mixed env.	29.5%	78	25.9%	27	28.6%
Typical Non-V2	40.5%	79	54.5%	22	43.6%

Table 3.12. Proportion Adv-Vfin in Guðrún Jónsdóttir’s letters in V2 and non-V2 environments, per recipient (n = total number of both variants).

To summarise, there seems to be very little if any clear evidence for individual letter writers changing their language use over time in the direction of standard norms. In general, the patterns of Adv-Vfin are also quite stable within speakers across clause type and across different recipients.

3.4.3 Student essays

3.4.3.1 Overall trends in the data

The student essays follow largely the same overall downward trend that we witnessed in the newspaper corpus, as shown in Table 3.13. The Adv-Vfin pattern receded from roughly 35% in the first period (1852-1874) down to approximately 9% in the final two periods combined (1895-1906). In comparison to the newspapers, the first two periods (1852-1884) exhibit on average considerably higher frequencies of Adv-Vfin in the essays, which is perhaps rather unexpected given the view that the Reykjavík Grammar School was a primary force in the implementation of the standard language. The rate of uptake of standard norms seems to suggest that the student essays are lagging somewhat behind the newspapers, although it should be borne in mind that some newspapers were shown to exhibit frequencies in the 20-30% range during (and beyond) this period as well (cf. Table 3.5). However, as we will see further below, the aggregated data in Table 3.13 may also conceal a potential interaction with the variable GRADUATION SCORE over and above the effect of the structural CLAUSE TYPE factor and the TIME factor.

	Vfin-Adv		Adv-Vfin	
1852-1874	64.7%	(44/68)	35.3%	(24/68)
1875-1884	72.1%	(62/86)	27.9%	(24/86)
1885-1894	84.4%	(38/45)	15.6%	(7/45)
1895-1900	94.1%	(48/51)	5.9%	(3/51)
1901-1906	87.5%	(49/56)	12.5%	(7/56)

Table 3.13. Proportion Vfin-Adv/Adv-Vfin orders in student essays (1852-1906).

3.4.3.2 Modelling the variation

This section presents a statistical analysis of verb-adverb placement in the student essays, expanding on a previous study of a near identical data set (Viðarsson 2017b). The data set is very limited ($n=306$), which makes it hard to study the effects of both

language-internal and language-external factors. For this reason, CLAUSE TYPE, the most important language-internal effect (cf. Sections 3.4.1.2 and 3.4.2.2), was selected in addition to the language-external effects GRADUATION SCORE, SCHOOL GRADE, TIME PERIOD and a random effect for the individual:

- (63) a. CLAUSE TYPE (three-level factor: *typical V2*, *typical non-V2*, *mixed contexts*)
 b. GRADUATION SCORE (three-level factor: high (honours, 1st), low (2nd, 3rd), none)
 c. SCHOOL GRADE (two-level factor: *grades 1-3*, *grades 4-6*)
 d. TIME PERIOD (continuous, periods 1-5: 1852-1874, 1875-1884, 1885-1894, 1895-1900, 1901-1906)

A forest plot of the estimates is provided as a basic overview of the model in Figure 3.12.⁷² Note that the language-external effects in the student essays model are all educational variables. This even applies to the three time periods (1852-1880, 1881-1895 and 1896-1906) subsumed under the factor TEACHERS, which correspond to changes in teachers responsible for teaching the subject Icelandic (see Viðarsson 2017b). Since TIME PERIOD in the original design of Viðarsson (2017b) was at the same time conceived of as a TEACHER variable, it made sense to treat it as categorical. However, the present model deviates from that model in that TIME PERIOD is now treated as a continuous variable, as in the previous sections.

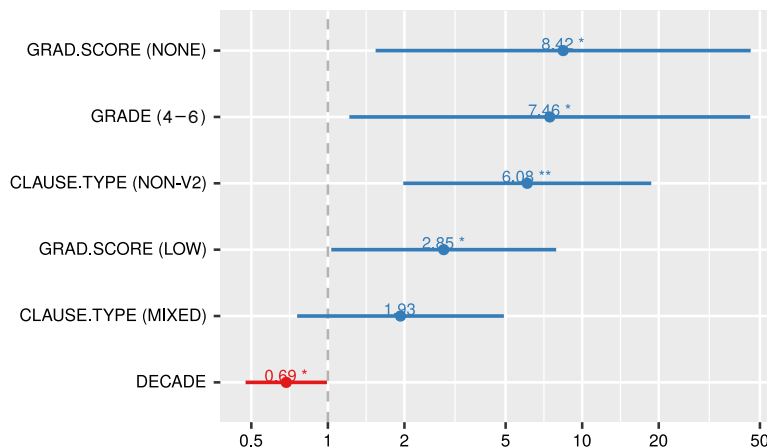


Figure 3.12. A forest plot of estimates with odds ratios of fixed effects: OR>1 (blue) favours Adv-Vfin, OR<1 (red) favours Vfin-Adv. (*' $p<0.05$, '**' $p<0.01$, '***' $p<0.001$.)

⁷²Model evaluation: C index of concordance = 0.9203958, Somers' D_{xy} = 0.8407916. Classification accuracy: 88.24% of the choices among variants predicted correctly (number of observations = 306).

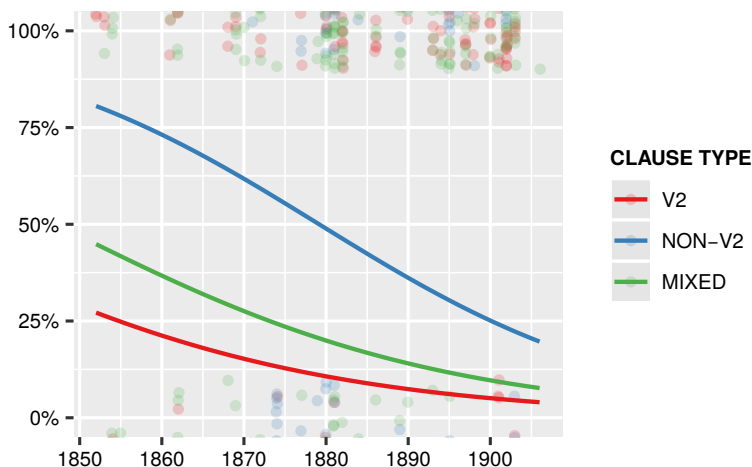


Figure 3.13. Essays: Predicted probabilities of `CLAUSE TYPE` over time (1852-1906), conditioned on random effects (lower values = less Adv-Vfin).

Unlike above, `CLAUSE TYPE` treats all adverbial clauses as mixed and declaratives as V2 (cf. also Heycock and Wallenberg 2013, Viðarsson 2017b), which means in effect that typical non-V2 environments are confined to relatives and indirect questions. This definition of `CLAUSE TYPE` corresponds to the basic types in Figure 3.15 below. Despite these slight differences in annotation, the distribution is largely the same: Adv-Vfin has the strongest foothold in non-V2 environments throughout the period and weakest in V2 environments, mixed environments being in between the two.

Interestingly, the present model shows a significant effect of not only `CLAUSE TYPE` but also `GRADE` and `GRADUATION SCORE`.⁷³ The `GRADUATION SCORE` effect is significant for both *low* and *no scores* as compared to *high scores*. In other words, students with a low graduation score have a stronger tendency to use Adv-Vfin than students with a high graduation score.⁷⁴ Not surprisingly given the small size of the

⁷³The fixed effect `GRADE` was left out in Viðarsson (2017b) since it did not improve that model. Additional linguistic variables were also tested as a part of the present model and the way these variables affect the choice of each variant was very similar to that reported for the newspapers in Section 3.4.1.2. Additional factors favourable to the choice of Adv-Vfin include modals as well as finite lexical verbs (compared to auxiliaries) and pronominal subjects (compared to nominals). In terms of adverb type, the negation favours the Vfin-Adv variant (compared to other sentence-medial adverbs), which begs the question whether the negation was perhaps more salient to the students in the Adv-Vfin order than other adverbs—prescriptive remarks tend to target the negation in particular.

Presumably due to the small corpus size, this increased complexity comes at a cost for the educational variable(s); incorporating the additional linguistic factors reduces the evidence for any statistical contribution of `SCORE` in the model. Nonetheless, the direction of the trends remains the same. Because the data are presently so very scarce, I have opted to go with the maximally minimal model with one internal and two external factors, in addition to the temporal aspect and the individual random effect.

⁷⁴A model with a continuous version of `SCORE` yields a similar result.

data set, the dropouts show a large standard error ($=0.8664$) in GRADUATION SCORE (*none*), but the std. error is lower in the case of *high* vs. *low* grades ($=0.5195$). GRADE, while significant, also has a large standard error ($=0.9259$). Surprisingly, perhaps, the effect of GRADE is the opposite of what one would expect if progression of study would correlate with a lower proportion of Adv-Vfin.⁷⁵

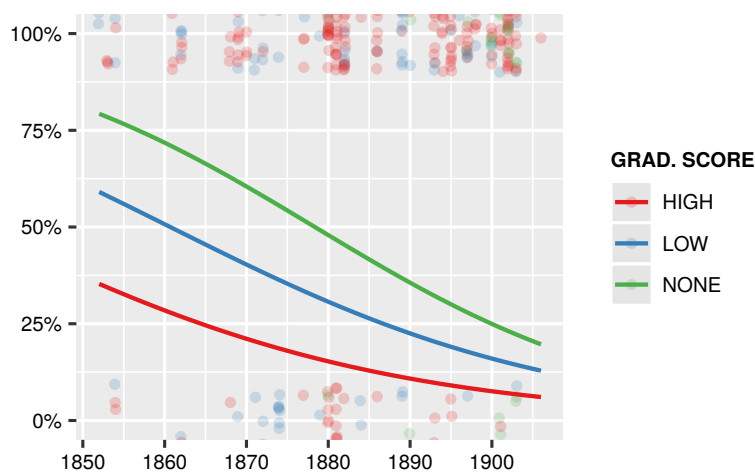


Figure 3.14. Essays: Predicted probabilities of GRADUATION SCORE over time, conditioned on random effects (lower values = less Adv-Vfin).

Social factors are another potentially relevant factor. While the students can be regarded as a privileged group of speakers mainly from the higher echelons of society, they are not a wholly homogeneous group. Ideally, one would like to take into account the social background of these students, especially with regard to the social status of the family as a whole, in part determined by the social background and occupation of the students' fathers (see Sections 2.1.3 and 1.3.5), but also based on factors such as geographical origin. However, there is little to suggest that the distribution reported in this section is greatly influenced by social ranks, e.g. in terms of officials vs. farmers.

I have attempted to enrich a subset of the corpus, confined to relative clauses, with the occupation of each student's father (data from ÍÆ I-VI). Zooming in on relatives across the whole period, it does not appear to be the case that officials' sons are any different from farmers' sons in the average use of the Adv-Vfin order: 53.8% ($n=13$), 50.0% ($n=20$) and 46.7% ($n=15$) for officials, farmers and rest, respectively. The rest

⁷⁵Study progress in terms of grade (or class) of course need not correlate with or translate into a lower proportion of non-standard features—score, indeed, might be a more likely candidate to do so. However, it should also be borne in mind that the data are not evenly distributed across the grades (see Section 2.1.3). In fact, data from the lower grades more often than not date from towards the end of the century, whereas the higher grades tend to date from earlier during the period. If time period is a factor as well, this will obviously skew the results with regard to this feature.

category refers to various occupations, including: teacher, county scribe, store manager, carpenter/captain and shoemaker. These data are admittedly very slender, but they are still suggestive in my view. I must leave a full analysis of social factors in this material to future research.

3.5 Discussion

The first question that we will now try to address is the effectiveness of standard norm implementation. From the perspective of the language history ‘from below’ approach, the results of the preceding sections can only suggest that the implementation of a standard language norm in terms of verb-adverb placement was ‘partly successful’. As we have seen, there is very little evidence of what we can refer to as ‘ordinary people’ shifting their language use in the direction of the standard norm, but this is further complicated by the fact that the stigmatised, non-standard Adv-Vfin variant was arguably never very frequent among speakers in this group in the first place. The scarcity of non-Adv-Vfin speakers suggests that Adv-Vfin was a rather widespread feature in 19th-century Icelandic (certainly compared to Modern Icelandic, cf. above), although speakers belonging to or sharing ties with the middle and upper classes used it more on average than those who were lower placed. Among the higher ranks, in contrast, where this variant is also found to a greater extent before and during the first half of the 19th century, we do see a sharp reduction in its use. Interestingly, female speakers associated with the officials/lettered group partake in this trend as well, albeit much more gradual (see Figure 3.10).

What also speaks in favour of standardisation playing a role is the fact that the proportion of Adv-Vfin for the males in the group of officials/lettered individuals during the last quarter of the 19th century is even lower than for females associated with this group, the former arguably being more successful in suppressing the variant than the latter due to their educational background. What this suggests is that elevated uses of Adv-Vfin were not only linked to the higher echelons due directly to these speakers’ educational backgrounds but rather through shared social networks, a key feature thus arguably being engagement in different *practices* rather than sex/gender per se.⁷⁶

Here, the evidence again suggests that women from the higher echelons actually pattern with their male counterparts to some degree, which in itself is very interesting and arguably a matter of social status above anything else. The fact that women also use Adv-Vfin to a considerable extent thus strongly suggests that it had a foothold in speakers (and/or writers) beyond the educated elite proper and/or those who went abroad to Copenhagen to study (cf. also Hróarsdóttir 1998). Defining style as “the locus of the individual’s internalization of broader social distributions of variation” (Rickford and Eckert 2001:1) leaves open the possibility that even unlettered individuals who are somehow associated with the educated ruling class, due to shared social backgrounds and perhaps “greater access to the situations and registers in which the features occur more frequently” (Finegan and Biber 2001:245), will make (more frequent) use of the Adv-Vfin variant.

⁷⁶I thank the audience at the Sociolinguistics Circle (2017) in Tilburg for valuable discussion on this point.

Interestingly, the situation in the group of peasants/labourers is quite the opposite of what we found among the officials/lettered group: there is little diachronic development to be seen and the proportion of Adv-Vfin throughout the latter half of the 19th century is actually higher for male peasants/labourers than for the female ones. As these women must have had a rather vague conception of written standard norms in comparison to their formally educated male counterparts in the higher ranks (and perhaps females associated with this class), it is likely that this apparent greater conformity to norms has less to do with the standard norms themselves and more to do with the severely demoted prestige of (seemingly) Danish features as a part of the standard language ideology and linguistic purism in general in the latter half of the 19th century onwards. The suggestion then would be that male peasants/labourers paid less attention to the social stigma attached to these features, much like the modern sociolinguistic literature proposes for stigmatised features. Strikingly, male participants in the Icelandic Dialect Syntax project were also significantly more likely to accept Adv-Vfin than female participants, whereas education levels were shown not to be significant (Þráinsson and Angantýsson 2015:322f.).⁷⁷

Recall the suggestion by Heycock and Wallenberg (2013) that the frequent use of Adv-Vfin in the early modern period is to be regarded as being “due to a process of calquing written Danish” (2013:130) and that it was “a phenomenon of the very literary written language, and never affected the spoken language” (2013:152). Two observations made above already cast some doubt on or at least call for further refinement of this hypothesis: (i) the Adv-Vfin variant is not confined to literary or formal texts, but also occurs in more colloquial written texts, as evidenced by private letters; (ii) Adv-Vfin is also used to a considerable degree by women, who were not highly educated, and, more generally, is attested in speakers of various social backgrounds.

Moreover, if variable verb-adverb placement was not simply a part of a speaker’s grammatical system like any other, artificial calquing of a Danish trait might lead us to expect either near perfect main vs. subordinate clause asymmetries according to Danish norms (for proficient scribes) or, since this hypothesis takes Adv-Vfin to be a contrived feature presumably with no basis in the spoken language, to exhibit a pretty random distribution within the individual, e.g. across clause types, over time or depending on the addressee/recipient of the letter. For those speakers who produced a fair amount of letters and/or wrote over a period of some years, even decades, we have been able to see that their usage is rather stable, both in time and with regard to different addressees. In my view, this state of affairs should be taken as potential evidence that the use of this feature ought not to be considered as imperfect imitation of the syntactic patterns of Danish, as (tentatively) suggested by Heycock and Wallenberg (2013), but rather as variable linguistic systems in their own right.

It should also be borne in mind that if speakers were consciously attempting to imitate Danish, surely cases of hypercorrection should abound where speakers e.g. violated the V2 constraint in main clauses and/or with sentence-initial elements other than subjects, such as topicalised objects. I have come across one instance of Adv-Vfin

⁷⁷As Þráinsson and Angantýsson (2015:323) point out, the use of Adv-Vfin appears to be on the rise in Modern Icelandic, but the change apparently is not being led by women. However, given the sociohistorical background and stigmatisation of this variable (cf. above), this need not come as a surprise (cf. Labov’s Conformity Paradox, see Section 1.2.2).

in a subject-initial main clause and a handful of cases involving adjunct-initial embedded clauses, which strike me not so much as V2 violations but rather instances of adverb stacking in the left periphery (see Viðarsson 2014:9f.).

Moving on to the other corpora, taking a view ‘from above’, there is clear evidence that the variable was targeted at the Reykjavík Grammar School and corrected in student assignments during the latter half of the 19th century. This is also partly reflected in the students’ language use such that a low GRADUATION SCORE was shown to correlate with higher odds of the Adv-Vfin variant in comparison to students with high scores. It would be very exciting to try to reproduce this finding on the basis of a much larger corpus of transcribed student essays, in the hope of making it more robust. A larger sample would potentially also allow us to better interpret other educational variables such as grade (class) and the teacher variable.

The effect of the grammar school is arguably immediately relevant to the newspaper corpus, since newspaper editors would typically have received their (secondary) education from that very school. It should therefore come as no surprise that the newspapers patterns in many ways not only with the student essays but also the letter-writers from the group of officials/lettered individuals. Still, the ‘smoking gun’ of standardisation, in my view, remains the evidence from the student essays, both in terms of the educational variable GRADUATION SCORE (cf. Figure 3.14) and the Neg-Vfin corrections in this material (cf. Figure 3.3 and footnote 39). Recall that until the middle of the 19th century, the proportion of Adv-Vfin in relatives is 65% on average (cf. Table 6.29) and still nearly 50% around 1850 in typical non-V2 environments (cf. Figure 3.5). Over the course of the next 50 years, the proportion of Adv-Vfin is more than halved in non-V2 environments (below 25%), and again towards the end of the first quarter of the 20th century.

These overall patterns suggest that the eradication of the Adv-Vfin pattern from the standard did not go quite as smoothly as one might expect based on the traditional narrative of the implementation of the Fjölnir norms. Since the Adv-Vfin feature persists in the newspapers for decades after it had been renounced in the standard, still more than marginally attested in the last period studied, and regularly occurs in the private letters, I suggest that extremely low frequencies as attested in *Ísafold* (1900-1924) be regarded as ‘artificially’ low—i.e. evidence of a suppression of a variable feature in the vernacular. In this last period, we even start seeing (near) categorical use of the standard Vfin-Adv variant in certain newspapers and Adv-Vfin has become a very rare feature in the papers during the 1950s onwards, even in relative clauses (cf. Figure 3.6). It would be difficult to explain the development we see in these public texts during the decades after 1850 if not at least partly due to standardisation.

In the rest of this chapter, our focus will shift to the syntax of verb-adverb placement, with a brief overview of previous analyses of verb-adverb placement in Icelandic and a suggestion for how the variation revealed in the previous sections can be analysed.

3.6 Previous analyses

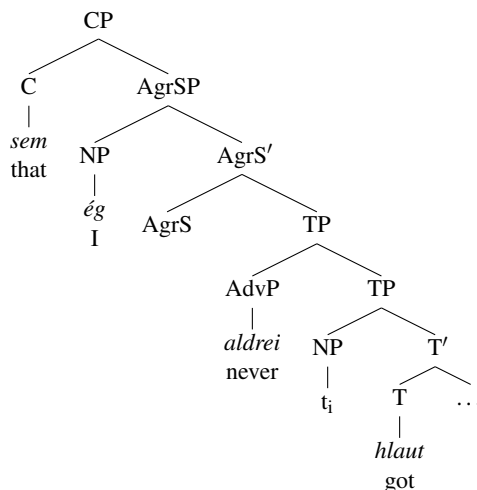
On standard assumptions in generative syntax, sentence-medial adverbs occur in a fixed location in the clausal structure based on which the position of other objects can be determined (see Section 3.3.2). In that sentence-medial adverbs, including the negation in the Scandinavian languages, are typically assumed to be attached to the *vP/VP*, these are regarded as diagnostics for verb movement: *VFIN* > *SADV* implies verb movement, *SADV* > *VFIN* implies *V-in-situ*. This means that the verb does not occupy the same position in (64) as it does in (65), the most straightforward analysis being that the former features verb movement (*V2*), whereas the verb in the latter is still in *VP*, just as in Mainland Scandinavian.

- (64) *já og svo skammdegið sem mjer er ætíð svo leiðinlegt*
 yes and then the.midwinter which me is always so dull
 (AnnGud-1879-10-26.xml)
- (65) *Eg þakka þjer fyrir góða vinsemd sem þú ætíð auðsýndir mjer frá því*
 I thank you for good kindness which you always bestowed.on me from the
fyrsta við kintustum við til þessa sýðasta er við sá ustum
 first we got.to.know_{MID} (we) until the last when we saw_{MID}
 (JonTho-1865-10-14.xml)

This view is indeed defended by Wiklund et al. (2007) and Sigurðsson (1986) also proposes an analysis, based on Platzack (1986), analysing Icelandic cases like (65) simply as in Mainland Scandinavian, i.e. “probably not due to an adverbial fronting” (Sigurðsson 1986:146). In contrast, the analysis provided by Bobaljik and Thráinsson (1998), further developed by Angantýsson (2001, 2007, 2011) and Thráinsson (2010), departs from the assumption that sentence-medial adverbs have a fixed position (in Icelandic) and instead regards cases like (64) and (65) as being primarily about variable attachment sites of the adverb, locating the verb in the inflectional domain—in other words, exhibiting *V-to-I* in both cases.

For Bobaljik and Thráinsson (1998), exceptional adjunction is a necessary departure from basic assumptions because if *Adv-Vfin* as in (65) were analysed on par with Mainland Scandinavian (as *V-in-situ*), Icelandic would be a counterexample to the Rich Agreement Hypothesis (RAH), a hypothesis they are trying to defend (for further discussion, see e.g. Vikner 1995, Rohrbacher 1999, Bobaljik 2002, Wiklund et al. 2007, Thráinsson 2010, Koeneman and Zeijlstra 2014, Heycock and Sundquist 2017). According to the RAH as formulated by Bobaljik and Thráinsson (1998), richly inflected languages such as Icelandic have a Split IP (=for Tense and Agreement), unlike Mainland Scandinavian which has a fused IP; only the former type of languages absolutely require verb movement to the inflectional domain (see also Pollock 1989, Thráinsson 1996). As such, the special TP-internal adjunction of adverbs, shown in a tree diagram in (66), to account for the *Adv-Vfin* order in Icelandic, being an apparent counterexample to the RAH, arguably carries a heavy burden of proof (from Thráinsson 2003:183):

(66)



All cases involving exceptional adverb placement on the analysis of Bobaljik and Thráinsson (1998) appear to be specific to deriving precisely these sorts of Adv-Vfin phenomena. As Hróarsdóttir et al. (2007) and Hrafnbjargarson and Wiklund (2010) point out, the fact that there seems to be no independent evidence for this makes such an analysis rather suspicious. Indeed, Thráinsson (2010:1077) acknowledges that he “cannot claim to have a clean and neat analysis” of this aspect of Icelandic syntax.

The main evidence for this kind of exceptional adjunction provided by Bobaljik and Thráinsson (1998) is the following:

- (67) a. Prosodic stress requirement on the adverb.
 b. Incompatibility with indefinite, non-specific subjects.
 c. Found where fronting is near impossible, esp. relative and adverbial clauses.
 d. (Low) frequency.

These points will be treated in turn below from a comparative 19th-century Icelandic perspective, except for (67c) which will be treated separately in Section 3.7.1 below). I claim that these points either do not hold for 19th-century Icelandic and/or call for a different analysis where V-in-situ is also an option.

As regards the first claim, Angantýsson (2011:76, fn. 49) already observes for Modern Icelandic that in interviews with participants in the IceDiaSyn dialect study, “the Adv-Vfin order was considered better if the sentence adverb was stressed” but this is clearly not the same as prosodic stress being required. What about 19th-century Icelandic? For obvious reasons, prosodic factors are not easily analysed in written texts in the absence of visible cues. However, there are, interestingly, a few instances involving the Adv-Vfin order where the speaker used visual means to specifically denote emphasis or stress. Interestingly, these never involve stress on the adverb alone, but always either on a non-finite lexical verb or *both* the adverb and the finite verb.⁷⁸

⁷⁸The critical edition, from which the following examples below are taken, uses *e m p h a t i c s p a c i n g*, presumably denoted by underlining in the original. First of all, we find emphatic marking on the lexical, non-finite verb form and not on the adverb:

Moreover, indefinite subjects did not appear to have a significant effect on the choice between variants in 19th-century Icelandic according to the statistical models of sections 3.4.1.2 and 3.4.2.2, for newspapers and private letters, respectively. However, there was an effect of subject type in terms of pronominal vs. nominal subjects, with the former slightly favouring Adv-Vfin over the latter. A closer scrutiny of the raw data only serves to corroborate the lack of effect with indefinites (see Table 6.30, Appendix A, page 245). Indefinite NP subjects are frequently found in the Adv-Vfin order, on par with or at times even exceeding the rate with definite NP subjects. Admittedly, the indefinite status of a subject does not automatically entail non-specificity, which is a crucial distinction to make with regard to the unavailability of indefinite subjects in Modern Icelandic argued for by Bobaljik and Thráinsson (1998) and Angantýsson (2011). However, in the 19th-century Icelandic data, there are numerous cases of indefinite subjects that are arguably also non-specific. Some examples are shown in (68) and (69):

- (68) *Þó eitthvað stundum ami að, eg ber mig að kasta því frá mér*
 though something sometimes bothers PRT I carry me to throw it from me
 ‘Although something is sometimes wrong, I try to push it away from me.’
 (AlfJon-1820-01-01.txt)

- (69) *verdur ad skjera skurd lángetis í góminn, sem altjent linar verkin, jafnvel*
 must to cut canal vertically in the palate REL at.least eases the.pain even
þó gróptur ecki altjent, komi út, því vid skurdir fær bólgan betra
 though pus not always comes out because with the.cut gets the.swelling better
rúm enn áður
 space than before
 ‘A vertical section must be cut in the palate, which at least eases the pain even though no pus always exits it, because by cutting, the swelling gets more room than before.’
 (Sunnanpósturinn 1835-02-01 (2. tbl. 1. árg.).txt)

A restricted distribution of Adv-Vfin with regard to subject status is thus a property that does not seem to hold for the 19th-century Icelandic data, considerably reducing much of the appeal that the IP-internal adjunction analysis might have had.

Finally, the claim concerning a low frequency of Adv-Vfin (in Modern Icelandic) is a complicated one in that the Adv-Vfin variant has long been considered non-standard,

- (1) *eg verd ad brúka þetta lióta Orðaittæki, því madur knékrypur atid þegar menn bidja um það*
 I must to use this ugly expression because man kneels always when men ask about that
sem menn ecki geta heimtala!!
 REL men not can demand
 ‘I have to use this ugly expression because one always kneels when people ask about that which they cannot DEMAND.’
 (BjaTho-1840-10-13.txt)

It would be pragmatically awkward to place the emphasis on *ecki* ‘not’ as the negation arguably does not constitute focus here. Second, in the following example, the same speaker does visually emphasise the negation, but then the finite verb is stressed as well:

- (2) *stundum er gott ad seigia satt, þegar madur vill ad hinir ecki trúi því sanna*
 sometimes is good to say true when man wants that others not believe the true
 (BjaTho-1829-09-11.txt)
 ‘Sometimes it is good to tell the truth, when one wants others NOT TO BELIEVE that which is true.’

This is not to deny that the adverb sometimes receives stress in the Adv-Vfin order, but it clearly need not.

as discussed in section 3.1. It might thus not be expected to occur frequently (if at all) in edited material, much limiting the relevance of the evidence provided by Bobaljik and Thráinsson (1998) in substantiating their claim, which is problematic even for Modern Icelandic (cf. Wiklund et al. 2007).

A related aspect of the analysis offered in Bobaljik and Thráinsson (1998) concerns restrictions on the high placement of an adverb vis-à-vis the subject alone.⁷⁹ The adverb-subject configuration that this sort of adjunction gives rise to is typically ruled out in Icelandic (see e.g. Jónsson 2002:82, Angantýsson 2011:66f.) but found e.g. in 19th-century Danish (cf. Sandersen 2003:371) and present-day Norwegian, Swedish and Övdalian (see e.g. Jónsson 2002:83, Angantýsson 2011:66f., Garbacz 2010:100ff., with references). Some potential examples of this type can be found in the 19th-century corpora but appear to be very rare. The phenomenon has not been studied systematically but sporadic examples do occur where the negation immediately precedes the subject, either in clause- or verb-initial structures:

- (70) *en það mun ördugt ef ei góðra og gagnlegra Manna Fylgi kémur til.*
 but it shall.be difficult if not good and useful men_{GEN} support comes PRT
 ‘But that proves to be difficult if the support of good and subservient men is not attained.’
 (BjaTho-1834-10-15.txt)
- (71) *enn reiddist þó ei Berns Hoggi heldur enn þeir*
 but angered_{MID} though not Berns Hoggi further than they
 ‘But Berns Hoggi did not get mad, any further than they did.’
 (HilJon-1836-06-03.xml)

It is not unlikely that the subject carries emphasis in most of these cases or is otherwise marked in terms of structural complexity, heaviness and so on. Another rare but potentially interesting pattern involves particle-like uses of temporal adverbs like *nú* ‘now’:⁸⁰

⁷⁹Here again, however, it turns out the prediction is not obviously carried out. Thus, Thráinsson (2010:1075, fn. 18) mentions that his analysis actually predicts that *??sem ekki einhver hafði lesið* ‘that not somebody had read’ should be possible, where the subject is adjacent to the verb, suggesting there may be a semantic story to block such cases (cf. *ekki einhver* ‘not someone’ vs. *aldrei neinn* ‘never anybody’, which sounds more natural). However, this issue arguably goes beyond SOME-phrases and negation, as the same applies to indefinite, non-specific phrase like *fólk* ‘people’, also unable to follow adverbs like *aldrei* ‘never’ or *oft* ‘often’, for instance:

- (1) *... *sem **aldrei/oft** fólk hafði lesið*
 REL never/finally people had read

Given that indefinite subjects are argued to occupy a lower subject position, shown e.g. by interpretational contrasts (see also Bentzen 2009) and the fact that they can co-occur with expletives, nothing should prevent these adverbs from preceding the subject in (1).

⁸⁰Such cases were frowned upon in the Reykjavík Grammar School as suggested by the student essays, where a comparable example is corrected (underlined) by the teacher in (1):

- (1) *Ef að **nú** maðurinn hefur þetta sama traust til guðs [...]*
 if C_{that} now man-the has this same trust to god
 ‘If man now has this same trust in God ...’
 (1855, 4th grade)

- (72) a. *En þótt nú fridurinn í Badajoz við Spán ekki gæfi Portúgísum stóra*
 but though now peace-the in Badajoz with Spain not gave_{SUB} Portuguese big
æru né ábata [...]
 reputation nor gain
 ‘But even though the peace in Badajoz with Spain did not give the Portuguese much reputation or gain ...’
 (Minnisverð tíðindi 1803-01-01 (1. tbl. 2. árg.).txt)
- b. *Og þar sem nú konungur samkvæmt alríkislögunum getur ekki [...]*
 and there REL now king according.to federal.law-the can not
 ‘And because the king according to the federal law now can not ...’
 (Eimreiðin 1899-01-01 (1. tbl. 5. árg.).txt)

If these cases were due to exceptional IP-adjunction along the lines of Bobaljik and Thráinsson (1998) and Thráinsson (2010), there would have to be two distinct adjunction sites in cases like (72a), one to get *nú* in front of the subject and another to derive the Adv-V_{fin} order. Unless further assumptions are made, the Adv-V_{fin} order may thus still arguably involve V-in-situ. However, it does not even seem to be possible to analyse *nú* as being IP-adjoined as Bobaljik and Thráinsson’s account would suggest, because the particle can definitely occur higher than IP. This is witnessed by the fact that it can precede the complementiser *að*:

- (73) a. *En enda þótt nú að höfundurinn hefði satt, að mæla [...]*
 but even though now C_{that} author-the had_{SUB} true to speak
 ‘And even if the author were right ...’
 (Norðri 1859-01-31 (1-2. tbl. 7. árg.).txt)
- b. *Þótt nú að vonir vorar tæki töluvert að deyrast [...]*
 Though now C_{that} hopes our took_{SUB} considerably to sadden_{MID}
 ‘Even though our hopes now are were becoming considerably weak ...’
 (Þjóðólfur 1860-02-15 (10-11. tbl. 12. árg.).txt)

The cartographic approach readily allows for multiple complementiser positions, which can be associated with fronted *wh*-items, relativisers, topics, foci and so on. In the case at hand, the right-most complementiser *þótt* is presumably in Force, whereas the lower *að* is realised in Top as a result of the fronting of the particle to TopP, which leaves Fin available for the finite verb.

A promising historical account of the rise of Adv-V_{fin} in Icelandic is provided by Heycock and Wallenberg (2013) in terms of competition between V-in-situ and V-to-T (for them, at least, synonymous with V-to-I).⁸¹ Observing on the basis of the IcePaHC corpus that the rate of Adv-V_{fin} is much higher in the same sorts of contexts that embedded V2 is typically ruled out in Mainland Scandinavian (cf. extensive discussion above), they suggest that certain speakers innovated a V-in-situ grammar alongside their V-to-T grammar. By combining the two grammars, an analysis along these lines is able to account for the historical variation attested.

⁸¹To avoid terminological confusion, I will avoid referring to verb raising in terms of V-to-T since V-to-T is the lower IP position on a Split-IP analysis, the higher IP position being AgrP. For all intents and purposes, V-to-I can be thought of as V-to-Fin, although FinP presumably corresponds to the higher IP position, AgrP, on a Split-IP analysis (for a potential link between V in Fin and agreement, see e.g. Hrafnbjargarson and Wiklund 2010). I therefore treat V-to-I (V-to-Agr) and V-to-Fin as being largely equivalent.

In my view, the V-in-situ vs. V-to-I grammar competition approach is the only tenable account to capture the 19th-century data and I will adopt and explore this suggestion in more detail below.

3.7 Towards an analysis

I propose to analyse variation in verb-adverb placement in 19th-century Icelandic in terms of variable chain resolution (see Bobaljik 2002, Landau 2006 and subsequent work) combined with a grammar competition model, more specifically competition between a V-to-I (or V-to-Fin if CP is split) and a V-in-situ grammar (cf. Heycock and Wallenberg 2013). While grammar competition alone goes a long way to account for major patterns in the data, some amendments are arguably necessary, depending on further assumptions made with regard to the structural analysis of adverbs as well as the scope and limits of competing grammars within one and the same individual.

3.7.1 Adv-Vfin as a low Spell-out and V2 violation

Recall that the version of the Minimalist probe-goal framework assumed here takes linearisation of hierarchical syntactic structure to be a property of the phonological component of grammar at spell out. During the derivation of a clause, the hierarchical structure of syntactic objects created by internal Merge form a chain of copies, resulting in a lower and higher copy (or copies) of an object such as the external argument and the finite verb; which copy is actually the one pronounced at PF is subject to variation. Agreement is established at a distance through c-command in the default case and overt displacement of an element can take place, subject to variation, but only if it additionally carries an “EPP feature” (or more generally an Edge Feature “that permits it to be merged”, cf. Chomsky 2008:139).

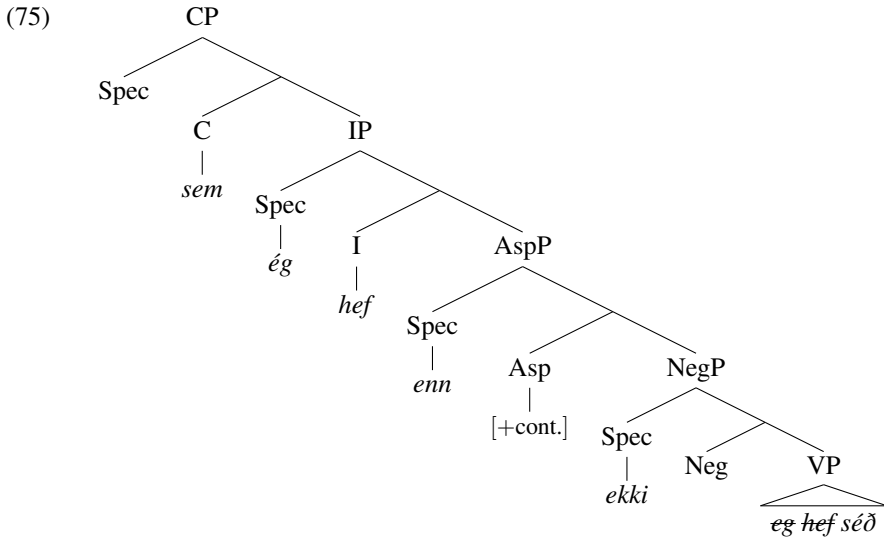
Following and partly adopting the cartographic analysis of Holmberg and Platzack (2005), Wiklund et al. (2007), Hrafnbjargarson and Wiklund (2010), with some necessary changes, I take the finite verb in verb-second configurations typically to be located in Fin (a part of the C-domain). For ease of representation, however, I will adopt the abbreviated cover terms for CP > IP > VP below when discussing the properties of the two grammars.⁸² Rather than assume that V-to-I is actually sufficient to derive subject-initial V2 sentences in Icelandic (e.g. Rögnvaldsson and Thráinsson 1990, Bobaljik and Thráinsson 1998, Thráinsson 2010), I take V2 to suggest that the verb is in the C-domain, but that the verb may alternatively stay low **in violation of verb-second**. As a result, it may be spelled out as high as in Fin (corresponding to I if CP is unsplit), as low as in V (*in situ*) or somewhere in between, along the Tense-Aspect-Mood regions typically subsumed under the traditional IP domain (see below).

An example Vfin-Adv (verb-second) construction is shown in (74), the proposed (simplified) structure of which is shown in (75):⁸³

⁸²An immediate consequence of the elaborated cartographic view of clause structure is that the Fin head, hosting finiteness, does not simply translate into the I/T head (of the I-domain) on more traditional conceptions of clause structure if we refer back to the abbreviated C-, I- and V-level.

⁸³Here and below, I adopt the traditional shorthand CP-IP-VP model, abstracting away from the richer

- (74) *Eg veit að sönnu ekki, hvernig sú er, sem eg hefi enn ekki séð, hún á
 I know at true not how DEM is REL I have still not seen she ought
 að koma þegar út hallar slætti.
 to come when out leans haymaking
 ‘I truly do not know how that one is, whom I still have not seen—she is supposed
 to come when the haymaking is well under way.’ (GudMag-1846-07-24.txt)*



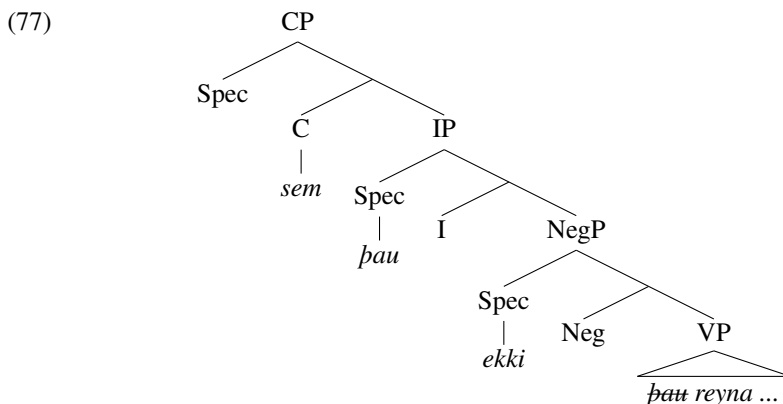
The relativiser is merged in C (\approx Sub), realised as *sem* ‘which’, which alternatively can take on the guise *sem að* (lit. ‘which that’) with *að* in Force. Unlike Mainland Scandinavian, where the finite verb is argued to occur in C (\approx Force) whenever merged higher than ν P/VP, the finite verb is taken to raise to IP (\approx Fin) in Icelandic, offering a potential account of how Icelandic allows embedded V2 across all clause types whereas in Mainland Scandinavian this is usually restricted to asserted clauses. The featural specification of finite verbs in Icelandic includes rich agreement with the subject (in person and number), completely lacking in Mainland Scandinavian, such as Danish, where there is only tense morphology. This means that there is an uninterpretable ϕ feature on I/Fin in Icelandic that needs to be eliminated. Importantly, the presence of these ϕ features means that Icelandic has more and not less potential landing sites for the finite verb than Mainland Scandinavian. Moreover, beyond the V2 constraint (targeting Fin/Force), these features also do not require spelling out the highest member of the V- ν -I-Fin-(Force) chain. I do not assume that the V2 effect has anything to do with these morphological features, but instead obtains independently. The V2 effect may be violated so long the Probe-Goal relation in question targets the lowest C-position (i.e. Fin), since higher C-positions (i.e. Force, Foc, Top) always require V2 in Icelandic.⁸⁴

The result of spelling out the lowest member of the chain gives rise to V-in-situ, as in (76), the simplified structure of which is shown in (77):

C-I-V levels. Note that although the subject is shown here in a VP-internal position, I assume that it is merged in ν P immediately above VP. Concerning AspP, see (11) in Section 1.1.2.

⁸⁴To account for the obligatoriness of V2 in main clauses, I must assume that the finite verb is located at the C-level above Fin.

- (76) *Okkur er báðum svo vel við hjónin, af því bæði eru þau góðar manneskjur*
 us is both so well with couple-the of it both are they good people
og geta ekki hugsað það upp, sem þau ekki reyna til að gleðja okkur með.
 and can not think it up REL they not try to C please us with
 ‘We both like the (married) couple so much because they are both good people
 and cannot think up what they do not try to please us with.’
 (GudMag-1845-07-17.txt)



The present analysis permits a third option: raising the verb but spelling it out lower than Fin, in the “IP” area that is a part of the functional sequence above VP (see (79) below). This option is arguably also attested. While verb-second is regarded as being triggered by features of the C-domain and typical violations of V2 as arising through V-in-situ, I do not follow Wiklund et al. (2007:211) in assuming that this exhausts the options available in Icelandic. The reasons for this are empirical. Contrary to what their account implies, the finite verb can be spelled out in the IP zone; this applies to present-day Icelandic (see Thráinsson 2010:1080f.) and 19th century Icelandic is no different in that regard:

- (78) *Höfudbrestur hans var að hann sparadi eingin Medöl til að hefna*
 main.fault his was that he saved no means PRT to take.revenge_{MID}
á hinum og vildi svo miög það hann vildi, að hann ei gáði
 on the.others and wanted so much it he wanted that he not watched.out
ætið Framqvæmdar Medalanna
 always action means-the
 ‘His main shortcoming was that he used every means to take revenge on others
 and he wanted what he wanted so badly that he did not always watch out with
 his means of action.’
 (BjaTho-1833-03-23.txt)

In (78) the negation *ei* ‘not’ precedes the finite verb *gáði* ‘watched out’, followed by the sentence-medial adverb *ætið* ‘always’. As a result, we must assume that the verb has been merged in a position higher than its base position. In order for this example to be parsed by a V-in-situ grammar that only permits V-to-C, the negation would have to be in an exceptionally high position in the C-domain.⁸⁵ The fact that variation in

⁸⁵It should be pointed out that this type of environment (consequence of degree) belongs to potential V2 context in Mainland Scandinavian and as such could be argued to have a rich left periphery.

19th-century Icelandic always exceeds (to a varying extent) the limits posed by a strict V-to-C/V-in-situ analysis already suggests that V-to-I (or V-to-Fin) is required alongside V-in-situ to account for the data. However, that option alone does not provide an account of (78) because *gáði* ‘watched out’ is arguably too low to be in I (or in Fin) but too high to be in V.

The problem with (78) disappears if, instead of assuming that NegP and AdvP are simply adjoined to VP, we adopt the Cinquean perspective of adverbs. Most previous discussion on verb-adverb placement assumes a traditional “adjunction approach” to adverbs where sentence-medial adverbs adjoin to the VP, thereby delimiting the IP-VP complex. This approach has been worked out in quite some detail by Ernst (2002, 2007), taking semantic interpretation to determine adverb placement and restrictions on the relative order of adverbs, with antecedents in much earlier work (e.g. Jackendoff 1972). The adjunction approach is adopted by most scholars working on V_{fin}-Adv/Adv-V_{fin} variation in Icelandic and related languages (for discussion, see e.g. Thráinsson 2010, Garbacz 2010, Angantýsson 2011).

On Cinque’s analysis, in contrast, adverbs are taken to involve rich hierarchical syntactic structure with both verb and argument positions, allowing observable differences in the position of verbal forms to be captured in terms of where in the adverb hierarchy the verb is spelled out. This account is thus not restricted to finite forms and languages can differ both in where finite and non-finite forms are realised in this structure, giving rise to subtle differences across languages and language varieties (cf. Cinque 2004:686f.). Cinque’s (1999, 2004) “functional-specifier approach” to adverbs modifies the traditional assumption that a finite verb preceding sentence-medial adverbs in non-V2 contexts is indicative of V-to-I. On Cinque’s view, where each subtype of adverbial phrases occurs as a specifier to empty functional heads (as many as there are subtypes), cross-linguistic variation in verb-adverb placement does not consist in whether or not a language allows V-to-I but where the verb can occur in the universal hierarchy of clausal functional projections, where each of his multitude of functional projections may be potential V_{fin} hosts.

Cinque’s hierarchy is shown in (79), where the Icelandic equivalents to the boldfaced adverbs could precede V_{fin} in subject-initial embedded clauses in the 19th-century private letter data:

(79) *Cinque’s hierarchy:*

[*frankly* Mood_{speech act} [*fortunately* Mood_{evaluative} [*allegedly* Mood_{evidential}
 [***probably*** Mod_{epistemic} [***once*** T_{past} [***then*** T_{future} [***perhaps*** Mood_{irrealis}
 [***necessarily*** Mod_{necessity} [***possibly*** Mod_{possibility} [*usually* Asp_{habitual} [***again***
 Asp_{repetitive(I)}
 [***often*** Asp_{frequentative(I)} [***intentionally*** Mod_{volitional} [***quickly*** Asp_{celerative(I)} [***al-***
ready T_{anterior}
 [***no longer*** Asp_{terminative} [***still*** Asp_{continuative} [***always*** Asp_{perfect} [***just*** Asp_{retrospective}
 [***soon*** Asp_{proximative} [***briefly*** Asp_{durative} [***characteristically*** Asp_{generic/progressive}
 [***almost*** Asp_{prospective} [***completely*** Asp_{SgCompletive(I)} [***tutto*** Asp_{PICompletive}
 [***well*** Voice [***fast/early*** Asp_{celerative(II)} [***again*** Asp_{repetitive(II)}
 [***often*** Asp_{frequentative(II)} [***completely*** Asp_{SgCompletive(II)}

Each of the boldfaced subtypes of adverbs can be exemplified on the basis of the Adv-

V_{fin} variant from the private letter corpus (see (182) in Appendix A, page 248). These data suggest that the structural position of the verb is extremely flexible and although the corpus may not be large enough for all of Cinque's (1999) subclasses to figure (in the appropriate context), it still seems likely that V_{fin} could be spelled out in any of the functional projections he identifies.

There is other potential evidence suggesting that the finite verb may occur in a VP-external position while not raising all the way up (neither to I nor to Fin), viz. Object Shift. Distributional co-occurrences of Adv-V_{fin} and Object Shift, traditionally taken to be contingent on verb raising, have been used to suggest that the finite verb always undergoes V-to-I in Modern Icelandic, even when strict subject-verb adjacency is violated:

- (80) *Mér fannst skrítið þegar hann oft lék hróknun ekki í tímahraki*
 me found strange when he often moved rook-the not in lack.of.time
 'I thought it was strange when he often didn't move the rook through lack of time.'
 (cf. Koenen and Zeijlstra 2014:580)

Again, despite violating V2, examples such as (80) indicate that the verb may still have left its base position. However, it should be borne in mind that the fact that there may be multiple landing sites for the verb does not necessarily entail that the Adv-V_{fin} order in Icelandic must always involve V-to-I. I have been unable to find instances in 19th-century Icelandic corresponding directly to (80) above with a nominal object. With the Adv-V_{fin} order, I have only been able to find one potential case where a pronominal object shifts over an adverb. This example features two medial adverbs, *ekki héðan af* 'not henceforth' and *ekki* 'not'. It must be emphasised that due to the scarcity of data from this speaker ($n=3$, his remaining two examples being V_{fin}-Adv), we cannot know if he is representative of high-frequency Adv-V_{fin} speakers:⁸⁶

- (81) *jeg hef ekki látið hann vita af gleimsku minni og vona að hann ekki*
 I have not let him know of forgetfulness my and hope that he not
héðan af spurji mig ekki neitt um bréfið
 hereafter asks me not at all about letter-the
 'I haven't mentioned my forgetfulness to him and I hope won't ask me anything about the letter from now on.'
 (VilOdd-1875-03-18.xml)

Accepting (81) as is, the finite verb thus intervenes between two medial adverbs (or rather, a spell-out of two copies of the same adverb), with the direct object *mig* 'me' shifting over the lower copy. This, indeed, suggests that the verb is able to move out of the VP independently of V-to-C movement. However, this does, of course, not automatically entail that all instances of Adv-V_{fin} must be analysed as V-to-I movement, in cases where such movement would be string-vacuous.⁸⁷ All it shows is that Object

⁸⁶What makes this example rather dubious as evidence is the fact that the negation actually occurs twice. A more usual way of saying (81) would be to spell out either the higher or the lower instance only. Importantly, *héðan af* 'henceforth' on its own also does not suffice as a reliable diagnostic for movement, not being a sentence-medial adverb. If the example were amended, it would thus either lose the Adv-V_{fin} property in the narrow sense defined here or, alternatively, it would cease to be Object Shift. This is obviously an unfortunate property of a singleton.

⁸⁷Even in Northern Norwegian, objects can shift in this configuration (Kristine Bentzen, p.c.):

Shift is possible—in fact, it would be unexpected if it were not—and thus that the verb *can* occur in sentence-medial position.

The result sections above already provide ample quantitative evidence that embedded clauses that tend to disallow main clause phenomena exhibit a predisposition toward the Adv-V_{fin} order (see Sections 3.4.1.2, 3.4.2.2, 3.4.2.4 and 3.4.3.2). Just as Heycock and Wallenberg (2013) have independently observed on the basis of IcePaHC (mainly literary texts), we have seen on the basis of newspapers, student assignments and private letters that the frequency of the Adv-V_{fin} variant is always higher—and for some speakers particularly high—in roughly the environments where Adv-V_{fin} is obligatory in most Mainland Scandinavian varieties. Heycock and Wallenberg (2013) take this to suggest the existence of a V-in-situ grammar alongside a traditional V-to-I grammar, as the use of V_{fin}-Adv exceeds the potential of V-in-situ alone.⁸⁸ Similar to Heycock and Wallenberg (2013), my results were categorised on the basis of an approximation of environments allowing or prohibiting main clause phenomena (root vs. non-root or peripheral vs. central, cf. discussion above). If this is correct, we predict that declaratives should also pattern differently depending on whether or not they allow main clause phenomena. That prediction is carried out although, again, the V_{fin}-Adv variant is also found in contexts that cannot be parsed by a V-in-situ/V-to-C grammar, requiring V-to-I.

Even more importantly, if 19th-century Icelandic is somehow (partly) observing Danish constraints on word order, as Heycock and Wallenberg (2013) propose, we do not expect all declaratives to pattern alike as typical V2 environments, but for verb-adverb placement to be sensitive to whether or not the complement clause is asserted. Non-asserted declaratives would therefore count as typical non-V2 environments. Whatever machinery gives rise to these differences, I will assume with Rögnvaldsson and Thráinsson (1990) that Icelandic syntax in principle accommodates embedded V2 which is then restricted at the level of pragmatics and meaning.⁸⁹ Since word order constraints thus cannot be reduced to properties of clause structure alone, the definition of V2 vs. non-V2 contexts will always be imperfect unless pragmatic aspects are properly factored in as well. A recent proposal along these lines by Trotzke (2015) incorporates the notion of derivational layers (see Section 1.1.1) to account for such restrictions as a part of the syntactic derivation. Recall that with verbs of saying and believing ('bridge verbs'), only one event is asserted, SAYING or BELIEVING, whereas verbs like regret ('non-bridge verbs') are statements that involve two events at the syntax-discourse level, REGRETTING *in addition to the content of the embedded clause*, which is a factive

- (1) *Jeg vet hvorfor Jon sannsynligvis gir henne så ofte gaver.*
 I know why Jon probably gives her so often gifts

As Bentzen (2003, 2009) argues, North Norwegian has short verb movement, meaning that V_{fin} can precede certain adverbs, although it cannot move past the negation. As a result, the verb can precede *så ofte* 'so often', with a subsequent object shift of *henne* 'her' across the lower adverb.

⁸⁸See Gärtner (2016) for a discussion on how different types of V2 systems, so-called free or broad V2 systems (*fV2*, *bV2*), such as Old Norse, permit main clause phenomena in a far greater number of embedded environments, even relative clauses, than narrow V2 (*nV2*) systems, such as Mainland Scandinavian. These properties are arguably relevant with regard to the historical development of verb-adverb placement in Icelandic. I must leave this aspect to future research.

⁸⁹There have been various proposals on how to capture this contrast in syntactic terms, where typical V2 environments are regarded as having a more elaborated CP structure either by means of recursion (see e.g. Vikner 1995) or through a blocking effect that obtains with presuppositional and factive verbs as they spell out their meaning, rendering the higher C-positions Force, Topic and Sub defunct (see Hrafnbjargarson 2008).

	Until 1850				After 1850			
	Private letters		Newspapers		Private letters		Newspapers	
A	22.2%	(44/198)	27.0%	(43/159)	4.7%	(12/254)	4.2%	(20/480)
B	17.1%	(42/246)	27.0%	(24/89)	8.5%	(38/446)	7.4%	(15/203)
C	11.4%	(4/35)	44.0%	(11/25)	9.8%	(4/41)	9.6%	(8/83)
D	45.3%	(77/170)	37.8%	(14/37)	9.3%	(20/215)	7.7%	(11/143)
E	27.6%	(43/156)	13.0%	(7/54)	7.2%	(13/180)	4.8%	(8/165)
O	37.9%	(114/301)	35.2%	(69/196)	11.9%	(55/462)	7.4%	(45/608)
∑ A+B+E	21.5%	(129/600)	24.5%	(74/302)	7.2%	(63/880)	5.1%	(43/848)
∑ C+D	39.5%	(81/205)	40.3%	(25/62)	9.4%	(24/256)	8.4%	(19/226)

Table 3.14. Relative frequency of Adv-Vfin in private letters and newspapers up until 1850 vs. after 1850 across different types of declaratives.

statement (i.e. meaning is retained in the presence of matrix negation). According to Trotzke (2015:101), the fact that there are two events with a verb like *regret* requires a separate derivational layer which then leads to opacity at the syntax-discourse interface.

Splitting up the declaratives according to Hooper and Thompson's seminal analysis of complement clauses, we again see an interesting pattern emerge, cf. Table 3.14. Note first that Class C predicates have been defined exactly as in Hooper and Thompson (1973), thus including predicates denoting possibility and likelihood (corresponding to 'be possible' and 'be likely') as well as (inherently) negative verbs (corresponding to 'doubt' and 'deny'). The distinction between (inherently) negative verbs as opposed to the rest will be explored further below. Note also that Class O(ther) here represents clauses that are complements to verbs, adjectives and (pro)nouns (e.g. *vandamálið er* 'the problem is'), counter-assertive verbs like *vilja* 'want' and *vona* 'hope' and various other types which do not fit into Hooper and Thompson's categorisation. Class O also includes causative predicates (corresponding to 'the reason that') identified by Jensen and Christensen (2013), which patterns in their material as a non-V2 environment to roughly the same degree as factive/Class D predicates. According to Jensen and Christensen (2013:47), causatives are distinct from factive/Class D predicates in not offering any sort of evaluation.

As summarised in the bottom two rows of Table 3.14, the asserted Classes A, B and (semifactive) E, on the whole, have a stronger disposition for V2 as compared to Class C and D in the period up until 1850. Given that Classes A, B and E are assertion-friendly environments and the ones typically allowing embedded topicalisation, whereas Class C and D are assertion-hostile, typically disallowing embedded topicalisation, this distribution suggests a correlation in 19th-century Icelandic between having rich left periphery and a higher rate of V2, and conversely, an impoverished left periphery and a higher rate of Adv-Vfin. Being a mixed class, it is not unexpected that Class O appears as an in-between category as regards the distribution of Vfin-Adv/Adv-Vfin. The data even suggest that Class O may have more in common with the Class C+D/assertion-hostile environments than Class A+B+E/assertion-friendly environments, presumably due mainly to the counter-assertives and causatives. Table 3.14 presents only aggregated data but, interestingly, these contrasts also hold at the individual level for high frequency Adv-Vfin speakers, cf. Table 6.31, Appendix A, p. 248.

As presently defined, Class C fails to pattern consistently with Class D. In fact, in the private letter corpus, Class C has a lower proportion of Adv-Vfin than Class A and B in the period until 1850. This is not entirely unexpected, of course, given that it has been argued for Swedish (cf. Andersson 1975) that the non-negative predicates in Hooper and Thompson's Class C should be treated as Class B. A closer inspection of Class C shows that nearly all of the examples involve predicates of possibility and likelihood, such as *vera líklegt* 'be likely', *geta skeð* 'can happen' and *vera hætt við* 'be possible/be a risk'.⁹⁰ While the total number of examples in Class C is certainly rather limited, the evidence does suggest that the non-negative predicates may indeed not belong to this class. Out of 188 examples in total (both corpora, until/after 1850), merely 18 examples occur with negative predicates, i.e. either negative variants of possibility/likelihood predicates (corresponding to 'be impossible') or inherently negative verbs (such as 'doubt' and 'deny'). In the period until 1850, a total of 8 out of 14 (57.1%) negative Class C occurrences had the Adv-Vfin order. In the same period, in contrast, 7 out of 46 (15.2%) of the non-negative Class C occurrences had the Adv-Vfin order.⁹¹

If we repartition Class C based on the negative vs. positive dichotomy in both corpora combined, the proportion of Adv-Vfin in Class C+D environments raises to an average of 44.7% until 1850 but stays as low as 8.4% after 1850, cf. (82):

		UNTIL 1850	AFTER 1850
(82)	Assertion-friendly (A+B+E)	22.5% (203/901)	6.3% (112/1772)
	Assertion-hostile (C _{neg} +D)	44.7% (97/217)	8.4% (31/368)

In other words, there is a very clear contrast between the two environments, the Adv-Vfin variant being roughly twice as likely to occur in assertion-hostile vis-à-vis assertion-friendly declaratives in the former of the two periods, whereas there is no significant difference between assertion-friendly and assertion-hostile declaratives in the latter period.⁹²

If this distinction were the whole story, we would expect high-frequency Adv-Vfin speakers only to be able to have V2 in clause types that allow main clause properties. For example, they should be unable to have V2, for instance, in typical conditional *if*-clauses (event condition) or in strictly temporal *when*- and *while*-clauses (cf. section 3.3.1). Unsurprisingly, given the overall high rate of V2, this is not what we find—even high-frequency Adv-Vfin speakers can have V2 in these contexts:

(83) *Central adverbial clauses: Conditional if*

- a. ... *Utfall Heiskaparins var hid æskilegasta – svo nú er óhætt ef Vetur*
 result haymaking_{GEN}-the was the desirable_{SUP} so now is safe if winter
verdur ecki því lánvinnari.
 will.be not it_{DAT} prolonged_{CMP}
 'The result of the haymaking was most desirable, so now it is safe if winter

⁹⁰In addition, I have also included in this class adverbs such as *kannske* and *máske* 'maybe(it) may be' taking a clausal complement.

⁹¹Despite the numbers being small, this contrast is statistically significant according to a Pearson's chi-squared test (Yates' continuity correction) and Fischer's exact test for count data. For chi-square, $\chi^2 = 7.9503$, $df = 1$, $p = 0.004808$. For Fischer's exact test, $p = 0.003401$, odds ratio = 7.102792.

⁹²For the period until 1850, the difference is highly significant according to a chi-square test: $\chi^2 = 42.659$, $df = 1$, $p = 6.518e-11$. For the period after 1850: $\chi^2 = 1.8378$, $df = 1$, $p = 0.1752$ (Fischer's exact test, $p = 0.1676$).

will not be that much more prolonged.’ (BjaTho-1824-09-09.txt)

- b. *ef Embætti og Búskapr ætludu ecki ad giöra útaf vid mig og eg*
 if incumbencies and husbandry intended not to do out.of with me and I
væri Efninu vaxinn, hefði mig lánkad til ad klóra um það.
 were matter-the grown had me wanted PRT to scribble about it
 ‘If my duties in office and husbandry were not scuppering me and I were
 competent with regard to the matter, I would have wanted to scribble about
 it.’ (BjaTho-1836-04-22.txt)

(84) *Central adverbial clauses: Temporal when/while*

- a. *eg ætlast ecki til ámedan sama Höfud er ecki á ockur báðum, ad þú*
 I expect_{MID} not to while same head is not on us both that you
fallist á allt hiá mér
 agree on everything at me
 ‘I don’t expect you to agree with me on everything while the same head
 isn’t on both of us.’ (BjaTho-1828-03-20.txt)
- b. *sídan eg þurfti ei ad skipta mér af því, hefi eg eng[v]ann Mann spurt*
 since I needed not to meddle me of it have I no man asked
um það
 about it
 ‘Ever since I did not have to interfere with it, I have not asked any person
 about it.’ (BjaTho-1823-03-25.txt)

This state of affairs is not at all surprising given that we are dealing here with ongoing language change where there is an interaction between an older stage arguably lacking the Adv-Vfin variant and a linguistic innovation.

In the following section, I will argue that an interplay between language-internal and language-external conditions in Icelandic resulted in the permanence of Adv-Vfin, but in modified form. If my account is on the right track, this innovation in Early Modern Icelandic, while ultimately unsuccessful, was capable of driving a change that arguably was (at least partly) successful, building on the classic notion of TARGETED CHANGE (e.g. Labov 1972a, Guy 2011) in addition to the FAILED CHANGE hypothesis of Postma (2010).

3.7.2 Targeted change, failure and success

I take the historical data revealed by Heycock and Wallenberg (2013) on the basis of IcePaHC texts to suggest that language contact was a factor in the proliferation of the Adv-Vfin word order during the period 1600–1850. According to Heycock and Wallenberg (2013), the main source of this contact would be interference via translations as well as more generally through “the language of the local Danish colonial aristocracy” (Heycock and Wallenberg 2013:153). It is far from obvious, however, whether language contact should be invoked as a primary cause or, alternatively, only as an additional factor, increasing the frequency of Adv-Vfin by tapping into resources that may already have been in place to some extent and extending them. If we adopt the latter scenario, mild contact and/or translation effects then only have served to amplify an already

existing pattern, resulting in frequency changes of the type discussed by Thomason (2001:70) where borrowers may be reasonably bilingual but in minority.

The way some of the variation attested in the private letter corpus was shown to be socially conditioned, mainly based on social status, strongly suggests that Adv-Vfin was adopted as a prestigious feature in a formal, written register. This would be an instance of TARGETED CHANGE originating in the highest social class, i.e. “the borrowing of some external prestige norm” (Guy 2011:180), to be contrasted with unborrowed, spontaneous innovations that tend to arise within the working classes (see Labov 1972a:290, Guy 2011:180f.). The adoption of this external prestige norm even in private letter writing as well as the relative stability within speakers over time and across different recipients suggests that Adv-Vfin must have had some basis in the casual code of these speakers. Moreover, even Adv-Vfin at high frequencies was not a phenomenon confined to speakers likely to be directly calquing Danish. Thus, a female letter-writer like Stefanía Siggeirsdóttir, writing in the 1860s and 70s, with social ties to the clergy, not only uses Adv-Vfin in the majority of cases but overwhelmingly so in typical non-V2 environments. While inherently variable, her grammar shows tendencies towards the assertion-based system we find in Mainland Scandinavian—but in the variable guise characteristic of Faroese, for example.

Observe that despite the fact that the Adv-Vfin variant is clearly in remission after 1850, it never disappears completely. This can be seen by its negative evaluation in the 20th century, for instance by Smári (1920) and Böðvarsson (1992), as well as the extensive discussion in the linguistic literature since the 1980s (e.g. Maling 1980, Sigurðsson 1989, Bobaljik and Thráinsson 1998, Angantýsson 2001, 2007, 2011, Wiklund et al. 2007, Thráinsson 2010). In other words, the targeted change failed in the sense that it did not progress to completion but its effects were still not entirely undone. The main difference between the (failed) targeted change and the other (presumably spontaneous) change thus appears to lie in the (partial) sensitivity to the assertion feature on Force in the former, lacking in the latter. This is best seen in the different behaviour of *that*-clauses depending on the matrix verb class—environments least favourable to the Vfin-Adv variant in the “target” (i.e. Danish, though in practice more comparable to Faroese) tend to be the ones most favourable to the Adv-Vfin variant.

An insightful way to deal with targeted changes failing to catch on in this way is the FAILED CHANGE model of Postma (2010). Failed changes in Postma’s specific implementation of Kroch’s (1989) grammar competition model are typically adult/L2 innovations (e.g. arising through contact) that, while unsuccessful in the sense that they were neither adopted by the majority of speakers nor faithfully transmitted to subsequent generations, nonetheless managed to have a lasting effect. Postma applies his model to the rise and fall of English *do*-support in positive affirmative clauses—a failed change—which he argues fuelled another related change that actually was successful, the familiar polarity use of *do*-support in English. From his perspective, the successful change resulting from the variation attested in the historical record constitutes an L1 accommodation of a failed L2 change. Similarly, one could argue, the more restricted Adv-Vfin phenomenon after 1850, the focus-dependent uses of Adv-Vfin in (varieties of) Modern Icelandic (adopting the analysis of Angantýsson 2001, 2007, 2011) in particular, could be seen to have arisen out of the extended Adv-Vfin uses of 19th-century Icelandic. It is important to emphasise in this context the weak position of Adv-Vfin towards

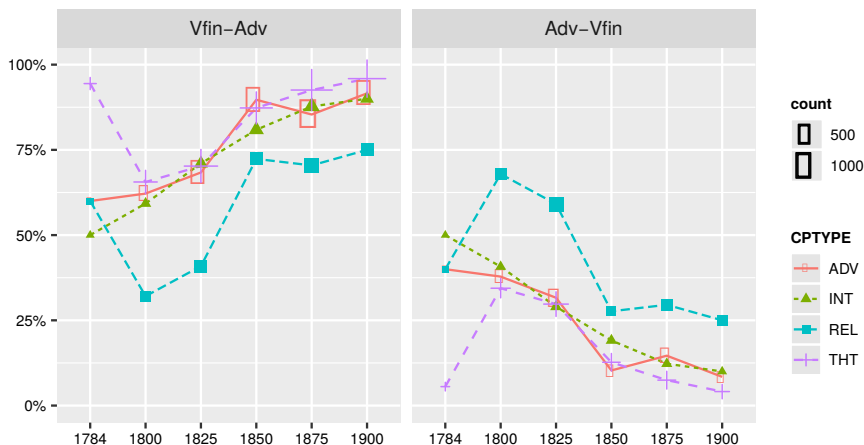


Figure 3.15. Verb-adverb placement in 19th-century Icelandic across four basic clause types. All data from private letters, newspapers and student essays combined (1784–1924, per 25 years).

the end of the period in similar contexts where Adv-Vfin is most restricted in Modern Icelandic, although the pattern is not perfectly clearcut.

To make this contrast maximally clear, let us focus our attention on declarative complement clauses and relatives, in all three types of corpora combined, cf. Figure 3.15.⁹³ Our data from the first period are thin but suggest an increase in Adv-Vfin beyond what Heycock and Wallenberg (2013) report on the basis of IcePaHC. Adv-Vfin peaks in the first half of the 19th century, showing an extensive increase in both relatives and declaratives, whereafter it takes on a strikingly different property. Whereas Adv-Vfin decreases sharply in declaratives in the period 1850 onwards, the frequency of Adv-Vfin in relatives remains largely stable between 1850 and 1924 around the 25% mark. Had the targeted change simply failed with no further repercussions, we would expect the relatives to follow a trajectory parallel to the declaratives, ultimately reaching levels below 10%.

The question now is why the Adv-Vfin phenomenon has the distribution it has in the post-1850 period and, correspondingly, why Adv-Vfin in Modern Icelandic is similarly most readily found in modifying/operator environments, in particular relative clauses (cf. Sigurðsson 1989, Angantýsson 2011). A part of the explanation for this distribution could be that the surviving/successful Adv-Vfin variant was modelled on Stylistic Fronting, typically found in the same sorts of operator environments in which Adv-Vfin has the strongest foothold in Modern Icelandic, a kind of accommodation that

⁹³In the last two periods (1875–1924), Adv-Vfin figures in relatives clauses at a rate of 27.7% (114/297) on average as opposed to 1.7% (4/233) and 5.9% (177/2810) in result clauses (*svo að* ‘so that’) and *að*-clauses (declaratives and consequence of degree clauses), respectively. The last two environments are among the ones least susceptible to Adv-Vfin in Modern Icelandic, in contrast to relatives where Adv-Vfin is most easily found (cf. Angantýsson 2011:72). Indirect *hvort* ‘whether’ questions, however, feature the Adv-Vfin variant in 13.6% (16/102) of the cases, similar to purpose clauses (16.7%, 18/90) and the rest of the adverbials (19.8%, 242/980), whereas *hvort*-clauses are hostile to Adv-Vfin in Modern Icelandic.

may have been facilitated by changes in the licensing of non-overt/null arguments.

The relevance of null subjects and Stylistic Fronting will be considered in a little more detail immediately below.

3.7.3 Stylistic Fronting, null subjects and reanalysis

Recall that the results above have only included cases where the initial position is occupied by an overt subject. However, earlier varieties of Icelandic, until the 19th century (Hjartardóttir 1993, Sigurðsson 1993a, Kinn et al. 2016), allowed arguments to be left unexpressed, usually referred to as *null subjects* or *pro drop* (see e.g. Hjartardóttir 1993, Hróarsdóttir 1998, Sigurðsson 1993a, Sigurðsson and Egerland 2009, Kinn et al. 2016). While null subjects are still allowed to some extent even in present-day Icelandic (for recent discussion, see Sigurðsson and Egerland 2009, Kinn et al. 2016), the difference is that null subjects in Modern Icelandic are mostly reserved for topic drop (discourse ellipsis) in a diary style narrative or in impersonals, either with a generic or an arbitrary referent:

- (85) *Við/e komum til London í gær. Við/e sáum ...*
we/e came to London yesterday we/e saw ...
 ‘Came to London yesterday. Saw ...’

(Kinn et al. 2016:55; Thráinsson 2007:477)

- (86) *Í þessari fjölskyldu má ___ bara ekki drekka áfengi*
in this family may just not drink alcohol
 ‘In this family, one is simply not allowed to drink alcohol.’

(Sigurðsson and Egerland 2009:160)

Null subjects in earlier varieties of Icelandic, however, could refer to definites, both controlled arguments and general arguments recoverable from discourse (cf. Sigurðsson and Egerland 2009).⁹⁴ What this means is that in many cases corresponding to the Vfin-Adv/Adv-Vfin variable as defined here, the subject may easily be and arguably often is simply left out. As a result, they fall outside the Vfin-Adv/Adv-Vfin radar since we have confined ourselves to subject-initial subclauses. Importantly, cases featuring a covert subject in the subclause will arguably be indistinguishable from the Vfin-Adv/Adv-Vfin construction proper. This situation is rather different from later stages when null subjects with a definite, non-generic referent are no longer licensed.

The similarity between these constructions is also noted by Angantýsson (2011:69), cf. (87) vs. (88) below, the latter indeed going by the name of Stylistic Fronting (SF) (see e.g. Maling 1980, Hrafnbjargarson 2004, Thráinsson 2007, Angantýsson 2011, 2017a, Þráinsson et al. 2015, and references therein):

- (87) a. *Það er ein bók þarna sem Haraldur hefur ekki lesið* (Vfin-Adv)
 there is one book there that Harold has not read

⁹⁴It is generally assumed that the early Icelandic *pro drop* requires an NP antecedent in the preceding discourse (cf. Sigurðsson 1993a), but this point is contested by Kinn et al. (2016), suggesting that this need not always be the case.

- b. *Það er ein bók þarna sem Haraldur ekki hefur lesið* (Adv-Vfin)
 there is one book there that Harold not has read
 ‘There is one book there that Harold hasn’t read’ (Angantýsson 2011:69)

- (88) *Haraldur var eini nemandinn sem ekki hafði lesið bókina* (SF)
 Harold was the only student that not had read the book
 ‘Harold was the only student that hadn’t read the book’ (*op. cit.*, fn. 47)

The adverb can also follow the verb in Modern Icelandic in this case:

- (89) *Haraldur var eini nemandinn sem hafði ekki lesið bókina*
 Harold was the only student that had not read the book

As discussed by Angantýsson (2011), it is necessary to keep examples like (88) separate from the Vfin-Adv/Adv-Vfin phenomenon proper because they are ambiguous between a structure where the adverb has undergone fronting to a position preceding the verb in I and having the verb *in situ* (cf. also Platzack 1988). This particular point is emphasised by scholars who take Stylistic Fronting to be an important factor in the emergence of the embedded Adv-Vfin order in Mainland Scandinavian. As Sundquist (2003:249) points out, the Adv-Vfin word order in Danish (as well as Swedish, citing Falk 1993) gradually increased while the frequency of Stylistic Fronting was still high, presenting the learner with ambiguous/conflicting evidence.

Previous research suggests *pro* drop was lost in/during the 19th century (Hjartardóttir 1993, Hróarsdóttir 1998, Kinn et al. 2016). Controlled subject *pro* gaps are not extremely frequent in my 19th-century material but they do occur:

- (90) a. *Sistur þinar biðja þaðar skiaelandi að heilsa þier og grata af því að ei gatu skrifað*
 sisters your ask both weeping to greet you and cry of it that not could write
 ‘Both your sisters give you their greetings weeping and cry because they could not write to you.’ (MalJen-1821-01-1X.xml)
- b. *enn eg hef haft svo mikid ad láta giera i vor ad ei hef getad leitad hans firr enn nú*
 but I have had so much to let do in spring that not have could searched his until now
 ‘But I have had so many things that need to be taken care of this spring that I have not been able to look for him (=the runaway horse).’ (SigPal-1840-06-15.xml)
- c. *Nú læt ieg þig þó vita um hægi mína þó ei sieu efni leigir*
 now let I you though know about affairs my though not are promising
 ‘Now, I let you know about how I am doing although it is not very promising.’ (JonJon-1858-11-03.xml)
- d. *framar vil jeg biðja yður að hjalpa um 2 Sálma bækur [...] og senda mjer þær með honum Jóakim ef ei fáir vissa ferð áður.*
 further want I ask you to help about 2 psalm books ... and send me them with him J. if not get_{2P.SG.} certain trip before
 ‘Furthermore, I want to ask you to provide two psalm books ... and send

them to me with Jóakim if you do get a guaranteed trip before.’
(BenJak-1855-04-06.xml)

Note that subject gaps could just as well occur with the Vfin-Adv order:

- (91) a. *Og fyrir hvað hnjúða Reikjavekur buar í eyar menn? máské*
and for what speak.badly Reykjavík_{GEN} inhabitants in island_{GEN} men maybe
þess vegna, að eru ekki klæddir perli og purpura lifa ekki á skínandi
there fore that are not dressed pearls_{DAT} and purpura live not on shiny
krásum hvurn dag dyrðlega?
delicacies each day gloriously
‘And why do the inhabitants of Reykjavík speak ill of the (Westmen?) island
people? Maybe because they are not dressed in pearls and bordeaux luxury
garments and don’t live gloriously on shimmering delicacies each day.’
(JonDan-1867-07-06.xml)
- b. *Hrædd er jeg um að komi ekki svo fljótt norður*
afraid am I about that come_{1P.SG.} not so soon north
‘I am afraid I will not come that soon to the north.’ (JakJon-1853-01-04.xml)

According to the results from IcePaHC reported by Heycock and Wallenberg (2013:143, Table 1), the Adv-Vfin order outnumbers Vfin-Adv in all types of subclauses lacking an overt subject, in declaratives and relatives by a factor of over 2:1.

It depends entirely on one’s conception of silent elements in syntax whether gaps of this kind really exist as (phonologically) ‘empty’ positions in the structure. However, once the option of leaving the subject features unexpressed in overt syntax (verbal agreement aside) is no longer an option for speakers, regardless of how the gap is represented, the subject must of course be realised somewhere. The choice will then presumably either be between a representation with the gap following the finite verb, i.e. with subject-verb inversion, or one in which the gap is in first position:

- (92) *af því að ei gátu <gap> skrifað*
because not could write
- (93) *af því að <gap> ei gátu skrifað*
because not could write

It seems plausible that ambiguous cases involving Stylistic Fronting could have contributed to the emergence of the Adv-Vfin order. This is especially so when taking into account that topics could be dropped in earlier varieties Icelandic, giving rise to the ambiguity noted in (93) vs. (92) over and above the effect of more typical cases of Stylistic Fronting where the subject gap was created by A’ movement such as in relative clauses, *wh*-questions and so on. However, the overall effect of this ambiguity in terms of contributing to the loss of V-to-T must be considered from a broader perspective, preferably by comparing and matching (un)ambiguous linguistic data to qualitatively different types of grammars along the lines of Heycock and Wallenberg (2013).⁹⁵

⁹⁵What Heycock and Wallenberg (2013) did was to circumscribe precisely the contexts that provide unambiguous evidence for a V-to-T and V-in-situ analysis in the output of V-to-T and V-in-situ grammars, respectively. In the case of Heycock and Wallenberg (2013), grammar fitness is determined on the basis of

As noted above, the availability of *pro* drop up until quite recently in the history of Icelandic will have served to increase the frequency of (suitable contexts for) Stylistic Fronting. From Sundquist's (2003) perspective, the Adv-V_{fin} adjacency due to Stylistic Fronting will in and of itself provide ambiguous/conflicting evidence to the learner. One could similarly imagine that the gradual loss of *pro* in Icelandic at roughly the same time that the Adv-V_{fin} pattern was gaining ground would only strengthen this potential source of structural ambiguity. The resulting instability within this domain would thus arguably make Adv-V_{fin} even more elusive to speakers, raising the likelihood that they were to settle on the representation in (93) over (92). As can be seen in Heycock and Wallenberg (2013:153, Figure 3), Stylistic Fronting in non-V2 environments (such as relatives), giving rise to the linear order Adv-V_{fin} as in (88), applies at a rate of approximately 85-90% during the 17th-19th century, after which it appears to be in free fall—dropping below 70% in the 20th-century material and below 25% in the 21st-century texts. Indeed, Angantýsson (2011:199f.) describes Stylistic Fronting as “an old phenomenon in Icelandic [...] considered to be relatively formal”, corroborated also by the fact that his youngest group of informants were significantly less positive towards it than his oldest group was (cf. Angantýsson 2011:153).

Heycock and Wallenberg (2013:146f.) point out that given that “a new innovation will almost always be introduced into the speech community at an initially low frequency”, it would be unlikely that V-in-situ would win out unless it had some selectional advantage over V-to-T. However, from their perspective, too, the reduction in Stylistic Fronting, due to the loss of *pro* drop and subsequent increase in overt subjects in the course of the 19th century, might be taken to affect the stability of V-to-T if contexts with null subjects, corresponding to (89), are crucial to the advantage of V-to-T over V-in-situ. At the same time, it must also be borne in mind that their Icelandic sample consists of diachronic evidence from IcePaHC, the periods 1150-1600 and 1850-2008.⁹⁶ Over that time period, the licensing mechanism of null subjects underwent dramatic changes and at the same time, we see a rise in expletive subjects and generic pronoun *maður* ‘one’. These innovations are arguably also relevant with regard to grammar fitness. Considering that both these phenomena can be seen as corresponding, at least in some cases, to

Icelandic and Swedish text samples, representing a V-to-T and a V-in-situ population, respectively. What they show is that V-in-situ has an advantage over V-to-T grammars, but only if possible EV2 contexts are taken into account, because here the input can always be parsed by a V-in-situ grammar as involving V-to-C (cf. Heycock and Wallenberg 2013:138f.).

If EV2 contexts are removed from the equation, the two grammars are more or less tied in terms of advantage, although V-to-T does receive a slightly higher fitness value. Importantly, that advantage is mainly due to Stylistic Fronting (Heycock and Wallenberg 2013:146), i.e. unambiguous contexts where there is a subject gap with V_{finite} > sentence-medial adverbs (provided that these occur in non-EV2 contexts, like relative clauses in (89) above). In other words, these are the cases where Stylistic Fronting may apply.

⁹⁶Icelandic texts from the period 1600-1850 were removed from the sample because they “show a mixture of V-to-T and V-in-situ grammars, rather than the uniform V-to-T system that one would expect from modern Icelandic and Old Icelandic” (Heycock and Wallenberg 2013:152).

A further complication concerns differences between EV2 systems. Gärtner (2016:11, fn. 19) raises the point that Heycock and Wallenberg (2013) predict that broad embedded V2 systems (*bEV2*), to which Icelandic is taken to belong (cf. discussion above), give rise to even lesser V-to-T fitness than on narrow-type systems (*nEV2*) such as Mainland Scandinavian: “the difference between determining these environments within *bEV2* as opposed to *nEV2* lies in speeding up the loss of V-to-T.” This is so because *bEV2* systems offer EV2 as an option in a wider range of environments, increasing the likelihood that learners settle on V-in-situ and V-to-C over V-to-T.

impersonal/generic *pro* (see Sigurðsson and Egerland 2009), the increasing overtiness of the subject (whether a noun, a personal pronoun, an expletive or the generic pronoun *maður*) results in an overall decrease in Stylistic Fronting. While this has no effect with regard to (91), as these already contributed to the fitness of V-to-T (in non-EV2 environments), once ambiguous contexts like (90) now require an overt subject, these, too, enter the pool of (potentially) unambiguous contexts, raising the fitness of V-to-T.⁹⁷

What about the frequency of unexpressed subjects? In order to be a likely candidate to facilitate changes in the spell-out position of the verb, this phenomenon (or these phenomena) would have to occur at a non-negligible rate. Two distinct studies of Icelandic, Heycock and Wallenberg (2013) and Kinn et al. (2016), both of which draw from sampled data from IcePaHC, give a rather different impression of the frequency of potentially relevant structural configurations. This is so because the former study deals with subject gaps caused by movement (as in (88)), whereas the latter deals with null subjects proper (as in (90)). Cases such as (88) are a relatively frequent feature of (19th-century) Icelandic, whereas the phenomenon in (90) is relatively rare.

Kinn et al. (2016) show that *pro* drop was never a very frequent phenomenon, albeit more frequent in the history of Icelandic than what has been reported for other related languages such as Old English and Old Swedish (Kinn et al. 2016:71). These scholars contrasted null with overt subjects according to a number of different factors, e.g. according to clause type (independent main vs. conjoined vs. embedded clauses, verb initial vs. verb non-initial), person and number, across time, and so on. The average relative frequency of null subjects vis-à-vis overt subject was merely 3.3%, with slightly higher values for Old Icelandic than for Modern Icelandic, for 3rd person than for 1st and 2nd person, and for embedded as opposed to main clauses. Despite these differences, the frequency in embedded clauses maxes out at only 5.5% in the period 1150-1250 and is down to 1.7% in the period 1902-2008 (cf. Kinn et al. 2016:62, Table 9).

In the data collected by Heycock and Wallenberg (2013), where the subject gap is due to movement, the rate of an unexpressed subject is much higher. As mentioned above, Heycock and Wallenberg only collected embedded clauses with sentence-medial adverbs, in effect a small subset of Kinn et al.'s data, analysing them based on the placement of the finite verb and the overtiness of the subject. In Heycock and Wallenberg's sample, actually no less than 61% of all relative clauses had a subject gap (Heycock and Wallenberg 2013:146). For other clause types the proportion is much lower, e.g. 14.2% and 11.9% in adverbials and declaratives, respectively, the average across all embedded clauses being 20.7% (Heycock and Wallenberg 2013:143, Table 1), which is still much higher than the 3.6% average reported for embedded clauses by Kinn et al. (2016:47, Table 4). Importantly, only the data collected according to Heycock and Wallenberg's criteria are able to tell us anything about the potential effect of null subjects in relation to Stylistic Fronting, giving rise to the structurally ambiguous contexts required for the kind of V-in-situ reanalysis argued for by Sundquist (2003). Their data thus suggest that while null subjects may not be very frequent on the whole (as shown by Kinn

⁹⁷One problematic aspect of testing this has to do with proscription. Since expletive *það* and the generic pronoun *maður* have been frowned upon by prescriptivists and are thus unlikely to be richly attested in formal writing (see Ragnarsdóttir and Strömquist 2005 on generic *maður*), material such as that found in the IcePaHC corpus may in fact underrepresent unambiguous clauses. This problem could be alleviated by turning to less formal, more spontaneous production data.

et al. 2016), these more relevant types of subject gap environments clearly occur at a non-negligible rate.

To sum up this section, we witnessed a steep increase in Adv-Vfin in a variety of contexts during the (late) 18th and early 19th century. It was argued that this was targeted change that eventually failed. Rather than reverting back to the earlier situation, the failed change was considered to have fuelled another change that was more successful—a type of Adv-Vfin that predominantly occurs in contexts with sentence-modifying operators. A potential link between these properties and Stylistic Fronting was tentatively proposed that could have been established via controlled null subject environments. The gradual loss of null subjects during the period under study, typically leading to obligatory overt subjects, would have increased the likelihood for these contexts to be equated with the Adv-Vfin variant.

3.7.4 Consequences for linguistic theory

The theoretical discussion on verb-adverb placement in Icelandic and various related languages has long centered on the RICH AGREEMENT HYPOTHESIS (RAH). According to the RAH, morphosyntactic factors prevent languages with rich verbal inflection (such as Icelandic) from developing the Adv-Vfin pattern as the unmarked case in subject-predicate constructions (see e.g. Vikner 1997, Rohrbacher 1999, Bobaljik and Thráinsson 1998, Bobaljik 2002, Koenenman and Zeijlstra 2014). The validity of this hypothesis has been a much debated issue, not only with regard to the proper definition of richness (e.g. Roberts 1985, Platzack and Holmberg 1989, Vikner 1997, Rohrbacher 1999, Bobaljik and Thráinsson 1998, Koenenman 2000, Koenenman and Zeijlstra 2014) but also serious doubts having been raised that any such direct link exists between verbal agreement morphology and the position of the verb in the clause in the first place (see e.g. Alexiadou and Fanselow 2002, Sundquist 2003, Hróarsdóttir et al. 2007, Wiklund et al. 2007, Garbacz 2010, Holmberg 2010, 2015).

Roberts (1985:31) appears to have been the first to explicitly argue for the Rich Agreement Hypothesis in print in the present context,⁹⁸ mainly on the basis of changes in the history of English, observing the following: “In languages with a variety of affixes marking agreement, we frequently find Verb-movement to INFL.” Roberts (1985:33, fn. 10) formulates this as a two-way implication:

- (94) a. “if there is rich agreement, there will be Verb-movement to INFL”
 b. “if there is little or no agreement, there will be no Verb-movement.”

While the richness of verb agreement seemed like a plausible candidate, accounts e.g. in terms of overt nominative case morphology, assigned by a finiteness feature, and whether negation was free or bound, leading to problems regarding adverbs lacking bound counterparts, were proposed as well (for an overview and critical discussion, see Rohrbacher 1999:94-106). More recently, the discussion has mostly revolved around two different flavours of the hypothesis which have come to be referred to as the ‘weak’ and the ‘strong’ version of the RAH, depending on whether it is formulated as a uni-

⁹⁸Citing an unpublished manuscript from a year earlier, scholars also credit Angelika Kratzer with first observing a potential link between morphology and syntax along these lines (see e.g. Bobaljik 2002:131).

or bidirectional universal statement (for extensive discussion, see e.g. Bobaljik 2002, Thráinsson 2010, Koenenman and Zeijlstra 2014).⁹⁹

Since Icelandic subject agreement is rich by any standards (proposed so far), the weak/strong distinction is not relevant for the purposes of this study. From the perspective of the RAH, therefore, low placement of the verb cannot arise in Icelandic if Adv-Vfin constitutes V-in-situ. Evidence from 19th-century Icelandic thus makes for an interesting comparison with Modern Icelandic regarding the evaluation of the RAH: How can the data presented above and subsequent V-in-situ analysis be reconciled with the RAH? Based on these findings, my hand is forced to reject both the weak (unidirectional) and the strong (bidirectional) version of the RAH, but only as an absolute linguistic universal. Even if this conclusion is accepted, there is still an interesting avenue for research, close to the weak version of the RAH but weaker in the sense that it can be ‘violated’. Due to its inherently violable properties, I propose to refer to this formulation as the **soft** version of the RAH—analysing it in terms of ‘laws of diachrony’ as argued by Alexiadou and Fanselow (2002). What their account of the RAH suggests is that the correlation is actually spurious, but in an interesting way.

Most approaches to the RAH have located the triggers for verb movement at the synchronic level, either with direct reference to inflectional paradigms of verbs (e.g. Rohrbacher 1999, Vikner 1995, 1997) or personal pronoun inventories, i.e. whether or not agreement is argumental (cf. Koenenman 2010, Koenenman and Zeijlstra 2014). On such a view, the correlation between morphology and syntax follows from universal principles of grammar. Having reviewed the empirical evidence for the RAH, Alexiadou and Fanselow (2002) reach the same conclusion as Bobaljik (2002), among others, that a bi-conditional (‘strong’) formulation fails. However, unlike others who have proposed a weaker unidirectional (‘weak’) formulation such that rich verbal agreement triggers verb movement, Alexiadou and Fanselow (2002) argue for an even more specific implicational relation:¹⁰⁰

(95) Suffixal rich inflection implies V-to-I-movement.

Why should the RAH necessarily hinge on rich suffixal agreement? To demonstrate why this would be, and also why the RAH might appear to exist independently as a universal principle of grammar, Alexiadou and Fanselow (2002) invite the reader to consider how inflectional agreement arises in languages. One hypothesis is to assume that inflectional suffixes arise via the cliticisation of a subject pronoun clitic, such that the sequence V + clitic is reinterpreted as V+AGR (see e.g. Bybee 1985, Siewierska 1999).¹⁰¹ For such a grammaticalisation process to be possible, the verb must occur in a position in front of the subject pronoun, in I (=Infl) or C (=Comp):

(96) a. [_{Infl} verb] [_{VP} subject]

⁹⁹Space does not permit an elaborate overview of the definitions of richness that have been proposed or the empirical consequences of each definition (see Rohrbacher 1999, Bobaljik 2002, Thráinsson 2010, Angantýsson 2011, Koenenman and Zeijlstra 2014, Tvica 2017).

¹⁰⁰The recent study of Tvica (2017), which greatly broadens the potential scope of the RAH, beyond Indo-European languages, suggests that prefixal agreement may actually also play a role. In my view, however, this in no way invalidates Alexiadou and Fanselow’s account with regard to suffixal agreement in Germanic.

¹⁰¹Alexiadou and Fanselow (2002) provide synchronic evidence from Irish, Standard Arabic and Moroccan Arabic of such a reinterpretation of subject clitics at three different stages of development.

b. [_{Comp} verb] [_{IP} subject]

The RAH, then, is not in any meaningful sense a synchronic principle of grammars but a consequence of diachronic developments whereby children acquiring the grammars in (96a,b) reanalyse the subject clitic as belonging to the inflection. Viewed in this way, rich affixal agreement requires strict V+S adjacency, which can only come about if the verb moves to a position preceding the subject, e.g. in an OV system with a V2 constraint, as arguably the case historically in Germanic and Indo-European in general. The assumption is that these languages later developed from OV-V2, attested in the Germanic and Romance languages historically, to the SVO system found e.g. in Mainland and Insular Scandinavian, English and the Romance languages.

They assume that Indo-European was richly inflected with V2 traits but a verb-final base order, which means that it had undergone V + clitic > V+AGR reanalysis. Such a system can develop into an SVO language and this will naturally result in a system with verb movement, i.e. deviating minimally from the strings generated by the earlier OV-V2 system: *XP verb adverb/negation*. A V2 system with rich (affixal) agreement will, all other things being equal, not develop into an SVO language without verb movement. The reason is that this will result in a completely different linear order of the verb with regard to adverbs/negation being generated: *XP adverb/negation verb*. Alexiadou and Fanselow (2002) claim that such a change will not occur in one fell swoop and that triggers for such a change, i.e. effectively a change from a system with verb movement and a system without it, will have to be sought elsewhere in the grammar. This aspect derives the historical stability of an SVO grammar that generates the *subject verb adverb/negation* pattern, viz. independent V-to-I movement.

V-to-I movement can be lost in an SVO language if triggered, for instance, by a reanalysis of structures where adverb/negation precedes the finite verb. It has often been argued that Stylistic Fronting provided the right type of environment for such a reanalysis in Mainland Scandinavian (cf. Section 3.7.3).

Alexiadou and Fanselow's recourse to 'laws of diachrony' is quite different from most generative theorising, where "the acquisition process is the product of the interaction of input data and UG and nothing else" (Bobaljik 2002:162, fn. 40). Furthermore, Bobaljik (2002) claims that on Alexiadou and Fanselow's account, the diachronic developments must be seen as having "an existence independent of UG and add additional constraints on the process of language acquisition." This is quite remarkable given that all they really propose is that language systems do not spring out of thin air but have histories (developmental paths) and the learners of these systems in the typical case will deviate from the input in small steps.¹⁰²

Assuming Greenberg's four-way typological distinction of linguistic universals (see Comrie 1989:17–23 for overview), I suggest that the traditional synchronic vs. the alternative diachronic approach to the RAH be seen as two distinct types of statements representing Greenbergian conditional/restricted universals that are either absolute (Type 3) or statistical (Type 4) in nature, cf. Table 3.15. Most research on the applicability and scope of the RAH has assumed that if the hypothesis holds, it is a Type 3 universal. This

¹⁰²This appears to be a similar point as stressed by Haspelmath concerning diachronic adaptations: "as in biology, we cannot understand synchronic language structure without taking into account its diachronic evolution" (Haspelmath 1999:189), cf. also his widely cited concluding remark: "A linguist who asks 'Why?' must be a historian." (1999:205).

	Absolute (exceptionless)	Statistical (tendencies)
Unconditional (unrestricted)	Type 1 Unrestricted absolute universals <i>All languages have property X</i>	Type 2 Unrestricted tendencies <i>Most languages have property X</i>
Conditional (restricted)	Type 3 Exceptionless implicational universals <i>If a language has property X, it also has property Y</i>	Type 4 Statistical implicational universals <i>If a language has property X, it will tend to have property Y</i>

Table 3.15. Logical types of universal statements following Greenberg (from Evans and Levinson 2009:437, adapted)

is falsified by the presence of Adv-Vfin in Icelandic, but also e.g. Övdalian and Faroese if verbal agreement is considered to be rich in these languages. Instead, I submit that the RAH must be re-formulated as a Type 4 statement. This makes sense from a historical, usage-based perspective and is arguably forced upon us, at least for consideration, by general minimalist guidelines such as “the simpler the assumptions, the deeper the explanatory force” Chomsky (2013:37).¹⁰³

Bobaljik (2002) asserts that if the option <+rich, –V-movement> were to exist, that would amount to a scenario where basically anything goes: “From a theoretical perspective, it is hard to consider such an outcome interesting, as there would be nothing to explain”—provided that “the differences do not correlate with any other variation.” (2002:163). On the present approach, nothing in principle prevents a poorly inflected language from exhibiting “V-to-I” or a richly inflected language from eventually losing it. This is an important property that allows for all logical possibilities of <±rich, ±V-movement>. However, at the same time there is a potential correlation to be expected between having rich agreement and spelling out the verb in a high position, a property that subsequently may (or may not) be lost over time.¹⁰⁴ From a typological perspective,

¹⁰³ I thank Noam Chomsky for valuable discussion about this point (28-07-2015, 29-07-2015). So as not to give the impression that Chomsky would necessarily endorse the view presented here, he points out that “there is no alternative to a pure synchronic approach” when studying individual I-languages from a biological perspective, to which he adds: “That of course does not mean that we should not, separately, investigate diachronically how it comes to be that the data presented to the child have the nature they do. Same when we investigate any other subsystem (“organ”) of the organism.” It seems fair to me to conclude that the diachronic view of Alexiadou and Fanselow (2002) is—at the very least—compatible with Chomsky’s general and most recent view that “externalization is an ancillary system” the establishment of which may perhaps best be regarded as “a cognitive problem to be solved by the child”. I fully agree on these points; we can try to explain how the patterns arose but we must also account for them in a synchronic way. The synchronic (morpho)syntactic properties, however, do not necessarily put any absolute bounds on possible linguistic systems in terms of word order. There is no independent reason to suspect that they would, given that no obvious interpretative issues arise depending on whether the verb is spelled out in situ or in a higher position and any such restriction therefore seems unlikely on a view that takes linear order purely to be a matter of externalisation.

¹⁰⁴ The loss of agreement markers in and of itself is not expected to trigger loss of V-to-I nor is there any expectation that rich agreement should somehow have an effect of preserving V-to-I (as suggested by Heycock and Wallenberg 2013). However, loss of V-to-I may occur on a number of scenarios. All things being equal, child acquisition tends to replicate the adult system rather faithfully, but with reanalysis as the result of structural ambiguity as a common source of innovation. One such scenario is the structural ambiguity taken to arise in Stylistic Fronting contexts, where the finite verb in SF is reanalysed as V in situ, contributing to

this arguably follows from relatively common assumptions about the way agreement markers are taken to arise through grammaticalisation from anaphoric pronouns to obligatory agreement affixes (see e.g. Bybee 1985, Siewierska 1999):

Anaphoric pronouns give rise to grammatical agreement markers. These commonly continue to perform an anaphoric function which over time may be lost, resulting in forms that only redundantly express person and number and/or gender. Such forms may undergo phonological erosion and subsequently be lost altogether.

(Siewierska 1999:225)

There is thus arguably still something “to explain”, viz. the proper linear configurations to give rise to agreement markers of the type relevant for the structural position (in terms of “height” as well as linearisation) of the finite verb. The dismissal of the possibility of <+rich, –V-movement> as a theoretically interesting outcome is related to generative theorising which typically engages in a constrained type of descriptive analysis predicting the existence of particular types of languages or constructions while at the same time ruling out others that ought to be impossible; as these are not generatable by the theory, explanatory adequacy is achieved (cf. Chomsky 1964, 2013; see also Haspelmath 2010 for critical discussion). From that perspective, explanation can then be reduced to principles of Universal Grammar, limiting the variation space of attested grammars and inquiring into how these UG principles arose or why they are that way, say from an evolutionary perspective, is ‘beyond’ explanatory adequacy (see e.g. Chomsky 2013). In the present context, Bobaljik (2002:163) specifically raises the point that questions of explanatory adequacy only arise on the outcome that “there is a definable class of languages which we will never find”, viz. Adv-Vfin (in the sense adopted here) in a richly inflected language like Icelandic.¹⁰⁵

Alexiadou and Fanselow (2002) essentially stress that suffixal agreement could never arise historically without verb movement. This is the same point as when Bybee (1985:201) observes that “the position of a morpheme before it becomes an inflection is important, since a morpheme cannot become fused with the verb unless it is immediately contiguous to the verb.” I take this to be largely uncontroversial and thus a viable explanation, albeit historical, for some of the effects that have been attributed to the RAH. This approach means that the Type 4 correlation is, in a way, an epiphenomenon—a byproduct of the accretion of grammatical agreement markers.

the possibility of embedded word order without verb raising (see e.g. Sundquist 2003). Another scenario is language contact which, depending on the type of contact involved, may not only lead to different patterns of use but also in the loss of case and agreement morphology (cf. Thomason 1997, 2001, Trudgill 2010, McWhorter 2011).

¹⁰⁵For Bobaljik and Thráinsson (1998), Bobaljik (2002), the explanation is partly in terms of the locality of checking relations, where non-local agreement ruled out, partly in terms of acquisition, where the language acquiring child deduces from the input whether there is evidence for a Split IP based on both morphological and syntactic evidence (see also Thráinsson 1996). Koenenman and Zeijlstra’s approach is similar, except that poor agreement markers are taken to be “the realization of some feature(s) that reside in v^0 ” (2014:601) with no higher position available for the verb. Richly inflected languages project a higher ArgP and the Stray Affix Filter, stating that an agreement marker must attach to the verb, ensures that the verb must move in this case.

3.8 Summary

The case study on verb-adverb placement reveals first and foremost that Adv-Vfin was a phenomenon that was much more widespread than made out to be in the traditional literature. Moreover, even in the language of people among the higher echelons, where Adv-Vfin had the strongest foothold—arguably constituting what the sociolinguistic literature refers to as targeted change—the use of this feature is nonetheless clearly grammatically conditioned. If Adv-Vfin had been ‘contrived’ in the sense suggested e.g. by Heycock and Wallenberg (2013), it is surprising that Adv-Vfin is relatively stable in terms of variation at the intra-speaker level. It was therefore concluded that the feature is simply to be treated as a variable aspect of these speakers’ grammatical systems like any other.

When verb-adverb placement is viewed ‘from above’, there are clear effects that seem to suggest an influence from the emerging standard language which prescribed Vfin-Adv. We witnessed effects in the newspaper corpus that indicate a suppression of Adv-Vfin towards the end of the 19th century, beyond the low average proportion attested in non-frequent users of Adv-Vfin in the private letter corpus. The student essays similarly show that Adv-Vfin was targeted at the Reykjavík Grammar School and there we find a significant effect of the educational variable graduation score. High-achieving students thus use the non-standard feature to a lesser extent than low-achieving students, which suggests that standardisation is at play.

The private letters furthermore show that the use of Adv-Vfin is greatly reduced over time among the higher echelons, which is an effect that interestingly applies to males as well as females. Among the peasants/labourers group, which in contrast may be regarded as providing a window to a view ‘from below’, there is a much lesser effect towards decreased use of Adv-Vfin and this effect among peasants/labourers is moreover confined to the females. The male speakers in the peasants/labourers group in fact exhibit a greater use of Adv-Vfin than the females.

The language-internal constraints on variation in verb-adverb placement indicate that the structure of the left periphery of the clause is a very important predictor for the use of Adv-Vfin. Beyond what has been suggested in the literature before, it was shown that not only were the odds of Adv-Vfin higher in certain relatives, interrogatives and (central, non-root) adverbials (cf. Bobaljik and Thráinsson 1998, Angantýsson 2011, Heycock and Wallenberg 2013), but also in certain kinds of (non-root) declaratives. As argued independently by Heycock and Wallenberg (2013), these properties strongly suggest an analysis in terms of V-in-situ and V-to-C, and I pursue such an analysis here.

A consequence of a V-in-situ analysis of Adv-Vfin in 19th-century Icelandic (see Section 3.7.1 for discussion of possible landing sites) is that the Rich Agreement Hypothesis must be rejected both in its ‘strong’ and ‘weak’ form (see Section 3.7.4). The data, however, do not contradict an alternative, diachronic formulation of the RAH which I referred to as the ‘soft’ formulation. This alternative view treats the RAH as a Greenbergian conditional/restricted universals that is not absolute but statistical in nature.

4 Free definite marker *sá* vs. *hinn*

The foreign guise of the language may appeal to some, as it appears to me based on the fact that some of those who hitherto have written pure Icelandic have all of a sudden adopted a Danicised language, gotten rid of the beautiful article of ours and adopted the Danish imitation *sá* instead (Magnússon 1870:67, my transl.)

The use of the definite article in 19th-century Icelandic is somewhat of an uncharted territory although numerous claims have been made to the effect that the free article *hinn/hin/hið*, used alongside the canonical bound definite marker *-inn/-in/-ið*, for masculine, feminine and neuter, respectively, was revived in the late 18th and 19th centuries on the model of the article system found in Old Norse; the preposed free article *hinn* is claimed to have eventually been superseded by *sá/sú/það* (see Ottosson 2003), originally a demonstrative. While it has been suggested that the free article was merely *sá* already by the 18th century, the free article *hinn* making its way back through language purism into the written language, from where it supposedly sifted into the spoken language eventually as well (e.g. Smári 1920, Ottosson 2003), it is not at all obvious that any research into the 19th-century distribution has actually been carried out to substantiate these claims.

The variation between *hinn* and *sá* is exemplified in (97) in almost the exact same context in two 19th-century letters by the same letter-writer:

(97) VARIABLE DEFINITE MARKER (*hinn* vs. *sá* ‘the’):

- a. *veduratta hefur vérid hinn æskilegasta*
weather has been the_{HINN} desirable_{SUP}
‘The weather has been (the) most desirable.’ (SigPal-1830-03-07.xml)
- b. *veduráttan hefur verid sú æskilegasta i vetur*
weather-the has been the_{SÁ} desirable_{SUP} in winter
‘The weather has been (the) most desirable this winter.’
(SigPal-1841-03-07.xml)

The impression that 19th-century Icelandic NP/DP structure is a domain of great interest was further strengthened by the following passage that washed ashore, so to speak, while in the process of studying verb-adverb placement in the previous chapter, as a part of a comparative study of the revised Icelandic Bible published in 1841 (see Viðarsson 2016; Hebr. 12:12-13):

- (98) *Réttid því hinar magnvönu hendurnar og þau máttþrota knén, og stígid*
Extend thus the exhausted hands-the and the powerless knees-the and step
óskeiðt með fótum ydrum, svo að sá hinn fatladi limur ekki vindist
uncrooked with feet your so that DET the disabled lim not bends_{MID}

heldur læknist.
 but cures_{MID}

The variation found in such a short passage is very striking from a modern perspective. First of all, notice the doubling of the definite article in pre- and postnominal position, much as in Norwegian, Swedish and Faroese (but unlike Danish), occurring twice in the passage cited: *hinar magnvönu hendurnar; þau máttþrota knén*. Second, observe the lexical variation where both the free form articles *hinn* (*hinar magnvönu hendurnar*) and *sá* (*þau máttþrota knén*) are used simultaneously at close range as if in free variation. This intriguing pattern is completely out in Modern Icelandic to the best of my knowledge (cf. also e.g. Magnússon 1984:91, Julien 2005:57, Sigurðsson 2006:205, Pfaff 2015:33). Third, see how *hinn* in *sá hinn falladi limur* co-occurs with the determiner *sá*, whereas in Modern Icelandic, the two, along with possessive pronouns, are usually in complementary distribution.¹⁰⁶ Note finally that in this specific case, no bound definite marker appears on the noun, unlike the two preceding examples.

Definite marking in Modern Icelandic and related languages has been subject to extensive study in much recent work on the syntax of nominal structure (see Julien 2005, Roehrs 2009, Pfaff 2015, Ingason 2016, Harðarson 2017). Surprisingly, however, the theoretical literature completely disregards the alternative realisation of the free article in the form of *sá*. From an ideological perspective, it thus appears that the non-standard variant is (presumably undeliberately) being erased or invisibilised by focusing solely on the standard variant. Instead, the focus of these studies has been to establish the relationship between the canonical bound definite marker *-inn* (e.g. as in (97b) *veduráttan* ‘the weather’) and the free article *hinn*.

Both of these patterns featuring (*h*)*inn* are considered to be standard although the use of the free article is typically severely constrained, as discussed in more detail below. In fact, the Icelandic pattern with the preposed free article is often left out of consideration and typically claimed to be mostly confined to formal, written and/or poetic language. The free article *hinn* is then treated as belonging to a different ‘variety’, viz. literary Icelandic, and mainly used with nominals that denote abstract entities (see e.g. Julien 2005:47, Roehrs 2009:13, 47f.).¹⁰⁷ However, this traditional view is contested in recent work into Modern Icelandic (see Pfaff 2015:35-38). Most importantly, there are certain contexts in which the free article is actually the one preferred over the canonical bound article (see further Pfaff 2015, Ingason 2016, Harðarson 2017). A crucial property indeed appears to be that the free article is referentially deficient and typically either makes reference to abstract notions modified by an adjective or refers to a unique salient referent that is known in the discourse; in the latter case, the adjective provides an evaluation (see e.g. Pfaff 2015:100).

The goal of this chapter is to measure the uptake of the standard free morpheme *hinn* norm and suppression of the corresponding proscribed *sá* form quantitatively. Qualitatively, we will try to assess empirically the more general claim that the free

¹⁰⁶The co-occurrence of the determiners *sá* and *hinn* is at least very stilted in present-day Icelandic (somewhat permissively judged grammatical by Magnússon 1984:96). In that case, *sá* is usually treated as a demonstrative, resulting in the pattern *sá – hinn – adjective – (noun)*. Note that the opposite pattern, *hinn – sá*, is considered to be ungrammatical in Modern Icelandic (see Magnússon 1984:96) but does at least marginally occur in the 19th-century material alongside the more commonly attested *sá – hinn* pattern.

¹⁰⁷Roehrs (2009:41, fn. 8) takes this as far as referring to “the two dialects of Modern Icelandic”.

article *hinn* was actually revived on the model of Old Norse. Moreover, we will explore in more detail the patterns we find in the 19th-century material, both in terms of spell out of definite morphemes and the functional sequence within the extended noun phrase. Unlike much previous work, my focus will not be on this relationship between the bound and the free article but rather the alternation between the free article *hinn* and the non-standard variant form *sá* (with oblique/neuter *þ*-forms).

The chapter is structured as follows. The first section provides a basic theoretical background to the study of the definite article, especially concerning the structural make-up of basic NP structure and the determiner slot(s). Section 2 gives an outline of the establishment of the definite marker in 19th-century Icelandic from a sociolinguistic perspective, focusing on evidence provided by grammar(ian)s. Section 3 is on the methodological issues, both with respect to the selection of texts and the circumscription of the linguistic variable, as well as its social context. In section 4, the results of a corpus investigation are presented and discussed in detail. Section 5 concludes the chapter with a summary and conclusion.

4.1 The standardisation of definite marking

4.1.1 Evidence from grammars and published material

Let us begin our journey through the previous literature on the standardisation of definite marking, including traditional 19th-century sources, by first considering the brief historical sketch of Icelandic (prenominal) definite markers provided by Ottosson (2003:128ff.). In the oldest Icelandic manuscripts, dating from the 12th century, the free articles *enn* or *inn* are dominant and *hinn* is found only sporadically. Around the 1300s, we find that *hinn* is gaining ground (2003:130). In the 15th century, *hinn* is more frequent than *enn/inn* (citing Þórólfsson 1925:51) and towards the end of the 16th century, observed on the basis of the 1584 edition of the Bible, *hinn* is used over 90% of the time as compared to *enn/inn* (Ottosson 2003:130). At the same time, a separate development is found in the competition among *hinn* and the use of the free article *sá*, identical to the demonstrative *sá* ‘that’ in form. A few potential attestations of *sá* as an article are found in the second half of the 13th century and more certain examples from the 14th-15th century. This development continues such that the definite article *sá* strongly increases in the 16th century, being dominant in the 1584 edition of the Bible (2003:129). The competition between *hinn* and *sá* resolved in such a way that *hinn* was superseded by *sá* during the next couple of centuries, until the 19th century when—due to puristic efforts—*hinn* was adopted as the standard norm. As a result, *sá* became confined to the spoken language, where it gradually receded as well (op. cit.).

Two Icelandic grammars from the early modern period mention prenominal definite markers but provide conflicting evidence, especially as to the status of *hinn*. These are on the one hand Runólfur Jónsson’s (1651) *Grammaticæ Islandicæ rudimenta*, the first comprehensive grammar of Icelandic, written in Latin and published in Copenhagen, and Jón Magnússon’s *Grammatica Islandica* on the other, written in 1737-1738 but not published until the 20th century (see Harðarson 1997:XXXI). The former of these works demonstrates uses of the proposed free article with weak forms of adjectives employing

sá, not *hinn*, in examples like *Saa Magre* (1651:68). No translation is provided but the intended meaning is presumably restrictive ‘the thin (one)’, with ellipsis, or an epithet. Somewhat strikingly, Jónsson states that the bound article *-inn* derives from *hann* ‘he’, which he refers to as a demonstrative pronoun (1651:5f.). This must be based on the proprial article (on which, see e.g. Sigurðsson 2006), used with proper names, but Jónsson (1651:69) also provides examples of the personal pronouns with adjectives: *hun smaa* (lit. ‘she small’) and *suu smaa* (‘the small (one)’). Jónsson’s description thus suggests that he may not even have been aware of the existence of the preposed free article *hinn*, as expected based on Ottosson’s (2003) account.¹⁰⁸

In Jón Magnússon’s *Grammatica Islandica*, all three options are provided: *sá*, *hinn* and *hann* (op. cit., p. 158). The personal pronoun is used as the proprial article (*hann Olafur*) but, oddly, *hinn* is also considered an option (*hinn Olafur*). With adjectives, we find *sá* as well as *hinn* and the comment that both can be used simultaneously, *sá hinn göde* (i.e. ‘the/that good (one)’). Unlike Runólfur Jónsson’s grammar, Jón Magnússon thus appears to show familiarity with the free article *hinn* in an apparent contradiction to Ottosson (2003). In addition, the bound article derives from the demonstrative *hinn* according to Magnússon, not the proprial article (see also Harðarson 1997:XLV-XLVI).

Turning to 19th-century sources, Rask’s (1811:99) Old Norse-Icelandic grammar states that both *hinn* and *sá* can be definite articles, whereas in his later Swedish edition, Rask (1818:98) only shows *hinn* as the preposed free article (with adjectives). However, in his discussion on the demonstrative *sá*, Rask (1818:122) notes that only *sá* is used as the free definite article in contemporary Icelandic.¹⁰⁹ Rask’s grammar clearly and firmly codifies *hinn* as the standard norm for the free article. According to Friðriksson (1846:XV) in his Icelandic reader (including a grammar section), the definite form of adjectives with the preposed free article is thus simply *hinn*: (*hinn*) *ríki* ‘(the) rich’. In Friðriksson’s subsequent Icelandic grammars, published in 1859 and 1861, we also find no mention of *sá* as a definite article, only *hinn*. Friðriksson (1859:9), furthermore, states in a note that if a noun occurs with an adjective, the article can either occur on the noun or preceding the adjective:

- (99) *Magri hesturinn, eða: Hinn magri hestur*
 thin horse_{DEF} or the_{HINN} thin horse

No restrictions on either of these variants are given. However, in his 1861 grammar, Friðriksson now states:

When an attribute follows the noun, it is up to the feeling and will of the one who is writing whether the article is kept separate preceding the attribute, or whether it is appended to the end of the noun.

(Friðriksson 1861:30, my transl.)

In principle, the two variants are therefore perfectly acceptable in the written norm but the remark on the ‘feeling’ and ‘will’ of the writer clearly suggests stylistic factors are

¹⁰⁸Jónsson does, however, provide the declension of the demonstrative *hinn* (1651:103). A part of the declension shown, the variant form *Hid*, usually belonging to the definite article *hinn*, alongside *Hitt*, usually belonging to the demonstrative *hinn*, indicates that the two have been conflated in Jónsson’s own grammar. This makes sense given that he uses *sá* as the free definite article.

¹⁰⁹As regards the proprial article, see also Rask (1818:123), stating that it is a modern feature.

at play. Note, importantly, the complete absence of the variant *sá*. This silent dismissal—*erasure* in fact—is all the more striking given that the preposed free article in the spoken language is still claimed to be *sá* rather than *hinn* in a 19th-century source, some two decades later. In Ólsen (1882) we read:

Übrigens wird *hinn* als bestimmter Artikel vor Adjectiven in der Volkssprache nicht gebraucht; der neuisländische bestimmte Artikel ist in diesem Falle das Pron. demonstr. *sá*.
(Ólsen 1882:283)

Similar to Friðriksson (1861), Bjarni Jónsson's (1893:51) syntax of Icelandic states that the preposed free article *hinn* can be used but that it is 'more often than before' (*tíðara en áður*) appended to the noun. In contrast, Jakob Jóh. Smári (1920:58) remarks that the preposed free article is not used in "unpretentious" spoken language (*í tilgerðarlausu talmáli*) except with adjectives where the noun has been elided (of the type *the best*), in which case *sá* is used, not *hinn*. According to Smári (1920:58), the old form *hinn* (*inn*) is used in poetry and to some extent also in the written language, from where it "creeps into" the spoken language (*slæðist [...] inn í talmálið*), at least partly similar to Ottosson's claims above. Similarly, Valtýr Guðmundsson (1922:84) states that the definite article *hinn*, while used with adjectives in the written language, is not found in the common spoken language ("fíndes ikke i det almindelige Talesprog"), but the demonstrative pronoun *sá* is used in its stead.

The codification of the free article in Rask's (1818) Old Norse-Icelandic grammar is crystal clear, albeit prescriptive rather than descriptive from a contemporary perspective. Since Rask's grammar was taught (or reworked into other material for in-house purposes) at the grammar school, first at Bessastaðir and later in Reykjavík (see Möller 2017), *hinn* could in principle be expected to quickly become the norm of the emerging standard. Yet, still, it appears that in terms of implementation, the norms were still in flux over the next couple of decades even at the grammar school, at least up until around the time of the *Fjölnir* journal and their influential book reviews.¹¹⁰

Perhaps the clearest indicator of this is the language use of a translation of the epic poems of Homer by Sveinbjörn Egilsson (1791-1852) in connection with his teaching at the Bessastaðir Grammar School, especially with regard to the *Odyssey*, published in the school's yearbook over a period of roughly a decade between 1829-1840. These translations reveal striking developments in the use of the free definite marker—the gradual adoption of the norm, in particular towards the very end. Thus, Finnboði Guðmundsson (1960:177) remarks that in his oldest translations of Homer's *Iliad*, read in school between 1819-1830 (cf. Guðmundsson 1960:32), Egilsson usually (*að jafnaði*) uses *sá* as an article. However, in his translation of the *Odyssey*, read in class between 1830-1844, Egilsson starts using *hinn* in part, the use of which reportedly increases slightly over time until *hinn* is used nearly categorically in the last books of the *Odyssey* (cf. Guðmundsson 1960:177). Guðmundsson (1960) does not substantiate this claim in any way statistically. However, a cursory examination of the *Odyssey* does confirm this overall picture.

¹¹⁰The book reviews do not explicitly target the *hinn* vs. *sá* distinction, although the *hinn* norm is clearly adopted and thus tacitly assumed. The only explicit comment concerns a choice not among *sá* vs. *hinn* but among (*hinn*) variants themselves, where the related free form variant *enn* (as opposed to *hinn* and *inn*) is rejected as inferior, based on the assumption that (*hinn*) is older (Gíslason 1844:79).

By manually skimming through the first (approximately) 10 pages of each Odyssey translation by Egilsson, I get the result shown in Table 4.16. According to the sample used, albeit small, a number of striking developments in Egilsson's use of *hinn* and *sá* can, indeed, be discerned, more or less in accord with Guðmundsson (1960) claims above. The distribution does give the impression that the use of *hinn* increases somewhat over time during the period 1829-1839 (books I, V, IX, XIII, cf. Egilsson 1829-1840), with a steep increase in 1840 and near categorical use in the last book studied. However, the sampled distribution between 1829-1839 turns out not to be statistically significant.¹¹¹ In contrast, the changes in 1840, viz. the increase in *hinn* from 66% to 95.4% and the corresponding decrease in *sá* from 34% to 4.6%, are both strongly significant.¹¹² Note that books XVII and XXI, although both dating from 1840, were published separately in volumes comprising books XVII-XX and XXI-XXIV, respectively.

YEAR	BOOK	HINN	%	SÁ	%
1829	I	17	26.6%	47	73.4%
1835	V	16	21.3%	59	78.7%
1838	IX	16	34.0%	31	66.0%
1839	XIII	23	39.0%	36	61.0%
1840	XVII	35	66.0%	18	34.0%
1840	XXI	62	95.4%	3	4.6%

Table 4.16. Egilsson's use of the free articles *hinn* and *sá* in his Odyssey translation, 1829-1840.

What is perhaps rather strange is why, during the period 1829-1839, the frequency of *hinn* is consistently far below that of *sá*, although abundantly present from the beginning nonetheless. This immediately raises the question of speaker agency. Are we really witnessing the sort of deliberate act of employing *hinn*, as implicit on Guðmundsson's 1960 account, allegedly nearly absent in the oldest (unpublished) translations of the Iliad? The two forms are used so interchangeably in the period 1829-1839, often in the same sentence only millimetres apart, even within the same phrase, that Egilsson would, in fact, have to have been incredibly sloppy if he was really making an effort to use *hinn*. It is as if *hinn* is being used as a décor at first, alongside *sá*, while attempting to avoid using the form excessively. It seems more likely that it was only in 1840 that Egilsson actually tried to use *hinn* consistently in what appears to have been a sudden decision. This is also the only interpretation of the above results available given the fact that the observed increase from 1829-1839 was not significant in the sample. Egilsson's embracement of *hinn* thus appears less gradual than Guðmundsson's 1960 account leads one to expect.¹¹³

In any case, Egilsson did obviously not feel much need for using one or the other variant consistently until the last volume published in 1840. In (100) a number of super

¹¹¹The chi-square statistic is 5.6993, p-value = 0.127192.

¹¹²For book XIII (1839) vs. XVII (1840), the chi-square statistic is 8.1848, p-value = 0.004224. For book XVII (1840) vs. XXI (1840), the chi-square statistic is 17.1872, p-value = 0.000034.

¹¹³Admittedly, the distribution might turn out to be significant if the sample were larger. I leave this to further study.

tokens are provided from different periods, featuring variation at very close range.

- (100) a. *það er vilji hinna sælu guda, að sá vitri Odysseifur komist aptur til*
 that is will the_{HINN} pleased gods that the_{sÁ} wise Odysseus reaches again to
húss síns (1829:7, Book I)
 house his
- b. *en sá ógurlegi adstedjandi bylur hinna samblöndnu vinda*
 but the_{sÁ} frightful imminent storm the_{HINN} diffusive winds
 (1835:12, Book V)
- c. *hinn ástfólgni sonur þess ágæta Odysseifs* (1840:3, Book XVII)
 the_{HINN} beloved son the_{sÁ} great Odysseus

Example (100a) shows *hinn* used in a matrix clause, embedded within which we find *sá*. Similarly, in (100b,c), *sá* and *hinn* are found within one and the same complex, extended NP.¹¹⁴

Zooming out again, in addition to the two major patterns, we find that a third pattern is acknowledged, combining the two definite-marking strategies, often referred to as double definiteness in the literature. As far as I have seen, the pattern is first explicitly noted by Jónsson (1893):

The article may be appended after the noun, even though an adjective with an article precedes it, as here: *hann á hið góða skipið, er siglir gegnum það hið grunna sundið* ‘He who sails through that the narrow channel-the, has the good ship-the’. (Jónsson 1893:44, my transl.)

The same double definiteness pattern is mentioned e.g. by Guðmundsson (1922):¹¹⁵

The article may sometimes occur both before the adjective and after the noun, e.g. *hinn eldri maðurinn, hin yngri konan* [‘the old man-the, the young woman-the’]. (Guðmundsson 1922:84, Anm. 1, my transl.)

A similar, partly historical, partly contemporary observation is provided by Björn K. Þórólfsson (1925:52).¹¹⁶ While double definiteness is not explicitly mentioned by Friðriksson (1846, 1859, 1861), he often uses the pattern in his own text, seemingly unaware of it.¹¹⁷ The modern linguistic literature does not recognise double definiteness as a feature of Modern Icelandic.

Before moving on to look at the results of the corpus studies, let us first consider the emerging standard from the perspective of choice of definite marker in a little more detail.

¹¹⁴Among the three examples of *sá* in the last volume sampled here, which featured *hinn* nearly throughout, only one instance actually involves a full NP, viz. *inn í þau velsettu herbergi* ‘into the well-situated rooms’. The other two involve *einn* in *sá eini* ‘the only one’, which is rather idiomatic. Observe, for instance, that *sá eini* greatly outnumbers *hinn eini* in the tagged MÍM corpus of present-day Icelandic, the former yielding 680 hits, the latter only 249 (incl. inflected forms).

¹¹⁵Smári (1920:43) also observes a similar pattern with demonstratives, the following examples he claims to be taken from the spoken language: *sá maðurinn* ‘that man-the’, *þessi gæðin* ‘these qualities-the’. He cites further examples where the noun may be definite, e.g. with numerals.

¹¹⁶Cf. also Nygaard (1905:29) on Old Norse, e.g. *líkit Þoris* ‘Þórir’s corpse-the’ and *hit mikla skipit* ‘the great ship-the’. Similarly Helgason (1929:114), stating that double definiteness is often (‘oftsinnis’) found in Oddur Gottskálksson’s New Testament translation, 16th century.

¹¹⁷Some typical examples are given below:

4.1.2 A view from the Reykjavík Grammar School

Based on the above, standard reference works only present us with a rather fragmentary and incomplete view of the standard norm. What do the stylistic remarks on the (semantically?) felicitous use of *hinn* really mean? What were the appropriate contexts for the free article *hinn* as opposed to the bound marker *-inn* if both were standard and thus acceptable in principle? How, for instance, was the tension resolved between the free article *hinn* and the bound marker *-inn* in restrictive contexts? What about thorny issues like restrictive uses of nominalised adjectives or noun ellipsis where both the bound marker and the free article *hinn* are considered unavailable (presently at least, cf. above)? Was *sá* simply tolerated in these contexts as a definite determiner, perhaps as a demonstrative bordering on being a definite article? To what extent was *sá* to be abolished given that *sá* often occurs restrictively in Modern Icelandic (typically corresponding to English ‘the (one)’), as well as as a demonstrative pronoun “proper” (typically corresponding to English ‘that’).

Luckily, the grammars may be complemented by student assignments from the Reykjavík Grammar School. Because this material was explicitly corrected by the teachers, we can arrive at a more precise view of the emerging (or attempted) standard from these corrections, and in some cases perhaps even (systematic) absence of corrections. These corrections thus also complement other kinds of evidence of the standard through actual language use, which can be expected to be noisier/more variable and to lag behind these corrections (on the time lag, see e.g. Auer 2009:71).

Indeed, corrections of the definite marker *sá* in the Reykjavík Grammar School assignments turned out to be extremely valuable, adding much to incomplete descriptions found in grammar books. The corrections suggest that *sá* as a definite article was simply considered UNACCEPTABLE as a part of the written standard, whereas most 19th-century grammars simply do not address *sá* in functions other than as a demonstrative. Moreover, *hinn* as the standard norm reached very far and much farther, in fact, than in the present-day standard. That fact alone suggests that norm implementation was not nearly as successful as it may appear at first, based on statements concerning the use of *hinn* in formal, literary Icelandic; the Reykjavík Grammar School had a much greater ambition, viz. for *hinn* to replace *sá* not only in its non-restrictive uses but in other contexts as well, including referential contexts. This strikes the modern reader as very odd.

- | | | |
|-----|---|-----------------------|
| (1) | a. <i>einkum í hinni nýrri tungunni</i>
especially in the newer tongue-the
‘Especially in contemporary Icelandic’ | (Friðriksson 1861:26) |
| | b. <i>Í hinni fornu tungunni</i>
in the old tongue-the
‘In the old tongue (i.e. Old Norse)’ | (op. cit., p. 36) |
| (2) | a. <i>hinar sömu beygingarendingarnar</i>
the same suffixes-the
‘the same suffixes’ | (op. cit., p. 31) |
| | b. <i>Hinar styttri endingarnar</i>
the shorter suffixes-the
‘the shorter suffixes (viz. <i>ri</i> , <i>ra</i> ; <i>stur</i> , <i>st</i>)’ | (op. cit., p. 37, 38) |

Double definiteness is also found with demonstratives, e.g. *Í þessari beygingunni* ‘in this declination-the’ (ibid., p. 22, 29). However, the noun also occurs regularly without the suffixed article, e.g. *leifar hinnar fornu beygingar* ‘remnants of the old inflection’ (ibid., p. 24), *hina styttri endingu* ‘the shorter ending’ (ibid., p. 38).

Beginning with the former, we naturally find corrections of typical non-restrictive uses of *sá* with evaluative adjectives and/or in reference to abstract entities, going all the way back to the early days of the grammar school. Recall that these are the *hinn*-friendly environments in present-day Icelandic. The first two examples are taken from Icelandic essays (referred to as *ritgjörð*), the third example from an Icelandic translation from Danish (referred to as *stíll*) on an entry exam:

- (101) a. *til að koma því góða til leiðar*
 for to come the good to way
 ‘In order to do what is good.’ (1852, 3rd grade)
- b. *Aptur á móti er það eins og þeim þrifna verði allt drjúgt*
 again on opposite is it as and the cleanly become everything lasting
 ‘On the other hand, it is as if everything lasts for one who is tidy.’
 (1855, 2nd grade)
- c. *Kaffarnir álíta, að þeirra hörundslitur sje sá fegursti í heimi*
 the Kaffirs think that their skin.colour is the most.beautiful in world
 ‘[Africans] think their skin colour is the most beautiful in the world.’
 (1875, entry exams)

On a general note, *sá* as a definite article is typically underlined by the teachers, sometimes twice or even up to three times as in (101a), which suggests that it was considered a serious grammatical error. While there is no particular doubt about the interpretation of these corrections, underlining in and of itself obviously does not actually state what the proper form should have been. In this regard, (101c) is particularly revealing. Other assignments from the same 1875 entry exam confirm that the correction is really about the choice between *sá* and *hinn*—assignments featuring *hinn* instead were not corrected, e.g. *álíta [...] hinn fegursta* ‘consider ... the most beautiful’.

Moving on to canonical uses of *sá* from the perspective of Modern Icelandic in potential demonstrative and/or referential contexts, we see in (102) and (103) that these were, rather surprisingly, corrected as well:

- (102) *Potential demonstrative uses*
- a. *Móílæti er allt það móðræga, sem manningum mætir hjer í lífi*
 adversity is all the adverse REL the man meets here in life
 ‘Adversity is all the unfortunate that one encounters here in this life.’
 (1855, 3rd grade essay)
- b. *Það fyrsta er hann gjörði, var það, að hann [...]*
 the first REL he did was it that he
 ‘The first that he did was to ...’ (1852, 3rd grade translation)
- c. *hinn traustasti og beztu bautasteinn (minnisvarði), sem hægt er að*
 the most.solid and best memorial.stone (monument) REL possible is to
reisa manni, sem deyr, er einmitt það góða mannorð, sem hann sjálfur
 raise man REL dies is exactly the good reputation REL he self
hefur áunnið sjer
 has earned REFL
 ‘The most solid and best memorial stone that it is possible to raise for a

deceased one is the very reputation he has earned for himself.’

(1890, 4th grade essay)

(103) *Anaphoric/referential*

- a. [...] *hindrað framkvæmd (þessa) \hins/ góða ásetnings [...] Svo er að*
 hinder execution this the good intention then is to
athuga þær útvortis orsakir, sem geta hindrað þann góða ásetning.
 check the external causes REL can hinder DET good intention
 ‘[...] (can) hinder the execution of (this) the good intention. [...] Then there
 are the external causes can hinder that/the good intention.’

(1852, 3rd grade essay)

- b. *Ef menn spurja Kafan, hvaða litur honum þyki fallegastur, þá*
 if people ask the Kafir what colour him thinks most.beautiful then
svarar hann: sá, sem líkist minum eigin það er að segja sá
 answers he the.one REL resembles my own that is to say
svarti

the black

‘When asked about his favourite colour, the [African] replies: the one that
 resembles my own, i.e. (the) black.’ (1875, entry exams)

- c. *Sú jörð, sem mennirnir búa á heitir pláneta, að því leyti, sem hún*
 DET earth REL the people live on is.called planet at that respect REL she
fær sína birtu og hita sinn frá sólnni; fyrir utan hana þekkja
 gets her_{REFL} brightness and heat her_{REFL} from the sun besides her know
menn sjö stórar plánetur og margar smáar. Þær stærri eru Merkur,
 people seven large planets and many small the larger are Mercury
Venus, Marz Jupiter, Saturn, Uranus og Neptun
 Venus Mars Jupiter Saturn Uranus and Neptune
 ‘The Earth, inhabited by humans, is a planet in that it gets its brightness and
 heat from the sun; besides it, seven large planets and many smaller ones are
 known. The larger (ones) are Mercury, Venus, Mars, Jupiter, Saturn, Uranus
 and Neptune.’ (1890, 3rd grade essay)

Just as above, an example corresponding to (103b) occurs with the standard form *hinn*, thus uncorrected, in other entry exams (i.e. *hinn svarti* ‘the black’).

4.2 Grammatical factors

Previous accounts of the free-standing article in Icelandic emphasise that whereas all adjectives typically block N-to-D in the Scandinavian languages, only a subset of adjectives (usually evaluatives) does so in Icelandic. With regard to variation between *sá* and *hinn*, the relevant contexts in the modern standard language thus mainly concerns evaluative adjectives, as well as the niche that *hinn* additionally has with dates. The Reykjavík Grammar School norm, which appears to have prescribed *hinn* at the expense of *sá* in all definite environments, clearly runs counter to the scope of variation in Modern Icelandic definite-marker strategies. The main difference lies in the fact that *hinn* could occur even when the adjective was restrictive, as we have seen above.

A number of aspects are of potential relevance to general linguistic theory. Similar to the supposed role of morphology with regard to verb placement above, it has been suggested that the adjective blocks movement of the noun to the determiner position in languages with poor case morphology, a mechanism which otherwise derives the suffixal pattern in the absence of an adjective (cf. Vangsnes 2003). Another point of interest from a (morpho)syntactic point of view concerns the weak vs. strong inflection of the adjective (see Pfaff 2015) and the possibility of double definiteness featuring the prenominal and suffixal determiner simultaneously, found in Faroese, Norwegian and Swedish but not in Danish (see e.g. Delsing 2003, Julien 2005). Moreover, there is evidence for multiple possessor positions within the DP, one additional position to the two positions argued for by Pfaff (2015).

In present-day Icelandic, the free article may in principle surface provided that the adjective is external to the NP, whereas NP-internal modification leads to the bound/suffixed article *-inn* (see Pfaff 2015:107ff., Ingason 2016:138f.). For Pfaff (2015:109) the free article is spelled out because the article needs to “see” the prenominal modifier—invisible to D on the thematic reading as the adjective is then merged inside NP (his nP). Pfaff’s (2015) technical implementation of how it is exactly that the free article requires mediation of an element of an appropriate type in terms of structural position is left open. However, he notes that “it relies on the presence of lexical material as mediator in being assigned a value” (Pfaff 2015:133). The slightly different intervention analysis of Ingason (2016) in a way turns this upside down in the sense that there is no mediator licensing *hinn* in a positive way. Instead, this is formulated in a negative way where there is blocking. We can then simply assume there is a displacement property responsible for nP movement to D (an Edge/EPP feature) where the free article only arises if there is an **intervener**, not as a first resort but as a last resort (for further details, see Ingason 2016).

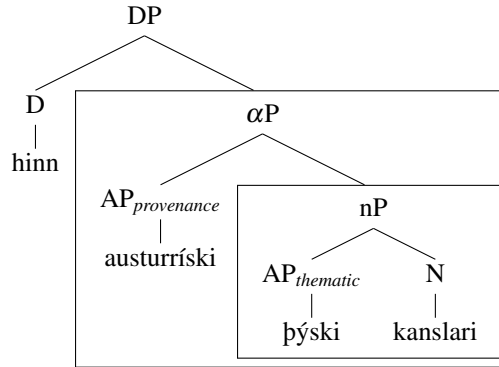
The structural contrast between NP-internal and NP-external modification may be illustrated on the basis of an example like (104) involving nationality adjectives where there is a clear semantic difference between the two types of environments, a thematic reading and a provenance reading respectively (cf. Pfaff 2015:107):

- (104) *franski forsetinn*
 French president_{DEF}
 ⇒ the president of France THEMATIC
 ⇒ the president who is French PROVENANCE

The former reading is argued to involve adjectival modification internal to the NP whereas the adjective on the latter reading, where the adjective is backgrounded (it just so happens that the president is French), is merged outside NP. The contrast can be made more explicit by accumulating distinct adjectives as in (105) (see Pfaff 2015:112):

- (105) *hinn austurríski þýski kanslari* (HINN)
 the Austrian German chancellor

(106)



The innermost subtree boxed in (106) does not require the presence of *hinn* because nothing intervenes, much in the way English periphrastic *do*-support does not arise in neutral affirmative declaratives such as *John saw* (see Ingason 2016 for an analysis of *hinn*, *hinn*-support, in precisely these terms). In contrast, the non-thematic (non-restrictive) reading of the structurally higher adjective *austurríski* ‘Austrian’ does count as an intervener, being nP-external. According to Cinque (2010:33), the comparative (indirect modification) will typically receive a restrictive reading and be deictic, whereas the positive (direct modification) will typically be non-restrictive and generic; superlatives, in contrast, can pattern either way.

As we will see below, the Modern Icelandic analyses of Ingason (2016) and Pfaff (2015) are both too restrictive to serve as accounts of the distribution of the free definite determiners *sá/hinn* in the 19th-century data, but they will serve as our point of departure in Section 4.5.

4.3 Results

Concerning actual *sá* vs. *hinn* frequencies reported on in this section, it is important to emphasise that the results were not disambiguated contextually to categorically exclude anaphoric or referential uses of the definite markers and may potentially include cases where *sá/hinn* could be considered demonstratives, such as the ones exemplified in the previous section. Therefore when observing frequency changes over time, our main focus below will typically not be on the specific frequency of either variant at a given point but overall trends and trajectories, i.e. whether or not the frequency of *sá* relative to *hinn* decreases, as indicative/result of the adoption of the standard norm (*hinn*). However, as we will see, some patterns are less open to interpretation or even unequivocal in this regard, being non-anaphoric and/or clearly feature the definite article.

4.3.1 Overall trends in the data

There is considerable variation in the use of *sá* vs. *hinn* in the newspaper corpus, indicative of quite a remarkably successful uptake of the standard norm over time. Raw data are provided in Table 4.17, split over three time periods and distinguishing between

	1784-1850		1851-1885		1886-1924	
NEWSPAPERS						
Evaluative	35.1%	(308/878)	12.8%	(72/561)	19.4%	(297/1534)
Descriptive/Other	35.3%	(823/2330)	9.6%	(145/1509)	13.3%	(454/3409)
Nationality/Origin	32.4%	(151/466)	5.2%	(6/115)	5.1%	(20/389)
Date	97.5%	(429/440)	35.7%	(46/129)	34.8%	(49/141)
PRIVATE LETTERS						
Evaluative	74.0%	(423/572)	54.1%	(238/440)	64.1%	(173/270)
Descriptive/Other	86.7%	(1013/1168)	70.2%	(649/924)	75.0%	(333/444)
Nationality/Origin	90.7%	(49/54)	50.0%	(11/22)	66.7%	(6/9)
Date	100.0%	(481/481)	95.8%	(182/190)	95.9%	(70/73)

Table 4.17. Percentage *sá* across adjective categories in private letters vs. newspapers over time (higher values = greater use of sub-standard *sá* norm).

four adjectival categories, viz. evaluatives, descriptives/other,¹¹⁸ nationality/origin and dates, respectively. The first and the last category, evaluatives and dates, are both pretty reliable indicators of the relevant non-restrictive reference we are mainly after, based on the discussion in the linguistic literature. The latter category only contains members that are neither deictic nor anaphoric and the former is widely regarded as typically being non-restrictive.

In terms of directionality, the private letter corpus does reveal an overall trend similar to the newspapers, albeit very subtle in the private letters in comparison to the newspapers. As Table 4.17 reveals, the effects are indeed very weak in the private letter corpus with a 10-20% decrease of *sá* after 1850 in the case of the evaluatives, whereas dates continue to occur nearly exclusively with *sá*. In the newspapers, in contrast, dates almost categorically occur with *sá* before 1850 but the rate of *sá* is down to approximately 35% after 1850, as measured against *hinn*.

Zooming further in than shown in Table 4.17 during the first half of the 19th century, the frequency of *sá* with evaluatives in the newspapers is 53% (98/184) between 1803-1818 and 58% (172/295) in the 1820s-1830s, but drops to mere 8% (29/366) in the **1840s**. Interestingly, this is quite parallel to the developments observed in Sveinbjörn Egilsson's *Odyssey* translation (1829-1840), cf. Table 4.16 above. Later periods show a slight 'setback' in evaluative adjectives in the newspaper corpus, occurring at an average rate of 17.6% after 1850, which still corresponds to a standard norm rate of 82.4% and reaching even higher levels with the other two adjective categories—as high as 95% with nationality/origin adjectives. This last group is severely underrepresented in the private letter corpus, making any comparison impossible.

Since cross-tabulations of this sort do not take other factors into account, such as the level of the individual or the interaction of social and (other) linguistic categories, let us now view these data through the lens of more advanced statistical modelling.

¹¹⁸ 'Other' mainly refers to relative/dimensional adjectives (*old/new/young, tall/short, big/small*), absolute adjectives (*famous/obscure, pure/impure, safe/dangerous*) and ordinals (cardinal numbers were excluded). Depending on the context, some of these can also be used 'evaluatively', see e.g. Pfaff (2015:104, 118f.) on *frægur* 'famous' and Ingason (2016:119f.) on *nýr* 'new'.

4.3.2 Modelling the variation

This section presents a statistical analysis of variability in free definite marker choice in the newspaper and private letter corpus, the former on the basis of six fixed effects, the latter adding social variables to the mix (first value listed = reference level for comparison, as above).

- (107) a. ADJECTIVE CATEGORY (four-level factor: *descriptive/other*, *evaluative*, *nationality/origin*, *date*)
 b. ADJECTIVE DEGREE (three-level factor: *positive*, *comparative*, *superlative*)
 c. PHRASE TYPE (two-level factor: *determiner-adjective-nominal*, *determiner-adjective*)
 d. RELATIVE (two-level factor: *independent* (simplex/unmodified structure), *relative* (complex/modified by a relative clause))
 e. TIME PERIOD (continuous: per decade)
 f. ORIGIN (three-level factor: *Southwest*, *rest of Iceland (countryside)*, *Copenhagen*)

The private letter model adds the social variable SPEAKER SEX/GENDER:¹¹⁹

- (108) SPEAKER SEX/GENDER (two-level factor: *males*, *females*)

A forest plot of the estimates is provided as a basic overview of the newspaper model in Figure 4.16.¹²⁰ To the exception of ADJ.CAT (*date*, *eval*), all factors reveal higher odds of *hinn* (OR <1) in the newspapers. Dates have 13.09 times the odds of occurring with *sá* than adjectives belonging to the descriptive/other category do, whereas evaluatives have 1.45 times the odds of occurring with *sá*. Complex DPs with relatives have 4.69 times the odds of occurring with *sá*. The significant factors with higher odds of *hinn* are adjectives of the category *nationality/origin* (OR=0.73), superlative adjectives (OR=0.54) and comparatives (OR=0.41), phrases with an overt nominal (OR=0.44) and newspapers published in Copenhagen (OR=0.37). All these effects are significant at the level $p < 0.001$, except *nationality/origin* which is significant at $p < 0.01$. The factors will be discussed in more detail below.

A forest plot of the estimates is provided as a basic overview of the private letter model in Figure 4.17.¹²¹ As in the newspapers above, dates exhibit very high odds of *sá*, 37.85 times the odds of adjectives belonging to the descriptive/other category, significant at the level $p < 0.001$. Complex DPs with relatives have 2.84 times the odds of occurring with *sá*. Female speakers, moreover, have 2.27 times the odds of *sá* in

¹¹⁹The factor SOCIAL RANK (cf. above) was tested but had to be left out because it did not significantly improve the model (see main text for further discussion). The factors ORIGIN REGION and PLACE REGION were tested as well, but they were not significant and/or did not improve the model. There was a statistical effect for higher odds of *sá* in the west (OR 3.17, std. error 0.50972, $p=0.02366$) in comparison to the Southwest. However, including the effect resulted in a much higher BIC value, a penalty for the increased complexity of the model. Simplifying the levels on par with the newspapers did not yield significant contrasts.

¹²⁰Model evaluation: C index of concordance = 0.8559190, Somers' D_{xy} = 0.7118381. Classification accuracy: 82.22% of the choices among variants predicted correctly (observations = 11893, titles = 46).

¹²¹Model evaluation: C index of concordance = 0.8424551, Somers' D_{xy} = 0.6849103. Classification accuracy: 83.85% of the choices among variants predicted correctly (observations = 4391, individuals = 155).

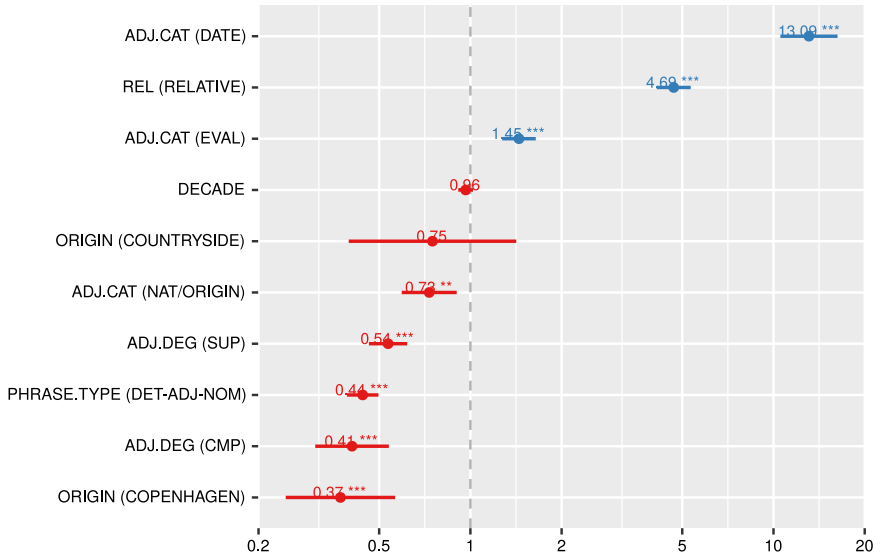


Figure 4.16. Newspapers: A forest plot of estimates with odds ratios of fixed effects. OR>1 (blue) increased odds of *sá*, OR<1 (red) decreased odds of *sá*. (*' $p<0.05$, ***' $p<0.01$, ****' $p<0.001$).

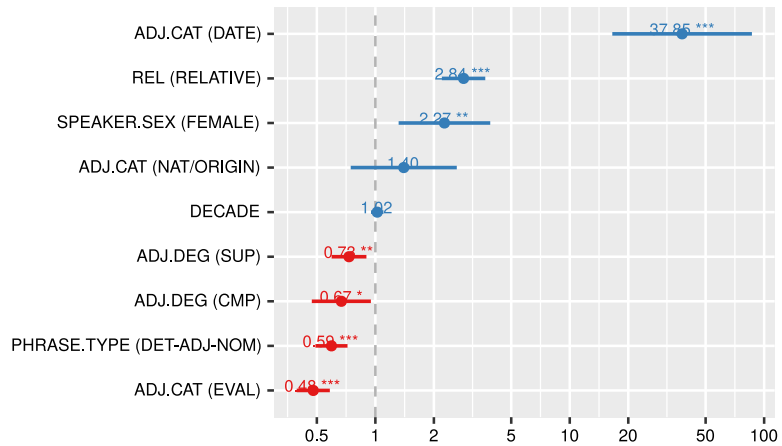


Figure 4.17. Letters: A forest plot of estimates with odds ratios of fixed effects. OR>1 (blue) increased odds of *sá*, OR<1 (red) decreased odds of *sá*. (*' $p<0.05$, ***' $p<0.01$, ****' $p<0.001$).

comparison to male speakers ($p=0.00312$), but it should be pointed out that this effect is only significant if speakers whose social rank is unknown are excluded (see below). All other factors exhibit decreased odds of *sá*, the significant ones being superlatives ($OR=0.73$, $p=0.00303$) and comparatives ($OR=0.67$, $p=0.02399$), evaluatives ($OR=0.48$, $p=2.01e-13$) and finally phrases with an overt nominal ($OR=0.59$, $p=8.34e-08$). Let us now look at these factors more closely in both corpora.

ADJECTIVE CATEGORY is an important factor from the perspective of Modern Icelandic in that evaluative adjectives have been argued to block the suffixed article; in other words, evaluatives constitute an environment with a strong tendency to be non-restrictive and to occur with the free determiner (see further discussion in Section 4.5). The other environments (descriptives/other, nationality/origin adjectives), in contrast, will not tend to be non-restrictive (although they can be). With regard to *sá*, there is thus a greater risk that these other environments will be restrictive, in which case *sá* may be better analysed as a demonstrative rather than a definite article. However, as we have seen in Section 4.1.2, the 19th-century standard norm appears to have been different from the present-day situation in that the free article *hinn* was also prescribed in case *sá* occurred with an adjective in restrictive contexts. For this reason, the scope of *hinn/sá* variation in 19th-century Icelandic arguably goes far beyond evaluatives.

As Figures 4.16 and 4.17 show, the trajectories of descriptives, evaluatives and nationalities/origins are similar, the proportion of *hinn* with evaluatives being lower than for the other three environments in the newspapers, but higher in the private letters. This seems to suggest that *hinn* has the strongest foothold in evaluatives in the private letters, which is interesting in that this continues to be the most typical environment for *hinn* in present-day Icelandic (cf. e.g. Pfaff 2015, Ingason 2016). The most striking difference, however, is the sharp decline of *sá* in the newspapers in all environments, although dates are lagging enormously behind the rest. There is also a gradual decrease of *sá* in the private letters but it is much more subtle than in the newspapers and dates do not participate in it at all, remaining around 100% throughout the whole period.

ADJECTIVE CATEGORY was added to the model mainly to distinguish between *hinn*-friendly and *hinn*-hostile environments, based on Modern Icelandic. DEGREE OF COMPARISON was included for a similar reason. Recall that comparative adjectives are typically restrictive and deictic, positive adjectives are typically non-restrictive and generic, and superlatives may pattern either way (see Section 4.2). From a present-day perspective, positive adjectives will thus tend to be *hinn*-friendly whereas comparatives tend to be *hinn*-hostile, superlatives tending to be mixed. As Figures 6.38 and 6.39 reveal (see Appendix B), this prediction is not borne out. On the contrary, in both corpora, positive adjectives more frequently occur with *sá* than comparatives. However, based on Figures 4.16 and 4.17, this result is not unexpected; if *hinn/sá* as free articles are not confined to non-restrictive adjectives in the first place, no such tendency is expected.

The nominal (DET-ADJ-NOM) value of PHRASE TYPE is associated with decreased odds for *sá* in both corpora, cf. Figures 4.16 and 4.17. However, the two corpora reveal strikingly different patterns as witnessed by Figures 4.20 and 4.21. Whereas the proportion of *sá* declines in a near identical fashion with the (nominalised) adjectival (DET-ADJ) pattern and the nominal pattern (DET-ADJ-NOM) in the newspapers corpus, the decline of *sá* in the private letters is mainly found with the nominal pattern (DET-ADJ-NOM); here, *sá* with nominalised adjectives (DET-ADJ) remains much stronger.

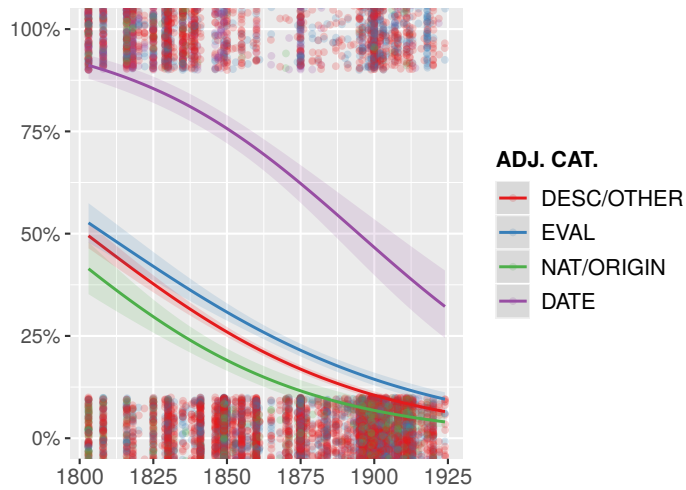


Figure 4.18. Newspapers: Predicted probabilities of fixed effect adjective category over time, conditioned on random effects (higher values = greater use of non-standard *sá* norm).

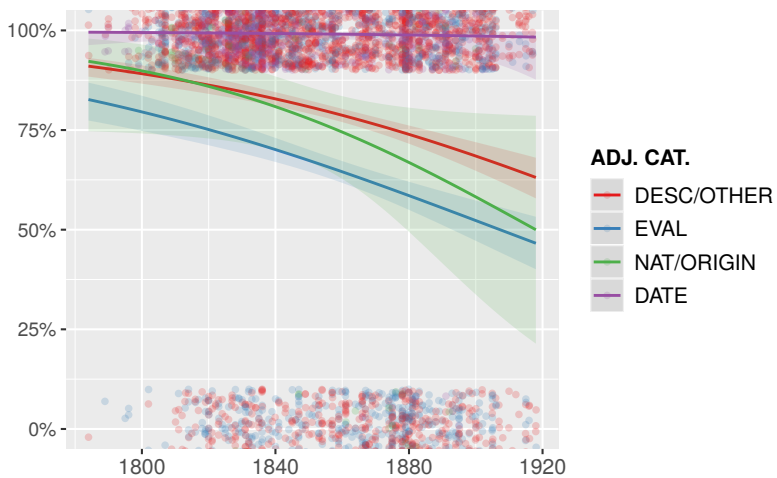


Figure 4.19. Letters: Predicted probabilities of fixed effect adjective category over time, conditioned on random effects (higher values = greater use of non-standard *sá* norm).

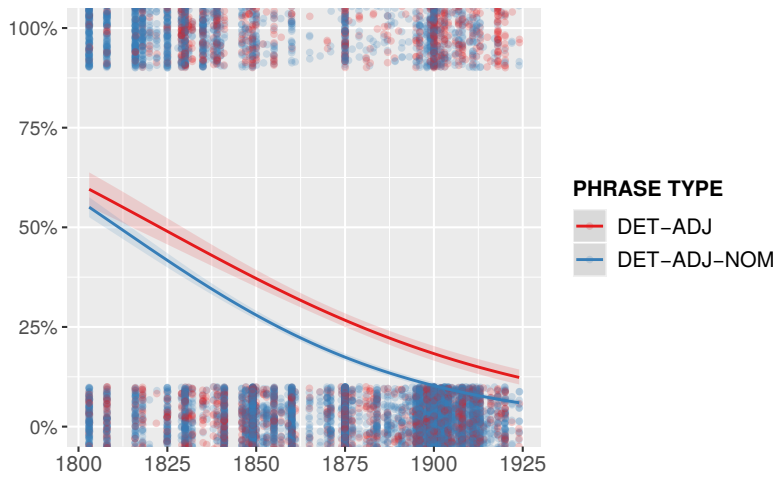


Figure 4.20. Newspapers: Predicted probabilities of fixed effect structural phrase type over time, conditioned on random effects (lower values = lesser use of non-standard *sá* norm).

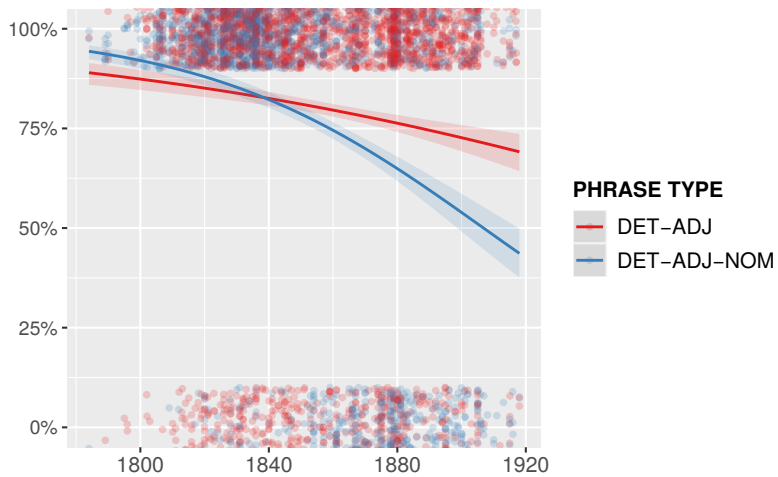


Figure 4.21. Letters: Predicted probabilities of fixed effect structural phrase type over time, conditioned on random effects (lower values = lesser use of non-standard *sá* norm).

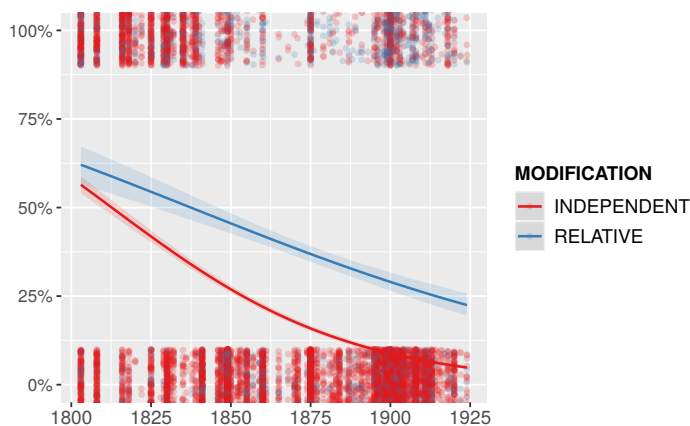


Figure 4.22. Newspapers: Predicted probabilities of fixed effect relative modification over time, conditioned on random effects (higher values = greater use of non-standard *sá* norm).

This effect of DET-ADJ in the private letters is likely to have an explanation in the spoken idiom. As observed by Ólsen (1882), *sá* was the preferred determiner with adjectives in the contemporary spoken language (see Section 4.1.1). It appears that the decline of *sá* in the DET-ADJ pattern found in the newspapers was rather short-lived. It is at least partly undone during the 20th century onwards, although *hinn* appears to remain strong with evaluatives (see Section 4.3.5).

The effect of complex phrases modified by a relative clause is factored in by the language-internal effect RELATIVE modification.¹²² Complex phrases that further define the reference of the DP in a relative clause have a predisposition for *sá* in Modern Icelandic, although the emerging 19th-century Icelandic standard was different from the present-day standard in this regard and aimed towards eliminating *sá* in this context as well (cf. Section 4.1.2). As a result, we expect dependent/complex as well as independent/simplex phrases to participate in the downward trend of *sá*. This is indeed what we find, although DPs that occur with relatives participate in this trend to a lesser extent than independent DPs, cf. Figures 4.22 and 4.23 for newspapers and private letters, respectively. As is to be expected based on the assumption that the private letters are closer to the spoken modality, the contrast between complex/dependent and simplex/independent DPs is much smaller in the letters than in the newspapers. Conversely, what this furthermore shows is that the broader anti-*sá* stance of the Reykjavík Grammar School norm is clearly being observed, especially in the newspapers.

Let us now zoom in on the private letters and consider **social aspects** as a potential factor. Figure 4.24 shows the overall distribution of SEX/GENDER over time. Males and females start out as nearly identical at the beginning of the period, both exhibiting a very

¹²²As there was no reliable way to automatically annotate the data based on relative clause modification, the results were filtered and annotated manually based on whether or not the relativisers *sem/er* occurred in word position 1-6, immediately following the determiner. This procedure resulted in 4062 potential hits, 2675 of which actually involved a DP with a relative clause.

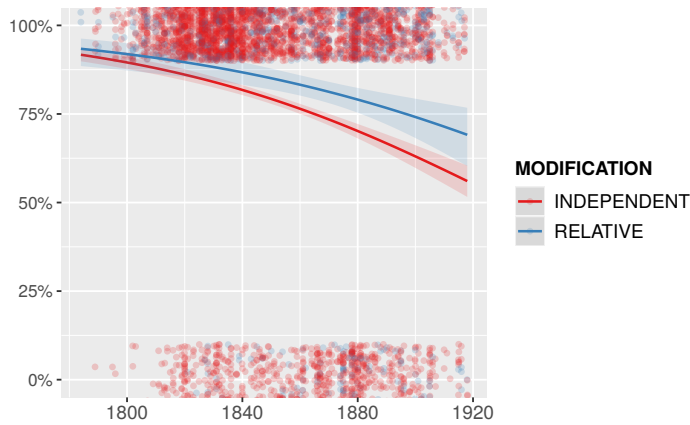


Figure 4.23. Letters: Predicted probabilities of fixed effect relative modification over time, conditioned on random effects (higher values = greater use of non-standard *sá* norm).

high proportion of *sá*. The genders start diverging around the middle of the century, when the male scribes greatly exceed the females in the adoption of the prescribed definite marker *hinn*. However, a more subtle decrease over time is also discernible among the female scribes. This might be taken to suggest standardisation effects, perhaps indirectly through exposure to the written standard where the *hinn* norm was quite successfully adopted (cf. Section 4.3.2 above).

Df	AIC	BIC	logLik	deviance	Chisq	Chi Df	Pr(>Chisq)
Model ^B (♂)	3544.7	3608.6	-1762.3	3524.7			
Model ^A (♀)	3538.2	3608.4	-1758.1	3516.2	8.535	1	0.003484**

Table 4.18. Model comparison with and without SEX/GENDER (*anova*).

Table 4.18 shows the effect of including SEX/GENDER in the model as an alternative means of quantifying the effect. The AIC value is considerably lower for the model with SEX/GENDER, suggesting that the inclusion of this factor is to be preferred over a model without this factor. The BIC value is slightly higher, which is a penalty for it being more complex. The difference between the models is statistically significant (p-value=0.003). As mentioned above, the contrast between male and female speakers importantly is only statistically significant in the subset of speakers whose social rank is known.¹²³ Since rank is quite often unknown, the number of individuals included in the model drops from 249 to 155, but most of these individuals did not produce a large number of data points (*n* only drops from 4620 to 4393). Even so, the two trajectories are basically the same in the non-significant 249-speaker superset as in the significant

¹²³Speakers with unknown ranks were automatically left out of the model in Chapter 3.4.2 since that study included RANK as a fixed effect and mixed-effects models do not work with missing data.

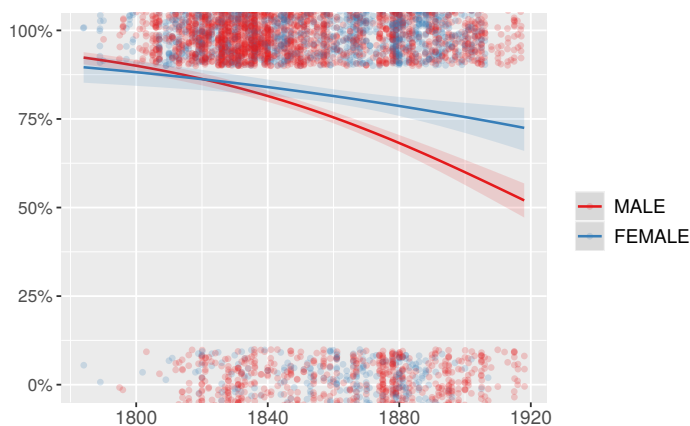


Figure 4.24. Predicted probabilities of *sá* across gender over time, conditioned on random effects (lower values = greater use of standard *hinn* norm).

155-speaker subset.

For comparison, the normalised frequency of the **unambiguous** DET-ADJ-NOM pattern with *hinn* is shown in Figure 4.25, split for gender over four time periods. Notice that the increase in *hinn* in the latter half of the 19th century observed above is clearly visible. The effect is relatively constant for males but dwindles sharply for the females in the subsequent periods. This state of affairs at least suggests a partially successful adoption of the norm, while at the same time it is clear that *hinn* is not a frequent feature, barely reaching six instances per 10,000 words during its peak in the latter half of the 19th century. In raw numbers, this amounts to 236 and 124 instances for males and females, respectively, in periods 2 and 3 combined and a grand total of mere 480 instances of this unambiguous pattern over all four periods.

While the model failed to implement RANK as a fixed effect, let us try to approach the issues at hand by means of traditional descriptive statistics using cross-tabulation, aggregating the data by SEX/GENDER, STATUS and TIME PERIOD (three periods) in two ways; on the one hand by considering the whole data set, on the other an identical set but with random sampling of 30 data points per individual if the data exceeded that number. This approach ignores any potential effect at the level of the INDIVIDUAL, but serves to minimise the individual bias by setting a maximum amount of 30 data points for each individual. Even regardless of this procedure, data scarcity is a real problem in a number of cells in Table 4.19. Among the males, this is particularly troublesome in the first period of peasants/labourers and other professions, and the third period of officials/lettered. As for the females, especially the third period of peasants/labourers is rather thin in terms of data. These issues aside, let us now go through the major contrasts, first for male speakers, then for females.¹²⁴

¹²⁴All findings below are significant at a $p < 0.05$ level based on a chi-square test. For further details, see: <https://github.com/heimirfreyr/ICENCC>.

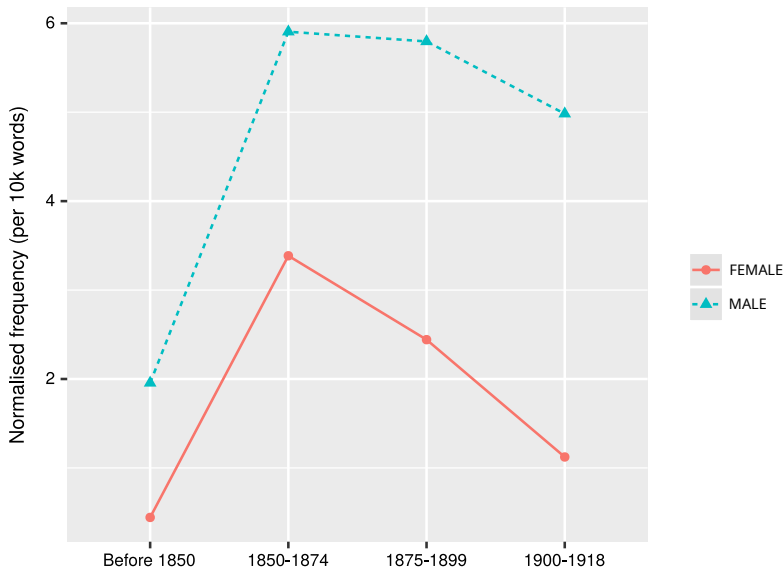


Figure 4.25. Normalised frequency of *hinn* in the unambiguous DET-ADJ-NOM pattern across SEX/GENDER over time.

	1784-1850		1851-1885		1886-1918	
MALES						
Officials/lettered	85.3%	(1398/1638)	66.0%	(101/153)	88.9%	(16/18)
Other professions	–	–	83.7%	(36/43)	52.6%	(90/171)
Peasants/labourers	88.9%	(8/9)	63.6%	(386/607)	73.1%	(209/286)
FEMALES (ranks mainly by association)						
Officials/lettered	89.3%	(484/542)	69.1%	(248/359)	84.9%	(101/119)
Other professions	83.3%	(20/24)	74.8%	(92/123)	88.7%	(110/124)
Peasants/labourers	88.4%	(38/43)	78.5%	(106/135)	78.9%	(15/19)
Unknown	70.0%	(7/10)	68.3%	(43/63)	63.3%	(19/30)
MALES						
Officials/lettered	81.0%	(201/248)	50.7%	(37/73)	87.5%	(14/16)
Other professions	–	–	85.4%	(35/41)	53.1%	(69/130)
Peasants/labourers	80.0%	(4/5)	68.1%	(271/398)	69.5%	(107/154)
FEMALES						
Officials/lettered	90.1%	(173/192)	68.8%	(77/112)	88.7%	(63/71)
Other professions	83.3%	(20/24)	72.2%	(39/54)	90.0%	(27/30)
Peasants/labourers	84.4%	(27/32)	75.3%	(61/81)	78.9%	(15/19)

Table 4.19. Upper: Percentage *sá* across social status and sex/gender in private letters over time using the whole data set. Lower: Alternative data set drawing a random sample of max. 30 observations per individual.

The uptake of the standard norm indeed appears to be most strongly present in the higher echelons, the group of officials/lettered speakers. We see this potential effect for males as well as (associated) females. Beginning with male speakers in the upper ranks, we may conclude that the temporal decrease in the use of *sá* from 85.3% in Period 1 (1784-1850) to 66.0% in Period 2 (1851-1885) is statistically significant, whereas the subsequent increase in Period 3 (1886-1918), unsurprisingly, is not. Similarly for male speakers in the other professions group, the decrease from 83.7% in Period 2 to 52.6% in Period 3 is statistically significant. As for peasants/labourers, Period 1 is too thin, but the increase from 63.6% in Period 2 to 73.1% in Period 3 is statistically significant.

Moreover, the contrasts between officials/lettered and peasants/labourers are significant when Periods 2 and 3 are taken together, but only for the smaller data set (lower part of Table 4.19). The contrast between other professions and peasants/labourers are statistically significant in both Periods 2 and 3. However, the group other professions is statistically non-distinct from officials/lettered (Periods 2 and 3 together). This means that there is an uptake effect present in the upper ranks which interestingly is not only absent for peasants/labourers, but goes in the opposite direction. This distribution is thus mostly in line with standardisation being ‘partly successful’ and only very modestly so.

Moving on to female speakers in the upper ranks, the temporal decrease from 89.3% in Period 1 to 69.1% in Period 2 is statistically significant. However, so is the temporal *increase* to 84.9% in Period 3. I have no other explanation to offer for these fluctuations other than that they be regarded as spurious correlations between status and sex/gender. The female speakers, in other words, fail to show a consistent effect in terms of the uptake of the standard norm.¹²⁵ For females in the other professions group, there appears to be a modest decrease from Period 1 to Period 2 but this is not significant. In contrast, again, the *increase* from 74.8% in Period 2 to 88.7% in Period 3 is significant, but only in the larger data set. As for peasants/labourers, Periods 1-3 are statistically non-distinct. The contrast between officials/lettered and peasants/labourers is statistically significant only in Period 2, the larger data set. However, neither the contrast between other professions and peasants/labourers (Period 1 vs. 2; Period 2 vs. 3) nor between other professions and officials/lettered is significant.

To summarise, similar to the males, a slight uptake effect appears to be present for the females in the higher echelons from Period 1 to Period 2, whereas such an effect is absent for females belonging to other professions. Moreover, while the females appear to show a similar effect for peasants/labourers as for the upper ranks, only less so, this did not turn out to be significant.

These diachronic contrasts between males and females also appear to be partly reflected in generational differences, cf. Table 4.20. Males born during the third quarter of the 19th century have a much greater predisposition for *hinn* than previous generations. While females appear to participate in the same trend a generation earlier, the proportion of *hinn* in the third quarter of the century is much lower than in their male counterparts. Differences in social status are likely to be at least a part of the reason behind these

¹²⁵ Another potential explanation is that a part of the speakers in Period 3 are Icelandic migrants in North-America and there may have been a lesser prescriptive force present in those communities than in Iceland (see e.g. Rögnvaldsson 2014). The fact that the variable PLACE WRITTEN did not have a significant effect may perhaps suggest that this is not the case with these data, at least. Most of these migrants will be from the lower ranks, but note that the same effect indeed appears in other professions and peasants/labourers as well. I leave this aspect to future research.

BIRTH YEAR	MALES			FEMALES			Σ
	HINN	SÁ	<i>n</i>	HINN	SÁ	<i>n</i>	
<1825	16.7%	83.8%	1838	12.3 %	87.7%	796	85.0%
1825-1849	17.9%	82.1%	480	37.7%	62.3%	337	73.9%
1850-1874	46.7%	53.3%	597	17.0%	83.0%	364	64.5%
>1874	44.1%	55.9%	34	–	–	–	55.9%

Table 4.20. Letters: Generational differences among males and females in the use of *hinn* vs. *sá*, per 25 years, where birth year is known (whole data set).

contrasts.

The discussion in this section about social background has inconveniently ignored the individual level for the most part. In the next section, we will move on to discuss the individual level in greater detail.

4.3.3 Individual variation

Table 4.21 shows the proportion of *hinn* vs. *sá* per individual title with at least 20 examples of the variable. Similar to the verb-adverb placement data in Chapter 3.4.1.3, the (later) non-standard variant *sá* is used more frequently than the (later) standard *hinn* determiner during Period I (1800-1824). During Period II (1825-1849), this appears to change rather abruptly in titles published after 1840, as witnessed by *Norðurfari* (1849), *Ný félagsrit* (1841) and *Þjóðólfur* (1848), where the proportion of *hinn* is in the 94-96% range, whereas *Reykjavíkurljósurinn* (1846-49) is lagging behind at 69%. *Ný félagsrit* and *Þjóðólfur* are the only newspaper titles during this period that have a low proportion of *sá* as well as the Adv-Vfin variant (cf. Table 3.5). In the other titles, the frequency of the (later?) non-standard variant is relatively high in both cases. As discussed in Section 3.4.1.3, these other titles were not edited by editors who were devoted purists in any sense, which begs the question why *Norðurfari*, in contrast, patterns so differently with regard *sá* than Adv-Vfin.

During Periods III-IV (1850-1874, 1875-1899), most of the newspapers exhibit a high proportion of *hinn*, around or above 80%, *Þjóðólfur* (1852-1860) now being an exception, but this need not come as a surprise as there were editorial changes during this period (cf. Section 3.4.1.3). In newspapers during Period V (1900-1924), it appears that *sá* is somewhat on the rise again. As we will see in Section 4.3.5, this is a trend that continues throughout the 20th century, at least in a subset of the environments under study.

Let us now return to the private letter corpus and consider the individual level in more detail, focusing on the first two periods. Before 1850, 98% of the cases (64 instances) belong to the officials/lettered group, the 2% remaining coming from the *unknown* group (a singleton). A fraction of these occur in letters by female writers (14%, 9 instances), all but one of which (=unknown) belong to the officials/lettered group. Note that peasants/labourers are severely underrepresented in this period. The distribution is shown in Table 4.22, excluding individuals with less than three examples of the unambiguous nominal pattern.

1800-1824	Time period	Place of publication	HINN	SÁ	<i>n</i>
Íslenzk sagnablöð	(1816)	DK	34.1%	65.9%	290
Klausturpósturinn	(1818-1822)	IS (W)	9.7%	90.3%	144
Margvíslegt gaman og alvara	(1818)	IS (W)	38.8%	61.2%	121
Minnisverð tíðindi	(1803-1808)	IS (W)	47.0%	53.0%	642
1825-1849					
Ármann á Alþingi	(1829-1832)	DK	53.2%	46.8%	327
Búnaðarrit Suðuramtsins ...	(1839)	IS (SW)	28.8%	71.2%	132
Íslenzk sagnablöð	(1825)	DK	49.3%	50.7%	217
Klausturpósturinn	(1825)	IS (SW)	22.6%	77.4%	31
Norðurfari	(1849)	DK	94.1%	5.9%	935
Ný félagsrit	(1841)	DK	95.0%	5.0%	261
Reykjavíkurbósturinn	(1846-1849)	IS (SW)	69.1%	30.9%	269
Skírnir	(1828-1840)	DK	45.1%	54.9%	335
Sunnanpósturinn	(1835-1838)	IS (SW)	19.1%	80.9%	267
Þjóðólfur	(1848)	IS (SW)	96.2%	3.8%	26
1850-1874					
Ársritið Gestur Vestfirðingur	(1850-1855)	IS (SW)/DK	79.8%	20.2%	198
Ingólfur	(1853-1855)	IS (SW)	93.8%	6.3%	48
Ísafold	(1874)	IS (SW)	89.7%	10.3%	29
Íslendingur	(1861)	IS (SW)	97.1%	2.9%	139
Norðanfari	(1865-1871)	IS (NE)	93.6%	6.4%	173
Norðri	(1853-1859)	IS (NE)	88.3%	11.7%	128
Ný sumargjöf	(1860)	DK	94.1%	5.9%	185
Þjóðólfur	(1852-1860)	IS (SW)	56.4%	43.6%	117
1875-1899					
Bjarki	(1896)	IS (E)	76.5%	23.5%	34
Dagskrá	(1896)	IS (SW)	88.8%	11.3%	80
Eimreiðin	(1895-1899)	DK	87.3%	12.7%	1106
Fjallkonan	(1884-1889)	IS (SW)	90.7%	9.3%	322
Fréttir frá Íslandi	(1875)	IS (SW)	99.5%	0.5%	185
Ísafold	(1875-1899)	IS (SW)	87.7%	12.3%	448
Íslendingur	(1875)	IS (SW)	80.7%	19.3%	145
Norðanfari	(1875)	IS (NE)	82.8%	17.2%	186
Norðlingur	(1875-1881)	IS (NE)	83.7%	16.3%	252
Skuld	(1877)	IS (E)	92.0%	8.0%	25
Suðri	(1883)	IS (SW)	84.6%	15.4%	39
Þjóðólfur	(1875)	IS (SW)	89.6%	10.4%	182
1900-1924					
Austri	(1900)	IS (E)	81.5%	18.5%	216
Austurland	(1907)	IS (E)	68.5%	31.5%	73
Bjarki	(1900)	IS (E)	77.0%	23.0%	61
Eimreiðin	(1900-1920)	DK/IS (SW)	87.4%	12.6%	1489
Fjallkonan	(1900-1911)	IS (SW)	85.2%	14.8%	547
Framsókn	(1900)	IS (SW)	87.0%	13.0%	69
Ísafold	(1900-1924)	IS (SW)	78.7%	21.3%	389
Kvennablaðið	(1900)	IS (SW)	88.0%	12.0%	50
Norðurland	(1902)	IS (NE)	74.5%	25.5%	51
Óðinn	(1908-1919)	IS (SW)	87.0%	13.0%	54
Reykjavík	(1900)	IS (SW)	72.7%	27.3%	22
Reykvíkingur	(1900)	IS (SW)	62.1%	37.9%	66
Stefnir	(1900)	IS (NE)	83.1%	16.9%	71
Þjóðólfur	(1900-1920)	IS (SW/S)	83.5%	16.5%	480
Þjóðviljinn	(1900)	IS (Wf)	90.5%	9.5%	105

Table 4.21. Proportion *hinn/sá* in newspapers and periodicals (1800-1924) with a minimum of 20 examples of either variant.

Individual	Gender	Status	D-A	D-A-N	Norm (DAN)	WC
<i>Period I: Before 1850</i>						
BalEin	Male	Officials/lettered	0	5	17.21	2906
BryPet	Male	Officials/lettered	5	4	3.21	12450
ArnHel	Male	Officials/lettered	65	12	1.30	92650
BjaTho	Male	Officials/lettered	79	21	1.65	127102
FinMag	Male	Officials/lettered	1	4	18.30	2186
GriTho	Male	Officials/lettered	6	4	1.64	24396
HogEin	Male	Officials/lettered	2	3	24.59	1220
IngJon	Female	Officials/lettered	27	6	0.88	68397
<i>Period II: 1850-1874</i>						
BjoOla	Male	Peasants/labourers	5	20	38.56	5187
EinAnd	Male	Peasants/labourers	1	10	46.95	2130
EinAsm	Male	Peasants/labourers	4	5	13.36	3743
JakJon	Female	Officials/lettered	23	53	9.56	55443
JohGud	Male	Peasants/labourers	0	4	6.65	6015
JohHal	Male	Peasants/labourers	8	5	2.20	22769
JonDan	Male	Peasants/labourers	0	3	43.42	691
KleBjo	Male	Peasants/labourers	9	9	5.10	17784
SigEir	Male	Peasants/labourers	0	7	27.80	2518
SigLyn	Male	Other professions	0	4	38.06	1051
SigPal	Female	Officials/lettered	4	3	0.30	100271

Table 4.22. Raw distribution of *hinn* in the adjectival and nominal pattern in Period I and II (only includes letter-writers with at least 3 occurrences of the nominal pattern).

However, *hinn* is not entirely absent from the other groups if we move beyond the nominal pattern, occurring in two female speakers belonging to the ranks of peasants (5 instances) and other professions (3 instances), respectively. These all involve clear-cut non-restrictive uses of superlatives (*hið besta/æskilegasta* ‘the best/most desirable’) or salient referents such as *hinn dauði/framliðni* ‘the deceased’. The earliest attestations of each of these speakers is shown below:

- (109) [...] *faðmaði mig að sér; bað mér alls hins bezta*
 hugged me to REFL wished me all the best
 ‘... hugged me and wished me all the best.’ (GudMag-1819-00-00.txt)
- (110) *Með uppstigningardegi batnaði tíðin og hefur síðan verið hin æskilegasta*
 with ascension day improved time-the and has since been the preferable_{SUP}
 ‘As of Ascension Day, conditions improved and have since been most preferable.’
 (SigOrm-1826-06-25.xml)

The letter-writer Guðríður Magnúsdóttir (GudMag) is associated with the peasants class and Sigríður Örum (SigOrm) with merchants, suggesting that *hinn* as a definite article was at least in limited circulation beyond the typical lettered spheres in the early-19th century. It is not clear whether this should be taken to mean that *hinn* did have a precedent in the vernacular after all (somewhat contradicting Ottosson 2003) or whether we are witnessing early effects of the emerging standard.

In period 2, more data is available from male and female peasants/labourers, but in contrast to the males, this is hardly visible at all in the use of *hinn* forms in the

female writers. Only two examples occur in the adjectival pattern in two distinct female speakers from this group, therefore not shown in Table 4.22, both towards the end of the period: *jeg vona als hins besta* ‘I hope all the best’ (1871) and *hið fyrsta* ‘the soonest’ (1873), thus similar to the examples above from period 1. HINN-ADJ-NOM is more regularly attested in males belonging to peasants/labourers during period 2, such as BjoOla, EinAnd and KleBjo, writing between 1857-1867, 1854-1858 and 1848-1862, respectively. For BjoOla, HINN-ADJ-NOM is in the majority, occurring with evaluatives adjectives and mainly in combination with abstract entities with adjectives like *gamall* ‘old’, *hjátrúarfullur* ‘superstitious’, *lærður* ‘learned’, *menntaður* ‘educated’ and *vísindalegur* ‘scientific’, with only three examples of SÁ-ADJ-NOM (two of which are dates). EinAnd is similar to BjoOla, *hinn* mainly used with evaluatives and abstract entities, *sá* mainly with dates, but *hinn* appears to alternate more frequently with *sá* KleBjo, as further reveal by the much lower normalised frequency.

Interestingly, there is a lot of variation at close range (super tokens) in the 19th-century material, occurring in letters as well as newspapers, the writings of female as well as male scribes.

(111) *Super tokens: SÁ vs. HINN*

- a. *svo verða þeir sýðustu sem hinir firstu*
so become the_{SÁ} last as the_{HINN} first
‘So the last will be like the first.’ (SigHal-1859-05-06.xml)
- b. *Börnin eru nú öll að læra; þau eldri 2 ögn í dönsku og gengur ekki illa, hinn yngri að lesa og skrifa.*
children_{DEF} are now all to learn the_{SÁ} older 2 a little in Danish and goes
not bad the_{HINN} younger to read and write
‘The children are all studying; the older two a little Danish and they’re not doing bad, the younger two to read and to write.’ (JakJon-1853-01-04.xml)

The female letter-writers Sigríður Pálsdóttir (SigPal) and Jakobína Jónsdóttir (JakJon) from the higher echelons are of particular interest, both writing over a long period. While the size of the latter subcorpus is nearly half that of the former,¹²⁶ the latter far outnumbers the former in the number of *hinn* attestations. Normalising the frequency to 10,000 words, the frequency of *hinn* in the nominal pattern in JakJon’s letters is about thirty times higher than in SigPal’s letters. The only nominal examples SigPal produces occur in a single letter from 1866 and all in identical phrases referring to the high governours (*hin háu stiptsyfirvöld*, i.e. the local governour and the bishop).¹²⁷ JakJon’s use of *hinn* in the nominal pattern appears to be evenly distributed over the period and the same appears to apply to her modest use of *sá* as a definite article. Against 53

¹²⁶SigPal produces 242 letters written between 1819-1871 (about 100,000 words) and Jakobína Jónsdóttir 82 letters between 1852-1889 (about 55,000 words).

¹²⁷SigPal uses *hinn* nine times in the adjectival pattern, similar to the above: *hið vanalega* ‘usually (lit. the usual)’ (1826), *hið/hin æskilegasta* ‘the most preferable (weather, seasons, etc.)’ (1823, 1830, 1832, 1869), *eg vona líka als hins besta* ‘I also hope all the best’ (1844), *bið þig hins sama* ‘ask you the same’ (1857), *breist til hins betra/verra* ‘changed for the better/worse’ (both 1864, two distinct letters, the former written *til ens betra*, using an archaic form of (*h*)*inn*). These alternate with non-standard *sá* forms more frequently used, e.g. *sú æskilegasta* (1841, 1853), *sá/sú/það besta* (1840 (2x), 1842, 1844, 1845 (4x), 1847, 1850 (2x), 1852, 1854, 1867 (2x), 1869, 1870), *það vanalega* (1863, 1870) and *sá/sú/það sami/sama* (1844, 1845 (2x), 1867 (2x), 1870 (2x), one undated letter), suggesting that *hinn* is a relatively marginal variant for SigPal.

instances of HINN-ADJ-NOM we have 19 instances of SÁ-ADJ-NOM, roughly four of which appear to involve demonstrative *sá* and additional five instances are dates which rarely occur with *hinn* in the private letter corpus. The definite article *hinn* is thus clearly a majority pattern in JakJon's letters, an indication that language standardisation was effective beyond newspaper texts. This is also suggested by the range of male speakers in the peasants/labourers group during this period.

4.3.4 Student essays

A basic overview of the four variables that were available for inclusion in the statistical model as fixed effects is provided in Table 4.23. Based on the contemporary description of Ólsen (1882), the low frequency of *sá* in this material is very striking. It thus seems that the students are actually rather successful in suppressing this non-standard feature in their essays. Note in this regard the following potentially interesting patterns in particular. First, the rising in frequency of *sá* in the last period, which is a 'setback' we were also able to observe in the newspaper corpus during roughly the same period. Second, the potential correlation between a high graduation score (high-achieving students) and a lower rate of non-standard *sá* in comparison to a low score (as well as the dropouts). Third, the higher rate of *sá* in the three lower grades (1-3) in comparison to the three higher grades (4-6). Fourth, the language-internal factor of phrase type, indicating a lower rate of *sá* in the adnominal pattern than in the adjectival one. Below, we will try to tease the effect of each of these variables apart by modelling the variation by means of a mixed-effects model.

	HINN	SÁ	%SÁ
TIME PERIOD			
1852-1880 (HKF teaching alone)	213	16	7.00%
1881-1895 (HKF among others)	216	11	4.80%
1896-1906 (post-HKF)	105	30	22.20%
GRADUATION SCORE			
High (with distinction, score 1)	339	25	6.90%
Low (scores 2-3)	132	25	15.90%
Did not graduate	63	7	10.00%
GRADE			
Grades 1-3	123	23	15.80%
Grades 4-6	388	34	8.10 %
PHRASE TYPE			
Nominal (Det-Adj-N)	312	23	6.90%
Adjectival (Det-Adj)	222	34	13.30%

Table 4.23. The use of the free article *hinn/sá* per period, graduation score, grade and phrase type.

In order to measure the weight of the dependent variables in accounting for the data, these were put in a mixed-effects model as fixed effects, with a random effect for the

individual. The fixed effects were as follows:¹²⁸

- (112) a. PHRASE TYPE (two-level factor: *adjectival* (DET-ADJ), *nominal* (DET-ADJ-NOM))
- b. GRADUATION SCORE (three-level factor: *high* (0='honours', 1='1st'), *low* (2-3); *none* ('did not graduate'))
- c. TIME PERIOD (continuous: 1-6 (1852-1864, 1865-1874, 1875-1884, 1885-1894, 1895-1900, 1901-1906))

By the same effect metrics for the inclusion of variables used above, the fixed effect GRADE had to be excluded in that it failed to provide a significant improvement to the model. However, based simply on the cross-tabulation of GRADE in Table 4.23, the contrast between grades 1-3 and grades 4-6 is statistically significant according to a Pearson's chi-squared test and Fischer's exact test for count data.¹²⁹

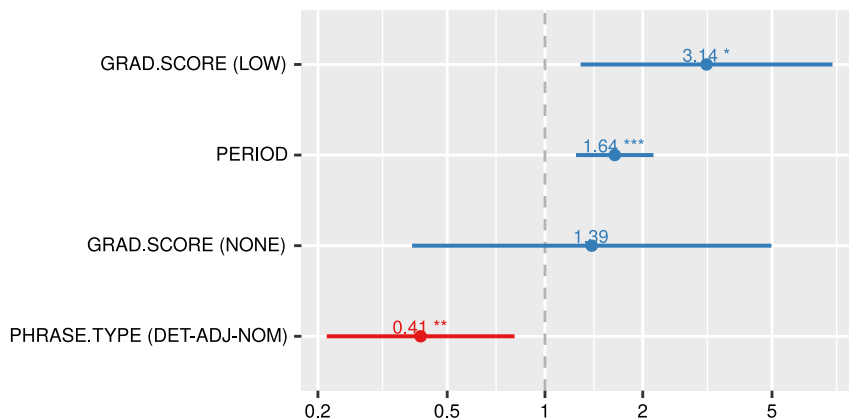


Figure 4.26. Newspapers: A forest plot of estimates with odds ratios of fixed effects. OR>1 (blue) increased odds of *sá*, OR<1 (red) decreased odds of *sá*. (*' $p < 0.05$, ***' $p < 0.01$, ****' $p < 0.001$).

A forest plot of the estimates is provided as a basic overview of the newspaper model in Figure 4.26.¹³⁰ The data set is near-identical to the one used by Viðarsson (2017b) and does not include all the language-internal factors of the models in the previous sections. In comparison to the letter and newspaper data, the types of examples in the essays are rather homogeneous, typically involving evaluatives and abstract entities,

¹²⁸Continuous versions of the variables SCORE and GRADE were also attempted, as well as a 5-period categorisation of the TIME period. While these models did converge, the resulting interaction of a more fine-grained partitioning of the data did not produce a significant effect.

¹²⁹For chi-square (with Yates' continuity correction), $\chi^2 = 6.2902$, $df = 1$, $p = 0.003401$. For Fischer's exact test, $p = 0.01036$, odds ratio = 0.4693313.

¹³⁰Model evaluation: C index of concordance = 0.8921578, Somers' $D_{xy} = 0.7843157$. Classification accuracy: 91.37% of the choices among variants predicted correctly (observations = 591, individuals = 137).

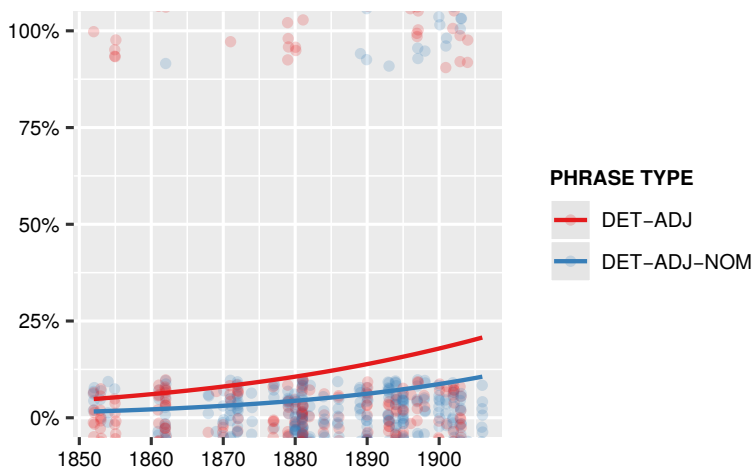


Figure 4.27. Essays: Predicted probabilities of PHRASE TYPE over time, conditioned on random effects.

never referring to dates, for instance. The other language-internal factors are therefore unlikely to play a large role in these data.

Figures 4.27 and 4.28 reveal a similar effect as we saw in Table 4.23 with regard to PHRASE TYPE and GRADUATION SCORE, now conditioned on random effects. The fact that all the educational variables do not exhibit a large contribution to the data is unsurprising given the fact that there is very limited variability within the data set to begin with—the students adopt the standard *hinn* variant *en masse*. Interestingly, the statistical analysis interprets TIME PERIOD as a relevant predictor in the model such that there are gradually *increased* odds of non-standard *sá* towards the end of the period (see also Viðarsson 2017b), which undermines the effect of GRADE.¹³¹ This outcome is not particularly unexpected given that the different grades are not evenly distributed over time (see Section 2.1.3) and the contrast in the last time period(s) is considerably greater than the contrast between the different grades. I hope to explore the effects of the educational variables in more detail in future work.

4.3.5 Beyond the 19th century

In the above we have seen what appears to be a relatively successful uptake of norms in the newspaper corpus from approximately 50-60% (later) non-standard *sá*, relative to *hinn*, in the early 19th century down to mere 8% *sá* with evaluatives in the 1840s. Recall that this effect was substantially reduced in the subsequent decades, although the overall frequency of *sá* continued to be relatively low in comparison both to the

¹³¹The result is the same if the time factor is split into the three periods of Table 4.23.

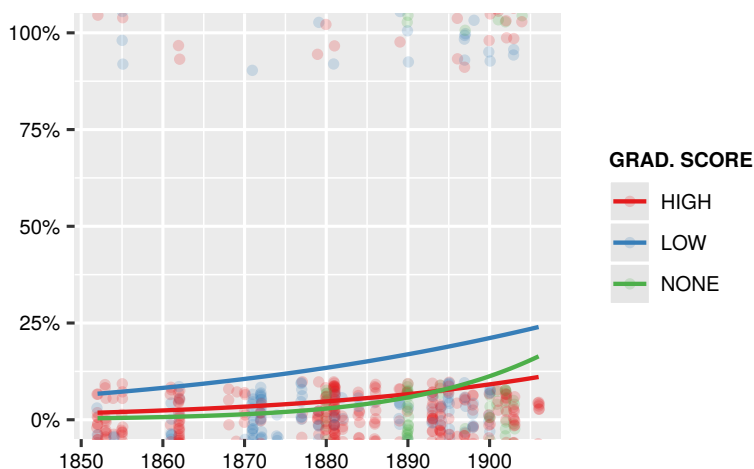


Figure 4.28. Essays: Predicted probabilities of GRADUATION SCORE (right) over time, conditioned on random effects.

first half of the century and to the private letter corpus, cf. Table 4.17. In this section we will follow up on this lead based on the larger *Tímarit.is* corpus, observing further developments in the use of the prenominal definite determiner up until modern times.

Figure 4.29 shows the trajectories of five different phrases over time, viz. *sá/hinn fyrrnefndi* ‘the aforementioned; the former (mentioned)’ and its corresponding antonym *sá/hinn síðarnefndi* ‘the latter (mentioned)’, followed by *sá/hinn svokallaði* ‘the so-called’, *sá/hinn frægi* ‘the famous’ and, finally, *sá/hinn látni* ‘the deceased’, all of which are richly attested in the corpus at least from the 1870s onwards. The first two obviously refer to entities that have already been established in the discourse and are presumably always referential in that they single out a member of a set. This is not so for the last three phrases. While referring to a deceased person as ‘the deceased’ may always involve D-linking of some sort, there is clearly no set membership involved. Such an item thus arguably relates to a salient entity in the given context, making it similar to what the literature describes for *svokallaður* ‘so-called’ and *frægur* ‘famous’.

The systematicity in these data is truly fascinating. Strikingly, there do not appear to be as many trajectories as there are phrases, but rather two different trajectories. Based on this major split, we can refer to these patterns as the FORMER/LATTER group and the SO-CALLED group, respectively. The two phrases in the former group thus pattern more or less alike throughout the whole period, with a gradual increase in the use of *sá* at the expense of *hinn* since the beginning of the 20th century, what appears to be a case of destandardisation. That trend continues at a steady rate until the present day, *hinn* being quite infrequent at approximately 26% for THE FORMER/LATTER combined during the first two decades of the 21st century. This clearly attests to the referential deficiency of *hinn* as a definite determiner, as noted in the literature.

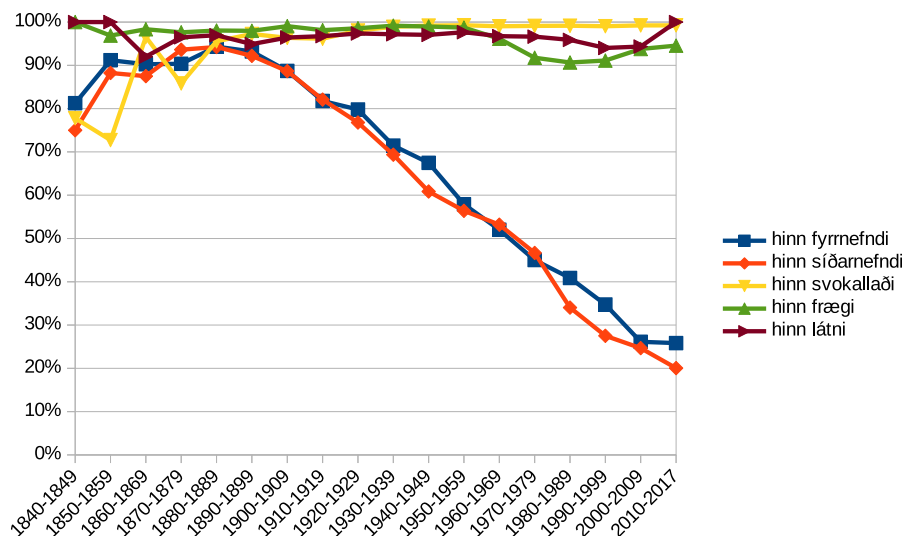


Figure 4.29. Proportion *hinn* in newspapers from 1840 onwards ($n=337,902$). Unanalysed data set containing the full paradigm (masc., fem., neut. in the sg. and pl.) of five distinct phrases, incl. common variants thereof (e.g. *sá/hinn fyr(r) ()nefndi*).

The trajectories of these two strands of data clearly start to diverge only around the turn of the 19th century. The SO-CALLED group, again comprising the three phrases THE SO-CALLED, THE FAMOUS and THE DECEASED, turns out to be very stable over time around the 90-100% mark in stark contrast to the historical decline witnessed with the FORMER/LATTER group. All three of these may probably be subsumed under salient referents, either universally (*so-called, famous*) or at least contextually (*deceased*). Towards the end of the period, at least during the 1970s onwards, there is apparently a slight dip in *hinn* even in the SO-CALLED group, with THE FORMER/LATTER at a historical low decade after decade. However, this trend is not enough to ‘tip over’ the stronger pattern, which continues to strongly favour *hinn* in the newspaper texts.

The journalists and the editors will most, if not all, have received their training at the Grammar School. From a standardisation perspective, the effectiveness of norm implementation in the latter half of the 19th century, as viewed ‘from above’, is abundantly clear from these data. The distribution in the newspapers shows that the standard norm revealed by teacher corrections in the student essays, in which *sá* appears quite simply not to have been tolerated in any form as a definite determiner, was being observed very closely in the newspapers during most of the time the school operated. With regard to the above patterns, note that in the 1880s, for instance, there is hardly any noticeable difference between the five phrases in Figure 4.29, nor is this difference statistical.¹³²

¹³²For the 1880s, X-squared = 6.4751, $df = 4$, p -value = 0.1664. The total n -sizes in 1880 are 264, 156, 202, 305, 142, respectively. In the 1890s, the differences between the patterns do become statistical: X-squared = 32.773, $df = 4$, p -value = 1.33e-06, $n=397, 256, 507, 939, 623$. For comparison, the n -sizes typically run in the thousands for each phrase per decade from the 20th century onwards, as the amount of data grows exponentially.

It is thus mostly in the period after Halldór Kr. Friðriksson retires from the Reykjavík Grammar School (1896),¹³³ after the Grammar School becomes a general high school (*Hinn almenni menntaskóli í Reykjavík*) in 1904 and with the introduction of compulsory education in 1907 that we see the Grammar School norm gradually fall into disuse, although the environment in which it still appears to have held some ground even in the early 19th century, based on evidence from the private letters, still stands rather firm (the SO-CALLED group). The status of *hinn* in the present-day standard norm is actually far from clear, but recent linguistic descriptions indeed suggest that *hinn* is mainly confined to these same (broadly) evaluative and universally salient contexts in colloquial speech (cf. above).

4.4 Discussion

The standardisation of the free definite marker is not quite what one would expect based on the linguistic literature. In being much more on the ‘surface’ than a highly abstract variable like verb-adverb placement, the largely lexical variation in the preceding results section should be more likely to be socially evaluated. For one, it would be much higher on an awareness scale such as Laycock and Mühlhäusler’s (1990) Degree of Interference Hierarchy (cf. Section 1.2.4). However, on the face of it, this is not at all what we find if we take a look ‘from below’. On the whole, the differences between the view from ‘from above’ in the newspapers and student essays and the view ‘from below’ in the private letters speak almost as strongly in favour of normative effects in the newspapers and the essays as against such effects in the private letters, with only rather subtle effects in the latter.

The sharp contrast between the newspapers and the private letters in Figures 4.16 and 4.17, respectively, reveals a fundamental difference between the two corpora with regard to the potential scope of *hinn*. The attempted (emerging) 19th-century standard prescribed *hinn* regardless of phrase type, i.e. in both Det-Adj and Det-Adj-Nom environments, and the newspapers eventually adopt this norm to a large extent. Note that the adjectival Det-Adj pattern will tend to be referential since the nominal element that is elided (or ‘understood’) is typically recoverable from previous discourse. Exceptions to this are cases such as *hið góða/illa* ‘the good/bad (in this world)’ that have more in common with Det-Adj-Nom with a universally salient referent. Mainly these different properties of Det-Adj and Det-Adj-Nom phrases is what is behind the contrast between the two phrases in the private letters (cf. Figure 4.17). The private letters thus pattern more closely with the situation as it is described for Modern Icelandic, *hinn* mainly being found with evaluatives and/or in non-restrictive contexts (alongside *sá*). This also means that the norm was partly descriptive, a selection of *hinn* in a variable context where both *hinn* and *sá* were attested (see Section 4.3.3).

What happens during the 20th century onwards is arguably a matter of destandardisation in what appears to be a striking return to *hinn* as the norm for evaluative/non-

¹³³ Anticipating the results in the next section somewhat, note in this regard that there appears to be a major increase in the use of non-standard *sá* in the student essays in the period after 1895, up from 4.8% between 1881-1895 to 22.2% between 1896-1906 (see Viðarsson 2017b:146, Table 5 and Section 4.3.4).

restrictive contexts or else *sá*. This is a completely different norm from what was taught at the Reykjavík Grammar School and appears to be observed rather closely in the newspapers during the latter half of the 19th century (see Figure 4.29 in particular). Later development towards present times thus reveals an ongoing shift in the standard towards a system that strongly prefers *hinn* in contexts that were already variable during the early 19th century, but otherwise tends to use *sá*.

A more accurate portrayal of the implementation of *hinn* as a norm that allegedly had gradual effects on the spoken language (cf. e.g. Ottosson 2003) is therefore the following. The attempted *hinn*-only standard and the variable casual code actually seem to have met somewhere ‘halfway’, typically resulting in *hinn* in the written language where this was already possible in the spoken language in the first place (=selection among existing variants), but *hinn* was not (or only temporarily) successful in contexts where this was not the case. It is probably no coincidence that the implementation of the Reykjavík Grammar School *hinn*-norm appears to have been more successful while a small, elite group had a monopoly on education and schooling had not yet spread to the entire population (cf. Haugen’s 1987 quote in Section 1.3.4).

A final remark concerns the potential factor of social rank with regard to the adoption of more standard-like language use in the private letters. Despite the fact that rank failed as a significant fixed effect in the mixed-effects model, some effects could be found based on cross-tabulations and chi-square tests. As with verb-adverb placement in the previous chapter, both males and females in the higher echelons exhibit slight effects in the direction of standard norms over time. For peasants/labourers, such effects were shown to be absent, non-significant or even went in the opposite direction, exhibiting a significant increase in *sá*. This serves to further underscore the contrast between the newspapers and the private letters and the relative lack of effect when taking a view ‘from below’.

4.5 Towards an analysis

In order to account for the main patterns in the preceding results sections, I take the Modern Icelandic analyses of Pfaff (2015) and Ingason (2016) as a point of departure. Since there are obviously differences to be found between the 19th-century Icelandic data and the later Modern Icelandic situation, it will be pointed out in the following where changes to these accounts are required to account for the differences.

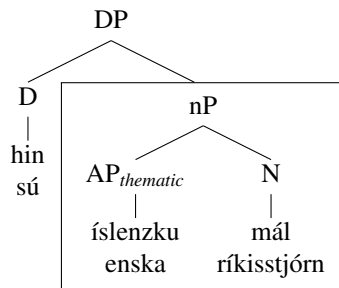
To account for 19th-century Icelandic we are forced to assume that even NP-internal modification could license the free-standing definite determiners *hinn* and *sá*, resulting in N *in situ*. In other words, even potentially restrictive adjectives could serve as interveners, blocking N-to-D much as in the rest of the Scandinavian languages.¹³⁴ There is clear evidence for this in a variety of different contexts.

¹³⁴Note that the suffixed article is more frequently employed in these contexts. In a subset of the ICENCC letter corpus (approximately 1784-1850), I have estimated based on automatic tagging that the suffix outnumbers the free-standing article *hinn* by roughly a factor of ten (250:26) in the configuration ADJ-N_{DEF} and DEF_{HINN}-ADJ-N, respectively. However, this grossly underestimates the frequency of the free-standing article because the private letters typically use *sá*. If we include *sá*, thus risking the inclusion of the demonstrative, the two patterns actually occur at the same rate (250:251).

The following examples clearly involve the NP-internal, thematic reading of the lower NP-internal AP_{thematic} position of (106), occurring with the definite determiners *hinn* and *sá*, respectively:

- (113) a. *Hin íslenzku mál eru lögð til annars hluta Rentukammersins*
 the_{HINN} Icelandic matters are put to another part Treasury-the
 ‘The Icelandic cases are placed in another part of the (Danish) Treasury.’
 (Ný félagsrit 1841-01-01 (1. tbl. 1. árg.).txt)
- b. *Nú hafa yfirvöldin á Íslandi eftir þessum peningum kallað, og sú*
 now have authorities-the on Iceland after these money_{PL} called and the_{SÁ}
enska ríkisstjórn hefir þeim skilað.
 English government has them returned
 ‘The authorities in Iceland have now requested this money and the British
 government has returned it.’
 (Sunnanpósturinn 1835-03-01 (3. tbl. 1. árg.).txt)

(114)



In the newspaper material, the overt expression of D in both NP-internal and NP-external environments is overwhelmingly the standard *hinn*, whereas *sá* is more commonly used in the private letters throughout the 19th century. The following serves as an example of a restrictive NP-internal entity in the private letter corpus with *sá*:

- (115) *fyrir utan prestinn er eg sú eina manneskja á bænum, sem ekki hefi lamazt*
 besides priest_{DEF} amI the sole person at farm_{DEF} REL not has paralyse_{MID}
af sóttinni, sem fólkið hefir verið að liggja í, bæði hér og annarsstaðar
 of illness_{DEF} REL people_{DEF} has been to lie in both here and elsewhere
 ‘Besides the priest, I am the only person at the farm who has not been paralysed
 by the illness that people have been getting, both here and elsewhere.’
 (GudMag-1847-08-15.txt)

The same restrictive meaning is found in (115) as typically with the suffixed article, corresponding directly to *eina manneskjan á bænum* ‘the only person at the farm’ with the definite suffix. However as we have seen, *hinn* does have the strongest foothold in evaluative contexts and with adjectives of nationality/origin (albeit rare in comparison to the evaluatives), reaching a relative frequency of about 40-50% towards the end of the century (see Section 4.3.2). In my view, it is not obvious at all that we are dealing with qualitatively different grammars in these corpora per se: the difference arguably lies in the relatively successful suppression of the *sá* variant in both NP-internal and NP-external contexts in the newspapers.

We have already seen that there is variation between *sá* and *hinn* for the phrases THE FORMER/LATTER, making it clear that the free-standing article could refer to a member of a set. It may also alternate with a proximal demonstrative like *þessi* ‘this’ as well as the definite suffix *-inn*:

- (116) a. *Þessi skóli á að verða í því frábrugðinn hinum fyrra, að [...]*
 this school is to be in it different the other that
 ‘This school is supposed to be different from the former in that ...’
 (ArnHel-VAR-1849-09-29.txt)
- b. *Jeg þakka þjer innilega fyrir 2 hin einkar hlýlegu og góðu brjef frá*
 I thank you dearly for 2 the particularly warm and good letters from
9. okt. og 19. nóv. síðastl. – Fyrra brjefið fjekk jeg 18. nóv. en hið
 9th Oct. and 19th Nov. last former letter-the got I 18th Nov. but the
síðara 13 des.
 latter 13th Dec.
 ‘I thank you dearly for the two particularly warm and good letters from 9th
 October and 19th November I got the former letter on 18th November and
 the latter on 13th December.’ (BalEgg-1905-01-03b.xml)
- (117) *ad því leyti sem latínu snertir, skulu atkvædin ekki samanlögð fyrir hid*
 at that way which Latin touches shall points-the not summed for the
munlega og skriflega prófid, heldur vera hvert fyrir sig, en í hinum
 oral and written exam-the, rather be each for itself, but in the
ødrum greinum skal leggja saman atkvædi fyrir hid munlega og skriflega
 other subjects shall add together points for the oral and written
próf og gjöra af eina einstaka einkunn.
 exam and make of one individual score
 ‘With regard to Latin, the points shall not be added together for the oral and the
 written exam, but rather be reported individually. But in the other subjects, the
 points shall be summed together for the oral and written exam and made into a
 single score.’ (Reykjavíkurbósturinn 1846-12-01 (3. tbl. 1. árg.).txt)

I interpret (117) the former *hid munlega* ... ‘the oral’ phrase as independent from *skriflega prófid* ‘the written exam’, viz. [_{DP} *hid munlega próf*], these could be interpreted as “double definiteness” at the level of NP/DP (see next section). The latter *hid munlega* phrase, however, clearly belongs to one and the same complex NP/DP. It is hard to see how *sá/hinn* could be used in such contexts if it were referentially deficient, suggesting that it must have had a very different status in 19th-century Icelandic than in accounts of present-day Icelandic.

Furthermore, as the examples above seem to indicate, *sá* may clearly alternate with the definite suffix, suggesting that it, too, corresponds to the definite article and not a demonstrative pronoun. This state of affairs suggests, moreover, that there is a (lexical) split in the use of free morpheme determiner such that restrictive uses favour *sá* whereas non-restrictive uses (typically evaluatives) favour *hinn*.

4.5.1 Double definiteness

As mentioned above, the “double definiteness” phenomenon is one aspect which sharply separates Icelandic from some of the other Scandinavian languages (see e.g. Vangsnes 2003:116 for an overview of Icelandic, Faroese, Norwegian and Swedish). To account for these differences the general idea is again one of intervention or blocking such that the adjective blocks N-movement to DP in all the Scandinavian languages except Icelandic. As briefly mentioned above, Vangsnes (2003) suggests that this contrast may have to do with the richness of adjectival inflection in Icelandic, which differs slightly from that of Faroese in terms of complexity. As the presence of double definiteness in 19th-century Icelandic suggests, and the use of the free-standing article more generally, such a relation between morphology and syntax can hardly be upheld.

Just as in Modern Icelandic, definite nouns with an adjective in restrictive contexts canonically undergo DP-movement, pied-piping the adjective. These are exemplified below:

- (118) *Þá sáum við nú fallega hjálminn eða ljósakrónuna.*
 then saw we PRT beautiful shade-the or chandelier-the
 ‘Then we saw the beautiful shade or chandelier.’ (AlfJon-1820-01-01.txt)

Interestingly, however, the double definiteness pattern we find in Faroese, Norwegian and Swedish is attested in 19th-century Icelandic as well, albeit a rare pattern in comparison to the single definiteness of (118)—and the free-standing article canonically selects an indefinite N as we have seen in the previous sections. About 125 doubly definite examples occurred in all three corpora: the newspapers (103 exx.), private letters (16 exx.), student essays (5 exx.).¹³⁵

The pattern occurs with *hinn* as well as *sá*. Some examples from the private letter corpus are provided in (119) and (120), respectively—notice the super token in (119c), featuring both variants of this pattern at close range:

- (119) a. *En þar eg býst ei við það verði annað en venjulegast er, að menn taki ölvæðir að gjörast og fáí þá hið lakara vínið, finnst mér eins gott að skrifa þér strax.*
 but there I expect not PRT it will.be other than usual_{SUP} is that men take intoxicated to do_{MED} and get then the inferior wine-the feel me as good to write you immediately
 ‘But since I don’t expect it to be any different than usual, that men will become drunk and then get the inferior wine, I think it is just as good that I write you immediately.’ (ArnHel-1822-01-26.txt)
- b. *allir grjetu hástöfum, medan vid leítudumst vid ad kalla hid dýrmæta*
 everybody cried loudly while we attempted_{MED}.PRT to call the precious

¹³⁵It should be borne in mind that the corpus is neither tagged nor parsed. The accurate retrieval of examples of double definiteness thus relies on manual annotation of them as such while in the process of collecting the *sá* vs. *hinn* data. Provided these numbers are more or less accurate, the rate of the double definiteness pattern in the whole data set is thus around 0.7%, assuming (perhaps wrongly) that it has semantics identical to the single definite, free morpheme determiners *sá/hinn* (see discussion below). The occurrence rate of double definiteness in the newspapers is over two times higher than in the private letters, occurring at 0.87% and 0.39%, respectively.

lífid til baka

life-the to back

‘Everybody was crying loudly while we were trying to educe the precious life back.’
(JakJon-1867-11-18.xml)

- c. *og við unnum af öllu hjarta elsku börnonum okkar hinnar eilífu gleðinnar á himnum^{uncl.}, þaug blessuð börninn okkar smá fjölga þar, þó þaug fækki á jarðríki*
and we appreciate of all heart dear children-the our the_{HINN} eternal joy-the on heaven the_{SÁ} blessed children-the our a.little grow.in.number there although they become.fewer on earh.realm
‘And we wholeheartedly appreciate the eternal joy for our dear children in heaven, our beloved children grow in number there little by little although they become fewer and fewer in the land of the living.’
(SteSal-1886-07-06.xml)

- (120) a. *hiýín^{uncl.} og þess vegna gagnid af þeirri málnita skepnuni sem en er heilbrigð, er hier mörg um meiri töpud- og skemd-REL still is healthy is here many about more lost or spoiled*
‘The wool(?) and therefore the use of the milk producing animals that are still healthy is to a large extent lost or spoiled [due to scab on sheep?–HFV]’
(TeiSim-1858-07-11.xml)
- b. *[...] maður vaknar ekki fyrr en i ótima, og sjer þá hvörsu liðið er á timan, og framhjá hlaupin \sú/ dyrmæta tíðin*
one awakes not until than in bad.time and sees then how passed is on time-the and by run the precious time-the
‘One does not wake up until it is too late and sees how what time it is and the precious time has passed by.’
(KleBjo-1858-11-15.xml)
- brit eg svo saman þenna lióta og fá orða miða í mesta hasti hripaðan, og fel ykkur að endíngu þeim algóða Föðurnum í bráð og leingð*
fold I so together this ugly and few worded note in most haste scribbled and entrust you at last the_{SÁ} all-good father-the_{DAT.} in short and long
á samt börnum ykkar
to gether children your
‘Then I fold this ugly and terse note, scribbled in haste, and finally leave you and your children in the care of the all-good father [in heaven–HFV].’
(SigEir-1865-06-18.xml)

It does appear that these examples are typically restrictive. However, they bear a resemblance to the (singly definite) free article in that it typically appears not to be discourse-linked, unlike the bound definite article. I assume that the double definiteness pattern is generated by spelling out both members of the D-chain, the high D-position and the low N-position. In that sense, it may be a sort of hybrid between the free vs. the suffixed article with potential interpretative consequences.

The double definiteness pattern is clearly more frequent in the newspaper corpus (cf. fn. 135), suggesting that it is sensitive to stylistic factors. A few examples are shown in (121) below:

- (121) a. *[kennimerkin] mismuna einúngis að því, að þau eru linari og fara hægar*
 symptoms-the differ only at it that they are weaker and go slower
í hinum lángrandi bólgusóttunum
 in the_{HINN} long.during inflammatory.diseases_{DAT}-the
 ‘The symptoms differ only in that they are less severe and progress slower
 in the chronic inflammatory disease [i.e. as compared to *bráðabólgusóttir*,
 inflammationses acutæ vs. chronicæ—HFV].’
 (Ný félagsrit 1841-01-01 (1. tbl. 1. árg.).txt)
- b. *Í hinum nedri bekkjunum verður þannig íslensk og dönsk túnga*
 in the_{HINN} lower grades-the will.be thus Icelandic and Danish tongue
adal-kénnslugreinirnar
 main-subjects-the
 ‘In the lower grades, Icelandic and Danish thus become the main subjects.’
 (Reykjavíkurbósturinn 1846-10-01 (1. tbl. 1. árg.).txt)
- c. *þeir sem hafa setið við kjötkatlana egyptzku geta eigi fengið að sjá hið*
 those REL have sat by fleshpot-the Egyptian can not get to see the_{HINN}
fyrirheitna landið.
 promised land-the
 ‘Those who have been at the Egyptian fleshpots will not get to see the
 promised land.’
 (Norðurfari 1849-01-01 (2. tbl. 2. árg.).txt)

As (121) clearly attests to, some of these examples are aligned with the bound marker *-inn* pattern, being clearly D-linked. Others feature the non-restrictive, salient referent typical of the free *sá/hinn* pattern, clear in example (121c) targeting a salient referent, (*hið*) *fyrirheitna landið* ‘the promised land’.

As regards the proper analysis of the free vs. suffixed article, one has to wonder whether the double definiteness pattern with evaluatives does not contradict analyses that treat evaluatives as interveners (cf. Ingason 2016). Provided that the free article emerges with evaluatives due to their being interveners, the evaluatives would thus be expected to block the lowering of DEF onto N just as in cases where the evaluatives occur in the more traditional singly definite pattern. Instead, what these data suggest is that (evaluative) adjectives in 19th-century Icelandic were not categorical interveners, neither categorically blocking D-to-N lowering nor necessarily enforcing N-to-D in D-linked contexts.

As with regard to the verb-adverb placement, Older Icelandic thus exhibits considerably more variation with regard to spelling out functional categories than Modern Icelandic. As we will see in the following sections, the same can be said about more complex NP/DP-internal structures, the apparent breaking of symmetry between definiteness and weak (also known as ‘definite’) inflection, and possible structural positions of numerals and possessors within the extended NP.

4.5.2 Complex NP structure

Compared to other related languages, the prenominal position of the Icelandic extended NP is surprisingly limited. For one, the possibility for NP complements to occur in prenominal position in Swedish as in (122) is completely ruled out, although some of

these may be expressed postnominally (cf. Thráinsson 2007:108, fn. 16, with references):

- (122) *en [sin hustru trogen] man* (Swedish)
 a her wife faithful man
 ‘A man, faithful to his wife.’

Interestingly, 19th-century Icelandic appears to have allowed for various types of elements in this prenominal position, typically involving AP modifiers such as temporal or aspectual adverbs, as shown in (123), and could even include nominal complements (124):

- (123) a. *Styrjöld þessi orðsakadiz af þeirri í fyrra umgetnu gripdeild Brasilú á*
 war this caused_{MID} of the_{SÁ} last year mentioned plunder Brazil_{GEN} on
því litla frílandi Montevídeó [...]
 the_{SÁ} little free.state Montevideo
 ‘This war was caused by the plundering mentioned last year in the little self-governing state of Montevideo.’
 (Íslenzk sagnablöð 1825-04-21 (10. tbl. 2. árg.).txt)
- b. *Og hinn blindaði unglíngur, sem hin of þunga byrði keisaradæmisins*
 and the blinded youngster REL the_{HINN} too heavy burden empire_{GEN}
hefur verið lögð á herðar [...]
 has been laid on shoulders
 ‘And the blinded youngster, on whose shoulders the too heavy burden of the empire had been laid, ...’
 (Norðurfari 1849-01-01 (2. tbl. 2. árg.).txt)
- c. *Þeim ber fyrst og fremst að annast um að brjóta ekki stefnuma. Brjóta*
 them is.duty first and foremost to take.care of to break not policy-the break
ekki hin enn óskrifðu, en sjálfögðu lög þjóðfélagsskipunar-réttarins
 not the_{HINN} still unwritten but self-evident laws regime.court-the
sjálfir
 selves
 ‘Their duty is first and foremost to make sure not to violate the policy. Not to violate the still unwritten but self-evident laws of the regime court.’
 (Ísafold 1916-02-12 (11. tbl. 43. árg.).txt)
- (124) a. *draumur einn og rádgáta [...] mót þeim oss öldungis óþecktu undrum,*
 dream one and mystery against the_{SÁ} us completely unknown wonders
sem vor skilníngr, lærdómur, æfilaung ransókn í þessu lífi aldrei fá
 REL our understanding learning lifelong inquiry in this life never get
tilnád.
 to.reach
 ‘A dream and mystery (set) against wonders, completely unknown to us, which our understanding, learning and lifelong inquiry in this live will never reach.’
 (Klausturpósturinn 1820-01-01 (1. tbl. 3. árg.).txt)
- b. *undir eins og eg nú aptur afhendi hið mér fyrtrúaða forseta*
 immediately as I now again hand.over the_{HINN} me entrusted president’s
embætti
 office
 ‘As soon as I will hand over the presidency that I was entrusted with.’

(Skírnir 1840-01-01 (1. tbl. 14. árg.).txt)

- c. *Mótsögnin getur reyndar hvergi verið nema í hinum allt*
 contradiction-the can actually nowhere be except in the_{HINN} everything
misskiljandi hug O. St.
 misunderstanding mind O. St.
 ‘The contradiction is actually nowhere except in the mind of O. St. that misunderstands everything.’ (Norðurfari 1849-01-01 (2. tbl. 2. árg.).txt)

Modern Icelandic does, of course, allow for incorporation forming complex adjectives such as *langþráður* ‘long-wanted’, *áðurnefndur* ‘aforementioned’ and *þáverandi* ‘former (lit. then-being)’, no doubt represented structurally as atomic, but the examples above featuring what appear to be adjuncts seem to be of a different nature.

It appears that complex NPs are more readily found in the more formal text types. They are not very frequent in the private letters, although they do occur there as well:

- (125) a. *en sé um Leid Rángárvalla Sýssla veitt þeim nú dána Sýsslumanni*
 but is_{SUB} simultan. Rángárvellir_{GEN} county given the now diseased county.sheriff
Jóni Gudmundssyni, þá sæki eg enn sem fyrri um hana.
 Jón Gudmundsson then apply I still as before PRT it
 ‘But if the Rángárvellir county is at the same time given to the by now diseased sheriff, Jón Gudmundsson, then I’ll apply for it once and again.’
 (BTh-BÞ-1820-07-23.txt)
- b. *Hjartanlega þakka eg þér, systir mín góð, tvö elskuleg tilskrifin og því*
 cordially thank I you sister my good two dear letters-the and the_{SÁ}
fyrri fylgjandi fingurbjörg mína.
 former accompanying thimble my
 ‘I thank you cordially, my dear sister, for the two lovely letters and my thimble accompanying the former.’ (RagnFinn-1797-09-27.txt)

Same as the above, but with an NP-internal object:

- (126) *Því miður er annað fráfall ein sú helzta mér viðvíkjandi fregn, nefnilega*
 sadly is another death one the_{SÁ} main me concerning news namely
minnar góðu tengdamóður, frú Frydensberg [...]
 my good mother-in-law_{GEN} Mrs Frydensberg
 ‘Sadly, another death is one of the major news concerning me, namely that of my good mother-in-law, Mrs Frydensberg.’ (FinMag-1832-03-31.txt)
- (127) *margt synist mjer hvorfa, og likt sem eyðast með öllu, og streima*
 much seems_{MID} me vanish and similar as erode_{MID} with all and stream
með ofljúgandi^{uncl.} hraða, út i hið mjer hulda haf skuggans og
 with ?rapid? speed out in the_{HINN} me hidden ocean shadow-the_{GEN} and
eyðingarinar
 distruction-the_{GEN}
 ‘Many things seem to me to be disappearing, and as if eroding completely, and streaming with rapid speed out into the ocean of the shadow and destruction that is hidden to me.’ (KleBjo-1857-01-10.xml)

Moreover, we find the following example featuring a complex possessive structure with a proper name, which must be postnominal in Modern Icelandic:¹³⁶

- (128) *að tillagðri minni og Magnúsar míns hjartanlegri ástarheilsan til þín og barna þinna*
 at to.put my_{GEN} and Magnús_{GEN} my_{GEN} cordial_{DAT} love.greeting_{DAT} to you and
 children your
 ‘In addition to a cordial (love) greeting from me and (my) Magnús to you and your children.’ (RagFin-1799-05-06.txt)

Similar examples featuring the proprial article have been shown to occur preminally in Scandinavian dialects, e.g. in Frostvikens Socken, Jämtland, Sweden (cf. Delsing 2003). As arguments, such possessive structures can become quite complex in the 19th-century Icelandic data:

- (129) *Eg er þín þig af hjarta elskandi systir*
 I am your you of heart loving sister
 ‘I am your sister, loving you with all my heart.’ (GudMag-1845-07-17.txt)

A lot of such examples are found where an internal argument (object) occurs NP-internally and/or an adverbial phrase. Most of these are found in the rather formulaic opening and closing sections of the letters. However, these examples also occur elsewhere in more ‘productive’ use, it appears:

- (130) a. *sú hér um sama leyti grasserandi slímfeber.*
 the_{SÁ} here about same time rampant mucosal.fever
 ‘The mucosal fever [bronchitis?–HFV] that was rampant here around the same time.’ (GeiVid-1811-05-09.txt)
- b. *Það er sannarlega guð elskandi manneskja með stilltu og blíðu hjarta.*
 it is truly god loving human.being with calm and gentle heart
 ‘That is truly a God loving human being with a calm and gentle heart.’ (GudMag-1844-06-13.txt)
- c. *Við þetta og annað hér af fljóttandi tjón eyðileggjast innbúar að líkindum*
 with this and other here of floating damage be.ruined_{MID} inhabitants at probabilities
 ‘Because of this [chaos due to volcanic eruption–HFV] and other damage that follows, the inhabitants will probably perish.’ (IngJon-1822-03-08.txt)

Some of these examples are reminiscent of incorporated relatives such as the reduced relatives giving rise to Pfaff’s Pattern IV above. The *-andi* type of examples appears to be akin to English *ing*-nominalisations such as *your loving sister*. Neither Modern Icelandic nor other present-day Scandinavian languages allow for verbal nominalisations of these sorts, unlike Dutch, for instance, where such examples do occur (see Julien 2005:152, fn. 7):¹³⁷

¹³⁶Regarding the strong form of the adjective, see Section 4.5.3.

¹³⁷Dutch nominalisations of this kind go much further than in English, allowing adverbs, larger adverbial phrases and objects within the NP, similar to some of the above 19th-century data (see Elffers et al. 2012:133):

- (131) [*Het [op konijnen jage-n [in de zomer]] is verboden.*]
 the on rabbits hunt_{INF} in the summer is forbidden
 ‘The hunting of rabbits is forbidden during summer.’

While the examples shown in this section are not frequently attested, they do occur in the private letters written not only by people from the higher ranks. This suggests that it would be too simplistic to write them off as belonging to some particularly learned style only, although more research into the extent and limits of these patterns is clearly desirable.

4.5.3 Strong adjectival inflection in NP-internal definite contexts

It is widely claimed that the form of the adjective, whether it is weak or strong, automatically follows from whether or not the noun is definite or indefinite; scholars therefore often refer to the weak pattern as the definite inflection form and the strong pattern as the indefinite inflection (see e.g. Friðriksson 1846:XIV, 1859:9, 1861:31, Smári 1920:57, Vangsnes 2003:134f.). We will see in this section that this is not entirely accurate and that we are better off referring to these two patterns as weak and strong, not as definite and indefinite, if we want to be faithful to the data. Interestingly, these data also contradict Pfaff’s (2014, 2015) analysis of the strong inflection of the adjective as always being NP-external. I take this to suggest that there is no direct link between syntax and morphology.

The strong and weak inflection of adjectives indeed strongly correlates with definiteness, the examples in (132) being typical of Modern as well as 19th-century Icelandic:

- (132) a. *Merkilegustu fréttirnar*
 most.noteworthy_{WK} news-the
 ‘The most noteworthy news.’ (IngJon-1831-04-22.txt)
- b. *hið kæra móðurmál*
 the dear_{WK} mother.tongue
 ‘The dear mother tongue.’ (ArnHel-1817-09-04.txt)
- c. *öflugt vopn*
 powerful_{STR} weapon
 ‘A powerful weapon.’ (GudMag-1845-07-20.txt)

Friðriksson (1861:36f.) explicitly states that the definite (weak) form of the adjective is found with the free as well as the suffixed article. As recently emphasised, and analysed, by Pfaff (2014, 2015), there is a whole range of data where this symmetry apparently breaks down, viz. in what Pfaff refers to as Pattern IV. Examples of this kind feature the

- (1) a. [*de [gisteren [bij mijn buurman] gestolen] voorwerpen*]
 the yesterday at my neighbour stolen items
 ‘The items that were stolen from my neighbour yesterday.’
- b. [*het [mij gegeven] boek*]
 the me given book
 ‘The book that was given to me.’

Dutch nominalisations are commonplace, also in the spoken language (cf. Elffers et al. 2012:130f.). However, the more complex these phrases, the more bookish they gets: “It seems to us that these more elaborate APP clusters occur only in written language” (*op. cit.*, p. 130).

suffixed article *-inn* with a strongly inflected adjective, which would then constitute an oxymoron of sorts: *the strong definite inflection*. Indeed, Pfaff’s strong definite pattern is amply attested in the 19th-century data:

- (133) a. *hvað nærri mér eg var búin að ganga til að bera mig að halda í það mér*
 how close me I was ready to go to C carry me to hold in it me
var mögulegt af fasteigninni vegna blessaðra barnanna.
 was possible of real.estate-the because blessed_{GEN,STR} children-the_{GEN}
 ‘How far I was willing to go in order to try to retain what was possible of
 the real estate because of the dear children.’ (GudMag-1845-07-20.txt)
- b. *Maður sér ekkert nema málaðan steinvegg á tvær hendur með ótal*
 one sees nothing but painted stonewall on two hands with countless
gluggum, og þar fyrir ofan uppi heiðan himininn.
 windows and there for above up.in clear_{STR} sky-the
 ‘One sees nothing except a painted stonewall on either side with countless
 windows and above that up into the clear sky.’ (BalEin-1826-08-14.txt)

Such examples have the same non-restrictive reference as noted for this pattern in Modern Icelandic, as if the adjective is contained within a non-restrictive relative clause. In (133a) *blessaðra barnanna* ‘the dear children’ then corresponds to: *because of the children, who are dear to me*. Similarly, in (133b) the phrase *heiðan himininn* ‘the clear sky’ does not single out a member of set of skies, but instead refers to *the sky, which I perceive to be clear*. However, it is not entirely obvious that all the 19th-century Icelandic examples are necessarily fully non-restrictive. The following example is a case in point:

- (134) *Þó sýnist henni af bræðrunum Brynjólfur blíðlegast og allra saklauslegast*
 though seems her of brothers-the Brynjólfur sweetest and of.all innocent_{SUP}
góðmennið.
 kind.man-the
 ‘But still, she thinks Brynjólfur is the most affectionate and innocent-looking
 kind person of all.’ (GudMag-1845-12-25.txt)

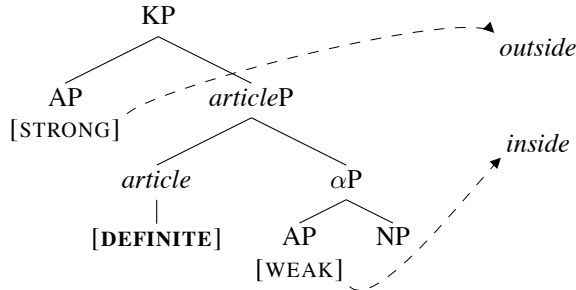
If the adjective in (134) were completely outside the reference of the definite N *góðmennið* in a non-restrictive meaning, similar to the non-restrictive relative representations above, that would actually imply that the other brothers referred to were not. However, this is clearly not the intended reading. In fact, the letter-writer explicitly states that all of Brynjólfur’s siblings are good children, just like him (“*En góð börn segir Sigríður mín, að öll hans systkini séu, og eins sjálfur hann.*”).

Pfaff (2015:54) formulates the generalisation regarding weak vs. strong inflection in such a way that the strong inflection is an elsewhere case that shows up when the weak inflection is not triggered. His generalisation is stated as follows (Pfaff 2015:54f.):

- (135) a. If the adjective is c-commanded by [DEFINITE], the weak inflection is triggered
 (→ [WEAK] must be licensed by a c-commanding feature [DEFINITE])
- b. If weak inflection is not triggered, the adjective occurs strongly inflected
 (→ there is no feature [STR])
 (→ strong inflection is not triggered; elsewhere condition)

A tree representation of the two forms of the adjective is shown in (136) (cf. Pfaff 2015:56):

(136) *Weak inflection if definite feature c-commands ADJ (or else strong)*



As far as Modern Icelandic is concerned, this generalisation accounts at least for the major patterns attested as Pfaff's (2015) thorough treatment of the Icelandic noun phrase shows. This is less obviously the case in 19th-century Icelandic, however. Below we will see that strongly inflected adjectives may occur in what appears to be DP-internal (*articleP*) position, in violation to Pfaff's generalisation in (135). Our evidence will mainly come from DPs involving PossP, NumP and DemP/ArticleP, which will be treated in turn.

In terms of pronominal possessives, Pfaff identifies two positions in the extended DP constellation (cf. Pfaff 2015:160, see also 157ff. for examples):¹³⁸

(137) N(-DEF) >> POSS₁ >> NUM >> ADJ >> N >> POSS₂

Importantly, both positions entail a weak form of the adjective, each being c-commanded by DEF. It is therefore unexpected on this account that POSS in 19th-century Icelandic may occur with ADJ_{STR}:

- (138) a. *með sinni margfaldri blessan.* (RagMag-1830-02-24.txt)
 with his_{REFL} multiple_{STR} blessing
- b. *að skila sinni auðmjúkri þakklætisheilsan til þín*
 to return his_{REFL} humble_{STR} gratitude.greeting to you
 (RagFin-1798-11-20.txt)
- c. *og er það mín einlæg hjartans ósk.* (KriEir-1812-08-25.txt)
 and is it my sincere_{STR} heart-the_{GEN} wish

¹³⁸The top-most (optionally definite) N obviously cannot simultaneously be filled twice by N in a high and low position. It is therefore greyed out in Pfaff's formulation, a detail which is left out here. However, as Pfaff (2015:159, with references) shows, both POSS positions, in contrast, may be filled simultaneously, indeed providing evidence for two positions:

- (1) [... *kapella* ...] en *nunnurnar* sungu [...] í *sínum hluta hennar*
 '... chapel₁ ... but the nuns₂ sang [...] in their₂ part of it₁'
 ⇒ *sinn hluti hennar*
 POSS.REFL part her NB: *kapella* 'chapel' is feminine

Here, *sínum* arguably occupies POSS₁, whereas *hennar* occupies POSS₂.

- d. *með hans þekkjanlegri hönd* (ArnHel-1845-14-01)
with his recognisable_{STR} hand

The weak pattern also occurs in this configuration, as in (139a). Presumably, the weak pattern outnumbers the strong one. I have not studied the relative frequencies of each of these patterns since my study focussed on the free definite determiners *sá/hinn*, typically absent in these examples.

- (139) a. *ad finna Sín heitt elskudu 2 börn* (HilJon-1836-01-12)
to find her_{REFL} hot loved_{WK} two children
'To find her two beloved children.'

Notice the post-adjectival position of the numeral in (139a), suggesting that the representation in (137) above is incomplete as regards these positions.

Second, we also find ADJ_{STR} where ADJ follows $NumP$ in the presence of $NDEF$. The following presents a minimal pair where ADJ_{STR} occurs in two sorts of definite contexts; one the one hand with a $[DEFINITE] N$, on the other with an $[INDEFINITE] N$ but preceded by $POSS$ (thus entailing definiteness):

- (140) a. *Hjartanlega þakka eg þér, systir mín góð, tvö elskuleg tilskrifin*
cordially thank I you sister my good two kind_{STR} letters-the
(RagFin-1797-09-27.txt)
- b. *Tvö yðar kærkomin elskuleg tilskrif* (RagTho-1790-08-23.txt)
two your welcome_{STR} kind_{STR}

Third, ADJ_{STR} may also occur with a definite determiner *sá* as in (141):

- (141) a. *Þrúða var eitthvert það greindarlegasta ársgamalt barn sem jeg hefi þekkt*
Þ. was some the/that most.intelligent_{WK} year.old_{STR} baby that I have known
(AnnGud-1880-02-08.xml)
- b. *Þetta verða nú víst allar þær almennar frjettir sem jeg skrifa þjer í þetta sinn*
this will.be now PRT all the_{SÁ} general_{STR} news REL I write you in this time
(EirJoh-1899-06-22.xml)

This is a less frequent pattern, the canonical agreement being the weak one. It is also not clear to me whether the absence of ADJ_{STR} when preceded by *hinn* in the private letters is merely a consequence of the much higher rate of *sá* in this material. This variation clearly merits further study.

As the data in this section suggest, the hierarchy in (137) requires amendments to properly account for the 19th-century data and it seems unlikely that weak vs. strong agreement on the adjective can be implemented in as straightforward a fashion in terms of c-command as Pfaff's (2015) account suggests.¹³⁹ The obvious way out is *not* to expect syntax and morphology to interact in such a direct way, but I must leave the consequences of such an approach to future research.

¹³⁹ An alternative would be to suggest that there are elements such as at least $POSS$ and DEM that appear higher than DEF , escaping c-command. Obviously, this also fails in (140a) with the definite suffix *-inn* spelling out DEF , and if we treat *sá* in examples like (141) as DEF .

4.5.4 Co-occurrence of determiners

Another aspect that we will briefly explore concerns the structural position of *sá* vs. *hinn* and their relationship to one another. In previous work, I have studied the grammaticalisation (or syntacticisation) of the definite article in the history of Icelandic based on criteria discussed by Van de Velde (2010), Van de Velde and Lamiroy (2017) and Lander and Haegeman (2014). What this research has revealed is that *sá* and *hinn* were not fully syntacticised definite articles, featuring a designated definite determiner slot at the level of DP, during the Old Norse period (see Viðarsson 2017a). 19th-century Icelandic still appears to preserve some of these characteristics. For the present purposes, the most important one is arguably the fact that *sá* and *hinn* may co-occur, typically (though not exclusively) in that order. Magnússon (1984:96) observes for Modern Icelandic that *sá* can co-occur with the definite article *hinn*, provided that *sá* precedes *hinn* and not the other way round. This fact raises the question whether *sá* (at least in that case) occupies DemP as a demonstrative with the definite article in DP (or ArtP).

Indeed, if the two determiners co-occur in the 19th-century data, the pattern is typically *SÁ* » *HINN*:

- (142) a. *En það hið svo kallaða Rithöfunda tal hans?*
 But the_{SÁ} the_{HINN} so called writers_{GEN} index his
 ‘What about his so-called bio-bibliographical dictionary?’
 (JohGud-1865-05-17.xml)
- b. [...] *og svo að lokum það hið þriðja, að mjer hefur aldrei hugkvæmst*
 and so at end the_{SÁ} the_{HINN} third that me has never come.to.mind_{MID}
að drífa mig í að setjast niður við skriftir nú leingi.
 to hurry me in to sit_{MID} down with writings no long
 ‘And then finally the third, that it has never occurred to me to sit down and
 write for a long time.’
 (FriBja-1878-01-30.xml)

However, *HINN/NDEF* » *SÁ* does occur in the data, albeit marginally attested:

- (143) *Ég þakka fyrir ostinn, hann er ágætr. Hinn þann fyrri hefi ég eigi*
 I thank for cheese-the he is excellent the_{HINN} the_{SÁ} former have I not
fengið, enn það gerir ekkert til.
 gotten but that does nothing to
 ‘I thank you for the cheese, it is excellent. The previous one, I haven’t received
 but that doesn’t matter.’
 (FinJon-1879-09-25.xml)
- (144) *Því áður fór eg að líta utan á umslagið, þegar sá hvernig bréfin*
 because before went I to look outside on envelope-the when saw how letters-the
þau innlögðu voru skrifuð.
 the_{SÁ} inlaid were written
 ‘Because before, I started looking at the envelope when I saw how the inlaid
 letters were written.’
 (ArnHel-1820-03-04.txt)

Similarly, demonstratives may precede *hinn* but apparently never precede *sá*:¹⁴⁰

¹⁴⁰Moreover, the demonstrative *þessi* and demonstrative uses of *sá* may co-occur with possessors.

- (145) *jeg óska þjer allra heylla og hamíngu á þessu hinu nybirjaða Sumri*
 I wish you all luck and happiness on this the_{HINN} newly.begun summer
 (AsgFri-1881-04-28.xml)

It also appears POSS can occur very high, in fact preceding DEF:

- (146) a. *Bréf þitt hið sidasta sem öll önur þakka eg hjartanlega*
 letter your the_{HINN} last as all others thank I cordially
 (KleBjo-1850-08-25.xml)

While this pattern is rare, it may explain how an adjective can show up with the strong inflection while dominated by POSS; since DEF does not c-command this POSS position, call it POSS_{HIGH}, weak inflection is not triggered. The presence of POSS_{HIGH}, moreover, may give rise to POSS_{HIGH} » DEF in the absence of N:

- (147) a. *hún átti þó sitt hið æðsta skart í augunum sínum fríðu.*
 she had though her the_{HINN} supreme_{SUP} jewellery in eyes-the her_{REFL} pretty
 (Eimreiðin 1895-01-01 (2. tbl. 1. árg.).txt)
- b. *Rómurinn var viðlíka þróttmikill og hljómfagur eins og þegar hann var upp á sitt hið bezta.* (Ísafold 1893-04-22 (22. tbl. 20. árg.).txt)
 up on his_{REFL} the_{HINN} best

To account for these data it appears we must assume not only POSS₁ and POSS₂, but also a higher POSS₃ position, following DemP:

- (148) DemP >> POSS₃ >> N(-DEF) >> POSS₂ >> NUM >> ADJ >> N >> POSS₁

As we will see in the next section, this hierarchy is still insufficient to account for the whole range of data once we take into account the available positions for NUM as well.

4.5.5 The functional sequence: The peculiar case of NUM

A widely observed restriction on numerals in the extended noun phrase is Greenberg's (1963) Universal 20 stating for the prenominal position the order Dem > Num > A (for discussion, see e.g. Cinque 2005, Harðarson 2017:105-109).¹⁴¹ So Magnússon (1984:95), for instance, observes that the free article may precede numerals and adjectives but the order Num > Dem > A > N is completely out:

- (1) a. *Eg fékk þetta þitt bréf úr biskupshúsinu í Reykjavík* (RagTho-1813-08-09)
 I got this your letter from bishop.house in Reykjavík
- b. *gleymt þessu sínu góða lofordi.* (IngJon-1835-03-05.txt)
 forgotten this his_{REFL} promise
- c. *fyrir þann minn slóðaskap* (IngJon-1845-11-10.txt)
 for that my slothfulness

For further discussion, see Viðarsson (2017a).

¹⁴¹Greenberg (1963:68f.) formulates the universal as follows:

- (1) Universal 20.
 When any or all of the items (demonstrative, numeral, and descriptive adjective) precede the noun,

- (149) a. *Hinar tvær stóru skurðgröfur*
 the two big excavators
 b. **Tvær hinar stóru skurðgröfur*
 two the big excavators

In the previous section, we have seen that the partial functional sequence for Modern Icelandic suggested by Pfaff (2015) had to be extended with a higher POSS position scoping over DEF. In this section we will similarly see that NUM may out-scope DEF as well.

As mentioned above, NUM is argued to be below DEF resulting in either (150) or, more typically, (151):

- (150) *en fullviss skuluð þjer um, að jeg mun ekki gleyma hinum þremur aðal*
 but certain shall you about that I will not forget the_{HINN} three main
regl um er þjer settuð mjer að fara eptir (AndAnd-1870-12-31.xml)
 rules REL you set me to go after
- (151) *Í beztu rúmunum tveimur kostar næturgisting 1 kr. á mann.*
 in best beds-the two costs overnight.accommodation 1 kr. per person
 (Ísafold 1876-11-27 (26. tbl. 3. árg.).txt)

However, besides these patterns, NUM may also precede definite noun phrases, indeed suggesting that the (partial/simplified) structural hierarchy is: NUM₂ >> DEF >> POSS₂ >> NUM₁ >> ADJ >> N >> POSS₂.¹⁴²

Due to a potentially confounding factor of little/big partitives (see below) in examples seemingly adhering to the same pattern, this partial structure is exemplified below on the basis of more than just a handful of examples to demonstrate that these arguably do not belong to partitives but truly to cardinals:

- (152) a. *Jú! víst voru Sr. J. Þorláksson, Gröndal og S. Petursson sannarleg*
 yes certainly were Rev. J. Þorláksson Gröndal and S. Petursson truly
Skáld, hvör uppá sína Vísu, og um tvo þá síðarnefndu var það
 poets each up.on their_{REFL} way and about two the_{SÁ} last.mentioned was it
því addáanlegra, sem þeir voru scientiv fice uppaldir í þeim Danska
 that.much admirable that they were scientiv fice raised in the_{SÁ} Danish
Vandperiode
 water.period
 ‘For sure! Rev. J. Þorláksson, Gröndal and S. Petursson were certainly poets,
 each in his own way, and with regard to the latter two, that is all the more
 admirable in that they were *scientiv fice* raised during the Danish water
 period.’ (BTh-1830-12-05.txt)

they are always found in that order. If they follow, the order is either the same or its exact opposite.

According to Cinque (2005), the first part of the statement still remains (virtually) unchallenged. On Cinque’s approach, Greenberg’s Universal 20 is derived by Merge projecting the functional sequence [Dem [Num [A [N]]]] and that basic order surfaces if there is no subsequent Merge. The typologically most common patterns arise simply by merging N with the elements higher up in the hierarchy, many more exceptional patterns arising depending on whether there is subsequent pied-piping of the dominating category (see Cinque 2005:321-325 for the derivation of attested patterns cross-linguistically).

¹⁴²The relative position of NUM₂ and POSS₃ is unclear.

- b. *Nú í dag gengu piltar upp í því seinasta, þriðju bekkingar í landafr.*
 now today went boys up in the_{sá} last third graders in geography
4. þeir efstu fengu 6. Mensi er einn af þeim.
 4 the_{sá} topmost got 6 Mensi was one of them
 ‘Today, the boys took the last exam, third graders in geography. The top four got six. Mensi is one of them.’ (AnnGud-1880-05-16.xml)
- c. *við eigum þrjú börn jeg sendi þjer mind af 2ur þeim elstu*
 we have three children I send you picture of two the_{sá} oldest
 ‘We have three children. I am sending you a picture of the oldest two.’ (AdaBja-1905-01-24.xml)

It ought to be clear from these examples (and there are many more) that the reading is not that of the partitive *two/three/four... of the...* but indeed the cardinal corresponding to *the two/three/four...*, as indicated in the gloss and translation. As demonstrated by Pfaff (2015:82-88), this is an important distinction to make because Modern Icelandic allows for NUM >> N-DEF, in principle, but only on the partitive reading (Pfaff 2015:87f.). These examples demonstrate once and again that the 19th-century Icelandic extended noun phrase is a lot less restricted than the Modern Icelandic one. The tighter organisation of elements within the Modern Icelandic NP/DP is presumably a consequence of the ongoing syntacticisation of definiteness (for a more detailed discussion, see Lander and Haegeman 2014, Viðarsson 2017a).

4.6 Summary

The case study on the free-standing definite determiner reveals that *hinn* appears to have existed alongside *sá* already in the beginning of the period under study, which begs the question whether the implementation of *hinn* as the standard norm really involved a *revival* of an extinct pattern rather than just *selection* among existing variants. That is indeed similar to what Heimisdóttir (2008) has independently observed for another widely-cited feature, the standard declination of *ija*-stems (see Section 1.3.4). Moreover, the uses of the definite determiner *hinn* in early 19th-century private letters are strikingly similar to the typical use of *hinn* in present-day Icelandic, usually appearing in contexts with a non-restrictive reading featuring an evaluative adjective. That thus furthermore suggests that the “revival” of *hinn* instead of *sá*, as attempted, actually failed miserably. Teacher corrections in the student essays clearly suggest that *hinn* was not merely intended to be used in non-restrictive contexts.

Similar to Adv-Vfin above, there are clear effects of standardisation in the newspaper corpus as well as in the student essays, but private letters much less so. In fact, the temporal endpoint of *hinn* in the private letters towards the end of the period roughly corresponds to the beginning point of the newspapers in terms of frequency in the early 19th century, almost a century earlier. Different types of phrases, moreover, exhibit strikingly different trajectories. This is especially so with regard to the definite determiner with dates. In these contexts, newspapers exhibit a very sharp decline of non-standard *sá*, an effect which is nearly absent in the private letters.

Unlike Adv-Vfin above, male scribes overall exhibit a higher adoption of the stan-

dard norm than women, but the effect is still visible for both groups. There is furthermore some evidence that social status is important. Females from the higher echelons show a similar trend to the male scribes, although the former appear to lag slightly behind the latter. In contrast, there is no indication that males from the group of peasants/workers contribute in any way to the uptake of the standard norm, where the use of the non-standard variant actually increases over time as opposed to a slight decrease among the females.

As regards the definite marker, we have seen that 19th-century Icelandic permitted much more variation at the NP/DP level than attested in Modern Icelandic. The free-standing article could be used in restrictive as well as non-restrictive contexts, the NP/DP structure could be modified by adjuncts and (presumably) complex applicative structures that are much more restricted in present-day Icelandic. Furthermore, a very striking feature of 19th-century Icelandic was the double definiteness pattern, attested at least marginally in all three corpora, as well as the possibility of co-occurring determiners. Moreover, the position of numerals in the 19th-century Icelandic NP does not fully match that of Modern Icelandic and requires a major elaboration of possible structural positions.

5 Generic pronoun *maður*

5.1 Basic properties

The generic (or ‘indefinite’) pronoun *maður* ‘one (lit. man)’, while relatively commonplace in the spoken language, is usually frowned upon in the written language (for discussion and further references, see e.g. Ottósson 1990, Ragnarsdóttir and Strömquist 2005, Sigurðsson and Egerland 2009).

- (153) *En þau tímabilin eru þyngrri, þegar maður getur ekki grátið.*
but those time.periods-the are heavier when one can not cry
‘But the times when one cannot weep are more difficult.’
(GudMag-1844-06-13.txt)

In a sense it could be argued that the generic pronoun quite literally fills a gap in Icelandic in that it may alternate with null arguments, although the two are only partially interchangeable (cf. Sigurðsson and Egerland 2009). For generic reference, the two are rather similar in distribution as the following examples show (Sigurðsson and Egerland 2009:159f.):

- (154) a. *Fyrst beygir maður til hægri.*
first turns.3SG one to right
‘First, one turns to the right.’
b. *Fyrst er _____ beygt til hægri.*
first is.3SG turned to right
‘First, one turns to the right.’

In (154) the *maður* construction and the impersonal passive construction are fully interchangeable (for more types of covert impersonals, see Sigurðsson and Egerland 2009:160, 166-172). Precisely this aspect is emphasised by language mavens, demonstrating that not only is the generic *maður* construction undesirable because it incorporates a foreign element, it is redundant, too.

The different senses of overt and covert impersonals can be captured based on the following typology (cf. Sigurðsson and Egerland 2009:161):

- (155) a. Generic: non-restricted +HUMAN reading, i.e., people in general
b. Arbitrary: a non-specific +HUMAN reading, excluding the speaker or the hearer
c. Specific: a specific +HUMAN reading, referring to a wholly or a partly specific set of individuals, most commonly including the speaker

The generic reading of (155) was already exemplified in (154) above. The two remaining ones will now be treated in turn.

Whereas null impersonals can have a speaker/hearer-exclusive arbitrary reading in Icelandic, *maður*, in contrast, cannot (Sigurðsson and Egerland 2009:174f.):

- (156) a. *Í Ódysseifskviðu er yfirleitt ferðast á báti.* *gen/^{OK}arb/*spec
 in Odyssey is generally traveled on boat
 ‘In the Odyssey they_{arb} generally travel on a boat.’
- b. **Í Ódysseifskviðu ferðast maður yfirleitt á báti.* *gen/*arb/*spec
 in Odyssey travels one generally on boat

The inability for *maður* to figure in contexts with arbitrary reference is a property that sets it sharply apart from typical uses of the corresponding overt impersonal *man* in Mainland Scandinavian, including Danish from where *maður* is supposedly borrowed, as shown further below. It moreover attests to the fact that *maður* and covert impersonals, while sometimes interchangeable, are not identical in their distribution.

Cross-linguistically, the specific reading may result in a speaker-inclusive reading with a plural interpretation (‘we’) as is the case in Romance, for instance (for examples and references, see Sigurðsson and Egerland 2009:163). In Icelandic, however, the central reading in specific contexts includes the speaker, corresponding most closely to the 1st person singular ‘I’ (example from Sigurðsson and Egerland 2009:163):

- (157) *Já, maður var óheppinn í gær.* Icelandic
 yes, one was unlucky in yesterday
 ‘Yes, I was unlucky yesterday.’ (specific / *arbitrary)

As (157) shows, the arbitrary reference (excluding the speaker) is ungrammatical in Icelandic. As Sigurðsson and Egerland (2009:163) point out, *man* in Swedish, again in contrast to Icelandic *maður*, allows the speaker-specific (‘I’) interpretation as well as the arbitrary (‘they’) one:

- (158) *Ja, man hade otur igår.* Swedish
 yes, one had bad-luck yesterday
 Interpretation 1: ‘Yes, I was unlucky yesterday.’ (specific)
 Interpretation 2: ‘Yes, they were unlucky yesterday.’ (arbitrary)

As Fenger (2018) shows, there is a further split among the Germanic languages that allow for the arbitrary reading (or ‘existential’ in her terminology) such that German, Danish and Norwegian only allow arbitrary readings with subjects, i.e. external arguments. Swedish and Dutch, in contrast, allow for arbitrary readings also with derived subjects, i.e. internal arguments raising to subject (such as existential readings with unaccusatives and passives).

Similar to the above examples, *man* in Danish canonically has generic reference, although it may also be used with specific reference similar to (157) and (158) above (see Jensen 2009). Each of these readings are shown in (159a,b), respectively (cf. Jensen 2009:86f.):

- (159) a. *man* *behøver* *bare* *at* *tage* *bussen* *for* *at* *høre* *at* *de* *unge* *taler* *utrolig*
 one needs only to take the-bus for to hear that the young talk incredibly
dårligt
 bad
 ‘you only need to take the bus to hear that the young people talk incredibly bad’
- b. *hvad* *har* *jeg* *været* *#* *ja* *der* *i* *midten* *af* *tresserne* *har* *man* *været* *en*
 what have I been yes there in the-middle of the-sixties have one been a
fem *seks* *år* *ikke*
 five six years not
 ‘how old was I yeah in the middle of the sixties you [I] were about five or six years, right’

While these two readings are possible in both languages, there is a very striking difference between Icelandic *maður* and Danish *man* in terms of syntax; whereas *maður* can occur in multiple syntactic positions, inflecting for any of the four morphological cases (acc. *mann*, dat. *manni*, gen. *manns*), Danish *man* not only has an invariant form but is barred from occurring in non-subject positions (see e.g. Jensen 2009:94). See Section 5.6 for further discussion and analysis.

Already in Gíslason (1851:95, 294), it is observed that pronominal *man* in Danish only has a nominative form, with *een* being used in particular in those cases for which *man* has no form (*einkanl. í þeim föllum, sem það orð [þ.e. man—HFV] vantar, ibid., p. 95*). As regards *man*, Gíslason (1851:294) notes that *maður* can be used in Icelandic, although the plural form of the noun, *menn* ‘men; people’, is used more frequently in such contexts. Furthermore, he points out that the singular form *maður*, unlike Danish *man*, cannot be used in arbitrary contexts (not his terminology), providing the example *man har sagt mig* ‘I have been told (lit. one has told me)’. That example, he remarks, cannot be translated using Icelandic *maður* “which would be something completely different” (*sem væri allt annað*), i.e. it would then have to refer to a specific individual.

While Gíslason (1851) does not particularly endorse Icelandic *maður* as a pronoun in the section on Danish *man* (*ibid.*, p. 294), he still uses it to translate the Danish generic pronoun *een*, i.e. where it occurs in non-subject contexts (*ibid.*, p. 95). Thus, *hvad der ikke vedkommer een* is translated *það, sem mann varðar ekki um* (i.e. ‘what does not concern one’), *naar een ikke har lært noget = þegar maður ...* (i.e. ‘when one ...’). For the latter example, Gíslason (1851) also provides *sá, sem ...* as a viable translation, using a demonstrative. Here Gíslason (1851) is far more permissive than Friðriksson (1857), writing just a few years later, where Icelandic *maður* as a pronoun is left completely unmentioned, with plural uses of the noun (*menn*) given as the one and only translation of Danish *man*.

As a generic pronoun, Danish *man* is attested since medieval times (see Jensen 2009:86, with references).¹⁴³ The literature is not clear on whether this also holds for

¹⁴³Jensen (2009) provides a nice overview of four different contexts in which *man* vs. *du* ‘you’ vs. *en* ‘one, sby’ can occur depending on reference type. I did not attempt to do this as it is quite labour-intensive and not obviously relevant to the distribution of *maður* alone. That said, contrasting patterns are bound to abound in this material, especially as regards the private letters vs. the newspapers, the former of which is likely to strongly favour implicit speaker-orientation. However, as my study did not include the personal 2nd person pronoun, which was non-existent during this period, I deemed this not to be interesting enough to justify

man with a 1st person reading or if that is a more recent phenomenon. For the historical development from Old to Modern Icelandic, Jónsson (1992) assumes at least three stages to be involved. For reasons of semantic plausibility, these are taken to begin with a stage lacking the generic pronoun in Old Icelandic, followed by pronominal *maður* only in a generic sense, followed by the innovation of 1st person *maður* due to the speaker-orientation of generic *maður* (cf. Jónsson 1992:1, 13f.):

(160)	I	II	III
	maður (“man”)	maður (“man”)	maður (“man”)
		maður (“one”)	maður (“one”)
			maður (“I”)

While it may be true that Old Icelandic lacked a full-fledged Stage II generic pronominal *maður*, there is metrical evidence from Old Icelandic poetry suggesting that the noun *maður* (Old Icelandic *maðr*) in contexts similar to generic uses could arguably take on *pronominal* properties by appearing in unstressed positions typically reserved for prosodically weak elements (cf. Kjartansson 2017; see also Pétursson 2005:1266 for a remark on pronominal uses of the plural form *menn* during the 16th century).¹⁴⁴ Fritzner (1891:618), furthermore, provides a variety of pronominal examples of *maðr* (“*i pronominal Betydning eller Anvendelse: en, nogen, man*”) from Old Norse-Icelandic prose as well as poetry, being the fifth sense of *maðr*. Admittedly, not all of Fritzner’s examples are unequivocally generic pronouns but they certainly could be conceived of as such. In addition to the metric evidence presented by Kjartansson (2017), it at least seems fair to assume that potential pronominal uses were ‘in the air’, so to speak, already in the medieval period.

Further steps in the development of generic *maður* in later centuries have not been documented to the best of my knowledge and it appears to be assumed to be a recent, 19th-century phenomenon in metalinguistic/prescriptive commentaries on its use (see e.g. Sigurmundsson 1996). This is probably not the case, although it could be that pronominal *maður* did not come into widespread use outside educated circles until relatively late. During the 16th to 18th century, the evidence for pronominal *maður* becomes more robust than in the Old Icelandic period, including the birth of the short-lived variant *mann*—strikingly similar to Danish/German *man*. Whereas pronominal *maður* simply follows the canonical declination of the homophonous noun from which it derives, the variant *mann* was an invariant form.¹⁴⁵ In the University Dictionary Written Language Collection (*Ritmálssafn Orðabókar Háskólans*), a diachronic dictionary with text excerpts covering the period from the 16th to 20th century, pronominal *maður* is listed as an entry separate from the noun, with examples already from the mid-16th century. Some of these feature the invariant form *mann*, others *maður* with the typical noun inflection. Based on the *Ritmálssafn* entries, the pronoun could apparently appear

ploughing through the data systematically with such distinctions in mind.

¹⁴⁴Note in contrast that the generic pronoun *maður* may not be stressed, being unacceptable both in contrastive and topicalised environments (see Jónsson 1992:18). However, *maður* may be emphasised by using the element *sjálfur* ‘self’ (i.e. *maður sjálfur* ‘oneself’).

¹⁴⁵It is unclear whether this form actually derives directly from Danish/German *man* or from the variant nominative singular form *mann*, which historically is also attested in the declination of the nominal alongside *maður*.

in that invariant guise (*mann*) as late as the late 18th century. Unsurprisingly, early attestations of pronominal *maður/mann* tend to occur in translated Biblical works and derivatives, potentially attesting to foreign origins.

Due to its supposed Danish and/or German origin (see e.g. Smári 1920:130), this construction never became a part of the emerging standard language and is still to this day typically not considered acceptable in the formal/written language. Kolbeinsson (1976:6), in a language column, writes that *maður* has been met with hard resistance in schools, language columns on radio programmes and elsewhere, a struggle he states “has gone rather well” so that the pronoun “is heard much more seldom now than a couple of years ago” (my transl.). It is thus probably with some optimism that he moves on to target the innovation of English-inspired generic *þú* ‘you’ as well, which “must be avoided” in his view. To the best of my knowledge, no actual quantification of this mild success exists, until after the turn of the century almost thirty years later.

In their contrastive study of *man* in Modern Swedish and *maður* in Modern Icelandic, Ragnarsdóttir and Strömquist (2005) observe that the use of *man* increases with age in Swedish, whereas the use of *maður* in Icelandic conversely drops drastically.¹⁴⁶ Based on both age and genre effects in their results, this is taken to reflect “the progressive acquisition of the appropriate style for formal or written language, constraining the use of *maður* to an infrequent use in mature, literacy-experienced, and well-educated speaker-writers.” (Ragnarsdóttir and Strömquist 2005:154) Interestingly, the rates of Icelandic *maður* and Swedish *man* are near identical in the youngest age group (10-11 yrs) with about 10% of the clauses featuring the generic pronoun (cf. Ragnarsdóttir and Strömquist 2005:152). There is a very striking statistically significant reduction in the use of *maður* across modalities (spoken vs. written) and genres (expository vs. narrative texts) from 10.1% on average to 4.9% in the junior college group (13-14 yrs) and the rate is consistently higher in the spoken than in the written modality across all age groups. These developments are totally absent in the Swedish data where the use of *man* increases slightly with age, up to 13.5% on average in the oldest group, compared to avg. 4% in Icelandic. Since there are very little differences between the junior college students on the one hand and high school students and university graduates on the other, it appears most of the effects on the use of *maður* have already been felt during the compulsory education level. In terms of standardisation effects, the attempted suppression of *maður* in present-day Icelandic thus fall at least in the *partly successful* category of Elspaß (2016) in (17) above.

5.2 Stigmatisation

Pronominal *maður* is mentioned specifically in relation to corrections of non-standard language use at the Reykjavík Grammar School (1846-1904) by the main Icelandic teacher, Halldór Kr. Friðriksson, based on what appears to be anecdotal evidence from student memoirs (see Ottósson 1990:96). Thus, reflecting on his own language use,

¹⁴⁶Their sample included 80 Icelandic and 78 Swedish subjects, equally distributed based on sex and divided into four age groups: Age 10-11 (grade school), 13-14 (junior high school), 16-17 (high school) and 26-40 (adult university graduates). The study consisted of a spoken and a written component, each divided into an expository and a narrative part.

having just made use of the pronoun, Thorsteinson (1946:80) states that few things bothered his teacher, Friðriksson, more than the pronominal use of *maður*, which actually led some to purposely tease him with it sometimes (*ibid.*). Besides these memoirs, there is at least a potential implicit dismissal of pronominal *maður* in the Danish grammar of Friðriksson (1857:45) where we find that the indefinite pronoun *man* in Danish is translated into Icelandic with the plural form of the noun *maður* (i.e. *menn* ‘men, people’) rather than using the pronoun (see also Viðarsson 2017b:137). Much in line with Halldór Kr. Friðriksson’s method of only showing standard norms, with no attention paid to listing variant forms, pronominal *maður* is not mentioned as an option in his Icelandic grammars such as for instance *Íslenzk málmundalýsing* (Friðriksson 1861), which was taught until 1885 (cf. Ottósson 1990:96f.). As mentioned in the previous section, Gíslason (1851) also gives the impression that the nominal plural *menn* is the conventional equivalent of Danish *man*, given alongside null arguments and the generic pronoun *maður*. The same strategy of employing the plural noun *menn*¹⁴⁷ for generic reference is also the typical advice in later prescriptive grammars, in addition to impersonal passive constructions, for instance, allowing for the indefinite/generic referent to be left out completely (cf. e.g. Blöndal and Stemann 1959:95).

With the ‘rediscovery’ of the Reykjavík Grammar School student assignments, we are in a position to study the basis for anecdotes of this kind in more detail. While a comprehensive study of this fascinating material still awaits, my preliminary snapshot overview of every decade since the grammar school moved (back) to Reykjavík in 1846 up until 1890 (see further Viðarsson 2014, 2016, 2017b) reveals an early awareness of this feature in Friðriksson’s corrections. His corrections go at least as far back as 1852 (single/double underlining as in original):

- (161) a. *svo maður yrði að ganga það* (1852, 2nd grade)
 so one must to walk it
 ‘... so one would have to walk the distance.’
- b. *reyna það jafnskjótt og manni dettur hann í hug* (1852, 3rd grade)
 try it as soon as and one falls he in mind
 ‘... (to) try it the minute one has the idea.’

The corrections indeed show that early in his career, Friðriksson was actively counter-acting this supposedly Danish-rooted phenomenon. Roughly forty years later, he was still at pains to eradicate pronominal *maður* in his students’ essays. For instance, in one fourth-grade essay from 1890 by Ólafía Jóhannsdóttir (1863-1924),¹⁴⁸ Friðriksson corrects the generic pronoun no less than seven times on one and the same page (for the published essay text, excluding corrections, see Ólafsson 2004:148f.).

What is not obvious from this, however, is that *maður* was necessarily a personal thorn in the side of Friðriksson, as previous scholarship seems to imply, as opposed to simply being his service towards implementing an emerging standard as his duties required. For one, Friðriksson himself does not seem to have hesitated to use pronominal

¹⁴⁷I thank Helgi Skúli Kjartansson for valuable discussions about the *maður* nominal-pronominal and singular-plural contrasts, with later also the noun *fólk* ‘people’ being used for generic reference.

¹⁴⁸The first female student to finish fourth grade since women were first admitted in 1886, although she never actually graduated (cf. Ólafsson 2004:29).

maður in a Fjöltnir book review in 1845, otherwise featuring a range of harsh metalinguistic remarks in the tradition of previous book reviews in the same journal. Similarly, the pronoun for instance appears in a letter of his to Jón Sigurðsson, dated 21st August 1868:

- (162) a. *Þegar maður skoðar hvernig greinarmerkin eru sett, þá er auðsjeð, að höfundurinn hefur enga hugmynd um, hvað hvert merki táknar; því þau eru sett út í bláinn*
 ‘When one considers how punctuation marks are used, it is clear that the author has no idea as to what each symbol signifies, for these are placed at haphazard.’ (Fjöltnir 8:74, signed H. F.)
- b. *Baldur getur maður eigi átt við í því máli*
 B. can one not own with in that matter
 ‘As for that matter, one cannot consult Baldur.’
 (ed. Halldórsson 1991:91; dated 21/08/1868)

No systematic study of Friðriksson’s language use has been undertaken, but pronominal *maður* was at least a part of his language, something even a figure of his caliber could produce in writing. Still, this raises the question whether these were slips of the tongue or whether Friðriksson’s supposed ‘contempt’ was perhaps purely professional—i.e. *maður*, like any other supposedly Danish trait, was simply not to be tolerated in the standard, which obviously did not necessarily include colloquial private letter writing. But then what about the attestation in Friðriksson’s review, published in 1844?

Further evidence, while inconclusive, does suggest that *maður* would not have been considered inappropriate in the emerging standard as this was conceived of in the Bessastaðir Grammar School, prior to its relocation to Reykjavík in 1846, nor shortly thereafter. Evidence to this effect is provided by the ceremonious speeches of Sveinbjörn Egilsson (ed. Sigurðsson 1968), albeit somewhat limited by the fact that while written, these speeches were read to the audience and thus never intended for print. Examples are shown in (163) below.¹⁴⁹

- (163) a. *og þessi barátta finst **manni** þó skémtileg, af því hún æfir kraptana,*
 and this struggle feels one though fun because she trains the powers
*og með æfingunni yfirvinnr **maðr** smám saman örðugleikann, so að*
 and with the practise overcomes one little together the difficulty so that
*það sem **manni** áðr var örðugt, finst **manni** léttara*
 it that one before was difficult feels one easier
 ‘And one likes this struggle because it trains the strengths and by practising, one overcomes the difficulties little by little, so that what one used to find difficult becomes easier for one.’ (p. 39, 1823 & 1832)
- b. *og dæmi félagsbræðra geta sýnt **manni** deginum ljósara, hvað mikið*
 and examples unionbrothers’ can show one the day lighter what much
sé unnið, og engu spillt, með því að halda sér gjörsamlega frá
 is done and nothing spoiled with it to keep oneself completely from

¹⁴⁹It should be mentioned in this context that Sveinbjörn Egilsson did not fully endorse the standard being set by the Fjöltnir book reviews, voicing criticism over how radical they were and stating his doubts he could live up to those standards (see Ottósson 1990:72, with references).

nautn áfeingra drykkja

pleasure alcoholic drinks.

‘And the examples of union brothers can show one clearly how beneficial it is, without spoiling anything, to abstain from alcoholic drinks completely.’

(p. 85, 1847)

- c. *Vekja má það hjá yður umhugsun um það, að það sem manni þykir vænt um, það sem manni er orðið dýrmætt, það skilur við mann*
 wake may it with you thought about it that it which one feels
 dearly about it which one is become precious it parts with one
 ‘It may get you to think about what one has come to care about and what has become precious to one, it parts with one.’

(p. 93, 1851)

5.3 Circumscribing the variable context

Recall from the discussion above that according to traditional wisdom, it ought in principle to be possible to distinguish between pronominal and nominal uses of *maður* in any stretch of text based on referential properties; if and only if *maður* is repeated when referring to a previous instance of the (alleged) generic pronoun, we can be certain both cases are indeed to be analysed as a generic pronoun, whereas if the third person singular pronoun *hann* ‘he’ is used, *maður* is not a generic pronoun but a noun.¹⁵⁰ Moreover, the former will be unstressed, similar to other pronouns, and the latter will typically receive stress, being a noun. For obvious reasons, phonological stress is a tricky metric in written prose. However, the referential metric is arguably also much less conclusive of nominal status than made out to be.

First of all, for practical reasons, if one were to exclude any instance of *maður* that does not explicitly act as a referent to another instance of a repeated, coreferential *maður*, the majority of the data will be rendered unusable. In and of itself, this is of course not an argument against widening the criteria to encompass single as well as repeated environments involving *maður*. Note that evidence from the Reykjavík Grammar School suggests *maður* ambiguous in such a manner was corrected just as well as (repeated) cases where *maður* was unequivocally a pronoun (see Viðarsson 2017b:137f.). More importantly, the referential properties do not necessarily preclude an analysis of *maður* as a pronoun even when serving as a referent of the third person singular pronoun *hann* ‘he’, at least not as far as 19th-century Icelandic data are concerned (see also Viðarsson 2017b:138, fn. 4). Thus, within one and the same discourse context, a speaker can use the first person pronoun *ég* ‘I’, followed by a generic pronoun *maður* whose main referent is the same entity as the first person pronoun, which again serves as an antecedent of a succession of repeated instances of *maður*. In an apparent break up of this *ég*–*maður* chain, the third person singular pronoun *hann* ‘he’ can surface, all without introducing a new referent to the discourse.

An example of precisely such alternation is shown in example (164), which for reasons of space is left untranslated, the point made arguably being sufficiently clear

¹⁵⁰Note that this use of unmodified *maður* serving as the antecedent of the third person pronoun is rare, the inclusion or exclusion of which does not materially affect the results reported on in this chapter (see also Viðarsson 2017b).

from the present discussion as well as the boldfaced elements (*eg* ‘I’, *mér* ‘me (dat.)’, *manni* ‘one (dat.)’, *hann* ‘he’, *elskan mín* ‘my darling’):

- (164) Þegar **eg** kom til bæjarins hérna, þóttu **mér** kofarnir æði háir, **mér** hafði aldrei dottið það í hug, að það væri eins. **Maður** sér ekkert nema málaðan steinvegg á tvær hendur með ótal gluggum, og þar fyrir ofan uppí heiðan himininn. [...] Á veginum [útá Friðriksbergsslot] hefur **maður** fyrst frá borginni einlæga húsaröð til beggja handa, svo **maður** skyldi hugsa að **maður** væri alltaf að ganga inní borginni á einhverju hennar stræti. [...] En þegar vindur blæs eru trén að hvíslast á fyrir ofan höfuðið á **manni** og beygja sig hvort til annars. [...] Þau benda **manni** til, að náttúran hafi ætlazt [til] að **maður** skyldi ekki ráfa þar einn í fánýtum grillum, heldur skuli **maður** ganga með það eitthvað undir arminn, sem helzt fær mýkt úr sorgunum, helzt aukið og eftt og hreinsað tilfinningarnar og gleðina, það sem rétt að segja gerir **mann** viðurskila við það jarðneska og lætur mannsins sálu sveima í einslags sætri, rólegri og áhyggjulausri unaðsemd, að **maður** veit varla að **hann** sé á þeirri jörðu, hvar **manni** beri að brjótast í gegnum allslags mæðu til að ná markinu, það sem í einu orði lætur **mann** líða uppí þriðja himin, hvar **maður** verður þess var, er **hann** aldrei gat ímyndað sér og sem er svo óútmálanlegt og fínt, að sálin getur aldrei tekið af því áreiðanlegt bílæti, hvað þessi óumræðilega unaðsemd verður í hvert sinn nokkuð nytt, sem **maður** ekki vel getur munað til að komið sé eins fyrir áður. Og hvað ætli þetta sé? Það er elskan **mín**!
(BalEin-1826-08-14.txt)

Example (164) first introduces the referent, viz. the speaker, recalling his exploring of his surroundings, who is clearly serving as a discourse antecedent of *maður* (in addition to being generic, i.e. <+speaker, +generic>). The succession of cases following that first instance of *maður* have the same properties, occurring multiple times in the following few lines. However, the ninth instance of *maður* now rather surprisingly figures as the antecedent of *hann* ‘he’, yet it does not in any clear way introduce a new referent not already implied by the generic pronoun. The personal pronoun is then followed by *maður*, still referring to the same referent, repeated twice, only to be followed again both by the personal pronoun *hann*; a little further on followed by *maður*, soon after which the speaker again uses the first person singular form. In this case the possessive *mín* ‘mine’ is actually referring specifically to the speaker’s own darling girlfriend.

In my view, examples of this kind clearly suggest that we should not (necessarily) regard reference to *maður* by a third person singular pronoun as implying that *maður* must be a noun.¹⁵¹ Rather, what we are witnessing is variation at the level of *reference form* (see also Gussenhoven 1987 on the same kind of variation in English *one*, either by *one* repetition or alternatively by *he*) and need thus not constitute a difference at the level of the referent or in terms of (pro)nominal status. That being said, even from this perspective, it is still conceivable that cases in which the personal pronoun *hann* ‘he’ is used to refer to *maður*, the personal pronoun is mainly targeting the generic referent and decidedly not the speaker alone, unlike reference by *maður* which is clearly impartial to this point, clearly allowing for the <+speaker, +generic> and <+speaker, -generic> readings (see Section 5.6 for a more basic decomposition).¹⁵² Note moreover, that the

¹⁵¹ Unless it carries stress, which is incompatible with the pronominal reading (see Jónsson 1992).

¹⁵² Note furthermore that if the instances of *maður* in this example, serving as the antecedent of the third

earliest attestation of pronominal *maður* that is provided by *Ritmálssafn Orðabókar Háskólans* has pronominal reference with *hann* ‘he’, dated 1546.

5.4 Results

5.4.1 Overall trends in the data

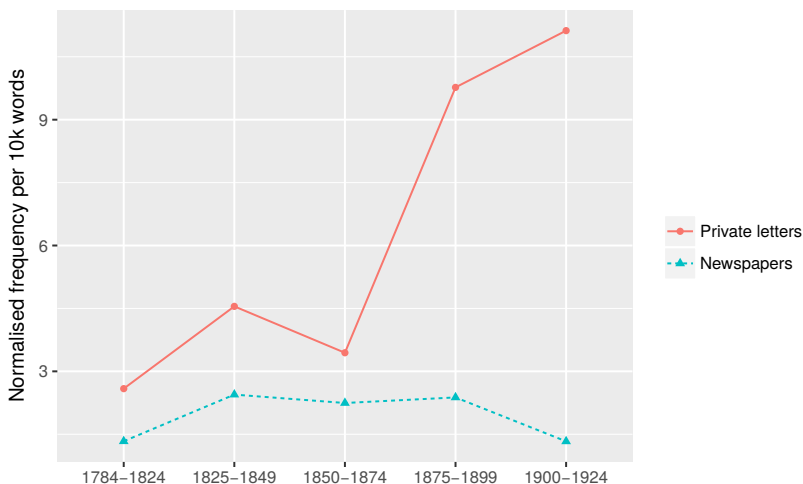


Figure 5.30. Overall frequency of *maður* in newspapers and periodicals, normalised per 10,000 words—contrasted against the private letter corpus for comparison.

Figure 5.30 presents the overall frequency of pronominal *maður* as a univariate variable in newspapers and periodicals, normalised per 10 thousand words.¹⁵³ The newspapers show a rather peculiar, flattened, trapezoid-shaped pattern with an endpoint identical to the starting point. Anticipating further discussion in Section 5.4.3 somewhat, Figure 5.30 also includes results from the private letter corpus for comparison. The main reason for including the letters here is the observable increase in the use of the generic pronoun over time attested in the letters, strikingly absent in newspapers and periodicals, which in my view represents a rather convincing case of the retarding effects of standardisation (Period 4) and subsequent near-elimination of *maður* (Period 5).

The first quarter of the century yields very few examples of pronominal *maður* in the newspaper corpus, occurring at a normalised rate of 1.3. Given the limited amount of data from that period, this rate amounts to merely 13 examples in total in Period 1.¹⁵⁴

person pronoun, really were the noun, one would at the very least expect the second mention to occur with a determiner such as *sá* ‘that’, contrary to fact.

¹⁵³This is a method similar to the one used by Ragnarsdóttir and Strömqvist (2005), who normalised the frequency of *maður* per clause.

¹⁵⁴Note that the private letter corpus extends further back, into the late 18th century, hence the axis label

The increase in Period 2 (1825-1849) to a rate of 2.4 (corresponding to 86 examples) is significant at the level $p < 0.05$ (LL 4.82). Periods 2, 3 and 4, i.e. from 1825 until the turn of the century, show a very stable pattern around the 2.2-2.4 mark. Such minor fluctuations in rate of occurrence are, unsurprisingly, not significant.¹⁵⁵ In Period 5 (1900-1924) there is a subsequent *decrease* to a rate of mere 1.3 (corresponding to 116 examples). The decrease in this period as compared to the previous Period 4 is strongly significant at the level $p < 0.0001$ (LL 19.80).

Granted, changes need not always be successful. However, the trajectory of change arguably attests not to any retrograde movement as concerns the change as a whole, but only in the newspapers. Briefly contrasting this with the private letters for comparison, there is an additional observation to be made regarding the stable period stretching over the first three quarters of a century in the newspapers. Both corpora have rather similar overall trajectories in these three periods. However, the corpora diverge sharply in Periods 4 and 5 where the newspapers simply do not participate in the vast increase observed in the use of *maður* in the private letters.

Pronominal *maður* on the whole is, indeed, consistently more frequent in the private letters than in the newspapers, but in no way does this undermine the points being made. Given the more colloquial nature of the letters, thus arguably closer to the spoken modality, this somewhat parallels the distribution in Modern Icelandic. As the reader will recall, *maður* is widely reported to be particularly frequent in the spoken language, more so than in the written language. While some of that effect may be due to standardisation, the speaker-orientation of *maður* in Icelandic is clearly also favourable to its use in ego-documents, such as private letters; a higher overall rate of occurrence is thus as expected. However, one would, all other things being equal, also expect the frequency in the newspapers to rise proportional to the increase in the private letters. In my view, the fact that the newspapers are lagging enormously behind the private letters in Period 4, with a further reduction in Period 5, strongly suggests standardisation effects are at play in the newspapers, at least in these periods—the last quarter of the century onwards.

The broad brush-strokes of Figure 5.30, neat and clear as they may appear, admittedly conceal a lot of detail that can be gleaned from the distribution in individual newspaper titles or by further zooming in on the present time periods. The problem with moving beyond the macro-level is obviously the relatively small size of the data set. At the risk of not being able to present much statistically significant results, let us now consider shorter time spans of decades before considering more micro-level aspects of the analysis when moving on to individual titles.

5.4.2 Newspapers and periodicals

1803-1824	Year(s)	Place of publication	RAW	WC	NORM
Minnisverð tíðindi	1803-1808	IS (W)	7	41343	1.69
Íslenzk sagnablöð	1816	DK	5	16986	2.94
Klausturpósturinn	1818-1822	IS (W)	0	16651	0.00
Margvíslegt gaman og alvara	1818	IS (W)	1	22593	0.44

1784-1824 (both corpora) as opposed to 1803-1824 (only newspapers).

¹⁵⁵The raw frequencies in these periods are 86, 42 and 121 examples, respectively.

1825-1849					
Íslenzk sagnablöð	1825	DK	0	12936	0
Klausturpósturinn	1825	IS (W)	0	5319	0
Skírnir	1828-1840	DK	4	35086	1.14
Ármann á Alþingi	1829-1832	DK	33	63996	5.16
Sunnanpósturinn	1835-1838	IS (SW)	2	34733	0.58
Búnaðarrit Suðuramtsins ...	1839	IS (SW)	15	21951	6.83
Ný félagsrit	1841	DK	12	41740	2.87
Reykjavíkurbósturinn	1846-1849	IS (SW)	10	43822	2.28
Þjóðólfur	1848	IS (SW)	0	5553	0
Norðurfari	1849	DK	2	86603	0.23
1850-1874					
Ársritið Gestur Vestfirðingur	1850-1855	IS (SW)/DK	8	51917	1.54
Lanztíðindi	1850	IS (SW)	0	2775	0
Þjóðólfur	1850-1860	IS (SW)	0	21789	0
Ný tíðindi	1852	IS (SW)	0	1981	0
Ingólfur	1853-1855	IS (SW)	1	8925	1.12
Norðri	1853-1859	IS (NE)	2	23865	0.84
Keilir og Krafla	1857	IS (NE)	2	1695	11.80
Íslendingur	1860	IS (SW)	10	15239	6.56
Ný sumargjöf	1860	DK	0	30396	0
Baldur	1868	IS (SW)	0	3656	0
Norðanfari	1870-1871	IS (NE)	8	19916	4.02
Ísafold	1874	IS (SW)	0	4901	0
1875-1899					
Fréttir frá Íslandi	1875	IS (SW)	0	27403	0
Íslendingur	1875	IS (SW)	2	18951	1.06
Norðanfari	1875	IS (NE)	8	28746	2.78
Ísafold	1875-1899	IS (SW)	22	84925	2.59
Norðlingur	1875-1881	IS (NE)	10	29591	3.38
Þjóðólfur	1875	IS (SW)	0	21901	0
Skuld	1877	IS (E)	6	4657	12.88
Suðri	1883	IS (SW)	0	10431	0
Austri	1884	IS (E)	0	4442	0
Fjallkonan	1884-1899	IS (SW)	16	69556	2.30
Akureyrarpósturinn	1885-1886	IS (NE)	0	5287	0
Eimreiðin	1895-1899	DK	30	179613	1.67
Bjarki	1896	IS (E)	1	10447	0.96
Dagskrá	1896	IS (SW)	3	10738	2.79
Ísfirðingur	1898	IS (Wf)	0	2097	0
1900-1924					
Austri	1900	IS (E)	4	25591	1.56
Bjarki	1900	IS (E)	3	20265	1.48
Eimreiðin	1900-1920	DK/IS (SW)	72	306565	2.35
Fjallkonan	1900-1911	IS (SW)	28	151102	1.85
Ísafold	1900-1924	IS (SW)	10	98761	1.01
Þjóðólfur	1900-1920	IS (SW/S)	11	116141	0.95
Framsókn	1900	IS (SW)	5	12709	3.93
Kvennablaðið	1900	IS (SW)	16	17284	9.26
Reykjavík	1900	IS (SW)	0	9108	0
Reykvíkingur	1900	IS (SW)	2	11840	1.69
Stefnir	1900	IS (NE)	2	18444	1.08
Þjóðviljinn	1900	IS (Wf)	3	29531	1.02
Norðurland	1902	IS (NE)	10	12102	8.26

Alþýðublaðið	1906	IS (SW)	4	3075	13.01
Austurland	1907	IS (E)	0	17301	0
Óðinn	1908-1915	IS (SW)	4	13245	3.02
Vísir	1910	IS (SW)	0	2344	0
Norðri	1913	IS (NE)	12	2373	50.57
Tíminn	1922	IS (SW)	0	4318	0

Table 5.24. Raw and normalised frequency of the generic pronoun *maður* in periodicals and newspapers 1803-1924, per quarter century.

Table 5.24 reveals the rate of use per individual title, per quarter of a century. No minimum threshold for attestations is set as this would obliterate categorical non-use of the (univariate) variable. A striking feature of Period 1 (1803-1824), in particular, is the relatively high rate of *maður* (2.94) in *Íslenzk sagnablöð*, published in Copenhagen, as opposed to much lower frequencies in the papers published in Iceland, with *Minnisverð tíðindi* ranking highest (1.69). While this might seem to attest to Danish influences, the difference between papers published in Iceland as opposed to Denmark taken together per period is far from being statistically significant and this result holds across the subsequent periods as well.¹⁵⁶

Let us now consider titles consisting of 20,000 words or more in the subsequent periods. *Ný félagsrit* in Period 2 (1825-1849) features a non-standard variant at a non-negligible rate (2.87),¹⁵⁷ rivalled only by *Búnaðarrit Suðuramtsins ...* (6.83) and *Ármann á Alþingi* (5.16), and immediately followed by *Reykjavíkarpósturinn* (2.28). The emerging standard norms, thus again (cf. Section 3.4.1.3), do not appear to have been adopted by the editors of *Reykjavíkarpósturinn* (on which, see Section 3.4.1.3). In contrast, the feature is much rarer in *Skírnir* (1.14), *Sunnanpósturinn* (0.58), and almost nonexistent in *Norðurfari* (0.23). Recall that *Norðurfari* featured non-standard Adv-Vfin at relatively high frequencies (cf. Table 3.5, p. 3.5). Given the frequently-made observation (or claim) that the emerging standard norms were catching on after 1840, it may be legitimate to emphasise that the time variable, already at this point, is arguably a factor in the (frequent) use vs. (absolute or near) non-use of the linguistic variable; the use of *maður* until 1840 is considerably higher than after 1840, occurring at a rate of 3.35 as opposed to 1.35 (LL 14.98, significant at the level of $p < 0.001$).

The following three periods continue to show mixed results. In Period 3 (1850-1874), the frequency of *maður* is low in certain newspapers at least, e.g. *Ársritið Gestur Vestfirðingur* (1.54), *Þjóðólfur* (0.00), *Norðri* (0.84) and *Ný sumargjöf* (0.00), whereas it is (still) relatively frequent in *Íslendingur* (6.56), published in Reykjavík in 1860, and *Norðanfari* (4.02), published in Akureyri in 1870-71. Both *Þjóðólfur* and *Norðri* were edited at least in part by the language purist Björn Jónsson, although one out of two examples appear to have ‘slipped in’ on his watch. Periods 4 and 5 (1875-1899, 1900-1924) reveal low or relatively low frequencies of *maður* in various papers, some of which, however, do not match our threshold of 20,000 words. Thus, *Norðanfari*,

¹⁵⁶It should also be borne in mind that while papers were published in Iceland, a sizeable portion of these would be translated from or loosely based on Danish originals, especially in the earliest material.

¹⁵⁷Interestingly, this is quite unlike the results in chapter 3.4.1.3 where *Ný félagsrit* stood out among most other largely contemporaneous titles in its rather limited use of Adv-Vfin. The more widespread use of *maður* might be suggestive of that feature being more deeply entrenched in the grammatical system (both as a spoken and written feature) than Adv-Vfin, which in contrast may have had a more limited distribution.

Ísafold, *Skuld*, *Fjallkonan* and *Dagskrá* clearly exhibit the feature in Period 4, whereas *Fréttir frá Íslandi* and *Þjóðólfur*, both of which are among the larger samples, exhibit categorical non-use of the variable. The paper *Þjóðólfur* is thus unique for not having any instances of *maður* through Period 2, 3 and 4 (1848, 1850-1860, 1875). In Period 5, the feature is quite prominent in *Eimreiðin* (2.35), *Framsókn* (3.93), *Kvennablaðið* (9.26) and *Norðurland* (8.26), although some of these samples are smaller than our 20,000 word minimum. In contrast to the previous periods, *Þjóðólfur* now has 11 examples of *maður* in Period 5, translating to a modest normalised rate of 0.95. As mentioned above, *Þjóðólfur* had several editors, Sveinbjörn Hallgrímsson between 1848-1852, with ties to the Grammar School, followed by Jón Guðmundsson between 1852-1874. During Periods 4 and 5, however, there is a succession of numerous other editors, of which space does not permit further analysis. It is thus in these latter periods, after Sveinbjörn Hallgrímsson and Jón Guðmundsson that we start to get examples of *maður*. Note furthermore that *Ísafold* exhibits a downward trend from Period 4 (2.59) to Period 5 (1.01), significant at the level of $p < 0.05$ (LL 6.61). During most of these two Periods, until 1909, the well-known language purist Björn Jónsson was chief editor of *Ísafold*, followed by his son Ólafur Björnsson between 1909-1919.

To briefly sum up this section, the generic pronoun *maður* thus still occurs occasionally in late 19th- and early 20th-century newspapers, but remains a rather infrequent feature both in comparison to the previous periods and to the private letter corpus. This result suggests that in certain newspapers, at least, we are witnessing a rather successful suppression of a non-standard feature in the written standard. However, the question remains whether this rather subtle temporal decrease had any lasting effect on the written record in the decades to come. We now briefly turn to this question, before moving on to *maður* in the Reykjavík Grammar School student essays.

5.4.3 Private letters

Table 5.25 shows the use of *maður* across speakers who produced the variable at least ten times or more.¹⁵⁸ Based on this distribution, one may observe that *maður* is not clearly confined to a particular social group, being produced by speakers of roughly the whole spectrum—from workers, to mistresses to officials. However, a potential generalisation over these data has to do with the birth dates; speakers born in the late 18th or early 19th century tend to produce the generic pronoun *maður* less frequently as compared to speakers born later in the 19th century. The result of a split between speakers born after 1840, based on Table 5.25, is shown in (165):

(165) *Splitting speakers based on birth year (two periods)*

	BIRTH YEAR	<i>n</i>	WC	NORM.
a.	Until 1840	264	540967	4.88
b.	After 1840	490	425121	11.53

For speakers born after 1840, the frequency of *maður* is over two times higher than for speakers born in 1840 or earlier, suggesting in turn that *maður* may be a novel and/or

¹⁵⁸Since *maður* is being treated here as a univariate variable, an (absolute) lack of attestations may, of course, be informative, provided the sample is large enough.

incoming feature that is increasingly spreading across new generations of speakers. This contrast is strongly significant, attaining a high value of LL 134.18, $p < 0.0001$, Odds Ratio 0.42. The generation effect will be considered in more detail below.

Scribe	<i>n</i>	Word count	Norm.	Born	Gender	Occupation
SigPal	13	100271	1.3	1809	f.	
AnnGud	10	36022	2.78	1828	f.	worker
BjaTho	37	127102	2.91	1786	m.	official
ArnHel	29	92650	3.13	1777	m.	bishop
SteSig	10	28102	3.56	1842	f.	
JakJon	28	55443	5.05	1835	f.	mistress, housewife
JohHal	13	22769	5.71	1851	m.	
AsgFri	11	18583	5.92	1860	m.	
RagDan	12	20081	5.98	1859	f.	
IngJons	41	68397	5.99	1784	f.	mistress, housewife
SofDan	29	46709	6.21	1858	f.	
KleBjo	18	17784	10.12	1829	m.	
GdrJon	90	85785	10.49	1856	f.	housewife (spinster)
ThoSte	12	9559	12.55	1861	m.	
AdaBja	21	16260	12.92	1860	m.	worker
GriThom	32	24396	13.12	1820	m.	scholar
BenHal	86	65317	13.17	1845	m.	worker
EirJoh	27	19427	13.9	1862	m.	
FinJon	58	35313	16.42	1858	m.	scholar
GunOdd	60	35053	17.12	1850	m.	carpenter, farmer
SteTho	22	10052	21.89	1890	m.	electrician
TorEgg	26	9952	26.13	1809	m.	student (DK)
LofJon	16	6044	26.47	1840	m.	carpenter
VilJon	27	9642	28	1870	m.	postman
BalEin	14	2906	48.18	1801	m.	student (DK)
JohSig	12	2469	48.6	1866	m.	

Table 5.25. Letters: The rate of the generic pronoun *maður* in letter-writers with 10 examples of the variable or more.

What most speakers with an exceptionally high rate of *maður* have in common are low figures for word count. This suggests that such rates are, indeed, exceptional and presumably artificially inflated due to chance or noise in the data. However, these speakers are strikingly also all males. Male speakers such as GunOdd, FinJon, EirJoh and BenHal, who were all born after 1840, all attest to the fact that *maður* is becoming a frequent feature for generic contexts, thus arguably taking over generic readings due to null arguments. The overall lower rate observed for individual female speakers is not unexpected, cf. Figure 5.31 above. The relatively high rate of *maður* in the letters of GdrJon, being the sister of FinJon, fits well in with the relatively high rate of the other non-standard variants, as observed in the previous chapters. In SofDan and RagDan's letter writings, both belonging to the same generation as GdrJon, *maður* is considerably less frequent and not much different from IngJons (b. 1784) and JakJon (b. 1835) in terms of frequency. It is not unlikely that as far as high frequency is concerned, 'early

adopters' such as GriThø (b. 1820) are exhibiting higher rates earlier than others due to Danish influences, having lived and studied in Copenhagen.

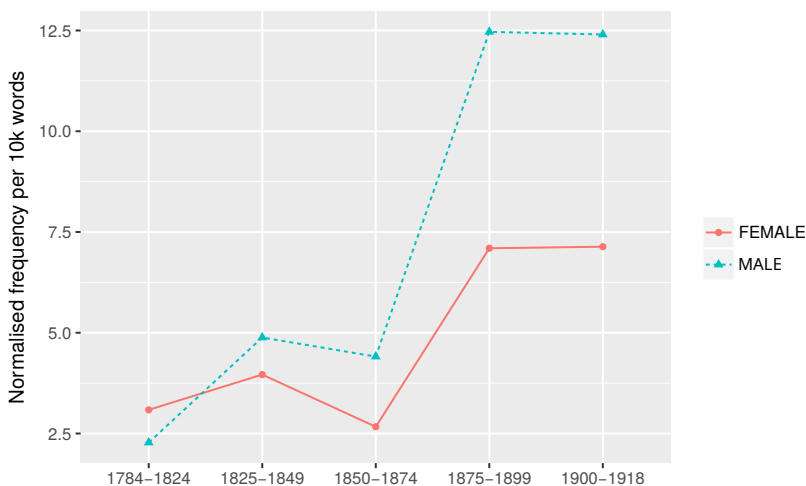


Figure 5.31. Private letters: The rate of occurrence of generic pronoun *maður*, split for time period and gender, normalised per 10,000 words.

The generation effect reported on above is even more compelling once we split the complete data set drawn from the private letter corpus ($n = 938$) for both genders into four periods instead of two, roughly per quarter century, cf. Table 5.26, shown graphically for ease of exposition in Figure 5.32. The spread of *maður* over generations fits nicely in with the common view that the generic pronoun is after all, by and large, a 19th-century innovation—at least in terms of its adoption beyond the educated niche.

BIRTH YEAR	MALES			FEMALES			Σ
	NORM	WC	n	NORM	WC	n	
<1825	4.12	407510	168	3.03	283814	86	3.67
1825-1849	9.16	179052	164	3.37	157229	53	6.45
1850-1874	12.50	215230	269	8.54	185033	158	10.67
>1874	14.33	16048	23	–	951	0	13.53

Table 5.26. Use of the generic pronoun *maður* over generations split for gender, normalised per 10,000 words.

The generational differences in Table 5.26, Figure 5.32 reveal that the while there is a steady rise in frequency of *maður* for both genders, the rate is consistently lower for females than for males across all four temporal groups. The exact same effect obtains if, instead of generations, we use the time of writing, split as above for both genders into three time periods (not shown).

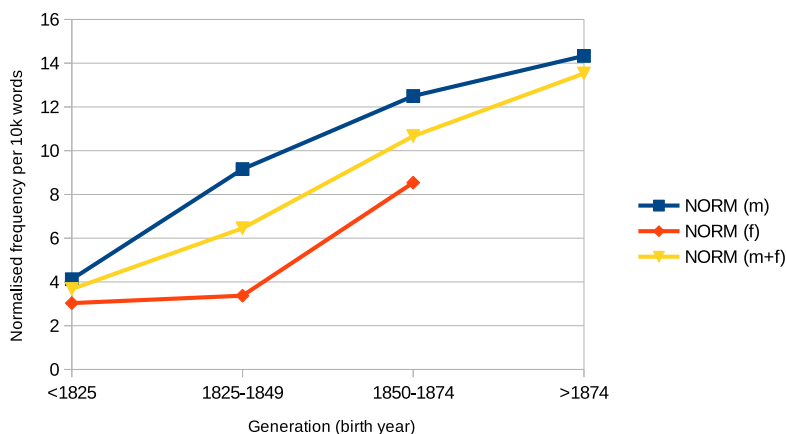


Figure 5.32. Use of the generic pronoun *maður* over generations split for gender, normalised per 10,000 words, cf. Table 5.26.

Table 6.32 further outlines the use of *maður*, with social status as an additional dimension (see Appendix C, page 253). Beginning with females, there is a downward trend for other professions over time, i.e. the variable period, based on year written. Moreover, females belonging to the category officials/lettered, again, by association through their fathers and/or husbands, use *maður* significantly less than females in the category peasants/labourers. This effect is significant for all three periods, P1 $p < 0.03$ (LL 6.08, OR 0.06), P2 $p < 0.01$ (LL 8.55, OR 0.04), P3 $p < 0.001$ (LL 12.23, OR 0.03). A similar effect obtains if we consider generations instead of time period, such that females from the higher echelons, born 1850 or later, also use *maður* markedly less than their peasant/labourer counterparts. For both the variables Period and Generation, females in the upper/mid-category of other professions outrank their counterparts in the high social category officials/lettered in their lesser use of the generic pronoun. Summing up, a discernible standardisation effect appears to be present for females during the period after 1850. Rather surprisingly, however, this effect does not carry through over time, as the frequency is *not* reduced diachronically in any of the social status categories besides other professions.

Moving on to consider male speakers, we observe that the use of *maður* appears to be *increasing* for the most part, except partly for the last period (Period 3) and last generation (G4) of Table 6.32. However, these slots in Table 6.32 are precisely the ones where data is simply much too thin to be representative. Contrasting males in the categories officials/lettered vs. peasants/labourers reveals a statistically significant difference only in Period 2, significant at the level of $p < 0.0001$ (LL 29.57, OR 0.03). That effect is thus exactly opposite that of the females in the corresponding groups, suggesting gender-related differences in the uptake of the standard norm, again such that females (in the higher echelons) are rather more successful than men in suppressing the feature.

Not considering gender as a variable, speakers in the category officials/lettered

	RAW FREQ.	WC	NORM. FREQ.
TIME PERIOD			
1852-1880 (HKF teaching alone)	29	24465	11.85
1881-1895 (HKF among others)	20	26233	7.62
1896-1906 (post-HKF)	51	31714	16.08
GRADUATION SCORE			
High (with distinction, score 1)	49	51574	9.5
Low (scores 2-3)	23	22518	10.21
Did not graduate	28	8320	33.65
GRADE			
Grades 1-3	56	27210	20.58
Grades 4-6	44	55202	7.97

Table 5.27. Essays: Generic pronoun *maður* based on three variables, raw frequency and normalised frequency (per 10,000 words).

interestingly lag behind the rest in Period 3 (bottom of Table 6.32, time written 1886 onwards), which is consistent with a standardisation effect, albeit rather subtle. The effect officials/lettered vs. peasants/labourers is significant for Period 3 at the level of $p < 0.0001$ (LL 21.21, OR 0.01). The difference between officials/lettered vs. other professions is significant across all three time periods, P1 at the level of $p < 0.05$ (LL 4.81, OR 0.16), P2 at $p < 0.01$ (LL 10.61, OR 0.09) and P3 at $p < 0.0001$ (LL 22.68, OR 0.02).

5.4.4 Student essays

A total of 100 examples of pronominal *maður* were found in the essays. While this boils down to one example per every other essay, on average, the pronoun was found in only 39 essays by 38 different students, whereas 150 essays by 132 different students did not contain the pronoun at all. Table 5.27 shows the distribution based on time period (defined roughly based on who did the teaching), graduation score and progression of study (grade). The frequency was normalised per 10,000 words.

As above, a log-likelihood test was used to test for statistical significance. The difference between the two time periods 1881-1895 and 1896-1906 was statistically significant ($p < 0.001$; LL 14.50), but the difference between 1852-1880 and 1881-1895 was not. However, it should also be borne in mind that essays from the 1st-3rd grade are relatively greater in number during that very period than essays from 4th-6th grade (see further below). In terms of graduation score, there is not a statistical difference between students with a high vs. a low grade, although there is, interestingly, a statistically significant difference between students who graduated vs. those who did not; the dropouts used *maður* much more than the rest ($p < 0.0001$; LL 124.89 and 50.59, respectively).¹⁵⁹ There is also a significant difference between 1st-3rd grade vs.

¹⁵⁹See Viðarsson (2017b) for further details.

4th-6th grade, indicating that use of the pronoun diminished as their study progressed. Importantly, this effect of GRADE between 1st-3rd graders vs. 4th-6th graders still holds even if we confine ourselves to the period 1896-1906 ($p < 0.0001$; LL 16.04), during which the frequency of *maður* was shown to increase.

In Table 5.28, the results have been split over six periods based on this contrast, grades 1-3 vs. 4-6. As we can see, there is insufficient data for grades 1-3 from the early periods, with an opposite effect for grades 1-3 vs. 4-6 between 1852-1864 (less *maður* in grades 1-3 than 4-6), suggesting perhaps that the distribution in that time slot may be random. However, in contrast, there is a rather neat GRADE effect for grades 1-3 vs. 4-6 in the periods 1885-1894, 1895-1900 and 1901-1906. Figure 5.33 provides a visualisation of the information in Table 5.28. We can thus conclude that where there is sufficient data, the distribution largely follows the sort of pattern one would expect to find if the uptake of the standard norm was being acquired gradually over time.

	<i>n</i>	GRADES 1-3		<i>n</i>	GRADES 4-6	
		WC	NORM. FREQ.		WC	NORM. FREQ.
1852-1865	2	1684	11.9	13	6803	19.1
1865-1875	-	-	-	3	7921	3.8
1875-1885	-	-	-	11	16798	6.5
1885-1895	7	3965	17.7	11	10116	10.9
1895-1901	23	12079	19	0	5120	0
1901-1906	24	9482	25.3	6	8444	7.1

Table 5.28. Essays: Generic pronoun *maður* by grade over six periods; raw frequency and normalised frequency (per 10,000 words).

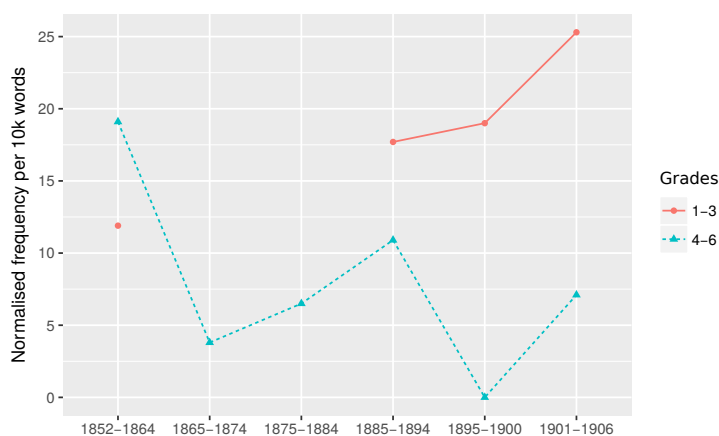


Figure 5.33. Essays: Generic pronoun *maður* by grade (year) over six periods, cf. Table 5.28.

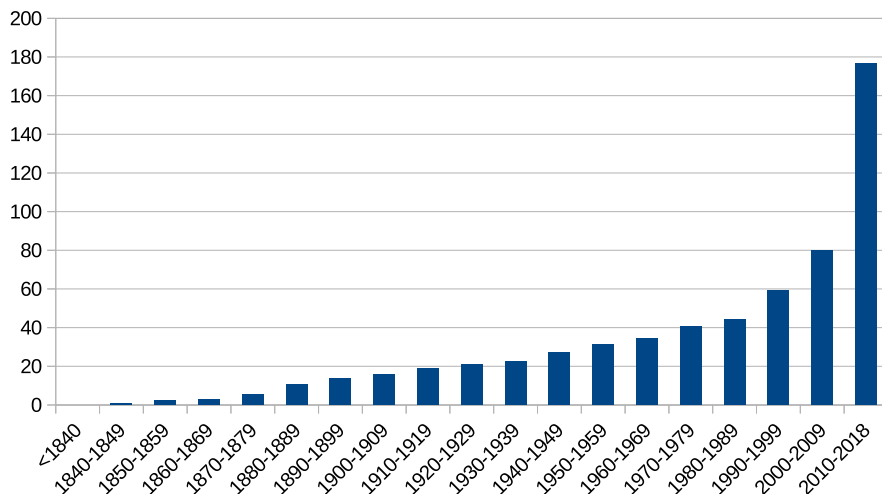


Figure 5.34. Frequency of *maður* normalised in newspapers from before 1840 onwards ($n = 511,991$) as a proportion of the raw frequency of the top 25 most frequent words (92,705,498 words in total) according to the Icelandic Frequency Dictionary. Unanalysed data set containing clauses introduced by *að* ‘that’, *sem* ‘which’, *hvort* ‘whether’, *ef* ‘if’ and *þegar* ‘when’, followed by *maður*, followed by an agreeing form of the verbs *geta* ‘can’, *hafa* ‘have’, *koma* ‘come’, *segja* ‘say’, *vera* ‘be’ and *verða* ‘must; will become’ (both IND. and SUBJ. forms).

5.4.5 Beyond the 19th century

As in the previous two case studies, the online *Tímarit.is* corpus was used for a rough approximation of the uptake of the norm, in this case the suppression of the univariate variable *maður*. Figure 5.34 shows the proportion of *maður* from before 1840 onwards based on a simple string search for patterns matching *að* ‘that’, *sem* ‘which’, *hvort* ‘whether’, *ef* ‘if’ and *þegar* ‘when’, followed by *maður*, followed by an agreeing form of the verbs *geta* ‘can’, *hafa* ‘have’, *koma* ‘come’, *segja* ‘say’, *vera* ‘be’ and *verða* ‘must; will become’ (both indicative and subjunctive forms). Only the nominative singular form of *maður* was considered, with finite verbal agreement both in the indicative and the subjunctive singular. The reader is thus warned that these data are unanalysed and may potentially contain instances of referential uses of the noun *maður*. Typically, the noun will occur with a modifier or determiner of some sort when a particular referent is singled out, greatly reducing the chance of false positives. Note furthermore that generic uses of the noun are always plural (*menn*), which are, therefore, excluded by design.

The normalised frequency of *maður* during the period under study (approximately 1920-1929 in Figure 5.34), is considerably at odds with the results reported for the smaller, disambiguated data set in the previous section. This trend is rather surprising given the fact that the generic pronoun appeared to be in clear remission during this period. The different method of collecting the data, in itself, is a potential source of these differences, at least in part, although it is unclear to me that it should translate into the observed pattern. A striking feature of this pattern is the exponential rise of

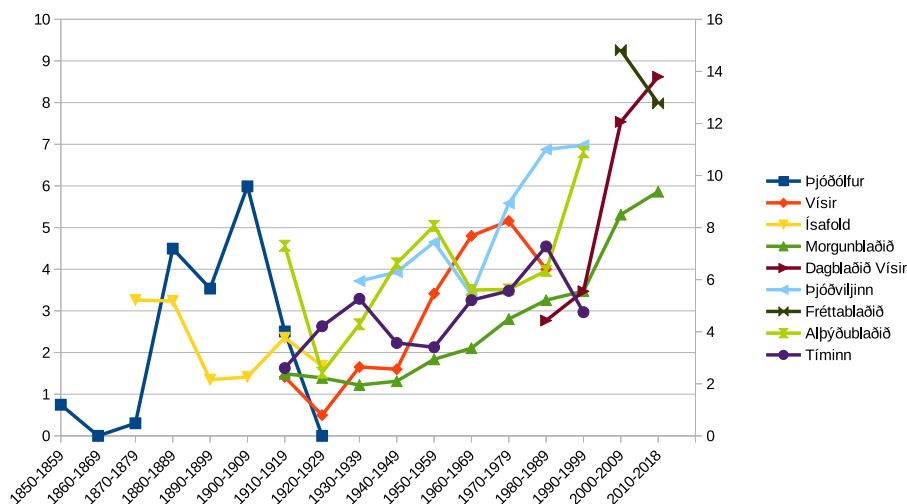


Figure 5.35. Frequency of *maður* normalised in selected newspapers from 1850 onwards ($n = 29,685$) as a proportion of the raw frequency of the top 25 most frequent words (51,720,463 words in total) according to the Icelandic Frequency Dictionary. Unanalysed data set containing clauses introduced by *að* ‘that’ and *sem* ‘which’, followed by *maður* (only NOM.SG.) and an agreeing form of the verbs *geta* ‘can’, *hafa* ‘have’ and *vera* ‘be’ (both IND. and SUBJ. forms).

maður in frequency from the 1840s until the 2000s and more than doubling in the 21st century from the 2000s to the 2010s. The distribution in the oldest newspapers is potentially skewed in unpredictable ways due to difficulties with the OCR used for automatic text extraction which have been corrected in the 19LCLV subcorpus. However, the most likely explanation for the gradual increase in the (supposedly) non-standard feature *maður* is the inclusion of a greater number of periodicals and newspapers in the *Tímarit.is* corpus, some of which may not adhere as strictly to the standard norms as in the 19LCLV subcorpus. The more recent growth in the late 20th and early 21st century may even suggest a shift in the norms in which the generic pronoun *maður* is becoming acceptable as a part of the standard, even in the relatively formal written modality.

Figure 5.35 presents the distribution of *maður* in selected newspapers over time from the 1850s onwards. Similar to 5.34, there is a clear upward trend to be seen in most of the newspapers and little evidence of standardisation effects, except as a potential retarding effect resulting in a temporary slowdown of the adoption of *maður*. The enormous fluctuations in frequency in *Þjóðólfur* are only partly statistical; the increase from a frequency of 0.3 to 4.5 from the 1870s to the 1880s is strongly significant (LL 15.70, $p < 0.0001$), but the further increase from the 1890s to the 1900s is not (LL 3.38). Neither is the decrease from the 1900s to 1910s from 6.0 to 2.5, presumably due to the rather limited size of the corpus.¹⁶⁰ A minor decrease can be seen in *Vísir* in the 1980s, *Tíminn* in the 1990s and *Fréttablaðið* in the 2010s. The rest of the titles reveal that the frequency of the generic pronoun is steadily rising, contrary to the prescribed standard norms.

¹⁶⁰ Also note that the paper only ran until 1920 so there is almost no data from the 1920s.

The rise in frequency in *Morgunblaðið* from 1990s to 2010s, for instance, is strongly significant (LL 305.13 and LL 10.99, respectively) and so is the contemporaneous rise in *Dagblaðið Vísir* (LL 578.67 and LL 12.92, respectively). Based on the assumption that strongly lexically-specified variables ought to be easier to manipulate and/or evaluate socially, this is clearly an unexpected result. Furthermore, the metalinguistic remark as to the effectiveness of norm implementation from the 1970s is nowhere to be seen in these data (see Section 5.2).

5.5 Discussion

The most natural interpretation of the data above is that pronominal *maður* was quickly catching on during the last quarter of the 19th, having long been a budding feature. This is further corroborated by clear generational differences. Based on the distribution across speakers of different social backgrounds, it also seems unlikely that *maður* was an innovation that had mostly been confined to educated circles (see also Section 5.6.4). The fact that *maður* was on the increase in the private letters during this period make the newspapers, which at the same time fail to show any signs of an increase, quite remarkable. The same effect continues in Period 5 (1900-1924), where the rate of *maður* continues to rise in the letters but actually decreases in the newspapers. While *maður* never got a chance to fully take off in this genre, it appears, unlike the private letters where it was flourishing, its rate of occurrence thus appears to have been successfully brought down to levels on par with what they had been a century earlier before *maður* had really caught on.

As concerns the female letter writes, it is striking that while the increase in the use of *maður* largely follows that of the males, the rate is considerably lower. It is possible that the lexical semantics of *maður* (lit. ‘man’) may be of some importance here, having obvious male connotations in terms of its referential properties. That being said, female speakers can and do use the generic pronoun in clear first person contexts, meaning ‘I’, despite the grammatical form obviously being masculine.¹⁶¹ The following examples exhibit this property:

- (166) a. *Maður er orðin svo gamall.*
 one is become so old
 ‘I have become so old.’¹⁶² (GdrJon-1900-04-09.xml)
- b. *Hjer þykir mjer mikið leiðinlegra að vera úti, en þegar jeg var heima;*
 here thinks me much boring_{CMP} to be outside than when I was home
maður er eitthvað svo ófrjáls, alltaf framaní gluggum eða þá fólki
 one is something so unfree
 ‘I think it is much more boring to be outside here than when I was at home;
 I am so unfree, always in front of windows or in people’s faces.’

¹⁶¹In my 19th-century data, I have only come across masculine agreement with *maður*. In Modern Icelandic, however, the agreement can sometimes/for some speakers be feminine in case the subject is female.

¹⁶²This speaker does not observe the standard spelling of *nn* for masc.sg. The apparent “fem.sg.” word form *orðin* is thus not unexpected for default/masc. gender agreement for this speaker, cf. also the unequivocal masc. form of *gamall* as opposed to fem. *gömul*.

(RagDan-1886-09-21.xml)

- c. *Gaman væri ad sjá einthvert af börnum ykkar með föður sinum í sumar*
 fun were to see some of children yours with father their_{REFL} in summer
ef maður lifði það!
 if one lived it
 ‘It would be fun to see some of your children with their father this summer
 if I would live to see it.’ (JakJon-1875-03-22.xml)

Potential interference of this type is thus clearly no absolute constraint, at most serving to limit the spread of this variant to speakers who tend not to associate themselves with or otherwise take issue with referring to themselves with an inherently/grammatically masculine referent. While fascinating, I must leave this aspect of variation to future research.

That *maður* was targeted at the Reykjavík Grammar School is confirmed here based on corrections in the student essays and it need thus not come as surprise that its rate of occurrence is low in the newspaper material. However, similar to what we observed for the *hinn*-only Reykjavík Grammar School norm in Chapter 4, there is a gradual increase in the use of *maður* throughout and beyond the 20th century. It is not unlikely that a part of this trend is explained by an increase in the proportion of quoted direct speech in interviews. Still, it seems unlikely that this explains the enormous rise during the 2010s, where *maður* more than doubles in frequency.

5.6 Towards an analysis

In this section I will first briefly discuss the structural analysis of *maður* in 19th-century Icelandic, adopting a feature decomposition analysis of the pronoun (cf. e.g. Holmberg and Roberts 2013). I will compare these properties to other Germanic languages following Fenger (2018), focusing on Danish. I then comment on potential uses with arbitrary reference, which is a marginal phenomenon in 19th-century Icelandic, at best. Finally, I offer a potential reconciliation between the use of *maður* in the singular and in the plural (*menn*) for arbitrary reference, seeking inspiration in ongoing (unpublished) work by Kjartansson (2017).

5.6.1 Feature decomposition

Let us assume with Holmberg and Roberts (2013) that definite pronouns such as the traditional inventory of personal pronouns are syntactically speaking feature bundles, composed of D, a definiteness feature, and ϕ -features, referring to person, number and gender. Being definite, these pronouns can be stated as [+D, ϕ], spelling out as e.g. *I/we, he/she/it/they, you* and so on. D-linked or anaphoric pronouns on this approach have an unvalued reference feature [*u*R], typically receiving its referential value from an antecedent. In other words, they are controlled by a referential NP/DP. The pronouns may also occur DP-internally, such as the proprial article in Icelandic (see Sigurðsson 2006). For pronominal elements to serve syntactically as subjects or objects, they additionally must have a nominal feature [N]:

- (167) a. *Maðurinn_i segir að hann_i viji panta vínarbrauð*
 the man_[N:+D, φ, R] says that he_[N:+D, φ, uR] wants order Danish-pastry
 ‘The man says he wants to order Danish pastry.’
- b. *Hann Jón vill panta vínarbrauð*
 he_[N:+D, φ, uR] John_[N:+D, φ, R] wants order Danish-pastry
 ‘John wants to order Danish pastry.’
- c. *Við Jón viljum panta vínarbrauð*
 we_[N:+D, φ, uR] John_[N:+D, φ, R] want order Danish-pastry
 ‘John and I want to order Danish pastry.’

I assume that *við Jón* ‘John and I (lit. we John)’ in (167c) is the output of a separate derivational layer, involving the complex *ég og hann* ‘I and he’. Adopting a traditional feature decomposition of personal pronouns, 1P.SG ‘I’ is [+participant] and [+speaker] (singular), 3P.SG ‘he’ being the absence of person, i.e. [-participant] (singular). The output of this derivation is then the input of another in a cyclic manner, resulting in [+participant, +speaker] (plural).¹⁶³

Generic pronouns, in contrast, are taken to be underspecified, not only in terms of lacking person features—recall that third person is no person, the absence of person—but lacking a D-feature as well (see Holmberg and Roberts 2013). In fact, Holmberg and Phimsawat (2017) argue that the generic pronoun in its inclusive sense (see Section 5.1) is the least specified and least restricted one, constituting a class of ‘truly minimal pronouns’.¹⁶⁴ In languages with agreement, the generic pronoun must have at least a minimal set of ϕ -features, triggering (default?) singular agreement despite being plural semantically. For reference to the speaker, the addressee and everybody else, we can assume underspecification of a ϕ -number feature: [\pm participant] (cf. Holmberg and Phimsawat 2017:27).

- (168) a. *Hér má *(hann) ekki borða vínarbrauð* (DEFINITE, SPECIFIC)
 here may he not eat Danish-pastry
 ‘He is not allowed to eat Danish pastry here.’
- b. *Hér má (maður) ekki borða vínarbrauð* (GENERIC)
 here may one not eat Danish-pastry
 ‘It/one is not allowed to eat Danish pastry here.’

In pro-drop languages such as Italian, definite pronouns as in (168a) can be dropped, presumably because agreement inflection has a definiteness component, an [μ D] feature, which get valued by the D-feature of the subject pronoun (Holmberg and Roberts 2013:120). In Icelandic, agreement does not carry a D-feature and, as a result, the subject pronoun in (168a) is obligatorily present. However, this property does not bar a null generic *pro* in (168b), which may ‘alternate’ with the overt generic pronoun *maður*. The lack of a D-feature in null/overt generic contexts like (168b) allows for either a null *pro* or an overt generic pronoun since the [$u\phi$] of agreement inflection on Fin (T

¹⁶³For the singular/plural distinction, one can adopt the feature [\pm group] (see Harley and Ritter 2002).

¹⁶⁴An alternative hypothesis is that (inclusive) generic pronouns are among the most richly specified of pronouns, specified for first, second and third person (for discussion and references, see Holmberg and Phimsawat 2017, Fenger 2018). Note that generic pronouns typically have an additional [+human] restriction (see also Sigurðsson and Egerland 2009).

a K-layer for case, a (featurally deficient) D-layer and a ϕ -layer, whereas the imp-N lacks this pronominal structure altogether. As is often assumed, going back at least to the work of Roman Jakobson, nominative case is treated as the absence of case.¹⁶⁵ This allows for imp-N to occur in nominative (i.e. caseless) contexts, whereas imp- ϕ can occur regardless of case.

Oblique forms of the pronoun occur since the beginning of the period under study. Most of these are internal arguments that raise to become oblique subjects, as in (171), but objective/non-subject surface uses are also amply attested in the data, cf. (172):

- (171) a. *Mundi ógjörlegt að grennslast eftir, hvort **manni** í þessu falli skeður*
 would.be impossible to inquire_{MID} PRT whether one_{DAT} in this case happens
rétt eða rangt? (IngJon-1812-08-15.txt)
 right or wrong
- b. *en ekki geingur altjend svo vel að ná verði bókkanna inn aptur; þó*
 but not goes always so well to get price books-the_{GEN} in again though
***manni** liggja á, efeithvað seldist* (BenJak-1855-04-06.xml)
 one_{DAT} lies_{SUB} on if something sold_{MID}
- (172) a. *Hvar fer betur um **mann** en heima*
 where goes better about one than home
 ‘What better place (for one) than home?’ (ArnHel-1816-04-30.txt)
- b. *En og svo börnin **manns** eru lánsfé drottins*
 but also children_{DEF} one’s are borrowed.funds lord_{GEN}
 ‘But one’s children are also the borrowed funds of the Lord.’
 (GudMag-1844-06-13.txt)
- c. *Eg vil ekki hafa það, að þú haldir líkræðu yfir mér, þú hrósar **manni***
 I will not have it that you hold eulogy over me you praise one
ekki.
 not
 ‘I do not want you to give my eulogy, you never praise me/one.’
 (ArnHel-1851-08-11.txt)

As expected based on these properties, *maður* also occurs in passives, where the generic pronoun receives dative case:

- (173) a. *hvert að hann gjeri rjett eða rángt með að fara fra yslandi seigi*
 whether that he does_{SUB} right or wrong by to go from Iceland say
*jeg ekkert um, **manni** er gjefið frjals ræði að velja og hafna*
 I nothing about one is given freedom to choose or refuse
 (BenHal-1883-05-20.xml)

Another important feature of *maður* that sets it apart from Danish *man* is the ability for *maður* to occur in ECM infinitives with raising to object of the embedded subject of an

¹⁶⁵See Neeleman and Weerman (1999:66, 69, 82) on how this hypothesis can be reconciled with languages such as Icelandic that (appear to) have a case suffix for the nominative. These authors suggest that the nature of case suffixes in Icelandic, being fusional, blurs the picture so that *-s* may be said to be associated with features such as <masculine declension group I, singular, genitive>, whereas there is no need to regard *-ur* as a nominative case suffix. Instead, these can be regarded as unspecified for case, the classification of *-ur* being only <masculine declension I, singular>.

infinitival complement. It should be mentioned that the private letter data set does not contain much *maður* data in ECM environments. However, as the data are rather limited in size to begin with, this is presumably not unexpected. Most of the examples occur as complements in *let*-infinitivals (i.e. with *láta*), an exhaustive list being provided in (174):

- (174) a. *það sem í einu orði lætur mann líða uppí þriðja himin*
 that REL in one word lets one_{ACC} glide up.into third heaven
 (BalEin-1826-08-14.txt)
- b. *Það er eins og að amíríku lífið ætli að láta mann gleima öllum*
 it is as if America life-the intends to let one_{ACC} forget all friends
vinum og vandamönnum heima á Froni
 and relatives home on Iceland
 (BenHal-1881-10-23.xml)

In addition, there are two examples where *maður* raises out of the infinitival clause to object position, occurring as the (dative) complement of *kenna* ‘teach’ in (175a) and the complement of the preposition *fyrir* ‘for’ in (175b):

- (175) a. *peningaleysið kjennir manni að vera nitin og hirðusamur*
 lack.of.money-the teaches one_{DAT} to be economical and orderly
 (AdaBja-1884-07-29.xml)
- b. *bandaríkin eru að jeg held það besta pláss í heimi fyrir mann að lifa*
 united.states-the are that I think the best place in world for one to live
og ala upp börn sín
 and raise up children his_{REFL}
 (AdaBja-1905-01-24.xml)

The examples in (171)-(175) thus attest to the fact that *maður*, also in terms of its syntactic properties, does not pattern with Danish and that it had (or had acquired) all of its present properties already in the 19th century.

5.6.3 Arbitrary reference?

Observe that in most, if not all, of the examples provided above, *maður* has had a straightforward inclusive reading. It is not entirely clear whether *maður* in 19th-century Icelandic was always inclusive (i.e. [+speaker]) or whether it could, perhaps marginally, be truly arbitrary in the sense above. Many examples are strictly speaking not quite unambiguous. Quite likely candidates occur in contexts where sketching a situation abroad, in which the speaker need not participate at all:

- (176) On the way off of the mountain that evening, we heard the mountain dairy girls from all the surrounding dairies “allure” their cows. It is namely tradition all over Norway to summon the cows, sheep, horses, goats and pigs, always walking ahead of the creatures and have them follow **one**(self). There’s a special song to “allure” each of the species and it is very fun to hear beautiful “alluring”.
- Á leiðinni ofan af fjallinu um kvöldið heyrðum við selstúlkurnar úr seljunum alt í kring „lokka“ kýrnar sínar. Það er nfl. siður alstaðar í Noregi, að það er kallað á kýrnar, kindurnar, hestana, geiturnar og svínin og altaf gengið á undan skepnunum og þær*

látnar elta mann. Það er sérstakt lag við að „lokka“ hverja tegund fyrir sig, og er mjög gaman að heyra fallega „lokkað“. (Kvennablaðið 1900-05-01 (5. tbl. 6. árg.).txt)

If these truly exclude the speaker, an example like (176) would thus qualify as [-speaker, ±participant]. There are also examples where the [±participant] reading is prominent, although these presumably always include the speaker as well:

- (177) *við höfðum sætt kaffe og bollur áðan, á eftir miðdegis matnum, sem var kjöt súpa [...].* *jeg held að maður megi annars hætta að drekka kaffe, því hvergi fæst nokkurs konar sikur.* (AnnGud-1879-01-19.xml)
- we had sweet coffee and buns earlier after midday food-the which was meat soup I think that one may otherwise stop to drink coffee because nowhere obtains any sorts sugar
- ‘We had sweet coffee and buns just a while ago after lunch, which was meat soup. I think (by the way / come to think of it) that one should stop drinking coffee, because there isn’t any sort of sugar anywhere.’

I take it that *maður* with arbitrary reference is a marginal phenomenon in 19th-century Icelandic, if it occurs at all. The typical way to denote arbitrary reference is with a null 3rd person impersonal subject or expletive constructions, as in Modern Icelandic (see Sigurðsson and Egerland 2009). As I will discuss in the following section, a third option may exist which arguably completes the overlap between *maður* and impersonal *pro* (contra Sigurðsson and Egerland 2009), involving the plural pronominal *menn* ‘(generic/arbitrary) people’.

5.6.4 A note on the origin of pronominal *maður*

Much traditional scholarship treats pronominal *maður* as a Danish borrowing. In previous sections, we have seen that Icelandic *maður* and Danish *man* are clearly grammatically distinct from one another, not only in terms of their form and function but also as in their reference. While that does not rule out foreign influences, facts of this kind certainly undermine such an account. Moreover, there is also recently discovered distributional evidence that pronominal uses of *maður* actually go back to the Old Norse period (cf. Kjartansson 2017). I will briefly comment on this below and offer a potential reinterpretation of Jónsson’s (1992) three-stage developmental path of *maður* and Sigurðsson and Egerland’s (2009) claim that pronominal *maður* cannot be treated as the overt equivalent of the covert (null) impersonal *pro*.

Kjartansson (2017) has recently pointed out that not only singular but also plural instances of *maður* are unstressed in their generic senses, suggesting that both may in fact be pronominal in nature in Icelandic. Using known metrical constraints to distinguish between nominal and pronominal status, Kjartansson (2017) shows that indefinite *maður* in the singular and plural could occur in unstressed metric positions already in Old Norse poetry, which nouns otherwise never do (metrical lifts here denoted by boxes, alliteration by underlining; cited after Kjartansson 2017):

- (178)

<u>Vin</u>	sínum
------------	-------

 // skal **maður**

<u>vinr</u>	vera
-------------	------
- friend his_{REFL} shall man friend be
- ‘Man/one_[unstressed] should be a friend to his/ones friend.’ (Háv.)

- (179) hrís það ið mæra // er **meðr** Myrkvið kalla
 forest that the famed REL men Darkwood call
 ‘The famed forest which men/people_[unstressed] call Darkwood.’ (Akv.)

As shown in (178)-(179) for singular *maðr* ‘man/one’ and plural *meðr* (= *menn*) ‘men, people’, respectively, the boldfaced elements both occur in dips, each in the upbeat of a second half-line. What this strongly suggests is, on the one hand, that *maðr* could serve as a pronoun already in Old Norse and, on the other, that the pronominal status may not have been confined to the singular but also the plural. Similar evidence has been provided for Old English indefinite *man*, which, too, has been shown to exhibit properties of a pronoun rather than a full nominal (see Van Bergen 2003; cf. also Walkden 2014:72, fn. 5).

Evidence of this kind strongly suggests that pronominal uses of *maður* cannot reasonably be regarded as a borrowed feature. Rather, the D-feature-less (=generic) pronoun will have been derived from the native nominal root $\sqrt{maður}$, presumably by extension of the plural to the singular. The arbitrary (or arbitrary/generic) reading of the unstressed *meðr/menn*_{PL} is closest to the original full nominal, at least in terms of its lexical meaning. Being in this sense the “nouniest”, I take it that the initial stage in the development of a generic pronoun would have been one in which *maður* was a full noun, with the pronominal *menn* denoting generic/arbitrary reading. Being an extension of the plural noun, one could imagine that Stage I lacks the singular generic pronominal *maðr/maður*. From here, I then tentatively assume a Stage II with the singular generic pronominal, deriving from the Stage I pronominal plural *meðr/menn*. Stage III further develops the singular generic *maður*, specified only for [+speaker], dropping [participant].¹⁶⁶

- | | | | | | | | | | | | | | | | | | | | | | | | |
|--|--|---------|------------------------|-------------------------------------|-------------------------------|--|--|----------|-------------------------|-------------------------------------|-------------------------------|--------------------------------|------------------|--|--|-----------|---|--|-------------------------------|---------------------------------|------------------|---------------------------------|-------|
| (180) | <table style="border-collapse: collapse; width: 100%;"> <tr> <td style="padding-right: 20px;">STAGE I</td> <td>[step 1 by hypothesis]</td> </tr> <tr> <td style="padding-right: 20px;"><i>meðr/menn</i>_{PRON.PL}</td> <td>(“people”, generic/arbitrary)</td> </tr> <tr> <td colspan="2"> </td> </tr> <tr> <td style="padding-right: 20px;">STAGE II</td> <td>[attested in Old Norse]</td> </tr> <tr> <td style="padding-right: 20px;"><i>meðr/menn</i>_{PRON.PL}</td> <td>(“people”, generic/arbitrary)</td> </tr> <tr> <td style="padding-right: 20px;"><i>maðr</i>_{PRON.SG}</td> <td>(“one”, generic)</td> </tr> <tr> <td colspan="2"> </td> </tr> <tr> <td style="padding-right: 20px;">STAGE III</td> <td>[attested in 19th-century Icelandic]</td> </tr> <tr> <td style="padding-right: 20px;"><i>menn</i>_{PRON.PL} (<i>fólk</i>)</td> <td>(“people”, generic/arbitrary)</td> </tr> <tr> <td style="padding-right: 20px;"><i>maður</i>_{PRON.SG}</td> <td>(“one”, generic)</td> </tr> <tr> <td style="padding-right: 20px;"><i>maður</i>_{PRON.SG}</td> <td>(“I”)</td> </tr> </table> | STAGE I | [step 1 by hypothesis] | <i>meðr/menn</i> _{PRON.PL} | (“people”, generic/arbitrary) | | | STAGE II | [attested in Old Norse] | <i>meðr/menn</i> _{PRON.PL} | (“people”, generic/arbitrary) | <i>maðr</i> _{PRON.SG} | (“one”, generic) | | | STAGE III | [attested in 19 th -century Icelandic] | <i>menn</i> _{PRON.PL} (<i>fólk</i>) | (“people”, generic/arbitrary) | <i>maður</i> _{PRON.SG} | (“one”, generic) | <i>maður</i> _{PRON.SG} | (“I”) |
| STAGE I | [step 1 by hypothesis] | | | | | | | | | | | | | | | | | | | | | | |
| <i>meðr/menn</i> _{PRON.PL} | (“people”, generic/arbitrary) | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | |
| STAGE II | [attested in Old Norse] | | | | | | | | | | | | | | | | | | | | | | |
| <i>meðr/menn</i> _{PRON.PL} | (“people”, generic/arbitrary) | | | | | | | | | | | | | | | | | | | | | | |
| <i>maðr</i> _{PRON.SG} | (“one”, generic) | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | |
| STAGE III | [attested in 19 th -century Icelandic] | | | | | | | | | | | | | | | | | | | | | | |
| <i>menn</i> _{PRON.PL} (<i>fólk</i>) | (“people”, generic/arbitrary) | | | | | | | | | | | | | | | | | | | | | | |
| <i>maður</i> _{PRON.SG} | (“one”, generic) | | | | | | | | | | | | | | | | | | | | | | |
| <i>maður</i> _{PRON.SG} | (“I”) | | | | | | | | | | | | | | | | | | | | | | |

This scenario arguably involves the smallest step conceivable in the development of a full-fledged generic pronoun, although the validity of each of these stages remains to be demonstrated. However, what is certain at this point is that 19th-century Icelandic is squarely situated at Stage III, featuring plural *menn* for generic/arbitrary reference, singular *maður* as a generic pronoun ([±speaker]) and pronominal *maður* in the singular referring only to the speaker ([+speaker, –participant]).

¹⁶⁶I am leaving out *maður* solely as a full nominal (assumed by Jónsson 1992), call it Stage 0, as there is no attested stage lacking a generic pronoun, judging by the evidence presented by Kjartansson (2017).

The inclusion of the pronominal plural *menn* above is of importance also to the claim that the generic pronoun *maður* cannot be analysed as the overt realisation of null impersonal *pro*, because the latter is not a proper subset of the former (Sigurðsson and Egerland 2009). If the plural form of pronominal *maður* is the one that corresponds to the truly generic and arbitrary readings of null impersonal *pro*, it becomes much less clear that pronominal *maður* cannot be regarded as the overt realisation of null impersonal *pro*. The main difference between pronominal *maður* (sg., pl.) and null impersonal *pro* is the first person ('I') reading of singular *maður*, which the null impersonal *pro*, to the best of my knowledge, does not allow. I would therefore like to suggest an account in terms of spell out of different feature bundles, consisting of one and the same root $\sqrt{\text{ma}\check{\text{d}}\text{ur}}$, spelling out either as the singular *maður* or the plural *menn* (later also *fólk*):

- (181) a. GENERIC/ARBITRARY \rightarrow *menn* 'people (in general)'
 b. GENERIC \rightarrow *maður* 'one'
 c. INCLUSIVE/SPECIFIC \rightarrow *maður* 'I'

One potential complication is that in addition to the referential properties with regard to the generic specific vs. generic arbitrary contrast (cf. above), the null impersonal *pro* can also refer to non-human referents in contrast to *maður* (see Sigurðsson and Egerland 2009).¹⁶⁷ However, as for the [+human] property of *maður*, I simply take this feature to be a direct consequence of the root ($\sqrt{\text{ma}\check{\text{d}}\text{ur}}$) and I would therefore not expect impersonal *pro* inherently to have this property. While I cannot pretend to have a full account of *maður* and null impersonal *pro*, for instance as regards the arbitrary reference of the alternative root $\sqrt{\text{fólk}}$, which obviously cannot be subsumed under *maður*, I still hope to have paved the way for a unified account of pronominal *maður* and impersonal *pro*.

5.7 Summary

The case study on the generic pronoun *maður* on the whole reveals a steady *increase* in the use of this non-standard feature for males and females in the private letter corpus, although the female scribes lag somewhat behind the males. As above, there are rather striking differences between the newspapers and the student essays on the one hand, exhibiting visible effects of standardisation, and the private letters on the other, where the effects are much more subtle. However, once social status is taken into account as well it becomes clear that the higher echelons participate in this increase to a far lesser degree, especially the females. This retarding effect might thus be interpreted as a potential standardisation effect, it is argued. Correspondingly, the fact that newspapers

¹⁶⁷This last point is perhaps contradicted by non-generic uses of *maður* in contexts like the following, that may be uttered in a context where, say, a new-born kitten is being breastfed by its mother:

- (1) *Æ, mikið er maður lítill!*
 o much is one small
 'Oh, look at how small (it is/I am)!'

This utterance sounds perfectly natural but requires viewing the world from someone else's perspective.

and periodicals do not participate in the temporal increase observed in the private letter corpus, coupled with a subsequent decrease in the newspapers towards the end of the period, suggests that the generic pronoun is being suppressed.

In other words, all three variables under study appear to exhibit social embedding and, similarly, for all three, there is at least some evidence that standardisation was partly effective. However, what is not corroborated by these results is the expectation of Labov's Interface Principle, that the least lexically-specified variant ought not to exhibit sociolinguistic patterns, but only (or mainly) the most lexically-specified one. On the contrary, these results reveal quite the opposite pattern, which I take broadly to suggest that even rather abstract linguistic/syntactic phenomena need not be below the level of conscious awareness.

The syntactic properties of the generic pronoun *maður*, finally, make it clearly distinct from the Danish generic pronoun *man*, as was to be expected based on its present-day distribution. It appears that the range of uses of *maður* mostly match that of Modern Icelandic and the pronoun typically has a straightforward inclusive reading, referring to the speaker, the addressee and everybody else. It is possible that there may be cases where the speaker is excluded, but if these occur, they are clearly marginal.

6 Summary and conclusions

The present study consisted of three case studies on language variation and change in 19th-century Icelandic, targeting the verbal/inflectional domain, the determiner system and the generic pronoun *maður*. The main points can be summarised as follows. The first variable, VERB-ADVERB PLACEMENT, revealed rather different social embedding of the Adv-Vfin variant than what is to be expected based on the traditional literature on the subject. Rather than being mostly confined to the higher echelons, the Adv-Vfin order was found to be attested in all social groups, and the absence of variation was typically not the norm, albeit true that educated writers belonging to the higher echelons tended to use this variant to a far greater extent than, say, the peasants/commoners. However, this was partly also true for the females, suggesting that we are dealing with variation not directly related to formal education, whether abroad or otherwise. Rather, the data suggest the relevance of social status more generally, presumably related to joint engagement in similar practices or styles.

Among the contributions of the present work is the unearthing of evidence of a social stigma attached to the Adv-Vfin order during the 19th century that goes far beyond a mention in passing in the *Fjölirnir* book reviews (Gíslason 1844) and Smári's (1920) Icelandic syntax. The feature was corrected by Icelandic teachers and examiners in student assignments at the Reykjavík Grammar School at least since the 1860s and tellingly characterised in a review as having "some rotten taste of Danish" (Gunnarsson 1878). The evidence provided by corrections in the student essays underscores in particular the importance of taking stigmatisation into account when studying the use of the Adv-Vfin variant, especially in edited, printed texts.

In terms of standardisation and sociolinguistic patterning, we have witnessed a reduction in the use of the non-standard Adv-Vfin variant in the letters of both males and females from the higher ranks, but interestingly only for the females in the peasants/commoners group. This partly also mirrors a corresponding sex/gender effect in Modern Icelandic where male participants were shown to accept Adv-Vfin more readily in the Icelandic Dialect Syntax project than female participants (cf. Þráinsson and Angantýsson 2015). Moreover, in good accord with the predictions of the Labov-inspired model assumed in the studies of Elspaß (2005b, 2007, 2012, 2016) and other related work, there is a stark contrast between the uptake of the standard norm in the private letters, both overall and among the educated, as opposed to the much greater effect visible in the newspaper corpus and the student essays. The newspapers show a very sharp decline of the non-standard Adv-Vfin order, a trend which was shown to continue into modern times. The student essays, furthermore, exhibit a clear temporal decrease in the use of Adv-Vfin towards the end of the 19th and beginning of the 20th century. There is also a significant effect of the educational variable GRADUATION SCORE such

that a low grade correlates significantly with a higher rate of Adv-Vfin, an effect which is even stronger for those who failed to graduate, the dropouts.

However, a voice of warning is also in order with regard to this variable. In addition to the sociolinguistic evidence, i.e. the social embedding of verb-adverb placement, there is also linguistic evidence suggesting that there were two systems of Adv-Vfin. On the one hand there was a grammar sensitive to an ASSERTION feature, but on the other hand we find a grammar where Adv-Vfin is mainly to be found in operator contexts such as relatives (the latter more similar to Modern Icelandic). The former of these is likely to be what in Labovian linguistics is referred to as TARGETED CHANGE originating in the higher classes, though not confined to those, presumably innovated on top of rather than alongside the latter type of grammar, that I suggest arose due to structural ambiguities related to Stylistic Fronting and the loss of null subjects. The targeted change is a likely candidate for Danish influences, especially given its sensitivity to the ASSERTION factor, similar also to Faroese. It seems likely that if the standard would have designated Adv-Vfin as the norm, and not Vfin-Adv, this targeted change could possibly have spread to a greater part of the population, as suggested by Pettersson (1988). That may indeed be what happened in Faroese where the Adv-Vfin variant has gradually been taking over.

As Heycock and Wallenberg (2013) have independently shown based on a different data set, the distribution of Adv-Vfin/Vfin-Adv in Icelandic between 1600-1850 coincides partly with the output of a grammar that lacks V-to-I. By adopting Heycock et al.'s (2012) more accurate operationalisation of embedded V2 vs. non-V2 contexts, we were able to not only corroborate but also strengthen this finding. Environments typically argued to have a restricted left periphery (here: indirect questions, relatives and conditionals) systematically exhibit greater use of Adv-Vfin than environments typically argued to have a more elaborate structure (here: declaratives, result, cause and consequence of degree clauses). Moving beyond these distinctions, we could furthermore show that this contrast also manifests itself within declaratives. Declarative clauses that occur in asserted contexts, i.e. when embedded under inherently negative verbs of Hooper and Thompson's Class C and factive Class D, have a much greater disposition for Adv-Vfin than in asserted contexts. This distribution strongly suggests competition between a V-in-situ/V-to-C grammar and a V-to-I grammar.

A final point regarding verb-adverb placement from the perspective of language standardisation that I would like to emphasise is the potential evidence for hypercorrection in 19th- to 21st-century newspapers based on the *Tímarit.is* corpus. As briefly discussed in Section 3.4.1.4, the rate of the non-standard Adv-Vfin variant decreases throughout the whole period, reaching near absolute zero during the 2010s. Recall that Adv-Vfin, as defined here, involves subject-initial embedded clauses with sentence-medial adverbs. A different variable, namely Stylistic Fronting, is similar to Adv-Vfin, except that it lacks an overt subject. Both may thus involve the configuration *C ... (ADV) VFIN (ADV)*, where '...' is either a subject gap (=Stylistic Fronting) or an overt subject (=our verb-adverb placement variable). At the same time that Adv-Vfin is receding in the newspapers, so does Adv-Vfin in Stylistic Fronting contexts as well. This sort of side-effect might be taken to suggest that some speakers are actually treating the syntactic variable purely as a linear surface phenomenon, independent of the different derivations giving rise to these structures. Whether the dramatic drop in the frequency

of Stylistic Fronting is due solely to other factors, I must leave for future research. If the effect is real, a more general conclusion might be that at least as long as we can treat variable aspects of syntax as ‘constructional’ at some level of representation (perhaps in the sense of Zwart 2015), these arguably can always be evaluated socially.

A working assumption of many approaches to language variation and language change that adopt the overall Chomskyan approach to syntax assumed in this work is that there are definable limits to variation, patterns that we never find (see e.g. Bobaljik 2002, Cinque 2005). A number of such aspects have been discussed with regard to the variation we find in 19th-century Icelandic, most importantly, perhaps, with respect to the existence of Adv-V_{fin} or verb-third (V3) constructions. The RICH AGREEMENT HYPOTHESIS has it that such patterns ought never to arise in a richly inflected language like Icelandic, yet for some speakers, it may even be regarded as the dominant pattern. For this reason, I have proposed to regard the RAH not as a universal statement but a Type 4 Greenbergian or statistical implicational universal. As a result, the RAH is neither to be formulated *weakly* (unidirectionally) nor *strongly* (bidirectionally), but rather *softly*, as a violable or “soft” constraint, the explanation of which must be viewed from a diachronic, usage-based perspective.

Importantly, this conclusion in no way invalidates a potential relation between morphology and syntax along the lines argued by Bobaljik and Thráinsson (1998). Various other syntactic consequences of having a Split IP (cf. Thráinsson 1996, Bobaljik and Thráinsson 1998, Bobaljik 2002, Thráinsson 2010), taken here to involve a uP_n feature on Fin, are arguably just as real and exciting. Rich (referential/argumental) agreement, then, arguably does have syntactic effects in the form of features in the numeration that will need to project, giving rise to additional (potential) spell-out positions. However, whether this will actually force the verb to be pronounced in the highest position or somewhere lower in the structure is a separate issue entirely, a matter of externalisation (cf. e.g. Chomsky 1995, 2008, 2013).

The second variable, the VARIABLE FREE-STANDING DEFINITE MARKER, is arguably most representative for the exceptional successes of the Icelandic standardisation enterprise in that the standard norm is argued to have introduced an element to the pool of variants—the supposed revival of an extinct pattern based on Old Norse. However, on closer inspection, it actually turns out that the standard free-standing determiner still appears to have existed before we would expect standardisation effects to become visible, and these happen to be found precisely in the sorts of contexts in which they still primarily occur in Modern Icelandic, viz. with evaluatives. Precisely where the use of *hinn* in the emerging standard does not appear to have been based on actual use, i.e. going against the spoken idiom such as where it has clear referential properties, usage in standardised works such as newspapers gradually fell back on a more vernacular-like norm featuring *sá* during the 20th century. In contrast, *hinn* still remains a standard norm that is frequently observed in evaluative contexts.

If the actual attempted implementation at the Reykjavík Grammar School is studied in more detail on the basis of the student assignments (the corpus of corrections), we see that the goal was much more ambitious than what the use of the standard free-standing article in Modern Icelandic leads one to expect. The Grammar School norm was an attempt to eradicate *sá* as a determiner in most of its contexts, even those where it refers to the member of a set and where the referent is established via a relative clause

following the determiner. These contexts would rarely be used with the Reykjavík Grammar norm in present-day Icelandic, suggesting in turn that the standardisation, as attempted, actually failed miserably. It is thus not so much language use that shifted towards the standard but rather the other way round: where *sá* and *hinn* were already in competition, *hinn* was the one selected and implemented in standardised texts. Where *hinn* as a part of the emerging standard appears to have had no backing in actual use, *sá* prevailed.

Looking at the private letters, there is but little effect to be detected in terms of the uptake of the standard norm as far as the definite article is concerned. However, it would be very misleading to suggest standardisation had no effect. The private letters show some effects towards the gradual uptake of the standard norm, discernible for the higher ranks, males and females, although the latter lag behind the former. For peasants/labourers, the males exhibit an increase in the use of the non-standard norm, whereas there is a slight increase among the females in the use of the standard norm. Most effects, unsurprisingly, are visible in the newspapers and the student essays, although there is also later evidence from the larger *Tímarit.is* newspaper corpus that the effect was largely temporary. A dramatic overall decrease over time was shown for the non-standard *sá* in the 19th-century newspapers.

As discussed in 4.3.5, the same strong effect obtained in the larger *Tímarit.is* corpus, again covering the 19th to 21st century. By contrasting entities that typically single out one member of a set corresponding to ‘the former/the latter’ (*sá/hinn fyrrnefndi/síðarnefndi*) with phrases that merely single out salient entities (or involve evaluatives) corresponding to ‘the so-called’, ‘the famous’ and ‘the deceased’, we were able to discern two striking patterns. While the latter group (i.e. ‘the so-called’, ‘the famous’ and ‘the deceased’) pattern together and predominantly occur with *hinn*, the former group (i.e. ‘the former/the latter’) did so in the 19th century, but increasingly occurs with *sá* throughout the 20th and 21st centuries. While all of these predominantly occurred with standard *hinn* during most of the latter half of the 19th century, in the 90-100% range, only the former group does in 21st-century newspapers, the latter group occurring with *hinn* roughly 20-25% of the time. As mentioned above, this suggests that there may have been changes in the norms, the 19th-century distribution being more similar to the 19th-century Reykjavík Grammar School norms which basically required *hinn* across the board.

The student essays also exhibit a clear temporal effect. Moreover, there is a significant correlation between the uptake of the standard norm and GRADUATION SCORE, with lower scores correlating with more use of the non-standard norm than the higher scores. Interestingly, in some agreement with the later 20th-century newspapers, the student essays towards the end of the period (1900-1906) suggests a temporal increase in the use of *sá*. The partly similar trends observed in these two types of material clearly merit further research.

In terms of syntactic aspects, we have seen that 19th-century Icelandic allowed for variation far exceeding what we find in the Modern Icelandic noun phrase, requiring extensions and adaptations of recent analyses proposed for variation at this level. The most interesting findings clearly pertain to the symmetry breaking between definiteness and the weak/strong inflection of the adjective, sometimes also referred to—misleadingly, as we have seen—as the indefinite vs. definite inflection. This antisymmetry is argued

to correlate with more positions being available in the NP/DP constellation, out-scoping DEF, while still occurring within the confines of the extended noun phrase, somewhat contradicting Pfaff (2015). The presence of double definiteness in 19th-century Icelandic with evaluatives, too, is unexpected on the analysis of these in terms of intervention (cf. Ingason 2016), although there appears to be no obvious alternative—an aspect which I must leave for future research.

The third and final variable, the GENERIC PRONOUN *maður*, reveals on the whole a steady increase in the categorically non-standard use of the generic pronoun over time. However, the social dimension does suggest a partial effect, similar to the above, revealing that the higher ranks participate in this increase to a lesser extent. That standardisation effect applies to both male and female writers, although less so for the males than for the females. Moving on to the newspapers and the student essays, we find some evidence that the increase observed in the use of the generic pronoun in the private letters over time is not present in these other text types, suggesting a retarding effect. Similarly, the student essays exhibit a significant effect for the GRADE variable, indicating that there is both decrease over time and a correlation such that the lower grades show more use of the generic pronoun than the higher grades.

While the main motivation for rejecting *maður* as a generic pronoun was found in the (alleged) fact that it was of Danish origin, this conclusion is not forced upon us by any means. For one, the distributional properties of the generic pronoun *maður* are very different from *man*. As far as its inflection and possible syntactic position is concerned, Icelandic *maður* inflects for case and may be a subject or an object, whereas the corresponding pronoun *man* in Danish is invariant in form and is reserved for subject position/nominative case. Only as regards arbitrary reference is there some potential convergence with Danish, but even in 19th-century Icelandic, arbitrary reference appears to be marginally attested, if at all. Moreover, based on recent research by Kjartansson 2017, pronominal uses of *maður* have old roots, going back to the medieval period. At best, Danish *man* will thus have served to reinforce pronominal uses of *maður* in the singular.

One of the interesting theoretical goals of the present study was to use the three linguistic variables studied in Chapters 3, 4 and 5, each targeting a different aspect of Icelandic syntax, to address Labov's INTERFACE PRINCIPLE. Ranging from a relatively abstract phenomenon such as embedded verb-adverb placement, to the more lexically specified definite determiner and the fully specified generic pronoun *maður* (at least as operationalised here), we could have expected there would be a continuum of sorts, with social embedding and deliberate top-down language standardisation most visibly present in the latter but not, or less robustly, in the former. However, taking Labov's hypothesis to an extreme, we could also have expected syntax not to be socially evaluated at all, linguistic awareness being limited to the surface—the words and the sounds of the language (cf. Labov and Harris 1986).

We have seen evidence that the Interface Principle must be overstated and that syntactic variation can be and is socially evaluated. Moreover, syntactic variation has social correlates and appears to operate in ways similar to what has been observed for other variables, including (morpho)phonological ones. This is precisely the sort of result expected on the socio-syntax approach to syntactic variability (cf. Cornips and Corrigan 2005a, Cornips 2015 and much related work) and in good agreement

with studies on variation in Icelandic over the past years and decades, where social background and/or stigmatisation have been shown to be relevant factors in variation at the syntactic and morphosyntactic level (see e.g. Sigurjónsdóttir and Maling 2001, Jónsson and Eyþórsson 2003, Þráinsson et al. 2013, Þráinsson, Eyþórsson, Svavarsdóttir, Blöndal 2015, Þráinsson and Angantýsson 2015).

To briefly recap, we have seen that high social status tends to correlate with more use of standard forms and vice versa, low(er) social status tends to correlate with more non-standard features (Adv-Vfin is partly an exception to this pattern, being targeted change to an extent). However, the most interesting result is that this correlation follows familiar gender-based patterns: where we see effects in the ego-documents, these hold for the higher ranks, males and females, but among the lower classes, typically only for the females. That result thus quite closely matches Labov's Gender/Conformity Paradox, a surprising result given the danger of anachronisms when applying the Uniformitarian Principle to earlier times, since females at the time typically received little or no formal education of the relevant sort. However, indirect effects through exposure to norms and standardisation ideology may still be expected to occur and if women tend to be more sensitive to deviations in a moderate range than males (cf. Labov et al. 2011), there need not be any contradiction here.

In terms of syntactic aspects, the distribution of variants among social groups is clearly not random but tends to follow familiar sociolinguistic patterns, especially in terms of the sex/gender of the speaker and their social status, but also with regard to generational differences and geographical distribution. This conclusion better aligns my findings with the socio-syntax tradition which hypothesises that syntactic variation *always* has sociolinguistic correlates, a central tenet of socio-syntax (Cornips 2015). Similarly, variationist (socio)linguistics regard the focus on phonology in traditional Labovian sociolinguistics more as a matter of convenience than necessity, featuring at least a variety of studies into socially conditioned variation at the level of morphosyntax.

On the face of it, the differences between the three variables go against our expectation for abstract linguistic structure not to be socially evaluated or easily malleable. However, these differences should arguably not be linked solely to their respective structural complexity. Instead, it appears that we are witnessing two entirely different processes. The gradual increase in the use of *sá* in referential contexts and the rise in frequency of *maður* fit in with a trend of destandardisation, featuring a relaxation of strict prescriptive norms in otherwise standardised usage. In these cases, the standard has been shifting more towards the spoken language. In the case of verb-adverb placement, on the other hand, the underlying situation is actually the reverse. Unlike *sá* and *maður*, the spread of Adv-Vfin had not reached the whole of the population, or to varying degrees, eventually grammaticalising rather differently, it appears, from the rest of the Scandinavian languages (cf. Angantýsson's claim that Adv-Vfin is focus-dependent in Modern Icelandic).

The view that presents itself rather strongly contradicts the traditional narrative both with regard to the scope and limits of variation. Even features traditionally associated (rightly or wrongly) with Danish influences in higher, educated circles were found to be attested in speakers of various social backgrounds, which is in vast disagreement with the alleged conservatism of Icelandic or the 'purity' of rural speakers. Moreover, the view that standard norms were implemented more or less successfully is only partly

corroborated by the data. Whether the same holds when more data and more variables are taken into account remains to be seen. Such an undertaking would greatly benefit from a broad range of 19th- and 20th-century corpora featuring ego-documents such as private letters that have been tagged and, preferably, annotated syntactically.

Bibliography

- Abney, Steven Paul. 1987. The English Noun Phrase in its Sentential Aspect. Doctoral dissertation, Massachusetts Institute of Technology. URL <http://www.vinartus.net/spa/87a.pdf>.
- Adger, David. 2003. *Core Syntax: A Minimalist Approach*. Oxford/New York: Oxford University Press.
- Adger, David. 2006. Combinatorial variability. *Journal of Linguistics* 42.3:503–530.
- Adger, David. 2007. Variability and modularity: A response to Hudson. *Journal of Linguistics* 43.3:695–700.
- Adger, David. 2014. Variability and grammatical architecture. In *Linguistic Variation in the Minimalist Framework*, ed. M. Carme Picallo, 179–196. Oxford: Oxford University Press.
- Adger, David, and Jennifer Smith. 2005. Variation and the Minimalist Program. In *Syntax and Variation: Reconciling the Biological and the Social*, ed. Leonie Cornips and Karen Corrigan, 149–178. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Adger, David, and Jennifer Smith. 2010. Variation in agreement: A lexical feature-based approach. *Lingua* 120.5:1109–1134.
- Alexiadou, Artemis, and Gisbert Fanselow. 2002. On the correlation between morphology and syntax: The case of V-to I. In *Studies in Comparative Germanic Syntax: Proceedings from the 15th Workshop on Comparative Germanic Syntax*, ed. Jan-Wouter Zwart and Werner Abraham, 219–242. John Benjamins.
- Anderson, Benedict. 2006. *Imagined Communities: Reflections on the Origin and Spread of Nationalism*. London/New York: Verso. [Revised Edition].
- Andersson, Lars-Gunnar. 1975. Form and Function of Subordinate Clauses. Doctoral dissertation, University of Göteborg. Gothenburg Monographs in Linguistics 1.
- Anderwald, Lieselotte. 2014a. Measuring the success of prescriptivism: Quantitative grammaticography, corpus linguistics and the progressive passive. *English Language and Linguistics* 18.1:1–21.
- Anderwald, Lieselotte. 2014b. *Burned, dwelled, dreamed: The evolution of a morphological Americanism and the role of prescriptive grammar writing*. *American Speech* 89.4:408–440.
- Anderwald, Lieselotte. 2016. *Language Between Description and Prescription: Verbs and Verb Categories in Nineteenth-Century Grammars of English*. Oxford Studies in the History of English. Oxford/New York: Oxford University Press.
- Angantýsson, Ásgrímur. 2001. Skandinavísk orðaröð í íslenskum aukasetningum. *Íslenskt mál og almenn málfraeði* 23:95–122.
- Angantýsson, Ásgrímur. 2007. Verb-third in embedded clauses in Icelandic. *Studia Linguistica* 61:237–260.
- Angantýsson, Ásgrímur. 2011. *The Syntax of Embedded Clauses in Icelandic and Related Languages*. Reykjavík: Hugvísindastofnun Háskóla Íslands.
- Angantýsson, Ásgrímur. 2017a. Framfærslur í aukasetningum. In *Tilbrigði í íslenskri setningagerð: III. Sérathuganir*, ed. Höskuldur Þráinsson, Ásgrímur Angantýsson, and Einar Freyr Sigurðsson, 11–37. Reykjavík: Málvísindastofnun Háskóla Íslands.
- Angantýsson, Ásgrímur. 2017b. Linguistic proficiency as cultural capital in school environments. *Beyond Philology* 14.1:163–179.
- Angantýsson, Ásgrímur. 2017c. Setningaratviksorð í íslensku og færeysku. *Íslenskt mál og almenn málfraeði* 39:75–86.
- Angantýsson, Ásgrímur. 2017d. The distribution of embedded V2 and V3 in Modern Icelandic. *Working Papers in Scandinavian Syntax* 99:54–82.
- Angantýsson, Ásgrímur. 2018. Verb-second in embedded clauses in Faroese. *Studia Linguistica* 72.1:165–189.
- Angantýsson, Ásgrímur. to appear. The distribution of embedded V2 and V3 in Modern Icelandic. In *Rethinking Verb Second: Rethinking Comparative Syntax*, ed. Theresa Biberauer, Sam Wolfe, and Rebecca Woods. Oxford: Oxford University Press.
- Angantýsson, Ásgrímur, and Dianne Jonas. 2016. On the syntax of adverbial clauses in Icelandic. *Working Papers in Scandinavian Syntax* 96:126–139.

- Anthony, Laurence. 2012. *Antcon* 3.2.2. Tokyo: Waseda University. URL <http://www.antlab.sci.waseda.ac.jp>.
- Auer, Anita. 2006. Precept and practice: The influence of prescriptivism on the English subjunctive. In *Syntax, Style and Grammatical Norms: English from 1500-2000*, ed. Christiane Dalton-Puffer, Dieter Kastovsky, Nikolaus Ritt, and Herbert Schendl, Linguistic Insights: Studies in Language and Communication 39, 33–53. Bern: Peter Lang.
- Auer, Anita. 2009. *The Subjunctive in the Age of Prescriptivism: English and German Developments during the Eighteenth Century*. Palgrave Studies in Language History and Language Change. Basingstoke: Palgrave Macmillan.
- Auer, Anita. 2014. Nineteenth-century English: Norms and usage. In *Norms and Usage in Language History, 1600-1900: A Sociolinguistic and Comparative Perspective*, ed. Gijsbert Rutten, Rik Vosters, and Wim Vandenbussche, Advances in Historical Sociolinguistics 3, 151–169. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Auer, Anita, Catharina Peersman, Simon Pickl, Gijsbert Rutten, and Rik Vosters. 2015. Historical sociolinguistics: The field and its future. *Journal of Historical Sociolinguistics* 1.1:1–12.
- Auer, Anita, and Anja Voeste. 2012. Grammatical variables. In *The Handbook of Historical Sociolinguistics*, ed. Juan Manuel Hernández-Campoy and Juan Camilo Conde-Silvestre, 253–270. Chichester: Wiley-Blackwell.
- Ármansson, Kristinn, Einar Magnússon, Guðni Guðmundsson, and Heimir Þorleifsson. 1975. *Saga Reykjavíkarskóla I: Nám og nemendur*. Reykjavík: Sögusjóður Menntaskólans í Reykjavík.
- Árnadóttir, Arngunnur, and Gyða Dögg Einarsdóttir. 2007. 21 aðferð til að beygja orðið hönd: Könnun á beygingu orðsins hönd í íslensku nútímamáli. *Skíma: Mál gagn móðurmálskennara* 30.1:18–21.
- Árnason, Kristján. 2003a. Icelandic. In *Germanic Standardizations: Past to Present*, ed. Ana Deumert and Wim Vandenbussche, Impact: Studies in Language and Society 18, 245–279. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Árnason, Kristján. 2003b. Language planning and the structure of Icelandic. In *Útnorður: West Nordic Standardisation and Variation*, 193–218. Reykjavík: Institute of Linguistics, University of Iceland Press.
- Baayen, R.H. 2008. *Analyzing Linguistic Data: A Practical Introduction to Statistics using R*. Cambridge: Cambridge University Press.
- Baayen, R.H., D.J. Davidson, and D.M. Bates. 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59.4:390–412.
- Barðdal, Jóhanna, and Thomas Smitherman. 2013. The quest for cognates: A reconstruction of oblique subject constructions in Proto-Indo-European. *Language Dynamics and Change* 3.1:28–67.
- Benediktsson, Hreinn. 1969. On the inflection of the *ia*-stems in Icelandic. In *Afmælisrit Jóns Helgasonar: 30. júní 1969*, ed. Jakob Benediktsson, Jón Samsonarson, Jónas Kristjánsson, Ólafur Halldórsson, and Stefán Karlsson, 391–402. Reykjavík: Heimskringla.
- Bennis, Hans, Leonie Cornips, and Marc van Oostendorp. 2004. *Verandering en verloedering: Normen en waarden in het Nederlands*. Amsterdam: Amsterdam University Press/Salomé.
- Bentzen, Kristine. 2003. V-to-I movement in the absence of morphological cues: Evidence from adult and child Northern Norwegian. *Nordlyd* 31.3:573–588.
- Bentzen, Kristine. 2009. Subject positions and their interaction with verb movement. *Studia Linguistica* 63.3:261–291.
- Bentzen, Kristine. 2014. Verb placement in relative clauses. *Nordic Atlas of Language Structures Journal* 1:240–249. URL http://www.tekstlab.uio.no/nals/system/archives/130/original/VerbPlaceIntro-KLAR_210613.pdf.
- van Bergen, Linda. 2003. *Pronouns and Word Order in Old English: With Particular Reference to the Indefinite Pronoun Man*. New York/London: Routledge.
- Bergs, Alexander. 2012. The Uniformitarian Principle and the risk of anachronisms in language and social history. In *The Handbook of Historical Sociolinguistics*, ed. Juan Manuel Hernández-Campoy and Juan Camilo Conde-Silvestre, 80–98. Chichester: Wiley-Blackwell.
- Bernharðsson, Haraldur. 2017. Jón Thoroddsen og málstöðlun níttjándu aldar: Nokkur málfarsatriði í skáldsögunni *Pilti og stúlku* 1850 og 1867. *Orð og tunga* 19:77–127.
- Bernharðsson, Haraldur. 2018. Spreading the standard: The nineteenth-century standardization of Icelandic and the first Icelandic novel. *Journal of Historical Sociolinguistics* 4.2:149–176.
- Bernharðsson, Haraldur, and Jóhannes Gísli Jónsson. 2012. Íslenskt mál á níttjándu öld — bréfasafn. Unpublished ms.
- den Besten, Hans. 1977. On the Interaction of Root Transformations and Lexical Deletive Rules. Instituut voor Algemene Taalwetenschap/University of Amsterdam. Unpublished ms.

- den Besten, Hans. 1989. *Studies in West Germanic Syntax*. Doctoral dissertation, Katholieke Universiteit Brabant. Amsterdam/Atlanta: Rodopi.
- Bickerton, Derek. 2014. Some problems for biolinguistics. *Biolinguistics* 8:73–96.
- Bjarnadóttir, Kristín, Aðalsteinn Eyþórsson, and Þorsteinn G. Indriðason. 1988–1989. Skrá um íslensk málfraeðirit til 1925: Mart finna hundar sjer í holum. *Íslenskt mál og almenn málfraeði* 10–11:177–257.
- Bjarnason, Halldór. 2006. The importance of being rich: Social mobility in two families in Iceland, ca 1800 to ca 1970. Paper presented at XIV International Economic History Congress, Helsinki, Finland, 21–25 August 2006. Session 39: Intergenerational Transmission of Occupation and Social Class. URL <http://rafhladan.is/handle/10802/7674>.
- Blöndal, Sigfús, and Ingeborg Stemann. 1959. *Praktisk Lærebog i islandsk Nutidssprog*. Ejnar Munksgaard Forlag, tredje oplag edition. Udgivet med Bidrag af Universitetets Fritryksskonto.
- Boas, Hans C. 2010. The syntax–lexicon continuum in Construction Grammar: A case study of English communication verbs. *Belgian Journal of Linguistics* 24:54–82.
- Bobaljik, Jonathan David. 1999. Adverbs: The hierarchy paradox. *Glott International* 4.9/10:27–28.
- Bobaljik, Jonathan David. 2002. Realizing Germanic inflection: Why morphology does not drive syntax. *The Journal of Comparative Germanic Linguistics* 6.2:129–167.
- Bobaljik, Jonathan David, and Höskuldur Thráinsson. 1998. Two heads aren't always better than one. *Syntax* 1.1:37–71.
- Bourdieu, Pierre. 1973. Cultural reproduction and social reproduction. In *Knowledge, Education, and Cultural Change: Papers in the Sociology of Education*, ed. Richard Brown, Explorations in Sociology 2, 71–112. London: Tavistock Publications.
- Bourdieu, Pierre. 1977. The economics of linguistic exchanges. *Social Science Information* 16.6:645–668.
- Bourdieu, Pierre. 1991. *Language and Symbolic Power*. Cambridge: Polity Press. [Translated by Gino Raymond and Matthew Adamson].
- Bybee, Joan L. 1985. *Morphology: A Study of the Relation between Meaning and Form*. Typological Studies in Language 9. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Bybee, Joan, and Sandra Thompson. 2000. Three frequency effects in syntax. In *Proceedings of the Twenty-Third Annual Meeting of the Berkeley Linguistics Society February 14–17, 1997: General Session and Parasession on Pragmatics and Grammatical Structure*, ed. Matthew L. Juge and Jeri L. Moxley, 378–388. Berkeley: Berkeley Linguistics Society.
- Böðvarsson, Árni. 1992. *Íslensk málfar*. Íslensk þjóðfræði. Reykjavík: Almenna bókafélagið.
- Cheshire, Jenny. 1987. Syntactic variation, the linguistic variable, and sociolinguistic theory. *Linguistics* 25.2:257–282.
- Cheshire, Jenny. 1999. Taming the vernacular: Some repercussions for the study of syntactic variation and spoken grammar. *Cuadernos de Filología Inglesa* 8.1:59–80.
- Cheshire, Jenny. 2005a. Syntactic variation and beyond: Gender and social class variation in the use of discourse-new markers. *Journal of Sociolinguistics* 9.4:479–508.
- Cheshire, Jenny. 2005b. Syntactic variation and spoken language. In *Syntax and variation: Reconciling the Biological and the Social*, 81–106. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Cheshire, Jenny, Paul Kerswill, and Ann Williams. 2005. Phonology, grammar, and discourse in dialect convergence. In *Dialect Change: Convergence and Divergence in European Languages*, ed. Frans Hinskens Peter Auer and Paul Kerswill, 135–167. Cambridge: Cambridge University Press.
- Chomsky, Noam. 1964. *Current Issues in Linguistic Theory*. London/The Hague/Paris: Mouton & Co.
- Chomsky, Noam. 1991. Some notes on economy of derivation and representation. In *Principles and Parameters in Comparative Grammar*, ed. Robert Freidin, 417–454. Cambridge (MA)/London: MIT Press.
- Chomsky, Noam. 1995. *The Minimalist Program*. Current Studies in Linguistics 28. Cambridge (MA)/London: The MIT Press.
- Chomsky, Noam. 2000. Minimalist inquiries: The framework. In *Step by Step. Essays on Minimalist Syntax in Honor of Howard Lasnik*, ed. David Michaels Roger Martin and Juan Uriagereka, 89–155. Cambridge (MA)/London: MIT Press.
- Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A Life in Language*. MIT Press.
- Chomsky, Noam. 2008. On phases. In *Foundational Issues in Linguistic Theory. Essays in Honor of Jean-Roger Vergnaud*, 133–166. Cambridge (MA)/London: MIT Press.
- Chomsky, Noam. 2013. Problems of projection. *Lingua* 130:33–49.
- Chomsky, Noam, and Morris Halle. 1968. *The Sound Pattern of English*. Studies in Language. New York/Evanston/London: Harper & Row, Publishers.
- Cinque, Guglielmo. 1999. *Adverbs and functional heads: A cross-linguistic perspective*. Oxford Studies in Comparative Syntax. Oxford: Oxford University Press.

- Cinque, Guglielmo. 2004. Issues in adverbial syntax. *Lingua* 114:683–710.
- Cinque, Guglielmo. 2005. Deriving Greenberg's Universal 20 and its exceptions. *Linguistic Inquiry* 36.3:315–332.
- Cinque, Guglielmo. 2010. *The Syntax of Adjectives: A Comparative Study*. Linguistic Inquiry Monograph 57. Cambridge (MA)/London: The MIT Press.
- Cinque, Guglielmo. 2013. Cognition, Universal Grammar, and typological generalizations. *Lingua* 130:50–65.
- Cinque, Guglielmo, and Luigi Rizzi. 2008. The Cartography of Syntactic Structures. *Studies in Linguistics (CISCL Working Papers)* 2:42–58.
- Clark, Eve V. 1987. The principle of contrast: A constraint on language acquisition. In *Mechanisms of Language Acquisition*, ed. Brian MacWhinney, 1–33. Hillsdale, New Jersey/London: Lawrence Erlbaum Associates.
- Comrie, Bernard. 1989. *Language Universals and Linguistic Typology*. Blackwell/University of Chicago Press.
- Conde-Silvestre, J. Camilo. 2016. A 'third-wave' historical sociolinguistic approach to late Middle English correspondence: Evidence from the Stonor Letters. In *Current Trends in Historical Sociolinguistics*, ed. Cinzia Russi, 46–66. Warsaw/Berlin: De Gruyter Open.
- Cooper, Robert L. 1989. *Language planning and social change*. Cambridge: Cambridge University Press.
- Cornips, Leonie. 2014. Socio-syntax and variation in acquisition: Problematizing monolingual and bidialectal acquisition. *Linguistic Variation* 14.1:1–25.
- Cornips, Leonie. 2015. The no man's land between syntax and variationist sociolinguistics: The case of idiolectal variability. In *Variation in Language: System- and Usage-based Approaches*, ed. Aria Adli, Marco García García, and Göz Kaufmann, *Linguae & litterae: Publications of the School of Language & Literature Freiburg Institute for Advanced Studies* 50, 147–171. Berlin/Boston: De Gruyter.
- Cornips, Leonie, and Karen Corrigan. 2005a. Toward an integrated approach to syntactic variation: A retrospective and prospective synopsis. In *Syntax and Variation: Reconciling the Biological and the Social*, 1–27. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Cornips, Leonie, and Karen P. Corrigan. 2005b. Convergence and divergence in grammar. In *Dialect Change: Convergence and Divergence in European Languages*, ed. Frans Hinskens Peter Auer and Paul Kerswill, 96–134. Cambridge: Cambridge University Press.
- Coveney, Aidan. 2002. *Variability in Spoken French: A Sociolinguistic Study of Interrogation and Negation*. Bristol/Portland (OR): Elm Bank.
- Delbrück, B. 1888. *Altindische Syntax*. Syntaktische Forschungen 5. Verlag der Buchhandlung des Waisenhauses.
- Delsing, Lars-Olof. 1988. The Scandinavian noun phrase. *Working Papers in Scandinavian syntax* 42:57–79.
- Delsing, Lars-Olof. 1993. On attributive adjectives in Scandinavian and other languages. *Studia Linguistica* 47.2:105–125.
- Delsing, Lars-Olof. 2003. Syntaktisk variation i nordiska nominalfraser. In *Dialektsyntaktiska studier av den nordiska nominalfrasen*, ed. Øystein Alexander Vangsnes, Anders Holmberg, and Lars-Olof Delsing, *Tromsø Studies in Linguistics* 22, 11–64. Oslo: Novus.
- Deumert, Ana. 2003a. Bringing speakers back in? Epistemological reflections on speaker-oriented explanations of language change. *Language Sciences* 25.1:15–76.
- Deumert, Ana. 2003b. Describing language standardisation: Models and methods. In *Úmorður: West Nordic Standardisation and Variation*, 9–32. Reykjavík: Institute of Linguistics, University of Iceland Press.
- Deumert, Ana, and Wim Vandenbussche. 2003. Research directions in the study of language standardization. In *Germanic Standardizations: Past to Present*, ed. Ana Deumert and Wim Vandenbussche, Impact: Studies in Language and Society 18, 455–469. John Benjamins Publishing Company.
- Ecay, Aaron William. 2015. A Multi-Step Analysis of the Evolution of English Do-Support. Doctoral dissertation, University of Pennsylvania. URL <http://repository.upenn.edu/edissertations/1049>.
- Eckert, Penelope. 1989. The whole woman: Sex and gender differences in variation. *Language Variation and Change* 1.3:245–267.
- Eckert, Penelope. 2008. Variation and the indexical field. *Journal of Sociolinguistics* 12.4:453–476.
- Eckert, Penelope. 2012. Three waves of variation study: The emergence of meaning in the study of sociolinguistic variation. *Annual Review of Anthropology* 41:87–100.
- Eckert, Penelope, and Sally McConnell-Ginet. 2013. *Language and Gender*. Cambridge: Cambridge University Press. [Second edition].
- Egilsson, Sveinbjörn. 1829-1840. *Skóla-hátíð í minningu Fæðingar-dags vors allranáðugasta Konungs Fridriks Sjötta* Videyjar Klaustri. [Bessastaðir Grammar School periodicals].

- Egilsson, Sveinn Yngvi. 1999. *Arfur og umbylting: Rannsókn á íslenskri rómantík*. Íslensk menning: Ritróð ReykjavíkurAkademíunnar og Hins íslenska bókmenntafélags. Reykjavík: Hið íslenska bókmenntafélag/ReykjavíkurAkademían.
- Ehala, Martin. 1998. How a man changed a parameter value: The loss of SOV in Estonian subclasses. In *Historical Linguistics 1995: Selected Papers from the 12th International Conference on Historical Linguistics, Manchester, August 1995*, ed. Richard M. Hogg and Linda van Bergen, Vol. 2: Germanic Linguistics. Current Issues in Linguistic Theory 162, 73–88. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Elffers, Els, Sies de Haan, and Ina Schermer. 2012. Attributive past participles in Dutch. In *Between West and East: Festschrift for Wim Honselaar on the occasion of his 65th birthday*, ed. R. Genis, E. de Haard, J. Kalsbeek, E. Keizer, and J. Stelleman, 125–166. Amsterdam: Pegasus.
- Elsaß, Stephan. 2005a. Language norm and language reality: Effectiveness and limits of prescriptivism in New High German. In *Language Purism in the Germanic Languages*, ed. Nils Langer and Winifred V. Davies, *Studia linguistica Germanica* 75, 20–45. Berlin: Walter de Gruyter.
- Elsaß, Stephan. 2005b. *Sprachgeschichte von unten: Untersuchungen zum geschriebenen Alltagsdeutsch im 19. Jahrhundert*. Reihe Germanistische Linguistik 263. Tübingen: Niemeyer.
- Elsaß, Stephan. 2007. ‘Everyday language’ in emigrant letters and its implications for language historiography – the German case. *Multilingua* 26.2/3:149–163.
- Elsaß, Stephan. 2012. The use of private letters and diaries in sociolinguistic investigation. In *The Handbook of Historical Sociolinguistics*, ed. Juan Manuel Hernández-Campoy and Juan Camilo Conde-Silvestre, 156–169. Chichester: Wiley-Blackwell.
- Elsaß, Stephan. 2014. Prescriptive norms and norms of usage in nineteenth-century German. In *Norms and Usage in Language History, 1600-1900: A Sociolinguistic and Comparative Perspective*, ed. Gijsbert Rutten, Rik Vosters, and Wim Vandenbussche, *Advances in Historical Sociolinguistics* 3, 303–320. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Elsaß, Stephan. 2016. Assessing prescriptivism and its (lack of) effect in the history of German. Paper presented at Effects of Prescriptivism in Language History, 21–22 January 2016, Universiteit Leiden.
- Elsaß, Stephan, Nils Langer, Joachim Scharloth, and Wim Vandenbussche (eds.). 2007. *Germanic Language Histories ‘from Below’ (1700–2000)*. *Studia Linguistica Germanica* 86. Berlin/New York: Walter de Gruyter.
- Erker, Daniel, and Gregory R. Guy. 2012. The role of lexical frequency in syntactic variability: Variable subject personal pronoun expression in Spanish. *Language* 88.3:526–557.
- Ernst, Thomas. 2002. *The Syntax of Adjuncts*. Cambridge Studies in Linguistics 96. Cambridge: Cambridge University Press.
- Ernst, Thomas. 2007. On the role of semantics in a theory of adverb syntax. *Lingua* 117.6:1008–1033.
- Erteschik-Shir, Nomi [Nomi Erteschik (Shir)]. 1973. On the Nature of Island Constraints. Doctoral dissertation, Massachusetts Institute of Technology.
- Evans, Nicholas, and Stephen C. Levinson. 2009. The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences* 32.5:429–492.
- Eythórsson, Thórhallur. 1995. Verbal Syntax in the Early Germanic Languages. Doctoral dissertation, Cornell University.
- Eyþórsson, Þórhallur. 1997–1998. Uppruni sagnfærslu í germönskum málum. *Íslensk og almenn málfraeði* 19-20:133–180.
- Eythórsson, Thórhallur. 2002. Changes in subject case-marking in Icelandic. In *Syntactic Effects of Morphological Change*, ed. David Lightfoot, 196–212. Oxford: Oxford University Press.
- Eythórsson, Thórhallur, and Jóhanna Barðdal. 2005. Oblique subjects: A common Germanic inheritance. *Language* 81.4:824–881.
- Falk, Cecilia. 1993. Non-Referential Subjects in the History of Swedish. Doctoral dissertation, Lund University.
- Fenger, Paula. 2018. How impersonal does one get? A study of *man*-pronouns in Germanic. *The Journal of Comparative Germanic Linguistics* 21.3:291–325.
- Finegan, Edward, and Douglas Biber. 2001. Register variation and social dialect variation: the Register Axiom. In *Style and Sociolinguistic Variation*, ed. Penelope Eckert and John R. Rickford, 235–267. Cambridge: Cambridge University Press.
- Frey, Werner. 2012. On two types of adverbial clauses allowing root-phenomena. In *Main clause phenomena: New horizons*, ed. Lobke Aelbrecht, Liliane Haegeman, and Rachel Nye, 405–429. Amsterdam/Philadelphia: John Benjamins Publishing Company.

- Freywald, Ulrike, Leonie Cornips, Natalia Ganuza, Ingvild Nistov, and Toril Opsahl. 2015. Beyond verb second – a matter of novel information-structural effects? Evidence from Norwegian, Swedish, German and Dutch. In *Language, Youth and Identity in the 21st Century. Linguistic Practices across Urban Spaces*, ed. Jacomine Nortier and Bente A. Svendsen, 73–92. Cambridge: Cambridge University Press.
- Fritzner, Johan. 1891. *Ordbog over Det gamle norske Sprog*. Vol. 2. Kristiania: Den norske Forlagsforening. [Omarbeidet, forøget og forbedret Udgave].
- Fríðriksson, Finnur. 2008. *Language change vs. stability in conservative language communities: A case study of Icelandic*. Gothenburg Monographs in Linguistics 37. Göteborg: University of Gothenburg.
- Fríðriksson, Halldór Kr. 1846. *Íslandsk Læsebog med Ordregister og en Oversigt over den islandske Formlære*. Kjöbenhavn: Jægers skandinaviske Forlagshandel / S.L. Möller.
- Fríðriksson, H. Kr. 1857. *Dönsk málfraedi*. Kaupmannahöfn: Egill Jónsson/Louis Klein.
- Fríðriksson, Halldór Kr. 1859. *Íslenzkar rjettritunarreglur*. Reykjavík: Hið íslenzka bókmenntafjelag.
- Fríðriksson, Halldór Kr. 1861. *Íslenzk málmyndaýsning*. Kaupmannahöfn: Hið íslenzka bókmenntafelag. Í prentsmiðju S. L. Möllers.
- Fríðriksson, Halldór Kr. 1871. Nokkrar athugasgreinir um íslenzku: verba valent usu. *Þjóðólfur* 23:105–111. Viðaukablæð við Þjóðólf.
- Garbacz, Piotr. 2010. *Word Order in Övdalian: A Study in Variation and Change*. Lundastudier i nordisk språkvitenskap. Serie A 70. Lund: Centre for Languages and Literature, Lund University.
- Gärtner, Hans-Martin. 2016. A note on the Rich Agreement Hypothesis and varieties of “embedded V2”. *Working Papers in Scandinavian Syntax* 96:1–13.
- Gíslason, K. [=Konráð]. 1851. *Dönsk orðabók með íslenzkum þýðingum*. Kaupmannahöfn: Bianco Luno.
- Gíslason, Konráð. 1844. Bókafregn. *Fjölnir: Árrit handa Íslendingum* 7:71–104.
- Gíslason, Vilhjálmur Þ. 1972. *Blæð og blaðamenn 1773-1944*. Reykjavík: Almenna bókafélagið.
- Goldberg, Adele E. 2006. *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.
- Greenberg, Joseph H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In *Universals of Language*, ed. Joseph H. Greenberg, 59–90. Cambridge (MA): The M.I.T. Press.
- Gregersen, Frans, and Inge Lise Pedersen. 2000. *A la recherche du word order not quite perdu: A methodological progress report*. In *Textual Parameters in Older Languages*, ed. Susan C. Herring, Pieter van Reenen, and Lene Schøsler, Current Issues in Linguistic Theory 195, 393–431. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Gries, Stefan Th. 2015a. Some current quantitative problems in corpus linguistics and a sketch of some solutions. *Language and Linguistics* 16.1:93–117.
- Gries, Stefan Th. 2015b. The most under-used statistical method in corpus linguistics: Multi-level (and mixed-effects) models. *Corpora* 10.1:95–125.
- Grondelaers, Stefan. 2009. Woordvolgorde in presentatieve zinnen en de theoretische basis van multifactoriële grammatica. *Nederlandse Taalkunde* 14.3:282–312.
- Guðfinnsson, Björn. 1940. Tilræði við íslenzkt mál. *Andvari* 65:73–81.
- Guðmundsson, Finnboogi. 1960. *Hómersþýðingar Sveinbjarnar Egilssonar*. Reykjavík: Bókaútgáfa Menningarsjóðs.
- Guðmundsson, Valtýr. 1922. *Íslandsk grammatik: Íslandsk nutidssprog*. København: H. Hagerups Forlag. [Rit um íslenzka málfraedi 1. Málvísindastofnun Háskóla Íslands, 1983].
- Gunnarsson, Sigurður. 1878. „Styrbjörn á Nesi“. *Ísafold* 5:9–10.
- Gunnlaugsson, Gísli Ágúst. 1988. *Family and Household in Iceland 1801-1930. Studies in the Relationship between Demographic and Socio-Economic Development, Social Legislation and Family and Household Structures*. Stockholm: Almqvist & Wiksell International.
- Gunnlaugsson, Gísli Ágúst, and Loftur Guttormsson. 1993. Household structure and urbanization in three Icelandic fishing districts, 1880-1930. *Journal of Family History* 18.4:315–340.
- Gussenhoven, C.H.M. 1987. Norm en variatie in het Standaard Engels. In *Variatie en Norm in de Standaardtaal*, ed. J. de Rooij, Publikaties van het P.J. Meertens-Instituut voor Dialectologie, Volkskunde en Naamkunde van de Koninklijke Nederlandse Akademie van Wetenschappen 7, 13–33. Amsterdam: P.J. Meertens-Instituut.
- Guttormsson, Loftur. 1983. *Bernska, ungdómur og uppeldi á einveldisöld: Tilraun til félagslegrar og lýðfræðilegrar greiningar*. Ritsafn Sagnfræðistofnunar 10. Reykjavík: Sagnfræðistofnun Háskóla Íslands.
- Guttormsson, Loftur. 1990. The development of popular religious literacy in the seventeenth and eighteenth centuries. *Scandinavian Journal of History* 15.1/2:7–35.

- Guttormsson, Loftur. 2003. Framleiðsla og dreifing ritaðs máls. In *Alþýðumenning á Íslandi 1830-1930: Ritað mál, menntun og félagsshreyfingar*, ed. Ingi Sigurðsson and Loftur Guttormsson, Sagnfræðirannsóknir - Studia Historica 18, 37–65. Reykjavík: Háskólaútgáfan.
- Guttormsson, Loftur. 2012. Udviklingen af læse- og skrivefærdighed i Norden 1650-1850: Historiografiske funderinger. In *Att læsa och att skriva: Två vågor av vardagligt skriftbruk i Norden 1800-2000*, ed. Ann-Catrine Edlund, Nordliga studier 3, Vardagligt skriftbruk 1, 19–37. Umeå: Umeå universitet, Kungl. Skytteanska Samfundet.
- Gutzmann, Daniel. 2017. Modal particles ≠ modal particles (= modal particles): Differences between German modal particles and how to deal with them semantically. In *Discourse particles: Formal Approaches to their Syntax and Semantics*, ed. Josef Bayer and Volker Struckmeier, 144–172. Berlin: De Gruyter.
- Guy, Gregory R. 2011. Language, social class, and status. In *The Cambridge Handbook of Sociolinguistics*, ed. Rajend Mesthrie, Cambridge Handbooks in Language and Linguistics, 159–185. Cambridge: Cambridge University Press.
- Haegeman, Liliane. 2004. The syntax of adverbial clauses and its consequences for topicalisation. *Antwerp Papers in Linguistics* 107:61–91. URL <http://uahost.uantwerpen.be/apil/apil107/file05.pdf>.
- Haegeman, Liliane. 2012. *Adverbial Clauses, Main Clause Phenomena, and the Composition of the Left Periphery: The Cartography of Syntactic Structures*. Vol. 8. Oxford Studies in Comparative Syntax. Oxford/New York: Oxford University Press.
- Håkansson, David. 2011. Bisatsledföljden i äldre svenska – variation eller förändring? *Arkiv för nordisk filologi* 126:93–140.
- Hálfðanarson, Guðmundur. 1993. Íslensk þjóðfélagsþróun á 19. öld. In *Íslensk þjóðfélagsþróun 1880-1990: Ritgerðir*, ed. Guðmundur Hálfðanarson and Svanur Kristjánsson, 9–58. Reykjavík: Félagsvísindastofnun Háskóla Íslands/Sagnfræðistofnun Háskóla Íslands.
- Hálfðanarson, Guðmundur. 1995. Social distinctions and national unity: On politics of nationalism in nineteenth-century Iceland. *History of European Ideas* 21.6:763–779.
- Hálfðanarson, Guðmundur. 2001. *Íslenska þjóðríkið: uppruni og endimörk*. Íslensk menning. Reykjavík: Hið íslenska bókmenntafélag, ReykjavíkurAkademían.
- Hálfðanarson, Guðmundur. 2005. From linguistic patriotism to cultural nationalism: language and identity in Iceland. In *Languages and Identities in Historical Perspective*, ed. Ann Katherine Isaacs, Clíoh's Workshop III X, 55–67. Pisa: Edizioni Plus – Università di Pisa.
- Halldórsdóttir, Erla Hulda. 2003. Af bréfaskrifum kvenna á 19. öld. In *Alþýðumenning á Íslandi 1830-1930: Ritað mál, menntun og félagsshreyfingar*, ed. Ingi Sigurðsson and Loftur Guttormsson, Sagnfræðirannsóknir - Studia Historica 18, 247–267. Reykjavík: Háskólaútgáfan.
- Halldórsdóttir, Erla Hulda. 2014. “Don’t you forget your always loving sister”: Writing as a social and cultural capital. In *Vernacular Literacies: Past, Present and Future*, ed. Ann-Catrine Edlund, Lars-Erik Edlund, and Susanne Haugen, Northern Studies Monographs 3/Vardagligt skriftbruk 3, 181–192. Umeå: Umeå University/Royal Skyttean Society.
- Halldórsson, Jóhannes (ed.). 1991. *Bréf til Jóns Sigurðssonar: Úrval. Bréfitarar: Guðmundur Ólafsson, Halldór Kr. Friðriksson, Jón Pétursson*. Vol. 3. Reykjavík: Bókaútgáfa Menningarsjóðs og Þjóðvínafélagsins.
- Halle, Morris, and Alec Marantz. 1994. Some key features of Distributed Morphology. *MIT Working Papers in Linguistics* 21:275–288. Papers on Phonology and Morphology.
- Han, Chung-hye, and Anthony Kroch. 2000. The rise of *do*-support in English: Implications for clause structure. In *Proceedings of the 30th North East Linguistics Society*, 311–325. Amherst, MA: GLSA.
- Harðarson, Gísli Rúnar. 2017. *Cycling through Grammar: On Compounds, Noun Phrases and Domains*. Doctoral dissertation, University of Connecticut.
- Harðarson, Jón Axel. 1997. Inngangur. In *Grammatica islandica: Íslensk málfræði*, XI–LXVII. Reykjavík: Málvísindastofnun Háskóla Íslands, Háskólaútgáfan.
- Harley, Heidi, and Elizabeth Ritter. 2002. Person and number in pronouns: A feature-geometric analysis. *Language* 78.3:482–526.
- Haspelmath, Martin. 1999. Optimality and diachronic adaptation. *Zeitschrift für Sprachwissenschaft* 18.2:180–205.
- Haspelmath, Martin. 2010. Framework-free grammatical theory. In *The Oxford Handbook of Linguistic Analysis*, ed. Bernd Heine and Heiko Narrog, 342–365. Oxford: Oxford University Press.
- Hasty, J. Daniel. 2014. We might should be thinking this way: Theory and practice in the study of syntactic variation. In *Micro-Syntactic Variation in North American English*, ed. Raffaella Zanuttini and Laurence R. Horn, Oxford Studies in Comparative Syntax, 269–293. Oxford/New York: Oxford University Press.

- Haugen, Einar. 1966. Dialect, language, nation. *American Anthropologist* 68.4:922–935.
- Haugen, Einar. 1987. *Blessings of Babel: Bilingualism and Language Planning. Problems and Pleasures*. Contributions to the Sociology of Language 46. Mouton de Gruyter.
- Heimisdóttir, Linda Ösp. 2008. Yfirlit yfir beygingu karlkyns ija-stofna frá nítjándu öld til nútímamáls. Unpublished B.A. thesis, University of Iceland.
- Helgason, Jón. 1896. II. Stúdentatal frá hinum lærða skóla í Reykjavík á 50 ára tímabilinu 1847–96. In *Míningarrit fímtíu ára afmælis hins lærða skóla í Reykjavík*. Reykjavík: Fjelagsprentsmiðjan.
- Helgason, Jón. 1929. *Málið á Nýja testamenti Odds Gottskálkssonar*. Safn fræðafjelagsins VII. Kaupmannahöfn: S. L. Möller. [Rit um íslenska málfræði 4. Málvísindastofnun Háskóla Íslands, 1999].
- Hellan, Lars. 1986. The headedness of NPs in Norwegian. In *Features and Projections*, ed. Pieter Muysken and Henk van Riemsdijk, 89–122. Dordrecht: Foris.
- Hendrickx, Els. 2013. De invloed van lexicale taaladviezen op Belgisch-Nederlandse krantentaal. *Over taal* 52.4:90–92.
- Hendriks, Jennifer. 2012. Case marking (accounts) in collapse: Evidence from Early Modern Dutch egodocuments (1572-1573). In *Proceedings of the 42nd Australian Linguistic Society Conference*, ed. Loan Dao Maña Ponsonnet and Margit Bowler, 123–151. Canberra Australia: Australian Linguistic Society. URL <http://hdl.handle.net/1885/9399>.
- Hermannsson, Halldór. 1919. *Modern Icelandic: An Essay*. Islandica: An Annual Relating to Iceland and the Fiske Icelandic Collection in Cornell University Library, Vol. 12. Ithaca, New York: Cornell University Library.
- Hernández-Campoy, Juan Manuel, and Juan Camilo Conde-Silvestre (eds.). 2012. *The Handbook of Historical Sociolinguistics*. Blackwell Handbooks in Linguistics. Chichester: Wiley-Blackwell.
- Heycock, Caroline, Antonella Sorace, and Zakaris Svabo Hansen. 2010. V-to-I and V2 in subordinate clauses: An investigation of Faroese in relation to Icelandic and Danish. *The Journal of Comparative Germanic Linguistics* 13.1:61–97.
- Heycock, Caroline, Antonella Sorace, Zakaris Svabo Hansen, Frances Wilson, and Sten Vikner. 2012. Detecting the late stages of syntactic change: The loss of V-to-T in Faroese. *Language* 88.3:558–600.
- Heycock, Caroline, and John D. Sundquist. 2017. Don't rush to rehabilitate: A remark on Koenehan and Zeijlstra 2014. *Linguistic Inquiry* 48.1:173–179.
- Heycock, Caroline, and Joel Wallenberg. 2013. How variational acquisition drives syntactic change. *The Journal of Comparative Germanic Linguistics* 16.2/3:127–157.
- Hilmarsson-Dunn, Amanda, and Ari Páll Kristinsson. 2010. The language situation in Iceland. *Current Issues in Language Planning* 11.3:207–276.
- Hinrichs, Lars, Benedikt Szmrecsanyi, and Bohmann Axel. 2015. Which-hunting and the Standard English relative clause. *Language* 91.4:806–836.
- Hjartardóttir, Þóra Björk. 1993. *Getið í eyðurnar: Um eyður fyrir frumlög og andlög í eldri íslensku*. Reykjavík: Málvísindastofnun Háskóla Íslands.
- Hofstadter, Douglas, and Emmanuel Sander. 2013. *Surfaces and Essences: Analogy as the Fuel and Fire of Thinking*. New York: Basic Books.
- Holmberg, Anders. 2005. Is there a little pro? Evidence from Finnish. *Linguistic Inquiry* 36.4:533–564.
- Holmberg, Anders. 2010. Parameters in minimalist theory: The case of Scandinavian. *Theoretical Linguistics* 36.1:1–48.
- Holmberg, Anders. 2015. Verb second. In *Syntax – Theory and Analysis: An International Handbook of Contemporary Syntactic Research*, ed. Tibor Kiss and Artemis Alexiadou, 342–382. Berlin: Walter de Gruyter. URL <http://ling.auf.net/lingbuzz/001087>.
- Holmberg, Anders, and On-Usa Phimsawat. 2017. Truly minimal pronouns. *Diadorim* 19:11–36.
- Holmberg, Anders, and Christer Platzack. 2005. The Scandinavian languages. In *The Oxford Handbook of Comparative Syntax*, ed. Guglielmo Cinque and Richard S. Kayne, 420–458. Oxford/New York: Oxford University Press.
- Holmberg, Anders, and Ian Roberts. 2013. The syntax-morphology relation. *Lingua* 130:111–131.
- Hooper, Joan B., and Sandra A. Thompson. 1973. On the applicability of root transformations. *Linguistic Inquiry* 4.4:465–497.
- Hornstein, Nobert, Jairo Nunes, and Kleanthes K. Grohmann. 2005. *Understanding Minimalism*. Cambridge Textbooks in Linguistics. Cambridge: Cambridge University Press.
- Hrafnbjargarson, Gunnar Hrafn. 2004. Stylistic Fronting. *Studia Linguistica* 58.2:88–134.
- Hrafnbjargarson, Gunnar Hrafn. 2008. Liberalizing modals and floating clause boundaries. *Working Papers in Scandinavian Syntax* 82:103–130.

- Hrafnbjargarson, Gunnar Hrafn, and Anna-Lena Wiklund. 2009. General embedded V2: Icelandic A, B, C, etc. *Working Papers in Scandinavian Syntax* 84:21–51.
- Hrafnbjargarson, Gunnar Hrafn, and Anna-Lena Wiklund. 2010. AGR and V2. *Theoretical Linguistics* 36.1:57–68.
- Hróarsdóttir, Þorbjörg. 1998. *Setningafræðilegar breytingar á 19. öld: Þróun þriggja málbreytinga*. Málfræðirannsóknir 10. Reykjavík: Málvísindastofnun Háskóla Íslands.
- Hróarsdóttir, Þorbjörg. 2000. *Word Order Change in Icelandic: From OV to VO*. *Linguistik Aktuell* 35. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Hróarsdóttir, Þorbjörg. 2009. OV languages: Expressions of cues. In *Information Structure and Language Change: New Approaches to Word Order Variation in Germanic*, ed. Roland Hinterhölzl and Svetlana Petrova, Trends in Linguistics. Studies and Monographs 203, 67–90. Berlin/New York: Mouton de Gruyter.
- Hróarsdóttir, Þorbjörg, Anna-Lena Wiklund, Kristine Bentzen, and Gunnar Hrafn Hrafnbjargarson. 2007. The afterglow of verb movement. *Working Papers in Scandinavian Syntax* 80:45–75.
- ÍÆ = *Íslenskar æviskrár frá landnáms tímum til ársloka 1940*. 1948–1952. Páll Eggert Ólason (ed.). Vol. 1–5. Reykjavík: Hið íslenska bókmenntafélag.
- ÍÆ = *Íslenskar æviskrár frá landnáms tímum til ársloka 1965*. 1976. Jón Guðnason and Ólafur Þ. Kristjánsson. Vol. 6. Reykjavík: Hið íslenska bókmenntafélag.
- Ingason, Anton Karl. 2016. *Realizing Morphemes in the Icelandic Noun Phrase*. Doctoral dissertation, University of Pennsylvania.
- Ingason, Anton Karl, Julie Anne Legate, and Charles Yang. 2013. The evolutionary trajectory of the Icelandic New Passive. *University of Pennsylvania Working Papers in Linguistics* 19.2:91–100.
- Ingason, Anton Karl, Einar Freyr Sigurðsson, and Joel Wallenberg. 2011. Distinguishing change and stability: A quantitative study of Icelandic oblique subjects. DiGS 13, University of Pennsylvania, June 2–5 2011. URL <http://linguist.is/skjol/digs13oblique-slides.pdf>.
- Ingason, Anton Karl, Einar Freyr Sigurðsson, and Joel Wallenberg. 2012. Antisocial syntax: Disentangling the Icelandic VO/OV parameter and its lexical remains. DiGS 14, University of Lisbon, July 4–6 2012. URL <http://linguist.is/skjol/antisocial2012.pdf>.
- Irvine, Judith T., and Susan Gal. 1995. The boundaries of languages and disciplines: How ideologies construct difference. *Social Research* 62.4:967–1001.
- Irvine, Judith T., and Susan Gal. 2000. Language ideology and linguistic differentiation. In *Regimes of Language: Ideologies, Politics, and Identities*, ed. Paul V. Kroskrity, 35–84. Santa Fe: School of American Research Press.
- Íslenskt textasafn. [No year]. The Árni Magnússon Institute for Icelandic Studies. URL <http://corpus.arnastofnun.is/>.
- Jackendoff, Ray S. 1972. *Semantic Interpretation in Generative Grammar*. Studies in Linguistics Series 2. Cambridge (MA)/London: The MIT Press.
- Jahr, Ernst Håkon. 1989. Limits of language planning? Norwegian language planning revisited. *International Journal of the Sociology of Language* 80:33–39.
- Jahr, Ernst Håkon. 2014. *Language Planning as a Sociolinguistic Experiment: The Case of Modern Norwegian*. Edinburgh: Edinburgh University Press.
- Jensen, Torben Juel. 2009. Generic variation? Developments in use of generic pronouns in late 20th century spoken Danish. *Acta Linguistica Hafniensia* 41.1:83–115.
- Jensen, Torben Juel. 2011. Ordstilling i ledsætninger i moderne dansk talesprog. *Ny forskning i grammatik* 18:123–150.
- Jensen, Torben Juel, and Tanya Karoli Christensen. 2013. Promoting the demoted: The distribution and semantics of “main clause word order” in spoken Danish complement clauses. *Lingua* 137:38–58.
- Johnson, Daniel Ezra. 2009. Getting off the GoldVarb standard: Introducing Rbrul for mixed-effects variable rule analysis. *Language and Linguistics Compass* 3.1:359–383.
- Johnson, Daniel Ezra. 2014. Progress in regression: Why natural language data calls for mixed-effects models. [Self-published manuscript]. URL http://www.danielezrajohnson.com/johnson_2014b.pdf.
- Jónasson, Jónas. 1909. *Íslensk málfræði handa byrjendum*. Akureyri: Bókavarslun og prentsmiðja Odds Björnssonar. [First edition].
- Jónasson, Jónas. 1920. *Íslensk málfræði handa byrjendum*. Akureyri: Bókavarslun og prentsmiðja Odds Björnssonar. [Third edition].
- Jónsson, Bjarni. 1893. *Íslensk málsgreinafræði*. Reykjavík: Prentsmiðja Ísafoldar.
- Jónsson, Björn. 1900. Viðbætur: Nokkur mállýti. In *Íslensk stafsetningarorðabók*, 62–64. Reykjavík: Ísafoldarprentsmiðja.

- Jónsson, Jóhannes Gísli. 1992. The pronoun *maður* in Icelandic. Unpublished manuscript, University of Massachusetts, Amherst.
- Jónsson, Jóhannes Gísli. 2002. S-adverbs in Icelandic and the feature theory of adverbs. *Leeds Working Papers in Linguistics and Phonetics* 9:73–89.
- Jónsson, Jóhannes Gísli, and Þórhallur Eyþórsson. 2003. Breytingar á frumlagsfalli í íslensku. *Íslenskt mál og almenn málfræði* 25:7–40.
- Jónsson, Runólfur [lat. per Runolphum Jonam]. 1651. *Grammaticæ islandicæ rudimenta*. Hafnia: Petrus Harius. URL http://baekur.is/is/bok/000337214/Grammaticae_Islandicae.
- Jónsson, Steingrímur. 1998. Upphaf Þjóðólfs 1848. *Ritmennt: Ársrit Landsbókasafns Íslands – Háskólabókasafns* 3:65–88.
- Julien, Marit. 2005. *Nominal Phrases from a Scandinavian Perspective*. Linguistik Aktuell 87. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Julien, Marit. 2008. Sá vanleg at det kan ikkje afveist: Om V2 i innfýdde setningar. In *Språk i Oslo. Ny forskning omkring talespråk*, ed. Janne Bondi Johannessen and Kristin Hagen, 159–171. Oslo: Novus Forlag. URL <http://lup.lub.lu.se/record/1272595>.
- Julien, Marit. 2009. Embedded clauses with main clause word order in Mainland Scandinavian. URL <http://ling.auf.net/lingbuzz/000475>.
- Kalm, Mikael, and Anna Sahlée. 2018. Lärarkommentarer i studentuppsatser från 1800-talets andra hälft. *Svenska språkets historia* 15, Tartu, 14th–15th June. Institutionen för nordiska språk, Uppsala universitet.
- Kaplan, Robert B., and Jr Baldauf, Richard B. 1997. *Lanugage Planning: From Practice to Theory*. Multilingual Matters 108. Clevedon/Philadelphia/Toronto/Sydney/Johannesburg: Multilingual Matters Ltd.
- Kinn, Kari, Kristian A. Rusten, and George Walkden. 2016. Null subjects in Early Icelandic. *Journal of Germanic Linguistics* 28.1:31–78.
- Kjartansson, Helgi Skúli. 2017. Jafnvel í forn máli? Um orðið „maður“ og fornafnseinkenni þess. Paper presented at Málvísindakaffi, 3rd March 2017, University of Iceland.
- Koch, Peter, and Wulf Oesterreicher. 1985. Sprache der Nähe — Sprache der Distanz: Mündlichkeit und Schriftlichkeit im Spannungsfeld von Sprachtheorie und Sprachgeschichte. *Romanistisches Jahrbuch* 36:15–43.
- Koch, Peter, and Wulf Oesterreicher. 2007. Schriftlichkeit und kommunikative Distanz. *Zeitschrift für germanistische Linguistik* 35.3:346–375.
- Koenenman, Olaf. 2000. *The Flexible Nature of Verb Movement*. LOT 37. Utrecht: LOT.
- Koenenman, Olaf. 2010. Verb movement in Germanic and Celtic languages: A flexible approach. *Lingua* 120.2:210–231.
- Koenenman, Olaf, and Hedde Zeijlstra. 2014. The Rich Agreement Hypothesis rehabilitated. *Linguistic Inquiry* 45.4:571–615.
- Kolbeinnsson, Guðni. 1976. Íslenskuþættir Alþýðublaðsins. *Alþýðublaðið* 57.172:6.
- Kristinsson, Ari Páll. 2007. Málræktarfræði. *Íslenskt mál og almenn málfræði* 29:99–124.
- Kroch, Anthony. 1989. Reflexes of grammar in patterns of language change. *Language Variation and Change* 1.3:199–244.
- Kroch, Anthony. 1994. Morphosyntactic variation. In *Papers from the 30th Regional Meeting of the Chicago Linguistics Society: Parasession on Variation and Linguistic Theory*, ed. Katharine Beals, 180–201. Chicago, IL: Chicago Linguistic Society. URL <https://www.ling.upenn.edu/~kroch/papers/morphosyntax.pdf>.
- Kroch, Anthony. 2005. Modeling language change and language acquisition. [Unpublished expansion of 2005 LSA Institute Forum Lecture]. URL ftp://babel.ling.upenn.edu/papers/faculty/tony_kroch/papers/lsa-forum.pdf.
- Krogull, Andreas. 2018. *Policy versus Practice: Language Variation and Change in Eighteenth- and Nineteenth-Century Dutch*. LOT 521. Utrecht: LOT.
- Krogull, Andreas, Gijsbert Rutten, and Marijke van der Wal. 2017. Relativisation in Dutch diaries, private letters and newspapers (1770–1840): A genre-specific national language? In *Exploring Future Paths for Historical Sociolinguistics*, ed. Tanja Säily, Arja Nurmi, Minna Palander-Collin, and Anita Auer, Advances in Historical Sociolinguistics 7. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Kusters, Wouter. 2003. *Linguistic Complexity: The Influence of Social Change on Verbal Inflection*. LOT 77. Utrecht: LOT.
- Kvaran, Guðrún. 2005. *Orð: Handbók um beygingar- og orðmyndunarfræði*. Íslensk tunga II. Reykavík: Almenna bókafélagið.

- Labov, William. 1965. On the mechanism of linguistic change. In *Report of the Sixteenth Annual Round Table Meeting on Linguistics and Language Studies*, ed. Charles W. Kreidler, 91–114. Washington, D.C.: Georgetown University Press/Institute of Languages and Linguistics.
- Labov, William. 1972a. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.
- Labov, William. 1972b. Some principles of linguistic methodology. *Language in Society* 1:97–120.
- Labov, William. 1990. The intersection of sex and social class in the course of linguistic change. *Language Variation and Change* 2.2:205–254.
- Labov, William. 1994. *Principles of Linguistic Change: Internal Factors*. Vol. 1. Oxford/Cambridge, MA: Blackwell.
- Labov, William. 2001. *Principles of Linguistic Change: Social Factors*. Vol. 2. Oxford: Blackwell.
- Labov, William. 2007. Transmission and diffusion. *Language* 83.2:344–387.
- Labov, William. 2010. *Principles of Linguistic Change: Cognitive and Cultural Factors*. Vol. 3. Chichester: Wiley-Blackwell.
- Labov, William. 2016. Afterword: Where are we now? *Journal of Sociolinguistics* 20.4:581–602.
- Labov, William, Sharon Ash, Maciej Baranowski, Naomi Nagy, Maya Ravindranath, and Tracey Weldon. 2006. Listeners' sensitivity to the frequency of sociolinguistic variables. *University of Pennsylvania Working Papers in Linguistics* 12. URL <https://repository.upenn.edu/pwpl/voll2/iss2/10>.
- Labov, William, Sharon Ash, Maya Ravindranath, Tracey Weldon, Maciej Baranowski, and Naomi Nagy. 2011. Properties of the sociolinguistic monitor. *Journal of Sociolinguistics* 15.4:431–463.
- Labov, William, and Wendell A. Harris. 1986. De facto segregation of black and white vernaculars. In *Diversity and Diachrony*, ed. David Sankoff, Current Issues in Linguistic Theory 53, 1–24. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Landau, Idan. 2006. Chain resolution in Hebrew V(P)-fronting. *Syntax* 9.1:32–66.
- Lander, Eric T., and Liliane Haegeman. 2014. Old Norse as an NP language: With observations on the Common Norse and Northwest Germanic runic inscriptions. *Transactions of the Philological Society* 112.3:279–318.
- Lavandera, Beatriz R. 1978. Where does the sociolinguistic variable stop? *Language in Society* 7.2:171–182.
- Laycock, Donald C., and Peter Mühlhäusler. 1990. Language engineering: Special languages. In *An encyclopaedia of language*, ed. N. E. Collinge, 843–875. Routledge.
- Leerssen, Joep. 1999. *Nationaal denken in Europa: Een cultuurhistorische schets*. Amsterdam: Amsterdam University Press.
- Leerssen, Joep. 2006a. *National Thought in Europe: A Cultural History*. Amsterdam: Amsterdam University Press.
- Leerssen, Joep. 2006b. Nationalism and the cultivation of culture. *Nations and Nationalism* 12.4:559–578.
- Leerssen, Joep. 2008. Þjóðernisstefna og ræktun menningar. *Ritið* 1/2008:190–214. [Translated by María Bjarkadóttir].
- Leonard, Stephen Pax, and Kristján Árnason. 2011. Language ideology and standardisation in Iceland. In *Standard Languages and Language Standards in a Changing Europe*, ed. Tore Kristiansen and Nikolas Coupland, Standard Language Ideology in Contemporary Europe 1, 91–96. Oslo: Novus Press.
- Levon, Erez, and Isabelle Buchstaller. 2015. Perception, cognition, and linguistic structure: The effect of linguistic modularity and cognitive style on sociolinguistic processing. *Language Variation and Change* 27.3:319–348.
- Lightfoot, David W. 1979. *Principles of Diachronic Syntax*. Cambridge Studies in Linguistics. Cambridge: Cambridge University Press.
- Lofsson, Hrafn, and Anton Karl Ingason. 2011. IceNLP: A Natural Language Processing Toolkit for Icelandic. URL <https://github.com/hrafnl/icenlp>.
- Maddeaux, Ruth, and Aaron Dinkin. 2017. Is like like like?: Evaluating the same variant across multiple variables. *Linguistics Vanguard* 3.1.
- Magnúss, Gunnar M. 1939. *Saga alþýðufræðslunnar á Íslandi: Háttíðarrit S. Í. B.* Reykjavík: Samband íslenzkra barnakennara.
- Magnússon, Eiríkur. 1870. [No title]. *Norðanfari* 9:67–68, 71–72.
- Magnússon, Friðrik. 1984. Um innri gerð nafnliða í íslensku. *Íslenskt mál og almenn málfræði* 6:81–111.
- Magnússon, Sigurður Gylfi. 1995. From children's point of view: Childhood in nineteenth-century Iceland. *Journal of Social History* 29.2:295–323.
- Magnússon, Sigurður Gylfi. 2010. *Wasteland with Words: A Social History of Iceland*. London: Reaktion Books.
- Magnússon, Sigurður Gylfi, and Davíð Ólafsson. 2012. Minor knowledge: Microhistory, scribal communities, and the importance of institutional structures. *Quaderni storici* 47.140-2:495–524.

- Mair, Christian. 2009. *Twentieth-Century English: History, Variation, and Standardization*. Studies in English Language. Cambridge: Cambridge University Press.
- Maling, Joan. 1980. Inversion in embedded clauses in Modern Icelandic. *Íslenskt mál og almenn málfræði* 2:175–193.
- McEnery, Tony, and Andrew Hardie. 2012. Support website for *Corpus Linguistics: Method, Theory and Practice*. Cambridge: Cambridge University Press. URL <http://corpora.lancs.ac.uk/clmtp/2-stat.php>.
- McWhorter, John H. 2011. *Linguistic Simplicity and Complexity: Why Do Languages Undress?* Language Contact and Biligualism 1. Boston/Berlin: De Gruyter Mouton.
- Meinunger, André. 2004. Verb position, verbal mood and the anchoring (potential) of sentences. In *The Syntax and Semantics of the Left Periphery*, ed. Horst Lohnstein and Susanne Trissler, 313–341. Berlin/New York: Mouton de Gruyter.
- Meyerhoff, Miriam, and James A. Walker. 2007. The persistence of variation in individual grammars: Copula absence in ‘urban sojourners’ and their stay-at-home peers, Bequia (St Vincent and the Grenadines). *Journal of Sociolinguistics* 11.3:346–366.
- Meyerhoff, Miriam, and James A. Walker. 2013. An existential problem: The sociolinguistic monitor and variation in existential constructions on Bequia (St. Vincent and the Grenadines). *Language in Society* 42.4:407–428.
- Milroy, James, and Lesley Milroy. 2012. *Authority in Language: Investigating Standard English*. Routledge Linguistics Classics. Routledge. [Fourth edition].
- MÍM = Sigrún Helgadóttir, Ásta Svavarsdóttir, Eiríkur Rögnvaldsson, Kristín Bjarnadóttir, and Hrafn Loftsson. 2012. The Tagged Icelandic Corpus (MÍM). In *Proceedings of the Workshop on Language Technology for Normalisation of Less-Resourced Languages – SaLTMiL 8 – AflaT 2012*, 67–72. Istanbul. URL <http://mim.hi.is>.
- Möller, Alda B. 2017. Íslenskukennsla í Bessastaðaskóla 1806–1846 og á fyrstu árum Reykjavíkurskóla. *Orð og tunga* 19:1–40.
- Müller, Max. 1861. *Lectures on the Science of Language: Delivered at the Royal Institution of Great Britain in April, May, and June, 1861*. London: Longman, Green, Longman, & Roberts.
- Müller, Max. 1862. *Lectures on the Science of Language: Delivered at the Royal Institution of Great Britain in April, May, and June, 1861*. New York: Charles Scribner. [From the second London edition, revised].
- Neeleman, Ad, and Fred Weerman. 1999. *Flexible Syntax: A Theory of Case and Arguments*. Dordrecht/Boston/London: Kluwer Academic Publishers.
- Nevalainen, Terttu, and Helena Raumolin-Brunberg. 2011. The diffusion of language change in real time: Progressive and conservative individuals and the time depth of change. *Language Variation and Change* 23.1:1–43.
- Nobels, Judith. 2013. *(Extra)Ordinary Letters: A View from Below on Seventeenth-Century Dutch*. LOT 329. Utrecht: LOT.
- Nowenstein, Iris Edda. 2017. Determining the nature of intra-speaker subject case variation. In *Syntactic Variation in Insular Scandinavian*, ed. Höskuldur Thráinsson, Caroline Heycock, Hjalmar P. Petersen, and Zakaris Svabo Hansen, Studies in Germanic Linguistics 1, 91–112. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Nygaard, Marius. 1905. *Norrøn syntax*. Kristiania: H. Aschehoug & Co. (W. Nygaard).
- Ólafsson, Bragi Þorgrímur. 2004. *Landsins útvöldu synir: Ritgerðir skólapilta Lærða skólans í íslenskum stíl 1846-1904*. Sýnisbók íslenskrar alþýðumenningar 7. Reykjavík: Háskólaútgáfan.
- Ólafsson, Davíð. 2012. Vernacular literacy practices in nineteenth-century Icelandic scribal culture. In *Att läsa och att skriva: Två vågor av vardagligt skriftbruk i Norden 1800-2000*, ed. Ann-Catrine Edlund, Nordliga studier 3, Vardagligt skriftbruk 1, 65–85. Umeå: Umeå universitet, Kungl. Skytteanska Samfundet.
- Ólsen, Björn Magnússon. 1882. Zur neuisländischen Grammatik. *Germania: Vierteljahrsschrift für deutsche Alterthumskunde* 27:257–287.
- Ordbog over det norrøne prosasprog*. [No year]. Aldís Sigurðardóttir, Simonetta Battista, Ellert Þór Jóhannsson and Þorbjörg Helgadóttir (eds.). Nordisk Forskningsinstitut, Den Arnamagnæanske Kommission, Den Arnamagnæanske Samling. URL <https://onp.ku.dk/>.
- Ottósson, Kjartan G. 1987. An archaising aspect of Icelandic purism: The revival of extinct morphological patterns. In *The Nordic Languages and Modern Linguistics*, ed. Pirkko Lilius and Mirja Saari, Vol. 6, 311–324. Helsinki: Helsinki University Press.
- Ottósson, Kjartan G. 1990. *Íslensk málhreinun: sögulegt yfirlit*. Rit Íslenskrar málnefndar 6. Reykjavík: Íslensk málnefnd.

- Ottosson, Kjartan. 2002. Research in Icelandic language history 1850-1950: An overview. In *The Nordic Languages: An International Handbook of the History of the North Germanic Languages*, ed. Oskar Bandle, Kurt Braunmüller, Ernst Håkon Jahr Jahr, Allan Karker, Hans-Peter Naumann, and Ulf Teleman, Handbücher zur Sprach- und Kommunikationswissenschaft 22.1, 95–101. Berlin/New York: Walter de Gruyter.
- Ottosson, Kjartan. 2003. Heimenorsk innverknad på islandsk språk i mellomalderen, særleg i morfologien. In *Úmorður: West Nordic Standardisation and Variation*, ed. Kristján Árnason, 111–152. Reykjavík: Institute of Linguistics/University of Iceland Press. Papers from a Symposium in Stockholm October 7th 2001.
- Ottosson, Kjartan. 2005. Language cultivation and language planning IV: Iceland. In *The Nordic Languages: An International Handbook of the History of the North Germanic Languages*, ed. Oskar Bandle, Kurt Braunmüller, Ernst Håkon Jahr Jahr, Allan Karker, Hans-Peter Naumann, and Ulf Teleman, 1997–2007. Berlin/New York: Walter de Gruyter.
- Paardekooper, P.C. 1961. Persoonsvorm en voegwoord. *De Nieuwe Taalgids* 54:296–301.
- Paolillo, John C. 2013. Individual effects in variation analysis: Model, software, and research design. *Language Variation and Change* 25.1:89–118.
- Pedersen, Inge Lise. 1996. ‘Der kan jo være nogen der kan itte tåle det’: Om hovedsætningsordstilling i bisætninger i danske dialekter. In *Studier i talesprogsvariation og sprogkontakt. Til Inger Ejsjær på halvfjerdsårsdagen den 20. maj 1996*, 242–251. København: C.A. Reitzels forlag.
- Percy, Carol. 2010. How eighteenth-century book reviewers became language guardians. In *Social Roles and Language Practices in Late Modern English*, ed. Arja Nurmi Päivi Pahta, Minna Nevala and Minna Palander-Collin, Pragmatics & Beyond New Series 195, 55–85. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Percy, Carol, and Ingrid Tieken-Boon van Ostade. 2016. Prescription and tradition: Establishing standards across time and space. In *Prescription and Tradition in Language: Establishing Standards across Time and Space*, ed. Ingrid Tieken-Boon van Ostade and Carol Percy, Multilingual Matters 165, 1–20. Bristol/Blue Ridge Summit: Multilingual Matters.
- Petersen, Hjalmar P. 2000. IP or TP in Modern Faroese. *Working Papers in Scandinavian Syntax* 66:75–83.
- Petersen, Hjalmar P. 2010. *The Dynamics of Faroese-Danish Language Contact*. Heidelberg: Universitätsverlag Winter.
- Petersson, Gertrud. 1988. Bisatsledföljden i svenska eller Varifrån kommer BIFF-regeln? *Arkiv för nordisk filologi* 103:157–180.
- Pétursson, Magnús. 2005. The development of Icelandic from the mid-16th century to 1800. In *The Nordic Languages: An International Handbook of the History of the North Germanic Languages*, ed. Oskar Bandle, Kurt Braunmüller, Ernst Håkon Jahr Jahr, Allan Karker, Hans-Peter Naumann, and Ulf Teleman, Handbücher zur Sprach- und Kommunikationswissenschaft 22.2, 1258–1269. Berlin/New York: Walter de Gruyter.
- Pfaff, Alexander. 2014. Inside and outside – Before and after. In *Adjectives in Germanic and Romance*, ed. Petra Sleeman, Freek Van de Velde, and Harry Perridon, Linguistik Aktuell 212, 217–244. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Pfaff, Alexander Peter. 2015. *Adjectival and Genitival Modification in Definite Noun Phrases in Icelandic: A Tale of Outsiders and Inside Jobs*. Tromsø: UiT The Arctic University of Norway.
- Platzack, Christer. 1986. COMP, INFL, and Germanic word order. In *Topics in Scandinavian Syntax*, ed. Lars Hellan and Kristi Koch Christensen, Studies in Natural Language and Linguistic Theory, 185–234. Dordrecht/Boston/Lancaster/Tokyo: D. Reidel Publishing Company.
- Platzack, Christer, and Anders Holmberg. 1989. The role of AGR and finiteness in Germanic VO languages. *Working Papers in Scandinavian Syntax* 43:51–76.
- Platzack, Christopher. 1988. The emergence of a word order difference in Scandinavian subordinate clauses. *McGill Working Papers in Linguistics* 215–238. [Special issue on Comparative Germanic Syntax].
- Pollock, Jean-Yves. 1989. Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 20.3:365–424.
- Pollock, Jean-Yves. 1997. Notes on clause structure. In *Elements of Grammar: Handbook in Generative Syntax*, ed. Liliane Haegeman, Kluwer International Handbooks of Linguistics 1, 237–279. Dordrecht/Boston/London: Kluwer.
- Poplack, Shana, and Nathalie Dion. 2009. Prescription vs. praxis: The evolution of future temporal reference in French. *Language* 85.3:557–587.
- Poplack, Shana, Lidia-Gabriela Jarmasz, Nathalie Dion, and Nicole Rosen. 2015. Searching for standard French: The construction and mining of the *Recueil historique des grammaires du français*. *Journal of Historical Sociolinguistics* 1.1:13–55.

- Postma, Gertjan. 2010. The impact of failed changes. In *Continuity and Change in Grammar*, ed. Anne Breitbarth, Christopher Lucas, Sheila Watts, and David Willis, Linguistik Aktuell 159, 269–302. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Ragnarsdóttir, Hrafnhildur, and Sven Strömquist. 2005. The development of generic *maður/man* for the construction of discourse stance in Icelandic and Swedish. *Journal of Pragmatics* 37.2:143–155.
- Rask, Erasmus Christian. 1818. *Anvisning till Isländskan eller Nordiska Fornspråket: Från Danskän öfversatt och omarbetad af Författeren*. Stockholm: A. Wiborgs förlag.
- Rask, Rasmus Kristian. 1811. *Vejledning til det Isländske eller gamle Nordiske Sprog*. Kjøbenhavn: Schubothes Forlag - J. R. Thiele.
- Rastrick, Ólafur. 2003. Lestrarkennsla og ritþjálfun í barnaskólum og alþýðlegum framhaldsskólum 1880–1920. In *Alþýðumenning á Íslandi 1830-1930: Ritað mál, menntun og félagshreyfingar*, ed. Ingi Sigurðsson and Loftur Guttormsson, Sagnfræðirannsóknir - Studia Historica 18, 93–113. Reykjavík: Háskólaútgáfan.
- Rayson, Paul. 2016. Log-likelihood and effect size calculator. University Centre for Computer Corpus Research on Language, Lancaster University. URL <http://ucrel.lancs.ac.uk/llwizard.html>.
- Rickford, John R., and Penelope Eckert. 2001. Introduction. In *Style and Sociolinguistic Variation*, ed. Penelope Eckert and John R. Rickford, 1–18. Cambridge: Cambridge University Press.
- Ritmálsafn Orðabókar Háskólans*. [No year]. The Árni Magnússon Institute for Icelandic Studies. URL <http://arnastofnun.is/page/ritmalssafn>.
- Rizzi, Luigi. 2004. On the cartography of syntactic structures. In *The Structure of CP and IP: The Cartography of Syntactic Structures*, ed. Luigi Rizzi, Vol. 2, 3–15. Oxford Studies in Comparative Syntax. Oxford/New York: Oxford University Press.
- Rizzi, Luigi, and Guglielmo Cinque. 2016. Functional categories and syntactic theory. *Annual Review of Linguistics* 2:139–163.
- Roberts, Ian G. 1985. Agreement parameters and the development of English modal auxiliaries. *Natural Language & Linguistic Theory* 3.1:21–58.
- Roehrs, Dorian. 2009. *Demonstratives and Definite Articles as Nominal Auxiliaries*. Linguistik Aktuell 140. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Rögvaldsson, Eiríkur. 1995. Old Icelandic: A non-configurational language? *North-Western European Language Evolution* 26:3–29.
- Rögvaldsson, Eiríkur. 1996. Frumlag og fall að fornu. *Íslenskt mál og almenn málfræði* 18:37–69.
- Rögvaldsson, Eiríkur. 2014. Setningarugl? Tengdar nafnháttarsetningar með frumlagi. *Íslenskt mál og almenn málfræði* 36:55–91.
- Rögvaldsson, Eiríkur, and Höskuldur Thráinsson. 1990. On Icelandic word order once more. In *Modern Icelandic Syntax*, ed. Joan Maling and Annie Zaenen, Syntax and Semantics 24, 3–40. San Diego: Academic Press.
- Rohrbacher, Bernhard Wolfgang. 1999. *Morphology-Driven Syntax: A Theory of V to I Raising and Pro-Drop*. Linguistik Aktuell 15. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Romaine, Suzanne. 1982. *Socio-Historical Linguistics: Its Status and Methodology*. Cambridge Studies in Linguistics 34. Cambridge: Cambridge University Press.
- Romaine, Suzanne. 1984. On the problem of syntactic variation and pragmatic meaning in sociolinguistic theory. *Folia Linguistica* 18.3/4:409–437.
- Rubin, Joan, and Björn H. Jernudd. 1971. Introduction: Language planning as an element in modernization. In *Can Language be Planned? Sociolinguistic Theory for Developing Nations*, ed. Joan Rubin and Björn H. Jernudd, xiii–xxiv. The University Press of Hawaii. URL <https://muse.jhu.edu/book/61287>.
- Rutten, Gijsbert. 2016. Diaglossia, individual variation and the limits of standardization: Evidence from Dutch. In *Current Trends in Historical Sociolinguistics*, ed. Cinzia Russi, 194–218. Warsaw/Berlin: De Gruyter Open.
- Rutten, Gijsbert. 2019. *Language Planning as Nation Building: Ideology, Policy and Implementation in the Netherlands, 1750-1850*. Advances in Historical Sociolinguistics 9. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Rutten, Gijsbert, Rik Vosters, and Wim Vandenbussche. 2014. The interplay of language norms and usage patterns: Comparing the history of Dutch, English, French and German. In *Norms and Usage in Language History, 1600-1900: A Sociolinguistic and Comparative Perspective*, ed. Gijsbert Rutten, Rik Vosters, and Wim Vandenbussche, Advances in Historical Sociolinguistics 3, 1–17. Amsterdam/Philadelphia: John Benjamins Publishing Company.

- Sandersen, Vibeke. 2003. „Jeg skriver dig til for at lade dig vide“: Skrivefærdighed og skriftsprog hos menige danske soldater i træårskrigen 1848–50. Vol. 1. Universitets-Jubilæets danske samfund nr. 558. C.A. Reitzels Forlag.
- Sandersen, Vibeke. 2007. Writing ability and the written language of Danish private soldiers in the Three Year's War (1848–50). *Multilingua* 26.2/3:247–278.
- Sankoff, David. 1982. Sociolinguistic method and linguistic theory. In *Logic, Methodology and Philosophy of Language VI: Proceedings of the Sixth International Congress of Logic, Methodology and Philosophy of Science, Hannover, 1979*, ed. L. Jonathan Cohen, Jerzy Łoś, Helmut Pfeiffer, and Klaus-Peter Podewski, Studies in Logic and the Foundations of Mathematics 104, 677–689. Amsterdam/New York/Oxford/Warszawa: North-Holland Publishing Company/Polish Scientific Publishers.
- Sankoff, David. 1988. Sociolinguistics and syntactic variation. In *Linguistics: The Cambridge Survey: Volume IV: Language: The Socio-cultural Context*, ed. Frederick J. Newmeyer, 140–161. Cambridge: Cambridge University Press.
- Sankoff, David, and Pierrette Thibault. 1981. Weak complementarity: Tense and aspect in Montreal French. In *Syntactic change*, ed. Brenda B. Johns and David R. Strong, Natural Language Studies 25, 205–216. Ann Arbor: University of Michigan.
- Siewierska, Anna. 1999. From anaphoric pronoun to grammatical agreement marker: Why objects don't make it. *Folia Linguistica* 33.2:225–251.
- Sigmundsson, Finnur (ed.). 1952. *Sendibréf frá íslenskum konum 1784-1900*. Reykjavík: Helgafell.
- Sigmundsson, Finnur (ed.). 1961. *Konur skrifa bréf: Sendibréf 1797-1907*. Reykjavík: Bókfellsútgáfan.
- Sigmundsson, Svavar. 1981. Íslensk málfræðiheiti á 19. öld. In *Afmæliskveðja til Halldórs Halldórssonar 13. júlí 1981*, ed. Guðrún Kvaran, Gunnlaugur Ingólfsson, and Sigmundsson Svavar, 285–300. Reykjavík: Íslenska málfræðifélagið.
- Sigmundsson, Svavar. 1987. Íslenska í samanburði við önnur mál. *Íslenskt mál og almenn málfræði* 9:95–109.
- Sigmundsson, Svavar. 1990–1991. Hreinsun íslenskunnar. *Íslenskt mál og almenn málfræði* 12–13:127–142.
- Sigmundsson, Svavar. 2002. Trends in the linguistic development since 1945 IV: Icelandic. In *The Nordic Languages: An International Handbook of the History of the North Germanic Languages*, ed. Oskar Bandle, Kurt Braunmüller, Ernst Håkon Jahr, Allan Karker, Hans-Peter Naumann, and Ulf Teleman, Handbücher zur Sprach- und Kommunikationswissenschaft 22.1, 1832–1839. Berlin/New York: Walter de Gruyter.
- Sigtryggsson, Jóhannes Bjarni. 2003. Hrein tunga. In *Þjóðerni í þúsund ár?*, ed. Jón Yngvi Jóhannsson, Kolbeinn Óttarsson Proppé, and Sverrir Jakobsson, 91–103. Reykjavík: Háskólaútgáfan.
- Sigtryggsson, Jóhannes B. 2011. *Málið á Evisögu Jóns Steingrímssonar*. Reykjavík: Hugvísindastofnun Háskóla Íslands.
- Sigtryggsson, Jóhannes B. 2017. Samræmdur úrvalsriðáttur fornboka – réttitrun Jóns Þorkelssonar. *Orð og tunga* 19:155–171.
- Sigurðsson, Halldór Ármann. 1986. Verb post-second in a V2 language. In *Scandinavian Syntax: Workshop at the Ninth Scandinavian Conference of Linguistics, Stockholm, January 1986*, ed. Östen Dahl and Anders Holmberg, 138–149. Stockholm: Institute of Linguistics, University of Stockholm.
- Sigurðsson, Halldór Ármann. 1989. Verbal Syntax and Case in Icelandic. Doctoral dissertation, Lund University.
- Sigurðsson, Halldór Ármann. 1993a. Argument-drop in Old Icelandic. *Lingua* 89.2/3:247–280.
- Sigurðsson, Halldór Ármann. 1993b. The structure of the Icelandic NP. *Studia Linguistica* 47.2:177–197.
- Sigurðsson, Halldór Ármann. 2006. The Icelandic noun phrase: Central traits. *Arkiv för nordisk filologi* 121:193–236.
- Sigurðsson, Halldór Ármann. 2011. Uniformity and diversity: A minimalist perspective. *Linguistic Variation* 11.2:189–222.
- Sigurðsson, Halldór Ármann. 2017. Stylistic Fronting in corpora. In *Syntactic Variation in Insular Scandinavian*, ed. Höskuldur Thráinsson, Caroline Heycock, Hjalmar P. Petersen, and Zakaris Svabo Hansen, Studies in Germanic Linguistics 1, 307–338. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Sigurðsson, Halldór Ármann, and Verner Egerland. 2009. Impersonal null-subjects in Icelandic and elsewhere. *Studia Linguistica* 63.1:158–185.
- Sigurðsson, Haraldur (ed.). 1968. *Skólaræður Sveinbjarnar Egilssonar*. Reykjavík: Almenna bókafélagið.
- Sigurjónsdóttir, Sigríður, and Joan Maling. 2001. Það var hrint mér á leiðinni í skólann: Þolmynd eða ekki þolmynd? *Íslenskt mál og almenn málfræði* 23:123–180.
- Sigurmundsson, Sigurður. 1996. Um notkun orðsins maður. *Morgunblaðið* 84.206:41.
- van der Sijs, Noline. 2004. *Taal als mensenwerk: Het ontstaan van het ABN*. Den Haag: Sdu Uitgevers.

- Silverstein, Michael. 1985. Language and the culture of gender: At the intersection of structure, usage, and ideology. In *Semiotic Mediation: Sociocultural and Psychological Perspectives*, ed. Elizabeth Mertz and Richard J. Parmentier, Language, Thought, and Culture: Advances in the Study of Cognition, 219–259. Orlando: Academic Press, Inc.
- Skautrup, Peter. 1953. *Det danske sprogs historie: Fra Holbergs komedier til H. C. Andersens eventyr*. Vol. 3. København: Gyldendalske boghandel/Nordisk forlag.
- Smári, Jakob Jóh. 1920. *Íslensk setningafræði*. Reykjavík: Bókaverzlun Ársæls Árnasonar.
- Sneller, Betsy, and Sabriya Fisher. 2015. When GET got noticed: The emerging salience of GET-passives. *University of Pennsylvania Working Papers in Linguistics* 21.1:1–9. URL <http://repository.upenn.edu/pwpl/vol21/iss1/34>.
- Steingrímsson, Steinþór, Sigrún Helgadóttir, Eiríkur Rögnvaldsson, Starkaður Barkarson, and Jón Guðnason. 2018. Risamálheild: A very large Icelandic text corpus. In *Proceedings of LREC 2018*, 4361–4366. Myazaki. URL <http://malheildir.arnastofnun.is>.
- Sundquist, John D. 2003. The Rich Agreement Hypothesis and Early Modern Danish embedded-clause word order. *Nordic Journal of Linguistics* 26.2:233–258.
- Svavarsdóttir, Ásta. 1982. „Þágufallssýki“: Breytingar á fallnotkun í frumlagssæti ópersónulegra setninga. *Íslenskt mál og almenn málfræði* 4:19–62.
- Svavarsdóttir, Ásta, Gísli Pálsson, and Þórólfur Þórlindsson. 1984. Fall er fararheill: Um fallnotkun með ópersónulegum sögnum. *Íslenskt mál og almenn málfræði* 6:33–55.
- Svavarsdóttir, Ásta, Sigrún Helgadóttir, and Guðrún Kvaran. 2014. Language resources for Early Modern Icelandic. In *LREC 2014 Proceedings: LRT4HDA Workshop*. URL <http://www.lrec-conf.org/proceedings/-lrec2014/workshops/LREC2014Workshop-LRT4HDA%20Proceedings.pdf>.
- Szabolcsi, Anna. 1983. The possessor that ran away from home. *The Linguistic Review* 3.1:89–102.
- Tagliamonte, Sali A. 2012. *Variationist Sociolinguistics: Change, Observation, Interpretation*. Chichester: Wiley-Blackwell.
- Tagliamonte, Sali A., and R. Harald 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.
- Tagliamonte, Sali A., and Alexandra D’Arcy. 2009. Peaks beyond Phonology: Adolescence, Incrementation, and Language Change. *Language* 85.1:58–108.
- Terkourafi, Marina. 2011. The pragmatic variable: Toward a procedural interpretation. *Language in Society* 40.3:343–372.
- Thomason, Sarah G. 1997. On mechanisms of interference. In *Language and its Ecology: Essays in Memory of Einar Haugen*, ed. Stig Eliasson and Ernst Håkon Jahr, Trends in Linguistics, Studies and Monographs 100, 181–207. Berlin/New York: Mouton de Gruyter.
- Thomason, Sarah G. 1999. Speakers’ choices in language change. *Studies in the Linguistic Sciences* 29.2:19–43.
- Thomason, Sarah G. 2001. *Language Contact*. Washington, D.C.: Georgetown University Press.
- Thomason, Sarah G. 2007. Language Contact and Deliberate Change. *Journal of Language Contact* 1.1:41–62.
- Þorkelsson, Jón. 1870. Um nokkurar rangar orðmyndir eða orðskipanir í íslenzku. *Norðanfari* 9:82–83, 86–87, 89–90.
- Þorkelsson, Jón [=J. Þ.]. 1903. Haldór Kristján Friðriksson. *Andvari* 28:160.
- Þórólfsson, Björn K. 1925. *Um íslenskar orðmyndir á 14. og 15. öld og breytingar þeirra úr forn málinu: Með viðauka um nýjungar í orðmyndum á 16. öld og síðar*. Reykjavík: Fjeldagsprentsmiðjan. [Rit um íslenska málfræði 2. Málvísindastofnun Háskóla Íslands. 1987].
- Thorsteinson, Árni. 1946. Á víð og dreif. In *Minningar úr menntaskóla*, ed. Ármann Kristinsson and Friðrik Sigurbjörnsson, 77–88. Reykjavík: Ármann Kristinsson.
- Þorsteinsson, Helgi Már. 2012. „Að vanrækja móðurmálið... það er synd, sem hefnir sín.“: Íslenskukennsla í barnaskólum 1880–1920. Unpublished master’s thesis, University of Iceland. URL <http://hdl.handle.net/1946/13431>.
- Þráinsson, Höskuldur. 1980. Tilvísunarformöfn? *Íslenskt mál og almenn málfræði* 53–96.
- Þráinsson, Höskuldur. 1986. On auxiliaries, AUX and VPs in Icelandic. In *Topics in Scandinavian Syntax*, ed. Lars Hellan and Kirsti Koch Christensen, 235–265. Studies in Natural Language and Linguistic Theory. Dordrecht/Boston/Lancaster/Tokyo: D. Reidel Publishing Company.
- Þráinsson, Höskuldur. 1994. Icelandic. In *The Germanic Languages*, ed. Ekkehard König and Johan van der Auwera, Routledge Language Family Descriptions, 142–189. London/New York: Routledge.

- Thráinsson, Höskuldur. 1996. On the non-universality of functional categories. In *Minimalist Ideas: Syntactic Studies in the Minimalist Framework*, ed. Werner Abraham, Samuel David Epstein, Höskuldur Thráinsson, and Jan-Wouter Zwart, Linguistik Aktuell 12, 253–281. Amsterdam: John Benjamins Publishing Company.
- Thráinsson, Höskuldur. 2001. Um sagnbeygingu, sagnfærslu og setningagerð í færeysku og fleiri málum. *Íslenskt mál og almenn málfræði* 23:7–70.
- Thráinsson, Höskuldur. 2003. Syntactic variation, historical development, and Minimalism. In *Minimalist syntax*, ed. Randall Hendrick, 152–191. Oxford: Blackwell.
- Thráinsson, Höskuldur. 2007. *The Syntax of Icelandic*. Cambridge Syntax Guides. Cambridge: Cambridge University Press.
- Thráinsson, Höskuldur. 2010. Predictable and unpredictable sources of variable verb and adverb placement in Scandinavian. *Lingua* 120.5:1062–1088.
- Thráinsson, Höskuldur. 2011. Icelandic A, B, C, D ...? Or: How Long is the Icelandic Alphabet? A paper presented at the Annual Meeting of DGfS (Deutsche Gesellschaft für Sprachwissenschaft), 25th February 2011, Göttingen.
- Práinsson, Höskuldur, and Ásgrímur Angantýsson. 2015. Orðaröð í aukasetningum. In *Tilbrigði í íslenski setningagerð: II. Helstu niðurstöður. Tölfræðilegt yfirlit*, ed. Höskuldur Práinsson, Ásgrímur Angantýsson, and Einar Freyr Sigurðsson, 299–330.
- Práinsson, Höskuldur, Ásgrímur Angantýsson, and Heimir Freyr Viðarsson. 2015. Kjarnafærsla, stílfærsla, leppsetningar og frumlagseyða. In *Tilbrigði í íslenski setningagerð: II. Helstu niðurstöður. Tölfræðilegt yfirlit*, ed. Höskuldur Práinsson, Ásgrímur Angantýsson, and Einar Freyr Sigurðsson, 275–297. Reykjavík: Málvísindastofnun Háskóla Íslands.
- Práinsson, Höskuldur, Þórhallur Eypórsson, Ásta Svavarsdóttir, and Þórunn Blöndal. 2015. Fallmörkun. In *Tilbrigði í íslenski setningagerð: II. Helstu niðurstöður. Tölfræðilegt yfirlit*, ed. Höskuldur Práinsson, Ásgrímur Angantýsson, and Einar Freyr Sigurðsson, 33–76. Reykjavík: Málvísindastofnun Háskóla Íslands.
- Práinsson, Höskuldur, Sigríður Sigurjónsdóttir, Hlíf Árnadóttir, and Þórhallur Eypórsson. 2015. Um þolmynd, germynd og það. In *Tilbrigði í íslenski setningagerð: II. Helstu niðurstöður. Tölfræðilegt yfirlit*, ed. Höskuldur Práinsson, Ásgrímur Angantýsson, and Einar Freyr Sigurðsson, 77–120. Reykjavík: Málvísindastofnun Háskóla Íslands.
- Práinsson, Höskuldur, Ásta Svavarsdóttir, Eiríkur Rögnvaldsson, Jóhannes Gísli Jónsson, Sigríður Sigurjónsdóttir, and Þórunn Blöndal. 2013. Hvert stefnir í íslenski setningagerð? Um samfímalegar kannanir og málbreytingar. *Íslenskt mál og almenn málfræði* 35:47–117.
- Tímarit.is. [No year]. National and University Library of Iceland. URL <http://timarit.is>.
- Trotzke, Andreas. 2015. *Rethinking Syntactocentrism: Architectural Issues and Case Studies at the Syntax-Pragmatics Interface*. Linguistik Aktuell 225. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Trotzke, Andreas, and Jan-Wouter Zwart. 2014. The complexity of narrow syntax: Minimalism, representational economy, and simplest Merge. In *Measuring Grammatical Complexity*, ed. Frederick J. Newmeyer and Laurel B. Preston, 128–147. Oxford: Oxford University Press.
- Truckenbrodt, Hubert. 2006. On the semantic motivation of syntactic verb movement to C in German. *Theoretical Linguistics* 32.3:257–306.
- Trudgill, Peter. 1986. *Dialects in Contact*. Language in Society 10. Oxford: Basil Blackwell.
- Trudgill, Peter. 2010. *Investigations in Sociohistorical Linguistics: Stories of Colonisation and Contact*. Cambridge University Press.
- Tvica, Seid. 2017. *Agreement and Verb Movement: The Rich Agreement Hypothesis from a Typological Perspective*. LOT 446. Utrecht: LOT.
- Vandenbussche, Wim. 2007. Shared standardization factors in the history of sixteen Germanic languages. In *Standard, Variation und Sprachwandel in germanischen Sprachen. Standard, Variation and Language Change in Germanic Languages*, ed. Christian Fandrych and Reinier Salverda, Studien zur Deutschen Sprache 41, 25–36. Tübingen: Gunter Narr Verlag.
- Vangnes, Oystein Alexander. 2003. 'Identifikasjon' og morfologiens rolle i den nordiske nominalfrasen. In *Dialektsyntaktiska studier av den nordiska nominalfrasen*, ed. Anders Holmberg Oystein Alexander Vangnes and Lars-Olof Delsing, Tromsø Studies in Linguistics 22, 113–194. Oslo: Novus.
- Van de Velde, Freek. 2010. The emergence of the determiner in the Dutch NP. *Linguistics* 48.2:263–299.
- Van de Velde, Freek, and Béatrice Lamiroy. 2017. External possessors in West Germanic and Romance: Differential speed in the drift toward NP configurationality. In *Aspects of Grammaticalization: (Inter)Subjectification and Unidirectionality*, ed. Daniel Van Olmen, Hubert Cuyckens, and Lobke Ghesquière, 353–399. Berlin/Boston: De Gruyter Mouton.

- Vikner, Sten. 1994. Finite verb movement in Scandinavian embedded clauses. In *Verb movement*, ed. David Lightfoot and Norbert Hornstein, 117–147. Cambridge University Press.
- Vikner, Sten. 1995. *Verb Movement and Expletive Subjects in the Germanic Languages*. Oxford: Oxford University Press.
- Vikner, Sten. 1997. V-I movement and inflection for person in all tenses. In *The New Comparative Syntax*, ed. Liliane Haegeman, 189–213. Longman.
- Viðarsson, Heimir Freyr van der Feest. 2014. Þáttur málstöðlunar í afstöðu sagnar til neitunar í 19. aldar íslensku: „málgreinir, sem mjer fannst eitthvert danskt óbragð að“. *Orð og tunga* 16:1–24.
- Viðarsson, Heimir van der Feest. 2016. The Syntax of Others: ‘Un-Icelandic’ Verb Placement in 19th- and Early 20th-Century Icelandic. In *Prescription and Tradition in Language: Establishing Standards across Time and Space*, ed. Ingrid Tieken-Boon van Ostade and Carol Percy, Multilingual Matters 165, 152–167. Bristol/Blue Ridge Summit: Multilingual Matters.
- Viðarsson, Heimir van der Feest. 2017a. Grimm’s “floating” datives: Applicatives and NP/DP configurationality in Icelandic from a diachronic perspective. *Working Papers in Scandinavian Syntax* 99:30–53.
- Viðarsson, Heimir van der Feest. 2017b. Málnotkun sem mælikvarði á áhrif málstöðlunar: Skólaritgerðir úr Lærða skólanum í Reykjavík (1846-1904). *Orð og tunga* 19:129–153.
- Wackernagel, Jacob. 1892. Über ein Gesetz der indogermanischen Wortstellung. *Indogermanische Forschungen* 1.1:333–436.
- Walkden, George. 2014. *Syntactic Reconstruction and Proto-Germanic*. Oxford Studies in Diachronic and Historical Linguistics. Oxford: Oxford University Press.
- Walker, James A., and Miriam Meyerhoff. 2013. Studies of the community and the individual. In *The Oxford Handbook of Sociolinguistics*, ed. Roberts Bayley, Richard Cameron, and Ceil Lucas, 175–194. Oxford/New York: Oxford University Press.
- Wallenberg, Joel C. 2016. Extraposition is disappearing. *Language* 92.4:e237–e256.
- Wallenberg, Joel C., Anton Karl Ingason, Einar Freyr Sigurðsson, and Eiríkur Rögnvaldsson. 2011. *Icelandic Parsed Historical Corpus*. URL http://www.linguist.is/icelandic_treebank.
- Warner, Anthony R. 1993. *English Auxiliaries: Structure and History*. Cambridge Studies in Linguistics 66. Cambridge: Cambridge University Press.
- Weiner, E. Judith, and William Labov. 1983. Constraints on the agentless passive. *Journal of Linguistics* 19.1:29–58.
- Weinreich, Uriel, William Labov, and Marvin I. Herzog. 1968. Empirical foundations for a theory of language change. In *Directions for Historical Linguistics: A Symposium*, ed. W. P. Lehmann and Yakov Malkiel, 95–198. Austin: University of Texas Press.
- Wiklund, Anna-Lena, Kristine Benzten, and Gunnar Hrafn Hrafnbjargarson. 2009. On the distribution and illocution of V2 in Scandinavian that-clauses. *Lingua* 119.12:1914–1938.
- Wiklund, Anna-Lena, Gunnar Hrafn Hrafnbjargarson, Kristine Benzten, and Þorbjörg Hróarsdóttir. 2007. Rethinking Scandinavian verb movement. *The Journal of Comparative Germanic Linguistics* 10.3:203–233.
- Winford, Donald. 1984. The linguistic variable and syntactic variation in creole continua. *Lingua* 62.4:267–288.
- Worham, Stanton. 2008. Linguistic anthropology of education. *Annual Review of Anthropology* 37.1:37–51.
- Wright, Sue. 2004. *Language Policy and Language Planning: From Nationalism to Globalisation*. Basingstoke: Palgrave Macmillan.
- Yang, Charles D. 2000. Internal and external forces in language change. *Language Variation and Change* 12.3:231–250.
- Yang, Charles D. 2010. Three factors in language variation. *Lingua* 120.5:1160–1177.
- Zeijlstra, Hedde. 2012. There is only one way to agree. *The Linguistic Review* 29.3:491–539.
- Zwart, Jan-Wouter. 1997. *Morphosyntax of Verb Movement: A Minimalist Approach to the Syntax of Dutch*. Dordrecht: Kluwer.
- Zwart, Jan-Wouter. 2008. Verb-second als speelbal van de grammaticatheorie. *Nederlandse Taalkunde* 13:198–216.
- Zwart, Jan-Wouter. 2009. Prospects for top-down derivation. *Catalan Journal of Linguistics* 8:161–187.
- Zwart, Jan-Wouter. 2015. Top-down derivation, recursion, and the model of grammar. In *Syntactic Complexity across Interfaces*, ed. Andreas Trotzke and Josef Bayer, Interface Explorations 30, 25–42. Berlin/Boston/Munich: De Gruyter Mouton.

p → *TH*

Manuscripts

Private letter corpus

Íslensk sendibréf frá 19.öld (Bernharðsson and Jónsson 2012)

(<http://brefasafn.arnastofnun.is>)

National and University Library of Iceland

ÍB. 100 fol. a, ÍB. 101 fol. a, ÍB. 102 fol. b, ÍB. 93 fol. a, ÍB. 93 fol. b, ÍB. 94 fol. a, ÍB. 94 fol. b, ÍB. 95 fol. a, ÍB. 96 fol. a, ÍB. 96 fol. b, ÍB. 97 fol. a, ÍB. 97 fol. b, ÍB. 98 fol. a, ÍB. 99 fol. a, ÍB. 99 fol. b, ÍB. 100 fol. b, ÍB. 101 fol. b, ÍB. 102 fol. a, Lbs. 2414 a 4to, Lbs. 2415 a 4to, Lbs. 2619 4to, Lbs. 2748 4to, Lbs. 2844 4to, Lbs. 3511 4to, Lbs. 3514 4to, Lbs. 3515 4to, Lbs. 3520 4to, Lbs. 3522 4to, Lbs. 3523 4to, Lbs. 3524 4to, Lbs. 3526 4to, Lbs. 3527 4to, Lbs. 4728 4to, Lbs. 2409 b 4to, Lbs. 2412 a 4to, Lbs. 2413 a 4to, Lbs. 2413 a 4to, Lbs. 2415 b 4to, Lbs. 2755 4to, Lbs. 3029 4to, Lbs. 3078 4to, Lbs. 3081 4to, Lbs. 3092 4to, Lbs. 3093 4to, Lbs. 3097 4to, Lbs. 3109 4to, Lbs. 3175 4to, Lbs. 3910 4to, Lbs. 4415 4to, Lbs. 4416 4to, Lbs. 4417 4to, Lbs. 4419 4to, Lbs. 4533 4to, Lbs. 4941 4to, Lbs. 5020 4to., Aðföng 11.12.2000.

Héraðsskjalasafn Þingeyinga

E-728-5

Other material

National and University Library of Iceland

Lbs. 1238, 8vo. *Stuttur Leidarvijsir fyri Íslenska í þerra eigin Módurmaali*. [A short introduction for Icelanders in their native tongue, presumably written c. 1800-1805 by Guttormur Pálsson; preserved copy written c. 1810-1815 by Sigfús Árnason.]

National Archives of Iceland

Skjalasafn Menntaskólans í Reykjavík: M.R. skrifleg próf 1847-1848, 1852, 1860-1861, 1870, 1875, 1882, 1890. [See also: M.R. skrifleg próf 1847-1906, in: *Landsins útvöldu synir*, ed. Ólafsson 2004.]

Further information

Repositories: osf.io/bdsv6 and github.com/heimirfreyr

RLSS: Corpus of Reykjavík Grammar School Essays (1846-1904)

ICENCC: Icelandic Corpus of Early Nineteenth-Century Correspondence

Data sets, source files, R scripts, letter inventories, bibliography

Appendices

Appendix A: Verb-adverb placement

	Until 1850				After 1850			
	Private letters		Newspapers		Private letters		Newspapers	
Decl	29%	(324/1106)	30%	(168/560)	9%	(142/1598)	6%	(107/1682)
Result	8%	(6/75)	3%	(1/29)	1%	(1/221)	2%	(2/97)
ConsDeg	36%	(35/97)	23%	(21/93)	9%	(19/212)	6%	(16/265)
Cause	14%	(55/397)	25%	(47/192)	4%	(23/623)	7%	(26/380)
Adv	40%	(136/341)	50%	(147/294)	17%	(99/587)	16%	(86/535)
IndQu	27%	(19/71)	44%	(14/32)	15%	(26/170)	12%	(8/69)
Rel	60%	(185/309)	65%	(95/147)	20%	(51/258)	31%	(80/259)
Cond	43%	(68/157)	45%	(54/120)	22%	(68/308)	20%	(59/289)
∑ 1-4	25%	(420/1675)	27%	(237/874)	7%	(185/2654)	6%	(151/2424)
∑ 5-8	47%	(408/878)	52%	(310/593)	18%	(244/1323)	20%	(233/1153)

Table 6.29. Relative frequency of Adv-Vfin in private letters and newspapers up until 1850 vs. after 1850.

	Until 1850				After 1850			
	Private letters		Newspapers		Private letters		Newspapers	
Quantified	–	(1)	10.0%	(30)	–	(2)	3.5%	(86)
Indefinite	32.2%	(149)	36.6%	(183)	8.1%	(136)	8.5%	(483)
Definite	26.7%	(307)	38.3%	(386)	6.6%	(380)	7.3%	(987)
Proper names	22.6%	(124)	33.3%	(84)	7.1%	(184)	7.5%	(212)
Pronouns	34.1%	(1982)	38.4%	(784)	11.9%	(3295)	13.9%	(1809)

Table 6.30. Relative frequency of Adv-Vfin in private letters and newspapers up until 1850 vs. after 1850 across different types of subjects.

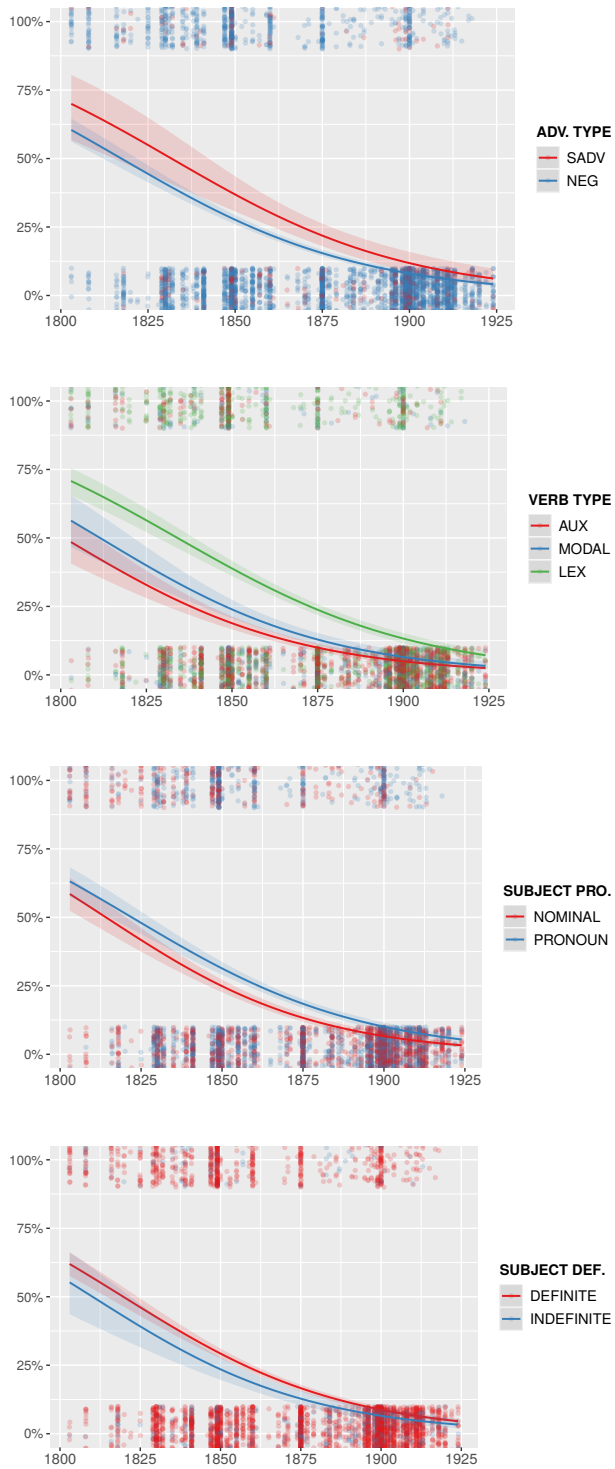


Figure 6.36. Newspapers: Predicted probabilities of fixed effects over time, conditioned on random effects.

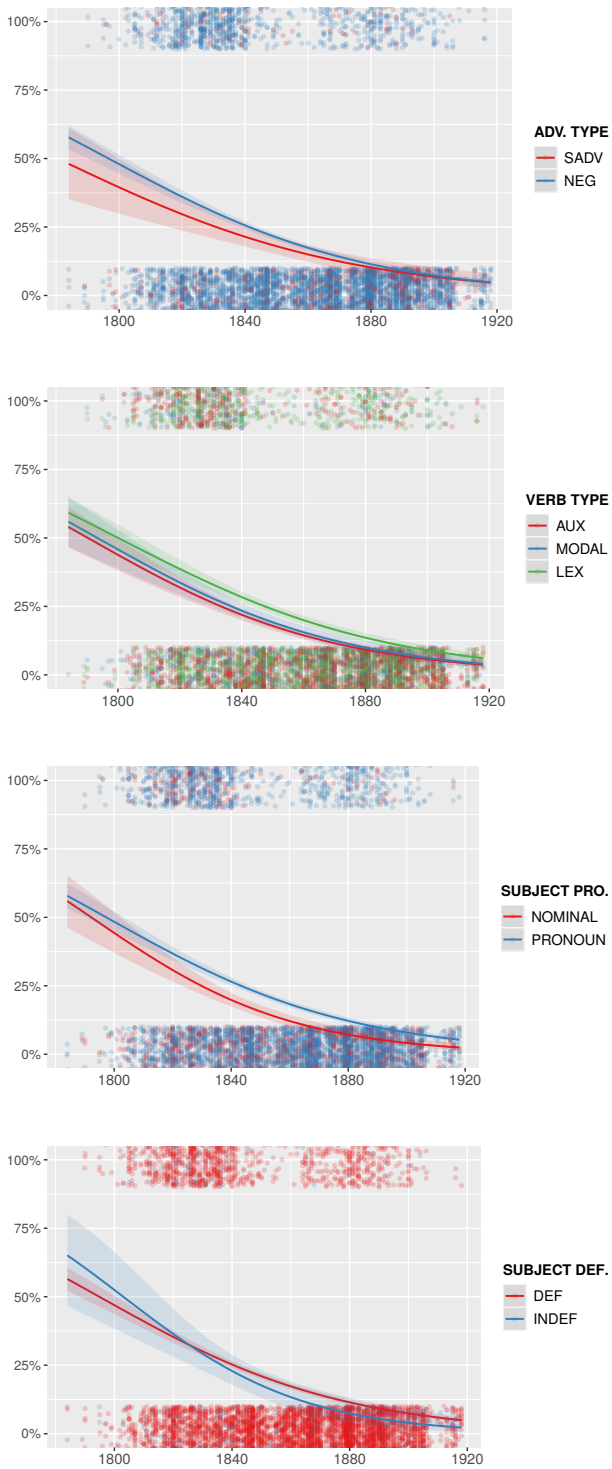


Figure 6.37. Letters: Predicted probabilities of language-internal fixed effects over time, conditioned on random effects.

	A+B+E		C+D		O	
GeiVid	40.7%	(24/59)	68.2%	(15/22)	44.8%	(13/29)
BjaTho	47.2%	(77/163)	70.1%	(47/67)	61.9%	(60/97)
SteSig	61.1%	(22/36)	100.0%	(2/2)	72.2%	(13/18)
IngJon	10.4%	(10/96)	52.4%	(11/21)	37.0%	(10/27)
GdrJon	14.1%	(13/92)	25.0%	(8/32)	19.7%	(12/61)
FinJon	16.0%	(4/25)	0.0%	(0/3)	0.0%	(0/12)
BenHal	1.6%	(1/63)	16.7%	(1/6)	3.8%	(1/26)
ArnHelg	3.0%	(4/135)	4.2%	(1/24)	14.5%	(8/55)
SofDan	1.6%	(1/63)	20.0%	(2/10)	6.2%	(2/32)
GunOdd	7.5%	(3/40)	0.0%	(0/3)	0.0%	(0/11)
SigPal	2.9%	(3/104)	0.0%	(0/22)	9.4%	(5/53)
JakJon	1.8%	(1/55)	0.0%	(0/13)	0.0%	(0/18)

Table 6.31. Percentage Adv-Vfin for letter-writers with at least 100 examples of the linguistic variable in different types of *that*-clauses. Arranged in order of appearance in Table 3.9.

(182) Cinque's Hierarchy by example: Adv-Vfin

- a. Mod_{epistemic} (*probably*)
því það er mál sem ykkur **líklega** langar að vita um greinilega.
(GunOdd-1890-04-01.xml)
- b. T_{past} (*the winter before last*)
Þó eg í **fyrirvetur** vissi ekki, hvað það hafði að þýða, sem fyrir mig bar laugardaginn fyrir páska, liggur það opið fyrir augum mínum nú (GudMag-1844-07-23.txt)
- c. T_{future} (*then*)
hvirtsem er kannski þjer **þá** væruð svo góður við okkur hjerna að sitja fyrir honum (SteSig-1867-03-05.xml)
- d. Mood_{irrealis} (*maybe*)
en þá Nafni minn sá ad eg **kannske mundi** ná Pluralitet féck hann allt fyrst um Sinn paralyserad [...]
(BjaTho-1839-09-07.txt)
- e. Mod_{necessity} (*necessarily*)
Þar á móti fæ eg ekki hér það, sem eg þó **nauðsynlega** þyrfti, ef til kæmi, til- dæmis hatt, frakka og fleira.
(IngJon-1810-10-18.txt)
- f. Mod_{possibility} (*possibly*)
Ég held líka mér sé það skást að hafa svo mikið sem eg **mögulega** get um að sjá (DomBri-1851-07-10.txt)
- g. Asp_{repetitive(I)} (*anew*)
í þessum raunum og mótlæti sem góðum Guði nú **á ný** hefur þóknast að láta mjer að höndum bera
(SteSal-1885-06-04.xml)
- h. Asp_{frequentative(I)} (*often*)
tala jeg nú ekki um, þó mig **opt** hafi lángað til að láta yður eitthvað fá af þessu tægi
(SteSig-1877-01-30.xml)
- i. Mod_{volitional} (*intentionally*)
ef hann vissi med siálfum sér ad hann ei **viljandi** hefði af þeim haft.
(BjaTho-1830-12-05.txt)
- j. Asp_{celerative(I)} (*soon*)
Það sem og fyrst er hrædd um er, að eg alltof **fljótt** megi sjá á bak þér til fulls
(IngJon-1809-01-01.txt)

- k. T_{anterior} (*already*)
það sem, jeg **nú þegar** hef sagt þjer (EirJoh-1892-03-14.xml)
- l. Asp_{terminative} (*no longer*)
þared eg **ecki lengur** er stórskuldugr í Kaupstad (BjaTho-1827-11-18.txt)
- m. Asp_{continuative} (*still*)
ef hún **enþá** skrifast á við B. <unclear>Schon.</unclear> (JohHal-1878-02-04.xml)
- n. Asp_{perfect} (*always*)
svo jeg kæmist úr þeirri skuld sem mjer **altaf** finst jeg ver í við yður siðan í firra (SteSig-1869-10-06.xml)
- o. Asp_{retrospective} (*just*)
Það er líka miklu auðveldara að sjá hið rétta en hið gagnlega, nema menn **aðeins** líti á hið næsta (ArnHel-1816-04-30.txt)
- p. Asp_{proximative} (*soon*)
eg hef nú að sönu í san frétt að han **bráðum** komi (SigPal-1842-06-02.xml)
- q. Asp_{prospective} (*nearly*)
Þar margir héðan eru nú farnir að heiðra vors konungs burðardag, svo eg **nærri því** sit ein heima, gríp eg þetta tækifæri að svara bréfi þínu með póstskipinu í haust. (IngJon-1822-01-28.txt)
- r. Asp_{SgCompletive(I)} (*completely*)
[...] með bréfinu [...], sem eg ennú iðrast og skammast mín fyrir, þareð það svo **algjörlega** stríddi á móti þeirri elsku og virðingu, sem hvert barn er föður um skyldugt (TorEgg-1835-04-20.txt)
- s. Voice (*well*)
Sú þín tilgáta er rétt, að það sem eg **vel** gat án verið lét eg dætur mínar fá (IngJon-1840-08-15.txt)
- t. Asp_{repetitive(II)} (*again*)
hugsa það bætist úr öllum raunum mínum í vor, þá eg **aftur** sé minn elskulegasta, vænsta, lærðasta kærasta (AlfJon-1820-01-01.txt)

(183) Potential Adv-Vfin during the Old Norse period (13th-14th century)

- a. *Kom hann eitt sinn á minn fund og sagði eg honum að eg ekki vildi hér í landi vistir hans af þeim sökum sem yður er áður kunnigt.*
land supplies his of the reasons which you are before known
'... I told him that I did not want his supplies here in this country for reasons already known to you.'
(MÍM: Egils saga)
- b. *[...] og margt annað fleira þótt þess eigi sé hér getið.*
and many other more although it not is here mentioned
'... and additionally many other things although it is not mentioned here.'
(MÍM: Sturlunga saga)
- c. *Snorri goði og Skafti löttu þess að leggja á þá hættu við Noregsmenn*
Snorri chieftain and Skafti discouraged it to lay on the danger with Norway.men
að allir senn færu af Íslandi og þangað þeir menn er mest réðu fyrir landi.
that all soon left from Iceland and there the men who most ruled for land
'Chieftain Snorri and Skafti discouraged putting the people of Norway in danger by having all soon leave Iceland and to there (N.) those men ruling most in that country.'
(MÍM: Heimskringla)

- d. ... *kjósa hvort eg vil hafa jarldóm fyrir því þá mun eg sjá hvort eg þá þykist*
 choose whether I want have earldom for it then will I see whether I then think
mega halda ríkið eða eigi.
 may hold state_{DEF} or not
 ‘... to choose whether I want to have earldom because then I will see whether I think
 to be able to hold the state or not.’ (MÍM: Svarfdæla saga)
- e. *En sem Oddi hafði úti verið slíka stund sem honum vel líkaði fór hann inn*
 but as Oddi had outside been such time which him well pleased went he inn
í rekkju sína og sofnaði þegar [...]
 in bed his and fell.asleep immediately
 ‘But since Oddi had been outside for as long as he pleased, he went inside to his bed
 and immediately fell asleep.’ (MÍM: Stjörnu-Odda draumur)
- f. *Tók hann þá svo mikla peninga sem honum vel líkaði og skildu þeir Hólmkell*
 took he then so much money as him well pleased and parted they Hólmkell
og Þorgrímur með mikilli vináttu.
 and Þorgrímur with great friendship
 ‘Then he took as much money as he pleased and Hólmkell and Þorgrímur parted with
 great friendship.’ (MÍM: Víglundar saga)
- g. ... *og skyldu sinn man hvorir til taka og kveða að hvorir betur hefðu og skulu*
 and should his_{REFL} man each to take and say PRT who better had and shall
þeirra atkvæði standa er til voru kosnir.
 their votes stand who to were chosen
 ‘And both parties were to choose someone to decide who had won and the verdict of
 those elected was to definite.’ (MÍM: Víga-Glúms saga)
- h. *Vil eg selja þér sjálfðæmi og ráð einn sætt þann veg sem þér best þykir.*
 want I sell you autonomy and decide alone settlement the way that you best think
 ‘I want to give you autonomy and you may alone decide on the settlement in the way
 you find most fitting.’ (MÍM: Finnboga saga)
- i. ... *sterkur að afli sem þeir menn margir sem fullkomnir voru að aldri og atgervi*
 strong of power as the men many which complete were of age and ability
nálega eftir því sem þeir best voru á sig komnir fyrir allra hluta sakir.
 nearly after it which they best were on REFL come for all things sake
 ‘... having the strength of many of the men who had attained the proper age and ability,
 nearly as those who were most fit in all regards.’ (MÍM: Stjörnu-Odda draumur)
- j. *En þér Arinbjörn er það að segja að þú svo megir vera hér í landi að ...*
 but you Arinbjörn is it to tell that you so may be here in land that
 ‘You, Arinbjörn, be told that you (so) may be here in this country so long you ...’
 (MÍM: Egils saga)
- k. *en tökum upp nýtt vinfengi og ef þér svo sýnist það ráð að eg taki við*
 but take_{IMP} up new friendship and if you so seem the advice that I take with
goðorði þínu þá ...
 chieftainship your then
 ‘But let us reinstate our friendship and if you deem it advisable that I takeover your
 chieftainship, then ...’ (MÍM: Þorsteins saga Síðu-Hallssonar)

Appendix B: Definite marker

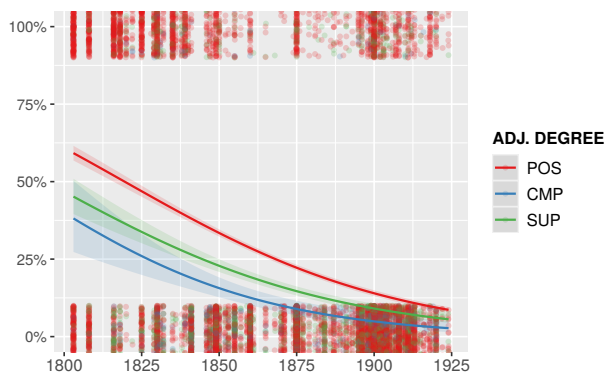


Figure 6.38. Newspapers: Predicted probabilities of fixed effect adjective category over time, conditioned on random effects (higher values = greater use of non-standard *sá* norm).

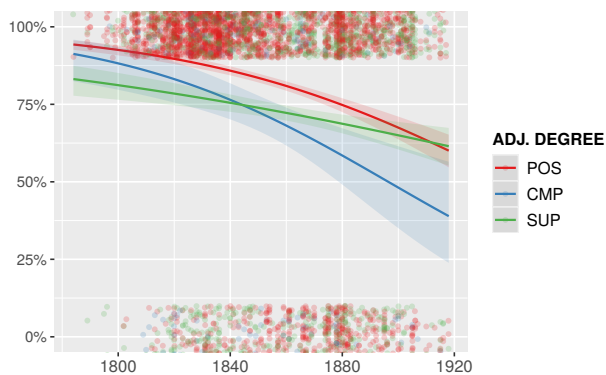


Figure 6.39. Letters: Predicted probabilities of fixed effect adjective degree over time, conditioned on random effects (higher values = greater use of non-standard *sá* norm).

Appendix C: Generic pronoun

