Abstract

Language Forecasting:
With Focus on Variation and Change in Icelandic

Sigríður Sæunn Sigurðardóttir
2024

Language forecasting, i.e., predicting the future state of a language, has long been regarded with a fair amount of skepticism. This is partly due to language change often being considered sudden, random, unpredictable, and viewed as the result of complex interacting factors that are not well understood (e.g., Keller 1994:72; Bauer 1994:25; Labov 1994:10; Croft 2000:3; discussion in Sanchez-Stockhammer 2015). Some have gone as far as to claim that "[d]iachronic linguistics is not a predictive science" (Bauer 1994:25). Nevertheless, more positive views on the possibility of language forecasting have emerged in recent years (Sóskuthy 2015; Sanchez-Stockhammer 2015; Van de Velde 2017).

In this dissertation I present arguments in favor of language forecasting, claiming that it can and should be practiced. I argue that forecasting can be used to test various expectations toward language change, including the understanding of the propagation of new linguistic variants through a language community. Using historical data in the form of regular time series, I produce short- to mid-range forecasts for two selected changes in Icelandic. The first change concerns the complex prepositions *á bak við* 'behind' and *við hliðina á* 'next to' which are occasionally encountered in a simplified form as *bakvið* and *hliðiná*, respectively. Although the change has been briefly mentioned before (Friðjónsson

2004, 2007 and Rögnvaldsson 2021), it has not been systematically documented. The second change involves subject case marking with the predicate *hlakka til* ‘look forward to’, where an oblique case (accusative or dative) replaces an earlier nominative case with subjects. This change has been extensively studied (e.g., Svavarsdóttir 1982; Jónsson & Eythórsson 2003; Nowenstein 2023), but the present work offers an original conception of its essence. The time series analysis and forecasting presented in the dissertation provide a novel type of documentation and a fresh insight into both types of changes. Since language forecasting is argued to require ample context to be comprehensible, efforts have been made to contextualize the changes under discussion to the extent possible.

Language Forecasting:
With Focus on Variation and Change in Icelandic

A Dissertation
Presented to the Faculty of the Graduate School
of
Yale University
In Candidacy for the Degree of
Doctor of Philosophy

by
Sigríður Sæunn Sigurðardóttir

Dissertation Director: Dr. Claire Bowern

May 2024

Contents

List of figures

List of tables

Abbreviations

1P	first person
2P	second person
3P	third person
SG	singular
PL	plural
NOM	nominative
ACC	accusative
DAT	dative
GEN	genitive
AMB	ambiguous case
NOPE	no P_1
P1	has P_1
ÍT	Íslenskt textasafn
ONP	Ordbog over det norrøne prosasprog
IGC	The Icelandic Gigaword Corpus

Acknowledgements

Nothing comes from nothing, and such is the case with this dissertation. I would first and foremost like to thank my parents, Margrét Vigdís Eiríksdóttir and Sigurður Örn Karlsson, for supporting me throughout all my studies. A long time ago, they promised to provide me with free room and board (at home) for as long as I continued being a student. With the submission of this dissertation, I gladly relinquish my student privileges – although I hope there is a grace period.

Many thanks to my advisor Claire Bowern and members of my dissertation committee, Jim Wood and Freek Van de Velde. Thank you for all the comments and the feedback. I am extremely grateful for having had the opportunity to work on this project. Allowing someone to study and write about things that have not happened yet is a risky business. Also, to Claire, thank you for your patience.

During my time in academia, I have met both wonderful people and not-so-nice people. In the former category, I am proud to count the following close friends and fellow students I originally met at The University of Iceland: Brynhildur Stefánsdóttir, Guðrún Steinþórsdóttir, Kolfinna Jónatansdóttir, Gunnhildur Jónatansdóttir, Katrín María Víðisdóttir, Kristján Sigurðarson, Heimir Freyr Viðarsson, Sigrún Gunnarsdóttir, Iris Edda Nowenstein, Ingunn Hreinberg Indriðadóttir, María Anna Garðarsdóttir and Mads Holm. Thank you for all your support and all the good times. I hope they will continue. To friends I met at Ghent University, I send my greetings and a big *dank u* to: Cindy Johnson, Jonah Rys, Inge Geerardyn, Michael Frotcher, Metin Bağriaçik, Elisabeth Witzenhausen and

Annemie Knockaert. I would also like to thank Haraldur Bernharðsson, Svanhildur Gunnarsdóttir, Jan Terje Faarlund, Anne Breitbarth, Liliane Haegeman, Nicole Dehé, Ellen Brandner and Alexander Pfaff for their support and help in various ways.

I am known to have a soft spot for medieval studies. I will be eternally grateful to Roberta Frank who taught me so many things about *Beowulf* and who introduced me to the medievalist community at Yale. To my friends who appreciate manuscripts, medieval literature and ancient languages, Kyle Conrau-Lewis, Justin Hudak, Mireille Pardon, Clara Wild, Evelyne Koubková and Jon Beltz, best of thanks: *nu-kán A-NA* DINGIR$^{\text{LIM}}$ KAŠ UDU *iš-pa-an-ta-aḫ-ḫi*.

To Yale linguistic graduate students, past and present, Rikker Dockum, Matthew Tyler, Luke Lindemann, Samuel Andersson, Jeremy Johns, Dani Katenkamp, Squid Tamar-Mattis, Comfort Ahenkorah, Catarina Soares, Ka Fai Yip, Michael Stern and Finn Amber, thank you for great conversations and for making the gradroom a nice place to be in. Also big thank you to Irene Sørensen (Ircha Gaming), DiskyDisk and Maeve for providing informative entertainment and the perfect background noise for working.

In the past few years, I have had the opportunity to present parts of my work. I would like to thank Nikolaos Lavidas and the participants at the Naxos summer school on diachronic linguistics (2021–2023) for taking the chance on language forecasting. I also thank the audience at ICHL in Heidelberg 2023 for helpful discussion and feedback on the diachrony of complex prepositions, in particular Frans Plank, Nigel Vincent and Mark Darling.

Special thanks go to Elly van Gelderen, Thor Aspelund and Þórhallur Eyþórsson who have helped me in so many ways and have been the source of engaging and informative conversations. I hope these will continue in the future.

Finally, for making the whole journey more bearable, Randi Martinez, Sarah Babinski & Marisha Evans, I cannot thank you enough for all the help and support you have provided me with – see you in the *Old ladies club*.

Ein sat hon úti,
þa er inn aldni kom
yggjungr ása
ok í augu leit.
Hvers fregnið mik?
hví freistið mín?

(from *Völuspá* 28.)

Alone she held séance out in the night,
when the old fellow came,
Æsir's Son of Dread,
And looked into her eyes.
'What do you ask me?
Why do you try me?'

(transl. Dronke 1997)

1 Introduction

1.1 The aim and scope of the dissertation

Like the weather, languages are constantly changing. Sometimes the changes are subtle and take place without us noticing them; at other times we are all too aware of new structures and trends. No matter how small they may be, or which domain of language they belong to, once in motion, they may trigger new changes. Thus, language is an ever-transforming system where the smallest ripple may cause a large wave. It has been claimed that making predictions about the future of language or a particular linguistic phenomenon is impossible due to myriad uncertainties and the complexity of the linguistic system. Despite this, some recent work has adopted a more optimistic view (Sóskuthy, 2015; Sanchez-Stockhammer, 2015; Van de Velde, 2017; Schneider, 2018). After all, uncertainties in fields like meteorology, economics, geology and epidemiology, have not prevented forecasting in these areas.

This dissertation is concerned with language change, or, more specifically, with whether language change, variation and the propagation of change can be predicted. It seeks to answer the question whether language forecasts are possible, and if so, how they can be made. As such, this dissertation is intended to serve three purposes:

i) Present arguments in favor of language forecasting, claiming that it *can* and *should* be practiced as it offers a novel way of studying language change.

ii) Explain how language forecasting might be approached, i.e., what kind of predictions can be made and which methods might be adapted from other areas of forecasting.

iii) Provide examples of language forecasting based on linguistic variation and changes in Modern Icelandic.

Focusing on variation and change within Modern Icelandic defines additional sub-goals of the dissertation. These include documentation of two types of change within Icelandic and predictions about the propagation of these in the next 10–20 years. The first of these pertains to changes in the complex prepositions *á bak við* ‘behind’ and *við hliðina á* ‘next to’ which often appear in a simplified form, i.e., *bakvið* and *hliðiná*. The second change is oblique case substitution with the predicate *hlakka til* ‘look forward to’. The predicate originally occurred with a nominative subject, but is now frequently attested with an accusative or dative subject. The sub-goals (or sub-contributions) can be summarized as follows:

iv) Provide a general overview of hitherto little discussed changes in the complex prepositions *á bak við* ‘behind’ and *við hliðina á* ‘next to’ in Icelandic; Present a documentation of these in form of regular time series and generate expectation towards the propagation of novel variants in the next few years.

v) Present novel type of documentation of changes in subject case marking with *hlakka til* ‘look forward to’ in the form of regular time series, and to generate expectation towards the situation of the change in the next few years.

New expectations towards the propagation of the relevant changes in the next few years are generated in the form of language forecasts that are made with formal forecasting models, relying on regular time series. Results from time series documentation and forecasting is contrasted with general expectations towards directionality in these changes based on what is known about grammaticalization processes and changes in subject case marking in Icelandic.

Unlike predictions about weather in the future or the potential number of tourists expected to visit a country in a given year, language forecasts do not serve an obvious practical purpose. Typically, forecasts “provide input to the planning and decision-making process from which the forecasting requirement arose” (Hoff 1983:35). While this may apply to language planning or prediction and prevention of language death, general language forecasting does not guide decision making or assist preparing for an upcoming language situation. Instead, they may be made for the sake of making them or for the sake of obtaining a new perspective on language change and language forecasting. In this sense, the overall claim of the dissertation is that language forecasting leads to better understanding of language change and language data across time (cf. Benediktsson 2002 for new methods leading to advancement in historical linguistics). The general philosophy adopted is that predictions make little sense out of context. Therefore, as much information as possible and appropriate is provided for every step of the forecasting, from defining the forecasting to gathering data and interpreting predictions. The hope is that the present work will lead to some advancement in the area of language forecasting and prompt more research into the topic.

1.2 Background and important concepts

Predictions are an integral part of linguistics. They may include expectations toward how language is structured, e.g., which utterances are predicted to be grammatical or ungrammatical, which rules are productive, and how languages can be systematically broken down into smaller units like morphemes, sounds and features. They can also be in the form of various types of (implicational) language universal which may be either synchronic and diachronic (Greenberg 1966:508; Greenberg, Osgood, Jenkins 1966:xxiii). An example of a synchronic implicational universal is “If a language has gender categories in the noun, it has gender categories in the pronoun” (Greenberg 1966:113), and an example of a diachronic universal is “A nasal syllabic phoneme, apart from borrowings and analogical formations, always results from loss of a vowel” (Ferguson 1966:59).

Despite much work within linguistics being predictive in nature, predictions tend not to be time dependent. They make assertions about what should be (im)possible in all languages or what should be the case for a particular language. As such they do not claim that given that if structure X is attested in a language at a particular period, structure Y will be attested at a given period in the future, nor do they predict how many individuals will find a particular linguistic variant grammatical in the future given how many find it grammatical at a current time. This is where language forecasting comes into play.

Most forecasts have a very clear practical purpose which involves preparing in one way or another for a future scenario. For instance, weather forecasts help make decisions on how to best dress for certain days (does one need rubber boots or sandals), how and where to plan holidays, or even where to build (or not build) a new house. Epidemiological forecasts can help workers in the healthcare system plan for the number of patients they may expect at a given time and evaluate how to best stop the spread of diseases. As such, these forecasts have the potential to influence the behavior of individuals.

At first sight, language forecasts do not seem to have an immediate practical purpose. Aside from their potential role in language planning and preventing language death, they do not necessarily provide information that helps to plan for the future, nor do they have the power to alter the behavior of individuals. Their role is thus different from other types of forecasts. In Chapter 2 of this dissertation, it is argued that language forecasting provides a novel way to study language change. It is claimed that it gives a novel perspective on various questions related to language over time, including future developments, the propagation of change, and what kind of data is most suitable for language forecasting. Naturally, language forecasting can only be systematically practiced provided the forecasting task, i.e., each step in the whole procedure, is identified and discussed. Doing so involves outlining the forecasting problem, gathering appropriate data, conducting an explanatory analysis, choosing, and fitting forecasting models and evaluating the outcome (adapted from Hyndman & Athanasopoulos 2021:22–23).

Defining the forecasting problem might be one of the hardest tasks as it involves defining the goal is, i.e., identifying which exact questions are to be answered. It also requires an understanding of how the forecast is to be used, who it is for, what the requirements for such a forecast is, and how far into the future predictions should be made (i.e., what the forecast horizon is). A second challenging part is identifying and gathering appropriate data for. Since language forecasting is relatively new, there is little understanding on exactly what kind of data is most appropriate for making good forecasts. The case studies in Chapters 8 and 9 rely on time series data (see e.g., Box, Jenkins & Reinsel 2008:1). A distinction is made between an *example*, i.e., a particular instance of a linguistic variant, and an *observation* which contains information on the proportion of innovative variants versus all relevant variants at a given time. For example, an observation for a single year might claim that an innovative variant was attested in 40% of the cases. Behind this observation could be a total of 2000 examples, of which 800 were instances of the innovative variant and 1200 instances of a traditional variant. Exactly how measurements like these are obtained might matter for the forecast (see Chapter 5) and figuring out what kind of measurements and data are most useful will likely only be determined through trials and error. In Figure 1.1 the first two steps are highlighted since these need to be identified specifically for linguistics. The remaining steps, methods and forecast-making, can be adopted (at least initially) from other sciences.

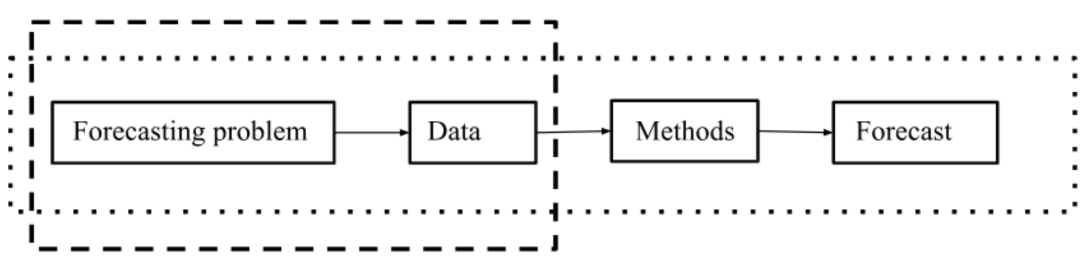


Figure 1.1. Language forecasting involves defining the forecasting problem, gathering data, selecting a method, and producing a forecast. All steps of language forecasting are currently understudied. Part I of the dissertation focuses on the first two steps, namely defining the problem and discussing appropriate data.

The forecasting in this dissertation relies on the use of popular methods that require time series data as input. These methods are designed to pick out emerging patterns in historical data over time and extrapolate them into the future. Models are chosen based on goodness of fit and how well they perform on a test data that has been set aside solely for the purpose of assessing the performance of the models (see further Section 1.3 and Chapter 7). Needless to say, the forecasts themselves cannot be properly evaluated until the time for which they were made is reached.

Predicting the situation of a language given a particular time in the future necessarily requires treating the time in a serious manner. For this reason, it is important to be explicit about the length of forecast horizons and what the situation is hypothesized to be at a certain point in time. Usually, forecasts are divided into short-, medium-, and long-range forecasting, depending on how far into the future predictions are made and how (un)certain predictions are likely to be. Additionally, different methods may be used for generating different forecasts of different lengths. What counts as short-, medium-, and long-range may depend on the phenomena of study. For weather forecasts, short-range tends to cover about 6 hours to a few days, medium-range from about 3 to 8.5 days and long-range anything above 8.5 days (Ahrens 2007:347). Naturally, one can assume the situation for language forecasting is slightly different, namely that short-, medium-, and long-range would be tied to years rather than days. The ranges proposed are as follows: short-range language forecasts might cover a period from one to up to 15 years; medium-range might go from around 15–30 years, and long-range forecasting might be regarded as dealing with everything above 30 years. These ranges are provisional and will likely need to be revised when more experience has been gained in language forecasting.

The general philosophy adopted here is that forecasting makes little sense out of context. For these reasons an effort has been made to contextualize the changes under discussion as much as possible. This involves providing a thorough background on the changes under investigation and any expectations towards how they may unfold. As already noted, the two case studies (Chapter 8 and 9) pertain to linguistic phenomena in Icelandic, a language spoken in Iceland by less than 400,000 people.[1] While Modern Icelandic is generally considered the stretch from 1540 to present times (1540 being the year in which the first book in Icelandic was printed), the case studies are mostly restricted to variation and changes within the last 20 years or so.

1.3 Data and methods

Forecasting methods employed in this dissertation rely on the use of historical data in the form of regular time series, i.e., observations ordered in time and taken at regular intervals (Makridakis & Wheelwright 1978:14–16; Box, Jenkins & Reinsel 2008:1). The methods were chosen based on popular forecasting approaches in other fields, e.g., economics, that assume that future observations are dependent on past observations. Some of the models used are very simple, such as hypothesizing that the future will be equal to the mean of the whole historical series, while others are more complex and are designed to pick out patterns that emerge in data over time. While there is no one single method that fits all time series, models from the same family of approaches can be chosen, for instance, models based on exponential smoothing or autoregression (see Chapter 7 for further information).

[1] In January 2023 the total population of Iceland was estimated to be 387,758 (https://hagstofa.is/). This figure includes both native Icelanders with Icelandic as a first language and immigrants who may or may not speak Icelandic.

In order to keep measurement of change consistent over an extended period of time, it was deemed necessary to use a data source that had the same type of material over an extended period of time. This allowed for the creation of regular time series with either quarterly or yearly frequency. An abundance of written data for Modern Icelandic exists both online and in printed format. In theory, it would have been possible to scrape the internet for data, e.g., by using Sketch engine (https://www.sketchengine.eu/) to construct a database from data obtained through a Google search. However, this was not done due to two issues arising from streaming search results into a database. First, Google does not always make a distinction between material originally written in Icelandic and material automatically translated into Icelandic from another language.[2] Although machine translated material can be interesting, it is not appropriate for the study of language change. Second, there was an issue with the timestamp of some data that was streamed automatically into a corpus. The timestamp did not always indicate the time of writing but instead the time of creation of the corpus or the time of day a search was made on the internet. Since the forecasting methods employed demanded regular time series data with proper time stamp, this was not feasible. As a result, the decision was made to use two sources for material, namely the Icelandic Gigaword Corpus (IGC, rmh=2019, rmh=2022 cf. Steingrímsson et al. 2018) and the social media platform X, formerly known as Twitter (https://twitter.com). Note that I will refer to X as Twitter throughout the dissertation as that was the name of the social media when this project started. The Icelandic Gigaword corpus is quite convenient as it contains a fair amount of written material with most of it dating from 2000 and later. It furthermore includes various types of registers such as

[2] This observation is based on my own experience looking for examples in Icelandic using Google.

formal, semi-formal and informal language (see further Chapter 8 and 9). Material from Twitter can be thought of as representing semi-formal or informal language.

Data gathering and annotation was carried out between early 2021 and mid-2023. Analysis was mostly done in the second half of 2023 although parts of it stretched into early January 2024. Data from IGC was gathered through the online interface hosted by The Árni Magnússon Institute for Icelandic studies (accessible at https://malheildir.arnastofnun.is/). Twitter material was obtained via the R-package *academictwitteR* (Barrie & Ho 2021) by sending requests to the Twitter API version 2, for which I obtained an academic license in early 2021. The license remained active until March 2023. All initial data annotation and cleaning was done using Microsoft Excel (version 2308, Microsoft Corporation 2022). Further data wrangling and arrangement into time series was done in R (R Core team 2021) and RStudio (Posit team 2023). For each time series that was constructed, measurements of the propagation of change were represented by the proportion of innovative variants at each given time as opposed to all potential variants.

Time series analysis and forecasting was done using R (R Core team 2021) and RStudio (Posit team 2023) with the package fpp3 (Hyndman 2023) which attaches a number of additional R packages relevant for forecasting and visualization, e.g., fable (O'Hara-Wild, Hyndman & Wang 2021) and ggplot2 (Wickham 2016). The forecasting textbook by Hyndman & Athanasopoulos (2021) was particularly helpful for understanding and implementing various forecasting methods and for describing how individual models work. The general forecasting workflow proposed there (and followed here) included gathering data and data wrangling, making sure the data was in the correct time series format, fitting various forecasting models to a training set, testing models on a test set and producing forecasts. Although there is no single model that is good for every time series, models in the model families that were used, that is the ETS family and the ARIMA family, are based on patterns in the time series themselves. It should be noted that sometimes simple models performed better and were chosen for forecasting, but in other cases models that rely on picking up trends or simple changes in the data were used. For further information on forecasting models see Chapter 7. For further information on data annotation, analysis, and forecasts, see individual case studies in Chapters 8 and 9.

1.4 Structure of the dissertation

The dissertation is divided into two parts. Part I deals with language forecasting in a general sense. It explains why it should be done (Chapter 2), what exactly can be predicted (Chapter 3), how the forecasting needs to be regular and systematic (Chapter 4), how frequently individual changes should be documented (Chapter 5), and what sort of patterns might emerge when language is viewed through time series data (Chapter 6). Part II presents two case studies, focusing on variation and change within Modern Icelandic. The first study (Chapter 8) deals with the grammaticalization of two complex prepositions, i.e., *á bak við* 'behind' and *við hliðina á* 'next to'. Although grammaticalization of these prepositions has been briefly mentioned in discussions on Icelandic (Friðjónsson 2004, 2007 and Rögnvaldsson 2021a, b), the changes have not been adequately documented or discussed before. The second study (Chapter 9) focuses on changes in subject case marking with the predicate *hlakka til* 'look forward to', a topic that has been widely discussed in the literature (e.g., Svavarsdóttir 1982; Jónsson & Eythórsson 2003; see further in Chapter 9).

2 Forecasting as a method to study change

2.1 Problems and pseudo-problems

Interest in language forecasting has existed for a long time, and has primarily been manifested through casual speculations on the future of a language or a linguistic variant. Despite this, systematic forecasting has typically been deemed impossible (Keller 1994:72; Croft 2000:3; for an overview see Sanches-Stockhammer 2015). The reasons for the negative views vary. They range from language being considered unsuitable for making predictions, for instance due to its nature or the nature of change (Entwistle 1953:41; Keller 1994:72; Labov 1994:10), to there being too much randomness and too many unknowns for predictions to be possible (Croft 200:2–3). While some of the reasons may reflect real problems, others are better described as distractions that unjustly present the goal as unfathomable.

In this chapter, I briefly review some of the key arguments for why language forecasting has been regarded as impossible (Section 2.2). I argue that lack of knowledge and uncertainties regarding the future are invalid reasons for avoiding predictions on language change, and that it is more fruitful to adopt an optimistic view (Section 2.3). By actively making predictions on the future development of language, new insights into processes of language change can be gained. Essentially, language forecasting may be treated as a method to study change. The future can then be thought of as a playfield for testing hypotheses and what is generally known about language change (Section 2.4). Predictions could, for instance, be based on which changes are expected to take place within a language or how changes might propagate through a language community. In case the development turns out to be different than predicted, answers may be sought as to what went wrong and how future predictions might be improved. Forecasting can thus give rise to new types of data (in the form of prediction and documentation of change) and provide new understanding on the process of change. In the end, many of the problems that have been claimed to make language forecasting impossible turn out to be pseudo-problems that disappear when forecasting is practiced.

2.2 The pessimist's view

Several arguments have been raised against language forecasting. While these vary in nature, they all lead to the conclusion that predicting the future course of a language is impossible. I name here five such arguments and briefly contextualize them. For a more optimistic view towards language forecasting, the reader is directed to Section 2.3.

The first argument against the possibility of predicting language change concerns some basic assumptions regarding the nature of change and the nature of the future. It seems reasonable to assume that the future will not work in a fundamentally different way from the present or the past. This view falls under *uniformitarianism* or the *continuity hypothesis* which is widely adopted in historical linguistics although it comes in various different flavors (e.g., Roberts 2017:338; Bergs 2012; Walkden 2019). In early discussions on the impossibility of predicting the future of a language the concept is built around the idea of natural laws and their possible parallels in linguistics. If a general law of nature (in the scientific sense) holds today, it should also hold tomorrow. This premise prevents the prediction of a future that is radically different from the present in every aspect. Laws impose restrictions on what can happen, and they make sure that predictions are grounded in stability. In the case of weather forecasting, it must be assumed that physics functions in fundamentally the same way irrespective of time and place. The question is then, what kind of laws are appropriate for language change?

One of the major contributions to the study of historical linguistics was the discovery of regular correspondence of sounds across related languages. An example of such a correspondence is the First Germanic Sound Shift which is typically stated in three parts. These are i) Indo-European voiceless stops changed into voiceless fricatives in Germanic, ii) Indo-European voiced stops became voiceless stops within Germanic, and iii) Indo-European voiced aspirated stops became voiced stops or fricative. Table 2.1 provides some examples of regular correspondence between two Germanic languages and Latin, a non-Germanic Indo-European language.

Change	**Latin**	**Icelandic**	**English**
*p > f	**p**ater	**f**aðir	**f**ather
*t > θ	**t**res	**þ**rír	**th**ree
*k > h	**c**or	**h**jarta	**h**eart

Table 2.1. Examples showing regular correspondence between Latin, a non-Germanic Indo-European language, and two Germanic languages, Icelandic and English.

The First Germanic Sound Shift is often referred to as Grimm's law, relating it to the German linguist Jacob Grimm (1785–1863) who was the first to discuss it systematically.[3] The use of the word ‘law’ is particularly noteworthy. Being anachronistic to Grimm, the term was prominent in discussions of historical linguistics in the 19th century. The

[3] The shift had already been pointed out by the Danish linguist Rasmus Christian Rask (1787–1832).

Neogrammarians famously claimed that sound laws were exceptionless (Osthoff & Brugmann 1878:XIII):

> Aller lautwandel, so weit er mechanisch vor sich geht, vollzieht sich nach ausnahmslosen gesetzen.
>
> '[E]very sound change, in as much as it occurs mechanically, takes place according to laws that admit no exceptions' (translation from Campbell 2003:92).

Interestingly, the word 'law' does not only imply that the changes are exceptionless, but it also links the study of language to sciences that incorporate the use of laws in theorizing (Robins 1997:198–199; Campbell 2003:92). However, the laws of language change are not equal to laws in other fields of study. It is not the case that sound laws *must* operate at all times. Instead, they describe trajectories of change that occur at a particular time, in a particular place, in a particular population. This has been thought to cause problems for prediction. Entwistle (1953:41), for instance, claims that because the laws "only operate for a period and then cease to work" predictions cannot be based on them as one simply does not know when the tendencies will work or cease to work.[4] Furthermore, even if common types of changes may be recognized and pointed out, for instance that prepositions may develop out of nouns (noted by Lehmann 1991:501 to be one of the most common types of grammaticalization), the knowledge is not predictive in nature. Not all nouns change into prepositions. For those prepositions that do, the time of change cannot be known before the change takes place. A dead end, or so it seems.[5]

[4] Although the works of Entwistle have been noted to be dubious, and, in some cases, based on misunderstanding (Hall 1954), his views on the possibility of predicting change is in line with the views of later scholars regarding the possibility of prediction.

[5] Uniformitarianism may have more in common with general laws of nature than sound laws do.

A second argument against language forecasting concerns recognizing premises for change. This point is discussed by Keller, who argues that the problem of predicting change is not whether laws operate or not, but whether the premises for the laws to operate will be fulfilled (Keller 1994:72). Importantly, Keller's (1994) views are grounded in a theory of change that draws on the metaphor of the invisible hand.

The invisible-hand metaphor, originating with Adam Smith (1776/1970), is typically used to explain how actions on an individual level may have unintended consequences on a larger scale. While the consequences may look like they were planned, they were not. Hence the invisible hand. Applying the invisible-hand metaphor to language, Keller (1994) argues that language change is an unintended macro-level consequence, brought about through intentional micro-level actions of individuals. The micro-level actions serve as explanatory prerequisites for the macro-level outcome. Keller (1994:70) explains this in terms of the formation of a path across a university lawn. A path emerges when multiple individuals cross the lawn in the same place. While the individuals may have certain intentions when crossing (for instance wanting to save time by taking a shortcut) the path itself is not intended. Rather, it emerges as a consequence of the cumulative behavior of individuals. An example of a language change could be the disappearance of a word from a language (Keller 1994:75). If a word stops being used by speakers of a language, the word will naturally vanish. The problem is that the premises, which serve as explanatory prerequisites for the change, cannot be predicted. We cannot foresee *why* a word should fall into disuse, and therefore we cannot predict the world's disappearance. Similarly, we cannot know whether a premises for taking a shortcut across a university lawn will continue to hold or not. Perhaps a fine will be imposed on those who cross the lawn, causing people to avoid doing so. The invisible hand theory, according to Keller, does not provide information on how things will develop. It simply explains how things have come to be the way they are.

Interestingly, the problem with predicting change for Keller is essentially that changes are unpredictable because the conditions that are required for them to take place are themselves unpredictable. Because situations that lead to language change cannot be foreseen, changes resulting from those situations can also not be foreseen. In fact, this view seems to be rooted in a real problem of forecasting, namely that predictions involve statements about the future, and the future is unknowable simply because it has not happened yet. If this view is maintained in all seriousness, no predictions about the future whatsoever should be attempted. Another dead end for forecasting. Yet, forecasts have been made and are being made in fields such as meteorology, economics, and epidemiology where the imminent future is not certain at all.

The third argument against forecasting can be linked to explanatory causes of language change. Under this scenario, circumstances that are known to have caused changes in the past might be hypothesized to cause changes in the future. It has been pointed out that language contact is often a driving force of language change (e.g, Thomason 2003). For instance, Icelandic youth are known to consume a fair amount of media in English, through extensive digital language contact (Sigurjónsdóttir & Rögnvaldsson 2018; Guðmundsdóttir, Sigurjónsdóttir & Nowenstein 2022). The extent of contact is perhaps best reflected in the fact that younger speakers sometimes code switch. In this situation, one might expect the contact situation to have an effect on Icelandic. However, according to Thomason (2003:689) even though language contact is a precondition for certain types of changes, it is not necessarily a sufficient condition, and (Thomason 2003:689):

> The general conclusion is obvious: as with internally motivated change, predicting when contact-induced change will occur is at best risky.

As it turns out, a definite causal relationship for changes in Icelandic due to contact with English have not been proven. Of the number of phenomena surveyed in a recent research project (MoLiCoDiLaCo, https://molicodilaco.hi.is/), only one was shown to correlate with the amount of speakers' exposure to English. This was the decline in the use of the subjunctive form of predicates. Nevertheless, even in this case it is dubious to infer a causal relationship between current contact with English and the decline in the use of the subjunctive form (see Sigurjónsdóttir & Nowenstein 2021:717). Variation and diminishing use of the subjunctive has been ongoing for several decades. This is not to say that English has had no influence on the structure or the use of Icelandic, it is simply that the causal link has satisfactorily proven. Thus, even though contact plays a role in certain language changes, Thomason (2003:689) is correct to point out that it is difficult to foresee exactly which effects language contact might have. A third dead end for forecasting has been reached. Even in situations which are known to cause changes, the changes themselves cannot be predicted.

The three arguments presented above can all be more or less linked to predicting the introduction of a new element into a language. But what if a change is already underway? Perhaps one could estimate its trajectory or determine when it will be completed. However, also under this scenario has forecasting been deemed impossible (Entwistle 1953:41, 53; Bauer 1994:25). The reason lies in the many uncertainties of how changes might propagate. Not all changes reach an endpoint. Some changes are reverted or never properly taken up. This might be due to new speakers failing to acquire the changes, or due to normative pressures within a society. An example of a sound change that was actively – and successfully – eradicated via the school system in Iceland, is the so-called *flámæli* (e. *'slack-jawed' speech*). The change involves a phonemic merger such that words like *sker* 'isolated rock in the sea' and *skyr* 'skyr, a type of yogurt' were no longer distinct in pronunciation. Around 42% of children in Reykjavík had this merger in 1929, but in 1935–39 the percentage had dropped down to 26–30%. A few decades later the change had been eradicated (Sigurjónsson 1960:98–99). Even in cases where changes are not actively eradicated, it is difficult to tell whether they will reach their endpoint or not (although, see Postma 2010 on the trajectory of failed changes). It follows that telling with any accuracy how quickly a change might proceed is also impossible. The only thing that can be known with some certainty is what has already happened or what the current situation is. As noted by Entwistle (1953:53):

> No particular development can be predicted, nor, once begun, can its end be foreseen, but we can give a continuous account of what has happened.

A fourth dead end for language forecasting has been reached.

Further arguments against the possibility of foreseeing change concern how changes are viewed. For instance, that they are in their nature unpredictable. To quote Labov (1994:10) "The phenomenon we are studying is irrational, violent, and unpredictable". Similarly, Bauer (1994:25) claims that "Diachronic linguistics is not a predictive science". If anyone thought the unpredictability was due to lack of knowledge, even this point has been allegedly refuted. Croft (2000:2–3) suggests that "...even with perfect knowledge of the initial state, we would not be able to predict a change" (Croft 2000:3). In fact, he discusses the possibility of predicting language change in terms of two views (Croft 2000); the optimistic view, where predictions cannot be made because we simply do not know enough about the state of affairs, and the pessimistic view, according to which there will always be some unpredictable randomness (some type of chaos) that makes predicting language change impossible. Croft leans towards the pessimistic view.

Strangely enough, Croft (2000:2) briefly compares language predictions with that of predictions in other areas, noting that "[i]n all probability we will not be able to make detailed predictions, any more than historical sciences of natural phenomena, such as meteorology, astrophysics or geology, are able to do." It is unclear what he means here by "detailed predictions". In any case, he does not mention a crucial difference between these areas of study. Namely that while forecasting is typically deemed impossible for language change, it is regularly carried out in the other areas, detailed or not.

An even more pessimistic view towards historical linguistics and prediction is found in Chomsky and Moro (2022) who claim that the mechanisms of change "remain essentially beyond our understanding" *because* changes cannot be predicted. Their view appears to convey that the ability to make predictions about a phenomenon is intrinsically linked to understanding that phenomenon. Since, "language changes themselves are almost unpredictable" (Chomsky & Moro 2022), it follows that the mechanism behind them cannot be understood. This somewhat odd view suggests that the possibility of forecasting is closely connected to the nature of the phenomenon under study.

To sum up. Language change has been deemed impossible to predict due to changes being unpredictable in nature. Even the existence of regularities in change does not facilitate predictions under the pessimists' view. The laws of language change do not operate in the same way as laws of natural phenomena. They are sometimes active and sometimes not. Predictions have also been claimed impossible because there is no way of knowing when preconditions for a change will be met, when a change might start, or if an ongoing change will continue, come to an end or be reversed. Even with a perfect knowledge of the state of affairs and factors that might influence change, there will always be some randomness, some chaos that makes predictions impossible. The outlook seems rather gloomy. It is almost as if the task at hand has been deemed theoretically impossible before it has been attempted.

2.3 Reasons for optimism

Despite various scholars having expressed negative views towards the possibility of predicting language change, the outlook may not be so bleak. In this section, I point out how the arguments against language forecasting are at best distractions. Before proceeding, it is fitting to consider the value of an optimistic view. In the words of Andersen (1990:1):

> It is often the case, when a certain sort of phenomenon evokes different attitudes in different observers, that some of these attitudes are more fruitful, more productive of understanding and insight, and others less so. In the case at hand, as in many other cases, there is no doubt that the optimists have contributed more than the skeptics or the pessimists; and no wonder: the optimists have after all accepted at face value observations that are in need of explanation and thus represent an intellectual challenge.

Although Andersen is mostly concerned with the structure of drift, it is easy to see how his comment might apply to forecasting.[6] That is to say, if a pessimistic view is adopted and

[6] The concept of drift originated with Sapir (1941) who inter alia uses it to describe a gradual change in language that, in the long run, seems to have a direction. Recently, the concept has been used in a more

things are deemed impossible, no progress will be made – not because progress is impossible, but because no attempts are made towards making progress. If, however, a more optimistic view is embraced, there is a chance (no matter how small it may be) that advancement can be made. This is the view adopted here.

Briefly summarizing the arguments discussed in Section 2.2, these involved predictions being regarded impossible due the five following factors: i) laws of language change are different from laws in the natural sciences, ii) premises for change cannot be predicted and therefore change cannot be predicted, iii) situations that have been known to lead to change are not informative on whether there will be a change or not, nor what kind of change might take place, iv) even if a change has started occurring, we cannot know if it will continue or be reversed, and v) there is simply too much randomness in language change for forecasting ever to be possible, even if we knew everything about a certain language at a particular point in time.

The claim that language change is unpredictable due to laws of change operating in different ways from laws in other sciences is at best misleading. The notion of 'laws' was originally meant to evoke parallels between the study of language and the study of natural sciences (Campbell 2003:92). They were not necessarily meant to suggest that some change or the other *must* always be operating. Rather, they capture regularities of change and reflect general tendencies. Fixating on the concept of 'law' is pointless if the goal is to figure out a constant in how language *must* always function. In this case, one should rather look towards a general form of uniformitarianism, the idea that language operates in fundamentally the same way at all times (e.g., Roberts 2017).

specific sense, i.e., for "any source of unbiased stochasticity, or sampling error, in the acquisition, processing, or production of language" (Ventura et al. 2022:2–3).

As discussed in detail by Walkden (2019:11), "uniformitarianism as understood in linguistics is not itself a uniform notion" but may be manifested in various ways. One such is uniformity of state which assumes that languages of the past are no different from languages of the present in any fundamental way. This view provides a basic premise that language will not function in a completely different way at different time periods. Typically, the view is tacitly assumed. Without it, different stages of a language could not be compared as one could not guarantee that the comparison was justifiable. Thus, uniformitarianism provides a convenient expectation towards language always operating in the same way, even in the future (although, see Bergs 2012; De Smet & Van de Velde 2017 for limitations and pitfalls of assuming strong uniformitarianism). Whether the word 'law' figures in or not is simply irrelevant.

Turning to the second point, namely that change cannot be predicted because the premises for a change cannot be predicted,[7] this too can be said to be misleading. In fact, the view seems to involve an intentional misinterpretation of the nature of forecasting. To be sure, it is not possible to know the future simply because the future has not happened yet. However, the goal of forecasting is not to *know*, but to make an informed "guess" (a prediction) about what the future *might* look like. Whether the guess turns out to be correct or not is another matter. Distinguishing between predictions and actual outcomes is tremendously important. It may even be claimed that the lack of differentiating between the two is the reason for language forecasting being deemed impossible. The fact that we cannot know with certainty what will happen acts as a barrier for making predictions. Simply put, fear of failure prevents forecasting.

[7] The premises in this context can be understood as a necessary but not sufficient condition for a change to occur.

Once a distinction has been made between a prediction and what the actual state of affairs turns out to be, a fresh perspective on forecasting can be obtained. Although predictions should ideally turn out to be correct, failure is not necessarily bad. Incorrect and inaccurate predictions can lead to revaluation or reassessment of various factors that were taken into consideration in generating the predictions. Keeping this in mind, arguments (iii)–(v) above seem to evaporate.

While situations that have given rise to changes in the past do not guarantee that changes will take place in similar situations the future (cf. Thomason 2003:689 on language contact), they nevertheless generate certain expectations towards what can happen. In the case of language contact, we might expect structures and patterns to be replicated from the source language into the target language (e.g., Matras & Sakel 2007). An instance of such replication in Icelandic is the farewell term *hafðu góðan dag* 'have a nice day' which has been claimed to be a loan translation from English (Sigurðardóttir 2019–2020). Although foreign influence was pointed out after the change took place, it stands to reason to assume that this similar type of change could be triggered in other contact situations. More generally, situations known to trigger change can provide hints as to what to keep an eye out for.

Once changes have started taking place, it may be hard to tell with certainty how they will proceed. Will they catch on and propagate through the community, remain stable for an extended period of time, or will they halt and disappear? Contrary to what has been claimed (see Section 2.2), the uncertainties do not prevent predictions. Expectations towards trajectories of change may serve as a baseline for hypothesizing about the future. Furthermore, if an ongoing trend has been noticed, it is unlikely that it will come to a sudden halt. Making predictions based on what has been observed to have occurred in the past allows for evaluation on how well the tendencies are understood. When predictions do not turn out as expected, explanations need to be sought. This may prompt further investigation into what causes certain developments or trajectories or even result in a new type of data needing to be gathered.

The final point, Croft's (2000:2–3) pessimistic view that even though we knew everything, there is too much randomness in change for making predictions, is also trivial. While randomness may exist in all processes, it does not prevent forecasting. In fact, some types of randomness can be factored into formal models and accounted for. They are foreseeable up to a certain extent and may be reflected in prediction intervals in forecasts, i.e., the lightly colored area around the point forecast. For example, if a temperature is predicted to be 15 °C (the point forecast) it is not surprising if it turns out to be a degree higher or lower in the area, reflecting minor uncertainties.

A second type of randomness involves real world events and factors that were not accounted for in the forecast. Of the more catastrophic nature are events such as epidemics, invasions, economical crashes, and mass migrations, all of which can affect the phenomenon one is interested in. As these are typically not foreseeable and do not occur regularly, there is no need to dwell on them here.[8] Other factors that can contribute to randomness are those that were incorrectly accounted for in a forecast or not taken into consideration when they should have. These might include the role of language acquisition

[8] While seemingly unpredictable, these types of events do not always occur completely out of the blue. Rather, they may have precursors that can be identified. Emigration, for instance, may be related to factors such as population size, level of education, status of the economy, employment opportunities, and the possibility for the emigrants to keep some aspects of their language or culture (see e.g., Hatton & Williamson 1998).

(Paul 1886; Andersen 1973; Lightfoot 1979 among others), social aspects such as gender and class (Labov 1966, 1990), the prestige of the variants under study and views toward individuals who use them, spoken and written language contact (Thomason 2003; Curzan 2009:1093–1094; Lavidas 2021), language standardization (Haugen 1966; Ammon 2015; on the standardization of Icelandic see Árnason 2002; Kristinsson 2019; on the loss of linguistic features due to standardization see Kikusawa 2012), and linguistic purism prescriptivism (Thomas 1991; Kristinsson 2006, 2007). Although these often figure in literature on language change, it can be difficult to foresee their effect (see discussion in Section 2.2) or formally account for them in forecasting. This does, of course, not mean that predictions cannot be made w.r.t these phenomena. It simply means that there is a lot to be learned about their effects and how they can be used in predictions. Instead of taking the stance that too little is known for language forecasting to be possible, a better strategy might be to start off from what ***is*** known and work from there. Perhaps forecasting can be used to gain more insight into these phenomena. Trials and errors in the area of forecasting might lead to better understanding of the whole process.

To conclude, there is no need to know everything in order to make predictions nor is there a need to accurately foresee everything. Forecasting is not about *knowing* what will happen. It is about making predictions on what *might* happen or what is *likely* to happen. Forecasting can be viewed as complementary to hindcasting. The latter focuses on making accurate predictions for an already known situation, while the former focuses on predicting unknown situations. Under this view, a new perspective on the study of language change can be obtained.

2.4 A new perspective on language change

Once a clear distinction has been made between a likely future (a prediction) and the true future (how things turn out), it is possible to view forecasting as a method to study change. Needless to say, the question arises what kind of insight it gives into the language change and how it fits in with other ways of studying language.

The study of language is traditionally divided up into synchronic and historical linguistics (de Saussure 1916/1959:79–83; for a brief overview see Sanchez-Stockhammer 2015). While synchronic linguistics deals with the structure of a particular language at a certain point in time, historical linguistics tends to focus on previous states of a language, relying on historical documentation and comparative evidence. Adding predictions for the future means adding a third dimension, i.e., the future. Sanchez-Stockhammer (2015), who discusses the possibility of forecasting, represents the three dimensions (language of the past, present and future) as in Figure 2.1. She furthermore suggests that what is currently known about language change might serve as a basis for making predictions. In her words, “it may be possible to extrapolate from the findings about the present and the past” (Sanchez-Stockhammer 2015).

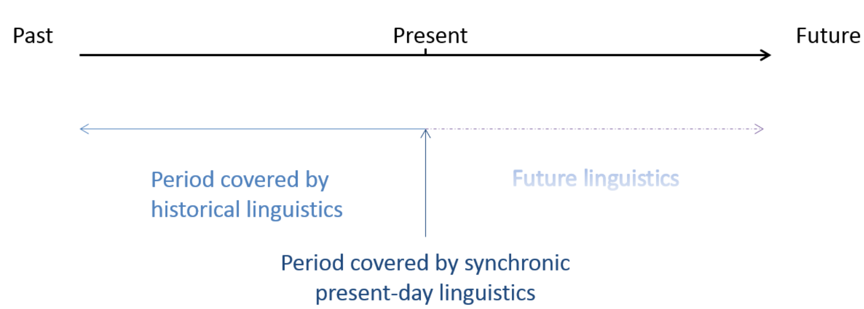


Figure 2.1. The period covered by historical linguistics is generally thought to stretch from the present back into the past. Synchronic linguistics, on the other hand, mostly focus on the present (from Sanchez-Stockhammer 2015).

Despite optimistic views, Sanchez-Stockhammer (2015) does not propose any concrete methods for how to approach the task at hand. In a review of her paper,[9] Schneider (2018) notes that the question whether forecasting is actually possible remains to be answered properly, but points out that "a few promising steps towards an answer" (Schneider 2018) have been taken.

To better understand the challenges of forecasting and how prior knowledge of historical linguistics can be of use in making forecasts, it is worth taking a second look at the three relevant points in time, the past, present and future, and observe how language change can be investigated by referring to or moving between these points in time, cf. Figure 2.2.

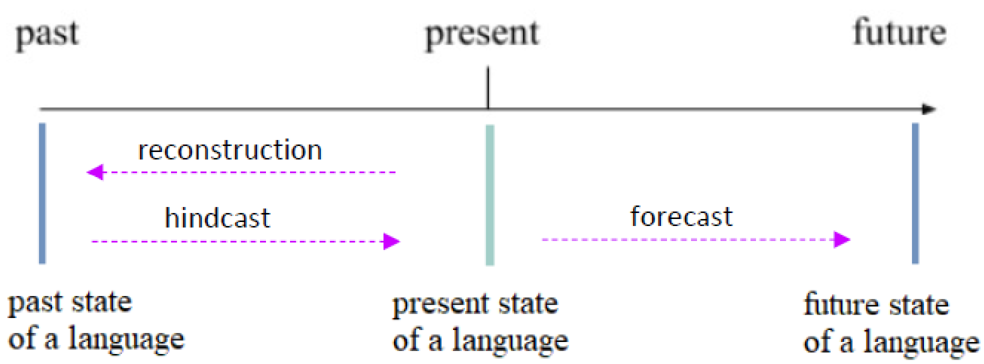


Figure 2.2. *Reconstruction* is when an unknown state of language is reconstructed from a known state, i.e., either the present or a known past. Predicting a known state of language, such as the present, from a known past can be referred to as *hindcasting*. *Forecasting* involves a prediction about a future state of a language.

[9] Sanchez-Stocammer's paper originally served as an introduction to a workshop at the 3rd ISLE (The International Society for the Linguistics of English) conference in 2014. The theme of the workshop was the possibility of predicting language changes. Papers presented in the workshop were eventually published together online. Schneider (2018) reviews the whole collection of papers, not just Sanchez-Stockhammer.

The most obvious way in which language change can be studied involves comparing a present state of a language to a known past state. This may be done in order to uncover which changes have taken place within a given language. Certain morphological forms in Old Icelandic (ca. 1100–1540) can, for instance, be systematically compared to corresponding forms in Modern Icelandic (1540 onwards). By observing Old Icelandic nominative singular forms like *hestr* ‘horse’ and the Modern Icelandic counterpart *hestur*, one can infer that an *u*-insertion (Ice. u-*innskot*) has taken place between the two stages (for an overview of changes from Old to Modern Icelandic see e.g., Noreen 1923; Karlsson 2000; Bernharðsson 2016). It is worth noting that changes are typically only discovered *after* they have taken place (Hockett 1958:444–445) and it is then that the condition for them can be identified. Knowledge that has been gained from observing changes and reconstruing conditions for them has provided valuable insight into ongoing changes in contemporary languages. In fact, it is through having identified changes in the past that ongoing changes may be recognized.

A second way in which changes can be studied is through reconstruction. This involves making inferences about a lesser known (or unknown) past through comparative evidence from a known past or a known present. The comparative evidence can be in the form of cognate words from related languages or in the form of related words and structures within the same language. For example, on the basis of the word *snjór* ‘snow’ in Old (and Modern) Icelandic and its cognates in Old Swedish *snior*, Old English *snāw*, and Gothic *snaiws*, the Proto-Germanic form can be reconstructed as **snaigwa-z* (Magnússon 1989 s.v.). In Figure 2.2, reconstruction is represented by a dotted line going from the present into the past.

A third way to study changes also involves moving from the past into the present. Modeling trajectories that gave rise to a known present (or a known past) is commonly referred to as hindcasting, retro predicting or backtesting. Backtesting can be used to test forecasting models before true predictions are generated. If a model gives less than optimal results through backtesting, it will likely perform poorly in true forecasting situations.[10] Although forecasting has not been systematically practiced in linguistics, backtesting has been attempted with promising results (Van de Velde, 2017). Van de Velde (2017) and Nijs and Van de Velde (under review) have, for instance, shown that by fitting an S-curve model to historical data it may be possible to make time-dependent predictions about the trajectory of a language change.

The fourth way of studying language change is to use what is already known about the topic and hypothesize what a future state of a language might look like. This falls under forecasting (see Figure 2.2). Despite views towards forecasting having been rather pessimistic (see Section 2.2), some attempts have been made, typically resulting in vague predictions or general statements about the future of a relevant language (see overview in Sanchez-Stockhammer 2015, and discussion in Chapter 5). As of yet, forecasting is a highly understudied area within linguistics that has great potential.

When attempting predictions, it is useful to start with something that is already known. In the case of language, it is safe to say that changes are never completely random. Observed regularities have been captured in various ways, for instance through the notion of laws of change (Osthoff & Brugmann 1878) or through paths of grammaticalization

[10] This does not need to be the case. A model might perform sub-optimally for a known situation but still manage to predict the future semi-accurately. Similarly, a model that performs extremely well for a known situation is not guaranteed to perform well for future predictions.

(e.g., Hopper & Traugott 2003). In short, there appears to be a structure to the way in which languages change (see discussion in Andersen 1990). Despite this, there is a slight problem. Language change tends to only be identified and discussed *after* it has taken place (Hockett 1958:444–445) and any reason for *why* or *how* it took place is generally reconstrued thereupon. There are disadvantages to only investigating changes after they have taken place as causative and explanations cannot be properly tested. Furthermore, there is no way of knowing if all relevant information has been taken into consideration or not. Putting it metaphorically: While evidence of various types can be connected in order to provide a holistic picture of what has happened, it can be difficult to evaluate whether everything was *correctly* puzzled together or even if *all* relevant factors were taken into consideration. Changes and explanations for them need to be – and are – constantly being reevaluated. Forecasting offers a way to do so. The future can be treated as a playfield for testing assumptions about language change, leading to better understanding of the factors at play. The question is then where to start?

Forecasting necessarily requires some understanding of where to look for signs that indicate what the future may hold. In the case of language change, this must be in the form of concrete data as well as methods that provide guidance in how to handle the data. Since forecasting has not been systematically practiced in linguistics, there is little understanding of what kind of data is best for this purpose. However, this is not a serious problem. Instead of waiting for the perfect type of data to appear, one may simply set out to systematically gather (or generate) data believed to be suitable. Worst case scenario, a better understanding will be gained of what is needed to make good predictions.

Once appropriate data has been obtained, questions and suitable methods for generating predictions can be applied. Potential questions may include extrapolating patterns of change into the future, attempting to predict how quickly changes might take place or hypothesize about which changes might take place in a given situation. Experimenting with causality, i.e., which factors influence change, is also possible. Note that whichever questions are tackled, and whichever methods are used, the results will lead to novel types of data for historical linguistics in the form of predictions.

Aside from obtaining new types of data, both in the form of predictions and systematically gathered data for forecasting, much can be learned about language change through making predictions about the future. The results from such investigations can, in turn, lead to better understanding of changes that have already taken place in the past. They may even contribute towards how reconstructions are done or how quickly changes are thought to take place under different conditions. Simply put, forecasting has the potential of generating knowledge that can feed into traditional studies of change. In case predictions are not born out, they may evoke new questions and prompt an investigation into why a change followed the path it did. Forecasting is thus not only a method to study what could happen in the future (cf. Sanchez-Stockhammer's 2015 discussion on future linguistics), but it can also provide insight into the past.

Essentially, language forecasting need not be treated all that differently from other types of predictions in linguistics. These include predictions such as what kind of structures are thought to exist (or not exist) given a certain framework of study, which phonetic category a sound belongs to, or which morphological rules are active in such-and-such situations. In all these cases, incorrect predictions help refine the approach or the analysis.

There is no reason to think that this should work differently for language change and forecasting.

Notwithstanding there being some similarities between language forecasting and other types of predictions in linguistics, it is worth noting that the nature of the former is slightly different from the nature of the latter. While forecasting is essentially time-dependent, time is usually not crucial to other types of predictions in linguistics.[11] Put slightly differently, when predicting the phonological category of a sound or whether a syntactic structure exists in a language or not, the predictions typically do not depend on the time of day or whether things were uttered yesterday or today or will be uttered tomorrow or even next year. These types of predictions are not proven or disproven with reference to time, and that makes them different from forecasts. The fact that the truth lies in the future does not affect the usefulness of forecasting. It might simply mean that progress in the area might be slower as predictions cannot be verified right away.

Demonstrating briefly that forecasting is feasible, it suffices to mention that researchers have already started looking into how trajectories and paths of change might be used in making predictions on language change (cf. Nevalainen 2015; Van de Velde 2017). As of yet, these methods have mostly been used for backtesting, but they have already resulted in new methodological knowledge for studying change and new knowledge on the nature of the data used.

Finally, it is worth mentioning how diachronic studies, either through forecasting or by other means, are relevant for synchronic ones. For instance, it has been argued that

[11] I am not claiming that time is of *no* importance in other areas of predictions in linguistics. There is always some kind of idea of a *current state of language*, *previous state of language* etc.

any restrictions on possible synchronic patterns must, at least to some extent, be linked to diachrony and the result of language change (see for instance Aristar 1991; Blevins 2006). Diachronic linguistics thus has a lot to offer when it comes to accounting for attested linguistic patterns and how common or uncommon they may be. Instead of patterns only being attributed to factors such as universal grammar or cognitive biases, history may also play a role. The separation between synchrony and diachrony, as suggested by de Saussure (1916/1959:79–83), may not be as important as often claimed.[12]

2.5 The outlook

It is safe to say that the outlook for language forecasting is not as poor as it has sometimes been claimed (cf. discussion in Section 2.2), and, by adopting an optimistic view (cf. discussion in Section 2.3), there are reasons to believe that it *can* and *should* be done (cf. discussion in Section 2.4).

Once a distinction has been made between predictions and true outcomes, language forecasting is both achievable and theoretically feasible. It may even be viewed as a new method for studying change. The starting point is, of course, knowledge that has already been gained. For instance, that there are regularities in language change that can be captured. Directionality is often observed, and even in circumstances where it is unclear whether any change will take place or whether an ongoing change will continue on the same path, there are restrictions on what can happen and how quickly things can happen. By using the future as a "playfield", assumptions and expectations towards change can be

[12] Paul Kiparsky (2019) brought up the importance of historical linguistics in a plenary talk given at NELS in 2019, noting that historical linguistics have implications for the study of language synchronically.

tested and, consequently, more knowledge gained in the areas of language forecasting and language change.

Since there is no tradition of forecasting in linguistics, it is currently not well understood what type of data is most suitable for the task. A considerable amount of language data is accessible through multiple online corpora, in scientific studies on language and on the internet, to name only a few sources. However, data has never been systematically gathered specifically for forecasting. Simply by attempting to make predictions on language change, a better understanding will be gained of the type of data needed for such a task. Forecasting can also put pressure on figuring out how to best quantify phenomena that influence change, leading to better overall documentation of language changes and processes at play.

Aside from learning about which kind of data is suitable for making predictions about the future, forecasting can provide novel information on the mechanisms of change. Knowledge that has been gained about changes in the past and the present has led to certain expectations on how changes should work. These expectations can be turned into hypotheses that can be tested through forecasting. The results could shed new light on changes in the past. Additionally, further information might be gained on propagation of change through a community of speakers, whether it always follows the same pattern or not, how quickly a change might be completed, and which factors affect language change. The information obtained through forecasting might feed into other areas of historical linguistics, e.g., what is known about rate of change, reconstruction, phylogenetics and language planning, to name a few.

Inaccurate forecasts may turn out to be both informative and useful. Ideally, an incorrect forecast should allow for re-evaluation. They should hold some clues as to figuring out why the predictions were wrong. With a systematic approach to forecasting, failure might be attributed to at least three different factors: i) the data that was used, ii) the forecasting model, and iii) randomness. Failure on the level of the data includes bad use of attested data, missing data and data that was not included in the model (or factored into the prediction). Problems with the model might involve an incorrect model being selected or changes in the premises for using a particular model. Perhaps a particular model fit the original data well, but as soon as more data was added the fit turned out to be sub-optimal. Finally, there might have been more randomness than originally estimated, or perhaps something happened that was not foreseen. Importantly, none of these need to prevent forecasting. While predictions may go wrong, much is to be gained in the domain of knowledge. Essentially, it is possible to learn about the future by studying the past and to learn about the past by studying the future.

3 Defining what to predict

3.1 Contextualizing the task

Language can be viewed as a complex system, manifested through the rule-governed behavior of individuals (see Searle 1969:12 on language as rule-governed behavior). Investigating this system is not straightforward as multiple factors may need to be taken into consideration.[13] In order for language forecasting to be useful in the sense laid out in Chapter 2, a systematic breakdown of the task is appropriate, including defining what the target of study is and which aspects of language can be predicted. A further consideration pertains to which kind of data is appropriate for the forecasting task. This is the goal of the current chapter.

Starting by distinguishing between internal and external language (I-language and E-language for short), Section 3.2 raises the question which one language forecasting should focus on. In theory, either one might be the target of predictions although this depends somewhat on what the exact goal of the forecast is and which kinds of data can be used to reach that goal. Naturally, different data provide different insights into language and change. Although data that is carefully generated with specific questions in mind (termed here *specialized language data*) tend to give the most holistic picture of a phenomenon that is being studied, such data may often be difficult or even impossible to obtain, especially when dealing with past stages of language. In these cases, data that exists

[13] To name only a few, these might include carefully framing what is being studied, explaining whether or how the phenomenon under investigation interacts with other aspects of language, or looking into variation among speakers and under what conditions such variation may arise.

independently of studies on language (referred to here as *convenient* (*E*-)*language data*) must suffice. While convenient language data leaves much to be desired, it is nevertheless an important source of information and sometimes the only source.

Transitioning from I-language, E-language, and various types of data to what it means to forecast language change, Section 3.3 deals with definitions of change and how the concept has been used to refer to both innovation and propagation. While it may seem logical to connect innovations with individuals and I-languages as is commonly done in mentalists approaches to language change individuals are necessarily always a part of a larger population and it can be beneficial to treat them as such in the context of forecasting.

Focusing on manifestations of language in a population, it can be claimed that language forecasting should be concerned with the following question:

(3.1) **Forecasting question** (to be revised in Chapter 4)

What will the situation *x*, of a phenomenon *p* be in the future?

The question in (3.1) can be approached in multiple ways, depending on types of changes one is interested in predicting, whether the focus is on language in an abstract sense or as a collection of attested utterances produced by a population, and whether we are concerned with a single individual or a population of speakers. As an example, one might ask how many individuals would accept a particular structure at a given time in the future, or hypothesize how frequently a particular structure is likely to appear in attested utterances at some time or other in the future. Furthermore, the concept of *phenomenon* in (3.1) may refer to different features of grammar and different layers of language change. It can, for instance, concern some aspect of a language system as a whole such as whether a language is analytic or synthetic or whether it uses prepositions or postpositions; it may also refer to a specific structure or variation within the language, for instance case marking with a particular predicate or a group of predicates. The various layers of change that predictions can be made about are discussed in Section 3.4.

Note that the current chapter is not so much about arguing for a single correct way of approaching language change or language forecasting. Rather, it highlights how a forecasting task may be affected by choices such as type of data used, whether the focus is on language change on an individual or a population level, and on which aspects of language change the predictions are intended to cover. Section 3.5 summarizes these factors and provides a reasonable path which might be pursued going forward.

3.2 Internal and external language

The distinction between internal mental grammar (I-language) and external language (E-language) features prominently in studies on language, especially in the context of syntax. Roughly speaking, the former refers to an abstract system which is capable of generating grammatical structures while the latter applies to whichever structures or output is generated. Although the distinction between mental grammar and external language is in principle already found in the writings of von Humboldt and de Saussure (Robins 1979:175), it is more commonly associated with Generative approaches to language, where a great emphasis has been laid on the study of language competence rather than performance (see for instance Chomsky 1965, 1986). Within the area of diachronic syntax much of the discussion has focused on problems related to abstract grammars of the past and the (im)possibility of syntactic reconstruction (e.g., Watkins 1976; Lightfoot 1979:154–166, 2002; Walkden 2014; for an overview of issues and approaches to syntactic reconstruction see Eythórsson & Barðdal 2016).

In light of the distinction between internal and external language, the question arises which one language forecasting should be concerned with. On the one hand, it is possible to make predictions about representations of language in the mind, including restrictions and flexibility of future I-languages, categorical boundaries in phonology or word-boundaries in morphosyntax, to name only a few. On the other hand, focusing on likely utterances in the future might be a viable option. Arguably both questions are of interest, and it is worth considering types of data accessible to serve as a basis for forecasting.

The focus on language as an abstract system in the mind of speakers, especially in the area of syntax, has given rise to considerable amount of linguistic data in the form of grammaticality judgments. While many structures tend to be judged either clearly acceptable or unacceptable, others fall into the gray area somewhere between the two. In theory, the judgments reflect some properties of the mental grammar of individuals, for instance showing how individuals impose categorization that "appears to correspond with the grammatical/ungrammatical distinction" (Sprouse 2007b:124, 2007a). In reality, these judgments may reflect complex statistical facts about types of utterances individuals are exposed to and expect (Bresnan 2007a, 2007b). The examples in (3.2) show how the expletive *það* in Modern Icelandic can only occur clause-initially (3.2a) and is considered ungrammatical clause-internally (3.2b) (for an overview see Thráinsson 2007:309–313).

(3.2)	a.	***Það***	*kom*	*til*	*mín*	*draugur*	í gær.
		EXPL	came	to	me	ghost	yesterday

‘A ghost came to me yesterday.’

b.	*Í gær*	*kom*	(****það***) til	mín	*draugur*.
	yesterday	came	EXPL to	me	ghost

‘Yesterday, a ghost came to me.’

Grammaticality judgments as in (3.2) provide an interesting piece of evidence for the study of language variation and change as they reflect which structures *can* and *cannot* be produced by speakers. The judgments can differ across time and space so that structures considered grammatical by speakers at one point in time will not necessarily be considered grammatical by speakers at a different point in time and vice versa. Although it is difficult to verify this with absolute certainty, it seems reasonable to assume that examples like (3.2a) were ungrammatical in Old Icelandic, based on the fact that the earliest attested examples of expletive *það* are found in texts in the 16th and 17th centuries (Rögnvaldsson 2002; Eythórsson & Sigurðardóttir 2016; for a different view see Booth 2018, 2019, 2020). Grammatical judgments may thus offer an insight into differences in I-languages of individuals at different points in time and provide clues about directionality of changes.

Grammaticality judgments are not the only type of data generated specifically in the context of studying language. Experimental data from perception and production tasks can also provide valuable insight.[14] For perception tasks, speakers might be required to

[14] The goal here is not to provide an exhaustive list of all possible types of studies that can generate data relevant for questions tied to specific linguistic phenomena.

complete tasks where information can be gathered on how they perceive a certain input. The tasks may focus on anything from categorical perception of sounds to assigning meaning to utterances or parsing syntactic structures. Production tasks might require participants to produce target utterances or structures, for instance through fill-in-the-blanks tasks which have been prominent studies of subject case marking in Icelandic (e.g., in the following studies: Svavarsdóttir 1982; Jónsson & Eythórsson 2003; Thráinsson et al. 2013). The type of data generated through experiments specifically for the study of language might be termed *specialized language data* and they can target either language in the mind (I-language) or focus on generated output (E-language).

Despite providing unique information on language related phenomena, especially synchronically, using specialized language data in diachronic studies of language presents several problems. First, taking older stages of a language into account, it is not possible to access data that reflects what was an impossible utterance according to previous generations. Provided a language has rich documentation, something can be said about types of grammatical utterances during a given period. For instance, take the Old Icelandic example in (3.3) which is from Snorri's Edda (preserved in a manuscript from around c1300-1350). Note the neuter, singular pronoun *það* 'it' which appears clause-initially and presumably refers cataphorically to the following infinitival clause.

(3.3)	***Það***	*var*	*eitt*	*sinn*	*er*	*hún*	*reið*	***að***	***vanir***	***nokkrir***
	it	was	one	time	when	she	rode	that	vanir	some
	sá	***reið***	***hennar***		***í***	***loftinu.***				
	saw	riding	hers		in	the.air				

‘One time it occurred, when she was riding, that some vanir saw her riding in the air.’ (*Snorra Edda* (Gylfaginning), Ch. 35)

The neuter singular pronoun *það* is homonymous with the non-referential expletive *það* found in Modern Icelandic. Although it is standardly assumed that the expletive *það* emerged in the 16th to 17th century (Rögnvaldsson 2002; Eythórsson & Sigurðardóttir 2016), Rögnvaldsson (2002) claims that examples of the type in (3.3) suggests a stage between a referential pronoun and an expletive. A similar view is adopted in the work of Booth (2018, 2019, 2020). In case *það* in (3.3) is taken to be an expletive, there is a need to explain why expletives are not used in other contexts in written records from the same time period. One might claim that such structures were indeed possible in Old Icelandic, but they happened to be never written down. On this view there would be little or no difference between Modern Icelandic and Old Icelandic grammars when it comes to expletives and whether the mental grammar of individuals allowed for them or not. The only difference would be that they happened to occur more often in attested language data in Modern Icelandic than Old Icelandic. Alternatively, *það* in (3.3) can be taken as a referential pronoun and structures with unambiguous expletives as in (3.4) might be assumed to have been ungrammatical in Old Icelandic.

(3.4) A hypothetical grammaticality judgment for Old Icelandic

****Það***	*kom*	*til*	*mín*	*draugur*	*í gær*
it	came	to	me	ghost	yesterday

‘A ghost visited me yesterday.’

Setting aside the impossibility of obtaining information on ungrammatical structure for past stages of languages, one might claim that generating such knowledge for present and future states of language is perfectly doable and reasonable. As one moves forward in time it should be possible to repeatedly conduct studies that generate specialized language data targeting a phenomenon one is interested in. However, generating specialized language data can be quite time consuming, especially when factoring in designs of experiments, and the time it takes to carry them out and processing the results. In the context of diachrony, using specialized language data imposes restrictions on how often it is realistic to repeatedly gather data. Often studies are not repeated, or when they are repeated it is not at regular time intervals. This can lead to problems related to sampling frequencies and the documentation of variation and change through time (see further discussion in Chapter 5).

Specialized language data can be contrasted with data that exists independently of language studies. These might be referred to as *convenient* (*E*-)*language data* and they can be in the form of natural spoken language, written material, or recordings. Naturally, the availability of spoken, written or recorded language depends heavily on which language is being studied as well as which state (past, present and future) of the language one is interested in. The availability of material will necessarily restrict possible forecasting questions.

Since convenient E-language data does not provide information about what the grammar of individuals ***cannot*** do, one might ask whether such data gives an adequate picture of language in the context of change. The answer is not simple. While it may offer a somewhat limited view on certain phenomena, it does indicate which utterances individuals ***can*** or ***might*** say, as attested data must be generated in accordance with someone's mental grammar.[15] Making predictions about future utterances (E-language) is therefore both reasonable and feasible. Such predictions entail postulating features about individuals' I-language. Furthermore, these types of predictions can be quite easily verified as one would only have to reference convenient E-language data to confirm or disprove that the relevant utterances are indeed being produced.

Making predictions on future E-language utterances does not exclude the option of making predictions on future I-languages. Under this scenario, predictions might concern which structure and utterances would be grammatical or ungrammatical at various times in the future. However, verifying these types of predictions might be more difficult than verifying predictions about future E-language, especially if the predictions concern ungrammatical structures. In the case one predicts a certain structure to be ungrammatical, the first step in verifying the prediction would be to observe attested data at the relevant point of time in the future. If the structure turns out to be attested, the prediction can be claimed to be incorrect. If, however, the structure turns out not to be attested it is not possible to immediately conclude that the prediction was born out. Rather, one would have to verify the correctness of the prediction by asking speakers for grammaticality judgments.

[15] Attested utterances and structures may, of course, have some errors in them. In case there is a doubt as to whether something should count as an "error" (presumed ungrammatical), one must make an informed stand on the matter and decide whether to view it as real data or not.

The reason for this is that a linguistic feature may be grammatical without ever happening to be produced.

Without arguing against predictions being made about I-languages of the future, it might be noted that predictions about utterances in the future are likely to be more easily verified than predictions about mental grammars in the future.

3.3 Change on the individual and the population level

Previous literature tends to make a distinction between two ways in which the word *change* is used. First, it can refer to the introduction of a new linguistic feature (often called either *change* or *innovation*) in one or more individuals. Second, it may be used to cover the spread of (sometimes called *diffusion* or *propagation*), or the disappearance of a particular feature within a given population (e.g. Milroy & Milroy 1985:347–348; Lass, 1997; Croft 2000).

Language change in the former sense, i.e., as an introduction of a new linguistic feature (or speaker innovation in the sense of Milroy & Milroy 1985:347–348), can be thought of as originating in the mental grammar (the I-language) of an individual during language acquisition (Andersen 1973; Lightfoot 1979). What happens is that the individual forms their mental grammar on the basis of available linguistic data (E-language) in their surroundings. Since the learner does not have direct access to the mental grammars of other language users, the individual's new mental grammar will be different in some respects (this being the basis for definitions of change see e.g. Hale 2007 who defines change as the set of differences between the new grammar and the source grammar). The new abstract grammar of the individual may then be able to produce novel output, a structure or a feature which was not attested in the input language and does not conform to the grammar of the previous generation. Figure 3.1, adopted from Andersen 1973:767, illustrates how an individual's mental grammar is formed on the basis of the output of a previous generation.

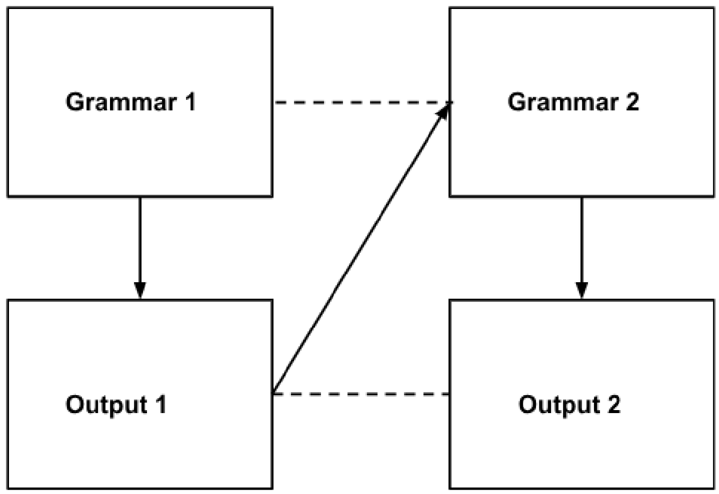


Figure 3.1. Transmission of grammar takes place indirectly in that grammar of a new generation (Grammar 2) is formed on the basis of utterances (Output 1) of a previous generation (see Andersen 1973:767).

Since there is no guarantee that individuals will produce all the structures that their internal grammar allows for, there may be a gap between a change being attested in an individual's I-language and the E-language of the whole population. A prediction focusing on innovation might either try to predict the existence of a speaker that *can* produce an innovative feature or predict the existence of a speaker that *does* produce an innovative feature. Typically, for the study of language change, one may be interested in identifying the first few instances of a novel feature, acknowledging that they are linked to change on the level of the individual, i.e., there must have been some individual that used language in an innovative way. However, as soon as there is more than one individual using a particular innovative feature, the second way in which the concept *change* is used becomes relevant.

The second sense in which the concept *change* is used is to refer to the propagation of a novel linguistic feature through a language community (Milroy & Milroy 1985 refer to this as linguistic change). In this case, a deviation from the standard language[16] may first be observed sporadically in written or spoken language by one or more individuals. Next, the new variant gains ground within the language community at the cost of an older variety which was previously used. Eventually the new variant becomes a part of the standard language, and the older variant disappears. This type of change might be considered to take place at the level of the population as opposed to at the level of the individual.

The difference between focusing on individuals versus population is large enough to matter for language forecasting as it determines whether one tries to make a prediction about what a given speaker might do in the future versus what the language community, treated as a whole, might do in the future. There is, of course, an overlap between making predictions on individual and the population level. A given population necessarily consists of multiple individual speakers. Therefore, whatever the individual speaker produces (or can produce) is part of what the community as a whole produces (or can produce).

Treating individuals as individuals that make up a population gives rise to a complex synchronic and diachronic picture. As depicted in Figure 3.2, which is a snapshot of a hypothetical language community at a particular point in time, each individual has

[16] The term *standard language* is simply used here to refer to a language feature (or a construction) that is generally accepted and widely used within a language community. Language standard may emerge through active language planning in combination with various social forces, such as how the language is written, who is considered a model speaker, who has the power to correct language use, and whether the community generally accepts relevant linguistic features (e.g., Haugen 1966; Ammon 2015; Árnason 2003; Kristinsson 2019). A language standard may be strengthened by official guidelines on how the language should be written and what is considered "good" or "appropriate" language use.

their own mental grammar and produces an output that ends up being a part of all attested utterances in the community. Keeping track of individuals as individuals can be quite challenging. One might want to document who contributes to attested utterances at any given time, noting the appearance of new individuals and the disappearance of older individuals. One might also want to keep a track on how much language data each person contributes, and whether the contributions stay consistent over time or change. For language forecasting, focusing on individuals as individuals in this manner appears to give rise to unnecessary complexity. Not only would one need to take into consideration linguistic variants at different times, but also who is producing those variants. Additionally, linking individuals back to innovation, it is likely difficult to accurately predict when exactly a particular linguistic feature that has not been observed before will be attested.

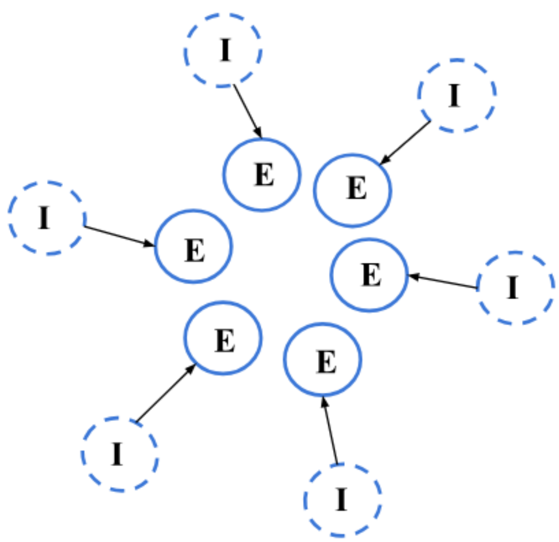


Figure 3.2. A snapshot of a language community at a particular time. Individuals, each with their own I-language, contribute to attested utterances (E-language). Tracking what each individual (or a certain individual) within the community does or can do at each given time gives rise to complexities in tracking features diachronically.

Despite individuals being an intrinsic part of the community, it is not necessary to always take them explicitly into account. One can abstract away from the single person and learn something about the status of the community, or a particular linguistic feature within the community, without needing to model language learning and transmission from an individual to individual and without knowing what a particular individual does. Treating individuals as part of a group makes them conveniently, although perhaps somewhat unrealistically, uniform. However, the "individuality" of speakers may show up in the form of how much variation is attested. Figure 3.2 shows a snapshot of a hypothetical language community at a particular point in time. Note that this approach does not care about what each individual does at the relevant point in time, but rather what the community does.

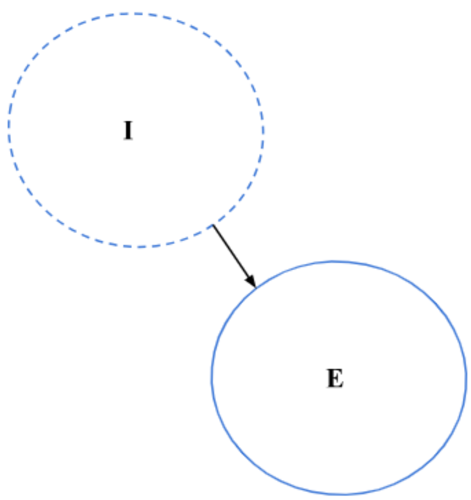


Figure 3.3. A snapshot of the state of a language at a particular time. By treating individuals as a part of a community, it is possible to ask questions about either E-language or I-language as a whole without tracking specific speakers. Thus, one might ask about the proportion of innovative variants in relation to all relevant variants (E-language focused) or about the proportion of speakers that allow for the innovative variant (I-language oriented).

One might think about change on a population level in a metaphorical way as tracking the temperature in a particular place over an extended period of time where the temperature represents a linguistic feature that is being measured and the various surfaces that give rise to the temperature are treated as individuals. There is no need to know how much each type of surface contributes to the overall temperature at the place in order to learn something about how the temperature changes over time.[17] It can be done, but it is not necessary. To take an authentic example, one might think of the spread of COVID-19 in 2020-2021. To learn something about the virus and how quickly it was transmitted, it was not necessary to model the transmission between individuals specially. It sufficed to track the number of cases each day over several days to gain some understanding of the virus' behavior and compute the *effective reproductive number*. In a similar way, one may learn something about language change and the propagation of change through tracking and predicting the trajectory it may take in a population of speakers.

It should be noted that Figure 3.3 is by no means a novel way for viewing change. When tracking changes that have taken place in the past, it is common to gather data that focuses on the proportion of an innovative variant versus all attested variants of interest. What is gained by explicitly stating the situation in Figure 3.3 is a more concrete definition and formalization of the forecasting task. It allows for putting snapshots of the language community into the perspective of a present state of a language, a past state and a future state, cf. Figure 3.4. In other words, we now have a better idea of how it is possible to move between points in time, from one state to the next (past - present - future). The past is

[17] This is not a claim about individuals or surface type being irrelevant or uninteresting for prediction. In fact, surface type and size is factored in in some forecasting models when predicting temperature. The moral of the story is that it is possible to learn about patterns in historical (and future) data by simply studying consecutive measurements without taking explanatory factors into consideration.

connected to the present and the present to the future. Needless to say, similar types of questions can be asked about each of these states. For instance, one may ask what the proportion of innovative variants (propagation) is in attested linguistic data (E-language), documenting the general pattern of propagation through time.

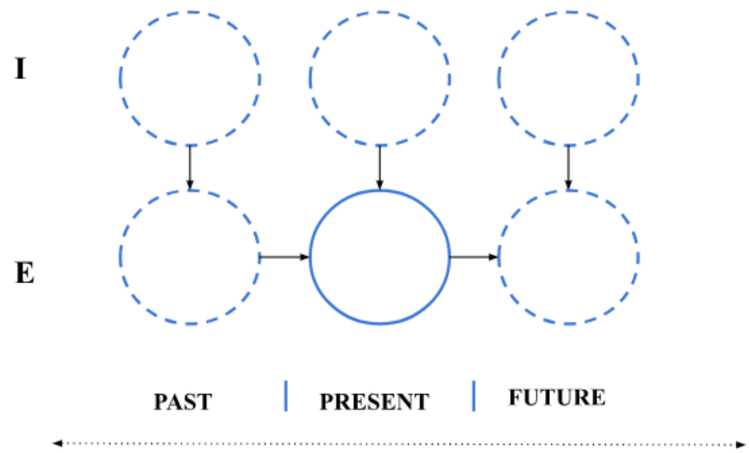


Figure 3.4. Snapshot of states of a language at a past, present and future time. Just like the past is connected to the present, so is the present connected to the future. Questions may be asked about each of these states.

Of course, Figure 3.4 offers a somewhat simplified view of various states of a language. For instance, the past is usually not treated as a single uniform "past-cloud" as in Figure 3.4. Rather, one can assume multiple pasts that are sequenced in time. Similarly, the future can be treated as consisting of multiple future periods sequenced in time. In Figure 3.5 there is a single present state of a language, two past states (P1 and P2) and two future states (F1 and F2). Just like with the previous figure (Figure 3.4) it is possible to move between adjacent states and study how one state is related to the next state.

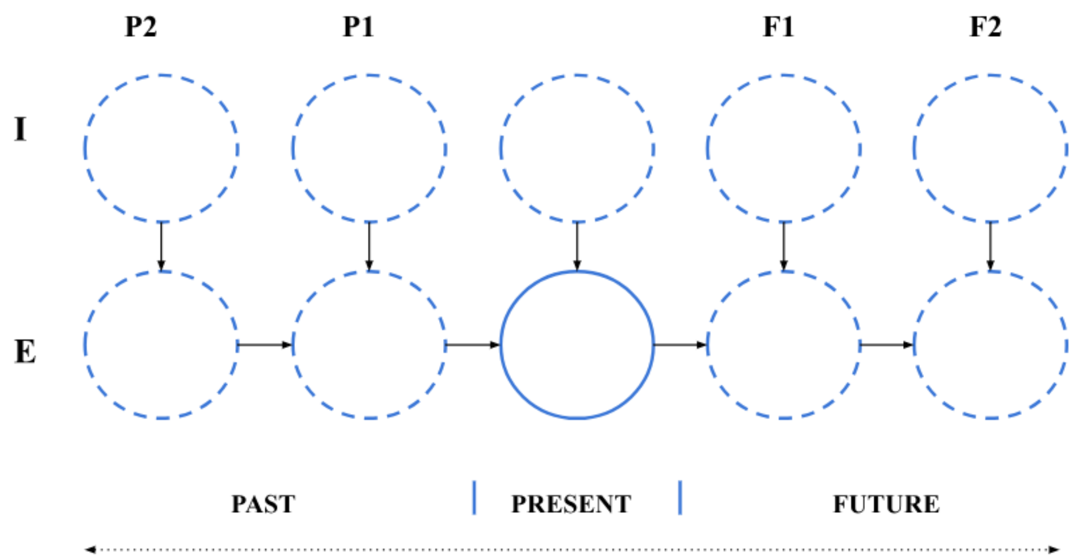


Figure 3.5. Snapshots of a language at different times. The past is not a single past, but multiple pasts sequenced in time. Similarly, the future can also be thought to consist of multiple futures that are sequenced in time.

Figure 3.5 is different from models describing transmission of language such as the one in Figure 3.1 above. Instead of focusing on transmission of language between generations of speakers, Figure 3.5 shows how various states of a language are connected in time. The states are in the form of “slices of language” at different times and they are completely independent of transmission. Past 1 (P1) could be Icelandic as it was on January 21st 2024 while Past 2 (P2) might be Icelandic on January 20th 2024. Arrows show the connection between the collection of I-languages at the relevant time to attested utterances, and how individual states are related, i.e., P2 is related to P1. The relationship between E-language of different states is meant to highlight how there can be no pure I-language approach to diachrony (cf. Walkden on syntax 2014:31, 37) and that E-language represents continuity in language transmission (Plank 2015:65). Thus, it is only possible to move between different points of time through E-language.

Summarizing the points discussed above, language forecasting might be concerned with changes on an individual level or changes on a population level. Focusing on the individual can be related to asking questions about innovation and predicting when in the future and under what circumstances a new variant will be attested. Although this is possible, a more feasible task might be to make predictions on language on the population level and the propagation of language change. Under this scenario, one can start with already attested variation and ask questions about the variation at different times. The different times are represented by different states of a language which are connected in a particular way through time. A good reason for focusing on E-language rather than I-language is that data is more readily available, and predictions can be more easily verified.

3.4 Layers of language change and what can be predicted

In theory, any aspect of language change – or any phenomenon in language – can be forecasted, leaving aside whether the forecast turns out to be correct or not. However, choosing to focus on one aspect over another may lead to a more or less complex forecasting task. It does not help that various aspects of language, as well as language transmission and change, tend to be interrelated. So, how does one choose *what* to forecast? Should something be said about a language as a whole, its status and future, or should the focus be on a smaller aspect of the language, perhaps the nature and development of a particular phenomenon? If the latter is chosen, does it involve instantiation of a linguistic feature in a certain environment or in various related but slightly different environments? As the following discussion suggests, it may be useful to distinguish between predictions about innovation, propagation, extension and typological aspects of a language, as well as the overall fate of a language. All of these can serve as potential topics for forecasting, and they call for slightly different approaches.

The distinction between innovation and propagation was already brought up in relation to the concept of *change* (see 3.3). Both innovation and propagation can be linked to populations of speakers, with innovation referring to the first time something is manifested within a population and propagation to how a community of speakers eventually converges on using (or eliminating) a particular feature. Although the distinction may appear clear-cut, the boundary between the two becomes fuzzy when viewed up close. Despite one easily morphing into the other in certain circumstances, there may still be a reason to try and keep the concepts separated, at least on some level as what exactly is being predicted may influence the complexity of the forecasting task. It might also be practical to be aware of the boundaries between propagation, extension and typological aspects of a language.

As already established (Section 3.3), ***innovation*** refers to a feature or a structure that was not previously present in a language. Under this definition, predicting innovation must be regarded as the most challenging and complex forecasting task imaginable; It involves foreseeing a change that has not been actualized. Instead of deeming the task impossible, it may be helpful to view the problem in light of what is known about impossible, possible and likely changes in language. To start with, language does not change randomly. The current state of a language (whatever it may be) necessarily imposes restrictions on which changes can take place (Wedel 2015). When predicting innovation, one may wish to assume that whichever change takes place, the change is conditioned by the linguistic environment or in the way individuals may perceive language data they are exposed to. Put more simply, it is necessary to assume that nothing emerges from nothing.

Exemplifying predictions about innovation, one might consider the expletive element *það* in Icelandic which is homonymous with the third person singular neuter pronoun *það* 'it'. Predicting its emergence would mean predicting non-referential use of *það*, as in (3.6), before such use was attested. For a speaker of Icelandic in the mid-15th century who was only familiar with referential uses of *það*, this might have been a difficult task. The speaker, always interpreting *það* as a referential pronoun, would need to understand the possibility of a novel use before such use was attested. They would furthermore need to be able to say at which time a new use might emerge. Expletive-like contexts as in (3.5), where reinterpretation is possible in light of ambiguity in the signal, might provide a hint about future existence of non-referential use of *það,* but it does not necessarily provide a clue about when such use will emerge.[18]

(3.5)	***Það***	*var*	*eitt*	*sinn*	*er*	*hún*	*reið*	***að***	***vanir***	***nokkrir***
	it	was	one	time	when	she	rode	that	vanir	some
	sá	***reið***	***hennar***		***í***	***loftinu.***				
	saw	riding	hers		in	the.air				

'One time it occurred, when she was riding, that some vanir saw her riding in the air.' (*Snorra Edda* (Gylfaginning), Ch. 35)

[18] Rosenkvist (2023) has argued, based on innovative usage of Swedish *fortsatt* as clause adverbial, that increase in the usage of ambiguous syntactic structure leads to higher probability of the emergence of new non-ambiguous structures. Once innovative structures have emerged, the use of ambiguous structures may decrease. Currently, the proportion of ambiguous structures required for innovation to appear (the "tipping point") is not known.

Similarly, it is difficult to predict *if* or *when* examples where *það* inverts with a finite verb (3.6b) could become grammatical in Modern Icelandic. One reason to think it might become grammatical at some point is that it would mirror a development that has already taken place (or is taking place) in other languages, for instance English and some varieties of Dutch.[19]

(3.6) a.

Það	*kom*	*til*	*mín*	*draugur*	í gær.
EXPL	came	to	me	ghost	yesterday

'A ghost came to me yesterday.'

b.

Í gær	*kom*	(****það***)	til	mín	*draugur*.
yesterday	came	EXPL	to	me	ghost

'Yesterday a ghost came to me.'

Although forecasting innovation may prove somewhat difficult, it is not necessarily impossible. As long as there is some understanding of where to look for potential changes and how to "read the signs", some changes may likely be accurately predicted. The problem with predicting innovations is thus not due to the nature of innovations, but rather due to poor understanding of how such a forecasting task might be approached.

[19] The difference between Icelandic *það* and expletives in other languages is that Icelandic *það* is restricted to initial position. In a sense it is like a placeholder or a filler for the left edge. Expletives in other languages may occur clause-internally and be interpreted as occupying a position often tied to subjects. This applies to e.g., *there* and *it* in English or *er* or *het* in Dutch (Zwart 2011:18–19).

Predicting ***propagation*** may constitute a somewhat more straightforward forecasting task than attempting to foresee innovation. Instead of needing to identify a potential change before it takes place, it is possible to start from a known linguistic variation and make predictions about the proportion of the innovative variant in the future. As an example, an age-related variation in the use of expletive *það* in the context of adverbial clauses with a subject gap has been documented in Icelandic (Angantýsson 2011). Angantýsson (2011), notes that for some types of embedded clauses, namely temporal *þegar* 'when' clauses, younger speakers accept *það*-insertion (as in (3.7) more readily (84.9%) than older speakers (67.7%). Angantýsson (2011:155) notes that this may suggest a development towards "increased use of the expletive in Icelandic". A prediction about propagation might involve trying to foresee what proportion of young speakers in 2050 will find utterances such as (3.7) acceptable.

(3.7)	*Þær*	*verða*	*opnaðar*	*þegar*	***það***	*fer*	*að*	*snjóa*
	they	will	opened	when	EXPL	starts	to	snow

'They will be opened when it starts to snow.'

(cf. Angantýsson 2011:155)

Note that if propagation is defined as a change in the proportion of speakers using a particular linguistic feature, the distinction between the transmission of a feature from one individual to another via E-language and multiple independent innovations is neutralized. In case two or more individuals start producing structures that were previously not a part of their linguistic environment, their adoption of a new variant can be viewed as multiple independent instances of innovation, provided the individuals had no interactions with each other and thus did not transmit the structure.[20] Although it may be theoretically interesting to distinguish between the two (multiple innovations and transmission), it is not easy to keep these apart in practice. In fact, the distinction might be argued to be irrelevant if the goal is simply to track the number of speakers using an older and newer variant.

In addition to innovation and propagation, there may be reasons to treat the spread of a change to a new environment as a separate type of a forecastable change, termed here ***extension***, although it might also be added under ***lexical diffusion*** or ***analogy***.[21] Like with propagation, extension arguably involves a feature or a structure that is already attested in a language. However, unlike propagation, predicting extension does not involve foreseeing the proportion of individuals that might have the relevant feature in the future. Rather, given that a feature occurs in a certain context, the prediction needs to capture *if* it will start occurring in a new context and *which type* of context that would be. Here we might note that the difference between innovation and extension is blurred as innovation was previously described as being concerned with the introduction of a new type of feature or structure in a certain context. Innovation and extension can still be distinguished. In the context of expletives in Icelandic, innovation can be used to refer to the first time a true expletive (non-referential element filling a particular place in the clause) occurred in a certain environment, for instance in clause-initial position with weather predicates (Rögnvaldsson 2002). Once expletives started occurring in a new environment, e.g., in

[20] Assuming they are unrelated in the sense of not in contact with each other and not moving around in the same "network" of speakers in the population.

[21] Although lexical diffusion and analogy are different types of changes, there is a sense in which they have something in common, i.e., an existing pattern or variant starts appearing in novel contexts. Extension might be taken to be a general term for extending a rule or a feature to novel context.

temporal clauses containing a subject gap as in (Angantýsson 2011), the term extension might be applied.

Naturally, there may be instances where it is difficult to determine whether something constitutes a new context or not. Take for instance subject case marking with the predicate *hlakka* til 'look forward to' in Icelandic. It has been noted that the first person singular is not affected by changes in case marking to the same extent as regular NP subjects (see discussion and references in Chapter 9). Thus, individuals who consistently use the nominative case when the subject is first person singular (*ég hlakka* 'I_{NOM} look forward to') may use other cases with other types of subjects. The question arises whether the first person singular should count as a different context than other types of subjects or if all subjects should be treated together.

When innovation, propagation and extension come together, they may give rise to long-term developments or general trajectories of change that affect aspects of the language system as a whole. In other words, they may lead to ***typological*** changes such as whether a language has prepositions or postpositions, whether a language is verb second or not, or how flexible a language is w.r.t. pro-drop. Thus, multiple related but distinct changes in a single domain might lead to a language becoming typologically different from what it was before.

Finally, language change that occurs on the level of the population does not need to be tied to particular features or aspects of a language. It may also concern the language as a whole, for instance diminishing or altered use of the language or even language death. Unlike innovation, propagation, extension and typological aspects of language, the overall ***fate of a language*** (whether it continues existing or not) is not an add-up of other types of changes. Instead, it is the result of various sociological factors, language transmission, number of living speakers, attitudes towards the language, etc.

Figure 3.6 attempts to capture the relationship between different layers of change. As already noted, there is not a clear distinction between these layers. Extension might be regarded as a type of innovation as it involves a linguistic variant appearing in a novel context. Propagation is either the result of multiple individuals innovating in the same manner or the result of a variant being transmitted in a regular fashion. Once a linguistic variant (or a rule) has become relatively widespread in a certain context (propagation), it may start appearing in a new context (extension). Together, innovation, propagation and extension may contribute towards changes on a typological level of a language.

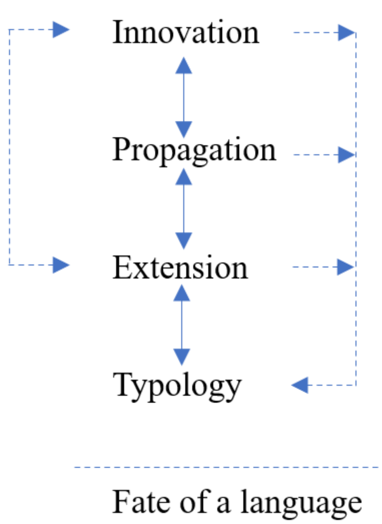


Figure 3.6. Distinguishing between different "layers" of change can be difficult. Innovation, propagation and extension overlap to a certain degree, and they may all contribute towards the typology of a language. The fate of a language does not directly hinge on any of these. Rather, it is linked to the number of speakers and domains of language use, to name only a few factors.

The distinction made above between innovation, propagation, extension, typological aspects, and the fate of a language is mostly done for practical reasons in relation to forecasting. Focusing on one of these rather than another is bound to give rise to different types of predictions which may require different approaches and different data. Furthermore, the awareness of interactions between various layers of change can be informative with respect to the overall direction of change and the situation of a language. A change may start in a particular linguistic context (innovation), and later be generalized to other contexts (extension). Along the way, a community of speakers may converge on using that feature (propagation).

3.5 The route forward

Since language forecasting is in its infancy, anyone who attempts to undertake a formal forecasting task is in a slightly uneasy position. Approaches to the topic are currently underdeveloped and the goals of language forecasting have not been fully fleshed out. On top of this, there is limited understanding of the type of data needed to carry out various types of language forecasting tasks. In this chapter, an attempt has been made to break down how questions related to language change and language forecasting can be formulated and approached. In particular, it was suggested that language forecasting should be concerned with the future of a language and seek answers to the question in (3.1).[22]

(3.1) **Forecasting question** (to be revised in Chapter 4)

What will the situation *x*, of a phenomenon *p* be in the future?

[22] The question presented in (3.1) does not need to only apply to forecasting language ***change***. It may also be applicable for forecasting language ***stability***, a task which is certainly worthwhile.

The question in (3.1) can be understood and approached on multiple levels. It is no doubt beneficial to be as explicit as possible about which levels the focus is on. For instance, whether predictions will be made about language in the mind of individuals (I-language) or language as it appears in the real world (E-language); whether the goal is to tie predictions to individual speakers or to the language community as a whole. Further considerations involve whether predictions should be made about innovations, i.e., the emergence of new linguistic variants, propagation of change, extension of particular features or rules to new contexts, typological changes or the future of a language.

Since language forecasting is an understudied subject, it is practical to make choices that make the forecasting task as straightforward as possible. This involves abstracting away from some of the complexities presented by language, for instance the notion of individual speakers and what they might be able to do in the future. A reasonable starting point might be to focus on the propagation of language change, relying on E-language data and predicting the proportion of attested novel features in the future. Convenient E-language data has the benefit of being relatively easy to access (compared to specialized language data) and predictions made about future E-language are certainly more easily verified than predictions about a future I-language. It is also worth keeping in mind that any claim about attested structures in the future is also a claim about I-languages of that period, i.e., in order for an utterance to be attested there must be a grammar that produced that utterance.

If language is viewed through time in terms of multiple "states" that are sequentially ordered in time (Figure 3.5), one may say that the forecasting task involves claiming something about a future state of a language. This may be done by using information from past states of a language, assuming that each state is a logical continuation of a previous state. Treating individuals as part of a community, there is no need to worry about what specific speakers might do. Rather, questions are asked about the language community as a whole. Tying this back to the idea that language forecasting can be used as a method to study change, the claim is that something can be learned about the propagation of change by attempting to predict a future distribution of a linguistic variant. The prediction can be treated as a testable hypothesis, i.e., given what is known about present and past states of a language, certain developments may be expected in the future and these expectations can be formalized and (at a later point in time) verified. While making predictions about the distribution of innovative variants in convenient E-language in the future may be likened to predicting human behavior, it is important to note that linguistic behavior plays a crucial role in language transmission and the construction of abstract mental grammars. In short, the relationship between I-language and E-language is not random.

4 From general predictions to systematic forecasting

4.1 Approaches to forecasting

Various approaches exist for forecasting across domains that use them. While there are no right or wrong ways to approach a particular forecasting task, there may be more or less appropriate ways of doing so. In short, not all forecasts are equal. This applies to all kinds of forecasting whether predictions are related to weather, economics, birth rate, epidemiology or even language. Anyone can make a prediction, but not all predictions are necessarily scientifically useful.[23] There is a difference between looking out the window, claiming "It will probably rain later today" and relying on an official weather forecast that predicts rain in the afternoon. Both predictions may turn out to be correct. However, only one of them involves a systematic approach such that it uses previously gathered data, applies a formal forecasting model, and can be reevaluated. In the case it does not rain in the afternoon, a forecast generated with a forecasting model can be revised to figure out *what* went wrong *where* and possibly *why*. Reevaluating intuitive forecasts, on the other hand, is more difficult.

Similar to the different approaches to weather predictions, language forecasting can be approached in various ways, with methods ranging from informal to more systematic and scientific. The approach that is chosen must be in accordance with the objectives of the forecasting while also aligning with the type of available data. A further factor that

[23] While all forecasts have the potential to be useful, not all forecasts lead to increased knowledge about the phenomenon of interest. As noted earlier (Chapter 2), one of the goals of language forecasting is to learn more about language change and therefore methods used for forecasting must be in line with the goal.

needs to be considered is how far into the future predictions are to be made. A distinction is typically made between short-, medium- and long-range forecasts, where the definitions are mostly based on the accuracy of the forecasts but may also depend on the general methods used. In the case of weather forecasting, short-range forecasts tend to cover about 6 hours to a few days, medium-range from about 3 to 8.5 days and long-range anything above 8.5 days (Ahrens 2007:347). Since language is linked to individuals who live for many years, it seems natural to assume that definitions of forecast range should be based on years, even decades, instead of days. This is further discussed below.

Turning back to forecasting methods, Castle, Clements & Hendry (2019:22), who discuss the fundamentals of forecasting in general, briefly describe seven commonly used methods.[24] Four of these are noted in (4.1), giving an idea of the range of available approaches, with the leftmost approach being the simplest and least quantifiable and the rightmost the most formal and complex one (see also Hyndman & Athanasopoulos 2018:12–14 and Makridakis, Wheelwright & Hyndman 2018:9 on the continuum between "intuitive or ad hoc methods, and formal quantitative methods based on statistical principles").

(4.1) Guessing : : naïve extrapolation : : simple models : : formal forecasting systems

The most simple method, guessing, will always be hard to evaluate as it does not contain any formal components leaving the way in which predictions are reached is unaccounted

[24] The methods Castle, Clements & Hendry (2019:22) mention are: i) Guessing, ii) 'Rules of thumb', iii) Naive extrapolation, iv) 'Leading indicators', v) Surveys of intentions and expectations, vi) Simple models and, vii) Formal forecasting systems. The list in (4.1) is only meant to give an idea about the range of methods from informal to formal.

for. Although slightly more formal than guessing, naïve extrapolation involves stipulating the continuation of a general trend into the future. This can be done either through a type of guesstimation or by using a simple trend line. Similar to naïve extrapolation although slightly more formal, simple models aim to capture regularities in the data and extrapolate patterns into the future. Simple models tend to assume that the phenomenon of interest behaves in a fairly regular way, so sudden or extreme developments may prove hard to predict with this type of approach. Finally, formal forecasting systems may incorporate multiple different assumptions and may sometimes attempt to explicitly model causal relationships (Castle, Clements & Hendry 2019:22)

In the following sections, some forecast-oriented work in linguistics is reviewed in relation to intuitive methods and formal approaches, and in light of how far into the future predictions are made (if predictions are made at all). Since language forecasting can be thought of as a tool to learn about language change and the process of predicting change (see Chapter 2), formal approaches are argued to be more appropriate and useful than less formal ones. In short, methods need to fulfill the condition in (4.2).

(4.2) The methods used to produce a forecast need to be formal, systematic and allow for reevaluation.

In addition to the forecasting methods being systematic and allowing for reevaluation, it is claimed that the forecast horizon, i.e., how far into the future predictions are to be made, should not stretch over multiple decades. Rather, forecasts should be concerned with the not-too-distant future. Making predictions a few years into the future involves incorporating time in a very explicit way and allows for predictions to be verified without having to wait for several decades.

The structure of the Chapter is as follows. Intuitive forecasting and naïve extrapolation are discussed in Section 4.2. In Section 4.3, more systematic approaches to studying language change are reviewed. Finally, Section 4.4 provides an outlook on the future of language forecasting, emphasizing the importance of time and further explaining why short- to mid-range language forecasting is considered practical.

4.2 Intuitive forecasting and naïve extrapolation

Interest in predicting the future of a given language, or a phenomenon within the language, has existed for a long time. Although some general predictions have been made, these are typically not based on systematic forecasting approaches. Rather, they are better described as "armchair predictions", that may fall under *intuitive forecasting* or *naïve extrapolations*.

The most famous intuitive forecast made for Icelandic is that of the Danish linguist Rasmus Christian Rask (b. 1787 d. 1832). Rask, believing Icelandic to be a rather conservative language, was convinced that Icelandic preserved many features of an older common nordic language. This was unlike Danish which had lost most of its inflection as well as the three-category gender system. While working on his essay on the origin of the Old Nordic or the Icelandic language (Rask 1818), he visited Iceland in 1813–1815. During his time there, Rask discovered to his dismay that Danish was spoken quite widely in Reykjavík.[25] In the countryside, however, the language still prevailed and Rask was able to learn enough to be able to both read and write proficiently. He wrote to one of his friends,

[25] Iceland was under Danish rule at the time and the capital of Iceland was Copenhagen.

Bjarni Thorsteinsson, expressing his worries about the status of Icelandic. He went on to predict a rather bleak future for the language, claiming that Icelandic would not be spoken in Reykjavík in 100 years, and in 200 years from that point hardly anyone in the country would understand it (Rask 1813, August 30th):

> Annars þèr einlægliga ad segia held ég ad íslendskan brádum mun útaf deyia, reikna eg ad valla mun nockur skilia hana í Reikiavík ad 100 árum lidnum, enn valla nockur í landinu ad ödrum 200 þaruppfrá, ef alt fer eins og híngad til og ecki verda rammar skordur vidreistar, jafvel hiá bestu mönnum er annadhvört ord á dönsku, hiá almúganum mun hún haldast vid leingst.
>
> In all honesty I tell you that I think Icelandic will soon become extinct. I suspect that hardly anyone in Reykjavík will understand it in 100 years and in 200 years thereafter hardly anyone in the country if things continue to develop the way they have, and nothing is done to hinder the development. Even among the best men every other word is in Danish. Among the general population it [Icelandic] will survive the longest.

Rask's 1813 prediction is an example of a rather general statement about the future of a language. It falls in the category of intuitive forecasting where no explicit model is used. Although the name may suggest that intuitive predictions are not based on meaningful knowledge, observations, or prior experience, this is not the case. These types of forecasts are not born out of nothing. In Rask's case, Rask had hands-on experience with the situation in Iceland when he was visiting. Observing that Danish was widely spoken in Reykjavík, he must have realized that with decreased usage of Icelandic, transmission of the language to the next generation was threatened. The fact that Danish was not as dominant outside of Reykjavík gave some hope that Icelandic would survive for longer there. Thus, the two-step language death was hypothesized, with Icelandic first disappearing in Reykjavík and then in the rest of the country. The rough time frame of 300 years (first 100 years and then

200 years) may have been linked to an idea of ‘generations of speakers’, although this is not clear. It could also be that these simply signaled a future that was somewhat far away but still within a perceivable distance.

The downside to Rask’s forecast (and to intuitive forecasts in general) is that it does not explicitly state which factors contribute to the prediction or to what degree these factors might affect the outcome. Forecasts of this type are therefore very hard to evaluate, and they provide little insights into language and language forecasting. For Rask, this was not a problem since the goal was never to learn about forecasting or language change. Rather, his prediction must be viewed in light of language purism and nationalistic movements within Iceland at the time.

The 19th century can be claimed to mark the beginning of a serious language purism in Iceland. Previously, there had been talk about Icelandic being worth preserving as it had changed very little from medieval times. In the 19th century, prominent individuals advocated for language purism and actively supported language planning (for an overview and discussion see Hilmarsson-Dunn & Árnason 2010; Leonard & Árnason 2011). Initially, the intention was to eradicate foreign words and influence. With Rask’s prediction, the focus was shifted from simply keeping the language “pure” to emphasizing the survival of the language. In other words, it functioned like a call for action to do something about the status of Icelandic. Rask himself played a key role there. Not only did he encourage friends in Iceland to respond by translating foreign material into Icelandic, he also participated in establishing the Icelandic Literary Society (Ice. *Hið íslenska bókmenntafélag*) whose goal was to support Icelandic literature, education, and the Icelandic language. Due to Rask’s efforts and actions taken to preserve and improve the status of the language, the imminent death of Icelandic was avoided. Even now (in 2024) more than 200 years later, Icelandic is still spoken in every part of the country, including in the capital city, Reykjavík. The question still remains what would have happened had the course not been reverted. Would Rask have been correct about Icelandic first disappearing in Reykjavík and then in the rest of the country? Would the disappearance of Icelandic have followed the timeline that Rask proposed? This is hard to estimate. One thing is clear, however. There were high hopes for the forecast to be wrong. As such, the intuitive forecasting method can be claimed to have been appropriate for the forecasting situation.

Forecasting situations where the future of a language needs to be evaluated with respect to whether the language is under threat of extinction or not occur regularly. The goal is usually to evaluate how endangered a language is and how to respond to the situation. For these purposes intuitive forecasting can certainly be used. Another option for assessing the situation is to use a slightly more formal method, for instance the one proposed by the UNESCO Ad Hoc Expert Group on Endangered languages (Brenzinger et al. 2003). Under that approach, the vitality of a language is evaluated based on factors such as in shifts in domains of language use, language attitude and policies, absolute numbers of speakers and whether the language is still being acquired by new speakers, particularly children. Each category is assigned a grade based on how well the language does in the relevant area. A language that scores very low in all categories is considered to be gravely endangered (Brenzinger et al. 2003; for a similar method with a scale of 13 levels see Lewis & Simons 2010). The method proposed by the UNESCO group provides a more systematic approach to estimating the vitality of a language than the intuitive method used by Rask.

However, it lacks explicit reference to time such as how long it will take for the language to move between a lower state of endangerment to higher one if nothing is done to reverse the trajectory of change.[26]

Another example of a prediction based on intuition can be found in the work of Sapir where the use of *whom*, an oblique form of *who*, is discussed (Sapir 1941: 156, 162). Being morphologically marked as an oblique form, *whom* is generally expected to appear in a position where another element assigns (or has assigned) case to it. For instance, it may appear in sentences such as ***Whom*** *did you see?* where *whom* is understood as the complement of *see* and in sentences such as *The man* ***whom*** *I referred to* where the preposition *to* is responsible for the case of *whom*. Since morphological case marking has been eroding away in English, it is unsurprising that *whom* is often replaced by the more unmarked form, *who*. Sapir, being aware of this tendency, hypothesized that *whom* would ultimately disappear from the language. The form would be gone in around a couple of hundred years from the time he was writing (Sapir 1941:156):

> It is safe to prophesy that within a couple of hundred years from to-day not even the most learned jurist will be saying "Whom did you see?" By that time the "whom" will be as delightfully archaic as the Elizabethan "his" for "its." No logical or historical argument will avail to save this hapless "whom".

Interestingly, Sapir predicted that *whom* would not cease to be used in all contexts at once. Rather, "locutions of the type *Whom did you see?* will be obsolete when phrases like *The man whom I referred to* are still in lingering use" (Sapir 1941:162). The assumption of

[26] Although no concrete length of time is provided, it is mentioned that languages can be in danger of being lost "in a short period of time" (Brenzinger et al. 2003:18).

stepwise disappearance of *whom* was based on the form occurring in various types of sentences and appearing in different positions within them. For instance, Sapir noted that using a morphologically marked (oblique) element at the beginning of the sentence was highly unusual.

While the predictions made by Rask and Sapir are in themselves interesting and can be informative (especially if the goal is to prevent language death), they do not meet the requirements laid out in Chapter 2 where language forecasting was viewed as a method to study language change. Although Rask's and Sapir's forecasts are based on experience and intuition, in neither case is the input data for creating the forecast explicitly encountered for. Moreover, the timeline for the expected development in both cases is relatively vague. Rask estimates 100 years followed by another 200 years and Sapir estimates a couple of hundred years. While these might be understood in absolute terms, it seems more likely they represent a semi-vague point in the foreseeable future. For predictions to be verifiable and their accuracy to be checked it is better to take the temporal element more seriously, for instance by treating it in a concrete way and by making predictions that are not so far off into the future.

4.3 Formal methods

Intuitive forecasting and naïve extrapolations, discussed in the previous section (Section 4.2), can be contrasted with more systematic methods that involve quantitative approaches and formal forecasting models. A distinction is generally made between explanatory models that take into account factors that stand in a causal relationship with what is being forecasted, and non-explanatory models that extrapolate from existing historical data without modeling a cause-and-effect relationship (e.g., Makridakis & Wheelwright 1978). The models can vary in both structure and complexity.

It should be noted that forecasts produced relying on formal forecasting models are not always accurate.[27] In some cases, formal forecasting models may yield less accurate predictions than intuitive or judgmental forecasts. However, the benefits of using formal approaches are several. The input data used for predictions needs to be accounted for and the models used to generate the forecasts must be well-defined. Provided this is the case, models and the forecasts can be validated, and both systematic and nonsystematic "errors" can be identified and dealt with. Thus, it is possible to learn something about the process of forecasting and the data needed to make the forecasts; it allows for figuring out *what* went wrong in the prediction, *where*, and *why*.

Within linguistics, few truly systematic attempts have been made at forecasting language change (for some discussion see Sanchez-Stockhammer 2015). This may be partially attributed to forecasting having been viewed pessimistically and deemed theoretically impossible (Chapter 2). Nevertheless, there is work in some areas that can be claimed to be both relevant and related to forecasting. As it is neither feasible nor possible to provide an overview of all work within historical linguistics that may relate to forecasting, only a few are mentioned here. These involve systematic approaches that incorporate, on some level, predictions about the trajectory of certain changes in relation to time, either by focusing on changes that are already completed or by looking at ongoing changes and their trajectory.

[27] Forecast accuracy depends on how far from an observed value a forecasted value is and whether or not the observed value is within estimated prediction intervals. There are various reasons for why a forecast may turn out to be inaccurate or incorrect. There may have been some issues with the data, the modeling may have been off, or, alternatively, something unexpected may have occurred that affected how events unfolded.

Work within linguistics that comes close to forecasting without generating actual forecasts includes sociolinguistic studies relying on real- and apparent-time studies (cf. Labov 1963, 1966; Bailey et al. 1991; Bowie 2005; Cukor-Avila & Bailey 2013). The methods applied in apparent time studies are often attributed to Labov (1963, 1966) who researched ongoing language change in areas such as Martha's Vineyard in the sixties. In addition to focusing on variation in well-defined areas, Labov also considered various demographic and social factors, including the age of speakers. Recording the age of speakers is considered particularly important as apparent time studies assume that trajectories of change may be studied synchronically by referring to linguistic variation in relation to speaker's age. Typically, results are grouped based on participants' age, making it easier to contrast what older speakers do with what younger speakers do. The distribution of age-related variation indicates how a change is propagating. If older speakers use or accept a linguistic variant to a higher degree than younger speakers, one can infer a "downward" trajectory of change, where the variant under discussion is losing ground over time. Importantly, this conclusion can only be reached provided that individual speakers are assumed not to change their language in any important way over time (e.g., Bailey et al. 1991:242; Chambers and Trudgill 1980:165). A group of speakers aged 40 who accept a particular linguistic variant at a very high degree is assumed to have accepted it at the same rate twenty years earlier when they were 20, thus being "comparable for diffusion research to the speech of 20-year-olds today " (Chambers and Trudgill 1980:165).

An example of a trajectory of change inferred from an apparent time study involves the use of indefinite nouns in possessive constructions in Icelandic which tend to be replaced by a definite noun by younger speakers, for example *hár hennar* 'her hair (indefinite)’ which becomes *hárið hennar* ‘her hair (definite)’. This particular type of variation was amongst phenomena tested in the project *Variation in Icelandic Syntax* which was active in the years 2005–2007 (Thráinsson, Angantýsson, Sigurðsson, Steingrímsdóttir & Eythórsson 2013:40). A target sentence with an indefinite noun in the dative case, *úlpuvasi* ‘coat pocket’, is provided in (4.3). Participants were asked to judge whether this was a grammatical sentence or not. The results for each age group that was tested is shown in Table 4.1.

(4.3)	Context:	*Vala*	*fann*	*farsímann*	*sinn*	*eftir*	*langa*	*leit*.
		Vala	found	cellphone	her	after	long	search
		‘Vala found her cellphone after a long search’						
	Target:	*Hann*	*var*	*í*	***úlpuvasa***	*hennar*.		
		he	was	in	coat.pocket	her		
		‘It was in her coat pocket.’						

(Thráinsson et al. 2013:40, test sentence T2040)

Target sentence: *Hann var í* ***úlpuvasa*** *hennar*.				
Age group	**9. Grade**	**20–25**	**40–45**	**65–70**
Deemed grammatical by:	48.50%	45.30%	77.40%	96.90%

Table 4.1. Proportion of speakers in each age group that deemed the target sentence grammatical in the project *Variation in Icelandic Syntax* which was active in the years 2005–2007 (Thráinsson et al. 2013:40). The 9th grade speakers were aged 15–16.

The results from the survey suggest that indefinite nouns in possessive constructions are becoming less acceptable. The results of each age-based category can be projected both backward and forward in time. Taking the oldest group as an example when projecting back in time, individuals who belong to the age category 65–70 in ca. 2005 would have been around 40–45 years old in 1980. Had a similar survey been conducted then, the results should have shown that 96.9% of individuals aged 40–45 would accept the target sentence in (4.3). Projecting 20 years forward in time, to the year 2025, the expectation is that about 45.3% of individuals between the age 40–45 will accept the target sentence. In this way, the apparent time studies provide implicit predictions about distribution of linguistic variants in the past and the future. Note, however, that predictions for future distribution are not always explicitly pointed out. Rather, the characteristics of each age category w.r.t. the linguistic variants under discussion are used to get a sense of the trajectory of the change over time, often in relation to sociolinguistic factors.

Although apparent time studies are often used in historical sociolinguistics, the methods are not free from problematic assumptions (e.g., Cukor-Avila and Bailey 2013; Bowie 2005). One issue that comes up is that the method implicitly (and in some cases explicitly) assumes that an individual's linguistics system, both the abstract I-language and the language use, remains stable from the individual’s teen years and onwards. Although this may be true for some speakers and, perhaps, certain linguistic features, this is not universal (Chambers & Trudgill 1980:165–166; Bailey et al. 1991:242–243; see also overview of longitudinal and apparent time studies in Sankoff 2013:261–279). Due to individuals’ linguistic systems not staying stable over extended periods of time, it is not possible to infer anything definite about age-related distribution of linguistic variants from a single study. As Bowie (2005:57) notes “... an apparent time analysis conducted in 2005 cannot reliably tell us about the precise state of a speech community in 1975 and 1945”.

Applying this to the case of indefinite nouns in possessive constructions in Icelandic noted above, the conclusion is that the study from 2005 does not provide reliable information on the distribution of acceptability rates in 2025, nor how they might have been in 1980. Essentially, the goal of apparent time studies is not to make predictions on what the situation will be in some years, or decades or what it must have been at a previous point in time. Rather, they hypothesize about trajectories of propagation of change in the context of age-grading and which social factors (class, education etc.) are relevant for those changes.

Another line of work focusing on language change and predictions makes use of so-called hindcasting or retro predicting to estimate parts of the trajectory of changes that have already taken place. Although this type of work looks promising for forecasting, it has so far not gained much attention. Using hindcasting, Van de Velde (2017) shows how S-curves can be fitted to historical data and used to predict observations not taken into account in the model fitting. S-curves have been noted to show up time over time when the propagation of change is examined (Weinreich et al. 1968:113; Bailey 1973; Kroch 1989; Denison 2003; Sanches-Stockhammer 2015; Nevalainen 2015; Pintzuk, Taylor & Warner 2017:221). The s-curve emerges due to new variants initially having a slow uptake, then being rapidly adopted by the language community and finally slowing down due some individuals not having picked the new variant up (Osgood & Seboek 1954:155). By fitting an s-type curve to historical data that covers roughly 100 years, i.e., from ca. 1830 to 1940, Van de Velde (2017) was able to predict the situation of the propagation of the change in the year 2000 relatively accurately. The results suggest that s-curves may be used to predict future trajectories. However, at least three things may be noted about the method. First, that the data used to fit the model must show when the change starts to take off, i.e., it is not enough to only have information about the propagation in its earlier stages (Van de Velde p.c.). Second, the historical data used to fit the S-curve model might need to cover a somewhat lengthy period in order to show when the propagation starts taking off. Of course this depends on the exact change under investigation. In the case of the change discussed by Van de Velde (2017), the period is roughly 100 years with predictions focusing on a time period 50–60 years after the historical data. Third, while relying on s-curves to predict the "completion" of a change may give relatively accurate results, the question is whether they can also accurately predict the situation closer in time. Future work may have answers to this question.

A second example of retro prediction is found in the work of Berdicevski, Coussé, Kopeling & Adesam (2024) who use both logistic regression and ARIMA models (on ARIMA models see also Van de Velde & Petré 20202 and Chapter 7 of this dissertation) to predict the proportion of examples lacking an infinitival marker in Swedish future constructions in data from various corpora. While some of the models Berdicevski et al. use take into account language internal predictors, others only rely on patterns in the historical data over time. The authors do not produce forecasts for future time periods, but rather predict already attested part of the data. Interestingly, they use monthly observations which means that one step ahead predicts one month ahead. Despite this they find changes in the patterns in various sources over time. They note that they cannot "reliably predict the presence or absence of *att* in an individual utterance and the proportion of *att*-omission in a given corpus in a given period of time" (Berdicevski et al. 2024:30). One of the questions the work of Berdicevski et al. (2024:30) raises is of course whether monthly observations should be used or not. A second question is how far into the future predictions can be made using monthly data based on how forecasting models work. Time series, time series analysis, ARIMA models and predictions for the future are further discussed in chapters 6, 7, 8 and 9.

In addition to work on retro predicting in linguistics, there have also been attempts to predict the situation of particular linguistic variants in the future. For instance, attempting to quantify the dynamics of language evolution, Lieberman et al. (2007) study how irregular verbs have been regularized throughout the history of English. They look at 177 predicates that remain a part of modern English, all of which were irregular in Old English. In Middle English, 145 of them remained regular and in Present Day English only 98 continue to be irregular. The other 79 predicates have regularized over time and now form a past tense using the dental suffix *-ed*. Using a model that takes into account the frequency of individual predicates Lieberman et al. estimate the regularization rate of irregular predicates, showing low-frequency predicates regularizing faster. Their modeling approach allows for predictions both backward and forward in time. They venture on to make a prediction for the future, saying that if "the current trends continue, only 83 of the 177 verbs studied will be irregular in 2500" (Lieberman et al. 2007:715).

Two remarks can be made about the approach and prediction of Lieberman et al. (2007). First, their prediction does not focus on propagation of change in the same sense as work relying on apparent time studies and the work implementing retro predictions. Rather, it focuses on an extension of the domain of a certain rule and is thus connected to both propagation of a change and typology (see Chapter 3). Second, the estimated time for when Lieberman et al.'s predictions will be borne out is set in the far future, i.e., more than

400 years from now. In that sense it appears to treat time in a similar way as Rask's and Sapir's intuitive predictions (see Section 4.2) which estimated completion of certain changes in around 100 to a few hundred years, making it unclear whether the time should be interpreted in a strict way (exactly in the year 2500) or if it is a general estimation about some very far off point in the future. For further discussion and critique on the study by Lieberman et al. (2007), for instance regarding assumptions about constant rate of change, see De Smet & Van de Velde (2019).

Finally, it is appropriate to mention work on language acquisition since changes are often thought to originate there (see Chapter 3, Section 3.3). This line of work typically deals with existing patterns in the language, i.e., patterns that children may pick up on and the rules they generalize. Furthermore, the focus tends to be on the acquisition of a "stable" linguistic system, based on UG and the primary linguistic input a child is exposed to. The acquisition process takes place over an extended period of time, over several years, and is generally treated as being deterministic. If two individuals receive the same input in the same order, their internal grammar is hypothesized to be identical. Although data obtained in acquisition research has been claimed to be "inappropriate for the study of change in progress" (Cukor-Avila & Bailey 2013; cf. also Van Hofwegen & Wolfram 2010), they nevertheless incorporate a form of prediction that concerns the nature of the linguistic system an individual will form, based on a given input. New and older systems are treated in terms of generations, i.e., a child acquiring a language is considered to be of a different generation than the individuals who produce the primary linguistic input the child is exposed to. While this may seem straightforward, it does introduce some issues in studying change over time. For instance, (to my knowledge at least) acquisition tends not to be directly linked to the exact time at which an individual was born, or to the time the acquisition period ended. Instead, the assumption is that children belonging to the same generation, defined over some period of time, will form roughly similar grammars. The linguistic input the children are exposed to appears to also be treated as coming from a semi-coherent generation.[28] Both assumptions greatly simplify real-life situations and lead to a somewhat vague time frame for studying ongoing changes.

If one were to use acquisition studies to predict a future trajectory of a change more than one generation ahead, i.e, predict what generations of the future will acquire, there is a need to assume that whatever system is acquired by a certain generation matches the output the generation will transmit onwards to the next generation. In other words, the relationship between language use and an abstract linguistic system must be assumed to be very close. This is the case for Ingason, Legate and Yang (2012) who use Yang's (2002) Variational Model to predict when the so-called New Passive (shown in (4.4c)) will be dominant in Icelandic, and the Canonical Passive (shown in (4.4b)) will have disappeared. Ingason et al. (2012) rely on data from The Icelandic Parsed Historical Corpus (Wallenberg et al. 2011) and model the evolution from 1950 to 2050 based on how children acquire and use language. According to their results "the first speakers who do not acquire the Canonical Passive will be born around 2050" (Ingason et al. (2012:98).

(4.4)	a.	*Álfurinn*	*lamdi*	*strákinn*	(Active)
		the.elf.SBJ	beat	the.boy.OBJ	

[28] As discussed elsewhere (Chapter 5), the concept of generation is problematic and should ideally not be used to measure language change over time.

b.	*Strákurinn*	*var*	*laminn*		(Canonical Passive)
	the.boy.SBJ	was	beaten.PASS		

c.	*Það*	*var*	*lamið*	*strákinn*.	(New Passive)
	EXPL	was	beaten.PASS	the.boy.OB	

‘The boy was beaten.’ (Ingason et al. 2012:91)

The prediction made by Ingason et al. is hard to verify since it concerns mental grammar of individuals instead of attested outputs. Additionally, the year 2050 is provided as an estimate. If the first speaker who does not acquire the Canonical Passive were to be born in 2046 or even 2056 instead of 2050, the prediction could still be claimed to be correct. If the year 2050 were to be understood in absolute terms, the prediction would be wrong if the last individuals acquiring the Canonical Passive were born in 2046 or 2056. While this observation may seem nit-picky, it is important to keep in mind that predictions need to be made in such a way that they can be evaluated in a way that is useful for the goal of studying language change and language forecasting.

From the discussion above, it can be gathered that various lines of research within linguistics involve some form of prediction without necessarily falling under systematic forecasting. These studies include formal approaches that use different types of data and different methods, sometimes resulting in tacitly implied predictions rather than explicit forecasts. In case of explicit forecasts for the future, time appears in many cases to be treated in a rather general way, often linked to generations or points in future that can be interpreted in a flexible way. It is unclear whether dates such as the year 2050 or 2500 should be understood in absolute terms or if they are estimates of some sorts with implicit margin of error. In this sense, the formal methods have something in common with the intuitive forecasts discussed in a previous section (Section 4.2), which also focused on time points in the far future.

4.4 Towards systematic forecasting

Even though language forecasting is in its infancy, there is work within linguistics that incorporates some form of future predictions about language change. Sometimes these predictions are implicit, as is the case in apparent time studies. In other cases, they are explicitly stated with reference to a future time, often a date in the far future. As already mentioned, the forecasting situation may affect which methods are appropriate, how far into the future predictions should be made, and how accurate the forecast needs to be. In some cases, multiple different approaches can be used to answer the same question, but there may nevertheless be reasons to choose one method over another.

Intuitive forecasts and forecasts based on naïve extrapolation tend to be rather general in nature. These are based on a forecaster's intuition and may be useful in certain circumstances, for instance when attempting to foresee and revert language death.[29] However, since they are not quantifiable and do not allow for reevaluation of the data or the method used for making the relevant predictions, they are not suitable for systematically studying trajectories of change or learning about the process of language forecasting. For this purpose, other methods must be employed.

[29] In some cases, intuitive forecasts may turn out to be more accurate than forecasts produced with formal methods. If the only goal is accuracy in predictions this may be appropriate.

The work discussed in Section 4.3 for studying trajectories of change and predicting the situation of a particular linguistic variant at specific times are more formal than intuitive forecasting methods. Importantly, the methods are quantifiable and allow for predictions to be verified. These include approaches such as fitting an S-curve to attested data (Van de Velde 2017), relying linear regression and ARIMA models (Van de Velde & Petré 2020; Berdicevski et al. 2024), using a variationist model based on language acquisition to derive an S-shaped curve (Ingason et al. 2012), or computing rate of change based on word frequency and projecting it into the future (Lieberman et al. 2007). Interestingly, some of these studies appear to more or less focus on when a change will be completed. The predictions target a date that is far away from the time of prediction (e.g., 2050 and 2500), resulting in the waiting time from when a forecast is generated and when it can be revisited and verified rather long. Sometimes, it is even unclear whether dates should be interpreted in an absolute manner or if they simply represent a rough estimate. Arguably, it is more beneficial to treat dates in an absolute manner since only in that way can the accuracy of the prediction be evaluated in a useful way. Additionally, it may prove useful to generate predictions that target time periods in the not-too-distant future. Of course, the concept of distance is relative to what is being measured or studied, but one may hypothesize that anything above 40–50 years into the future should be considered to be relatively far off in terms of predicting language change.

Dividing language forecasts into short-, mid- and long-range forecasts depending on how far into the future the predictions are made, it appears that most (if not all) of the predictions that have been made about language change belong to the long-range category. The long-range forecasts are likely the results of the forecasting questions being centered around when a change will be completed or when a certain stage will be reached. While the time at which a change is completed is certainly interesting, one may have to wait for a long time to verify such predictions.

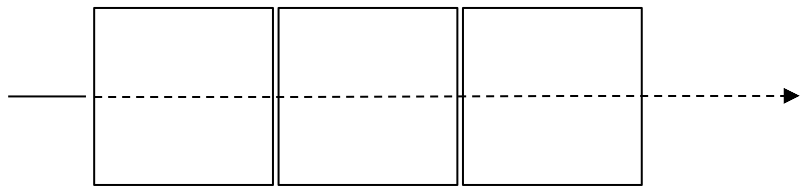


Figure 4.1. Visualization of short-, mid- and long-range predictions. Instead of attempting to predict when changes will be completed, resulting in medium- to long-range forecasts, it might be more informative to focus on what the situation will be in the near future by attempting short- to mid-range forecasting.

Attempting to make language forecasting more systematic, it may be proposed that forecasts should be set up in such a way that progress can be checked regularly without having to wait for decades or centuries. For doing this, one might choose formal forecasting methods, explicitly accounting for data that is used, and focus on producing short- to mid-range forecasts. In this way, the near future would be foregrounded, and the task would be to predict a situation at a given time in the immediately foreseeable future. Put more formally, the question might be framed as in (4.5).

(4.5) Forecasting question (Revised)

What will the situation *x*, of phenomenon *p* be at future times *t+1, t+2 ... t+ n,* given what is known about the situation of phenomenon *p* at times *t*, *t-1*, *t-2*… *t-n*?

Note that focusing on what the situation will be at a certain point in time is very different from focusing on when a certain situation will be reached. It does, for instance, not presuppose that changes will be completed, nor does it require changes to propagate in a particular way. It is a fairly neutral question that allows for studying and predicting both trajectories of change as well as stable variation in a language. The question is then, what kind of formal approaches should be used?

Systematic forecasting methods come in many different shapes and flavors, ranging from being relatively simple to quite complex. Methods can also be divided into two categories based on whether they incorporate explanatory factors or not. When simulating processes that give rise to certain states, explanatory factors are often part of the modeling. Perhaps the most ambitious attempt at large-scale simulation up to date is a project called Destination Earth (DestinE), led by The European Space Agency (Nativi & Craglia 2021). The goal of the project is to create a full-scale model of Earth, referred to as Earth's digital twin, to study weather and climate. The hope is that the model will assist researchers both in forecasting future weather and in providing insight into climate situations of the past. A project like Destination Earth has only been made possible through decades of work on weather and forecasting and through advancement in computer science. Note how weather and climate are treated as different entities. Similar distinction might be applied to language, in which case short-time language use can be compared to day-to-day weather and long-time trajectories of change to climate. Importantly, the everyday weather (or language) includes information about climate (language) and climate changes (language changes).

Although complex models and simulations can be constructed, there is often no need to do so in order to learn new things about forecasting and the phenomenon of interest. In many cases, models that abstract away from causal relationships can still provide new information on the phenomenon that is being studied. There is thus no immediate need to aim for the most complex models conceivable, intending to capture every aspect of language change. It may also be informative to rely on methods that focus on analyzing historical data and extrapolating observed patterns into the future. Time series analysis is one such approach.

Time series forecasting is a quantitative statistical approach that focuses on analyzing the relationship between observations that are consecutive in time. The approach demands access to a certain amount of historical data and focuses on analyzing the relationship between consecutive observations. The basic assumption is that past observations provide a clue for what future observations might be. In short, already observed patterns are extrapolated into the future. These types of models can range from being quite simple to elaborate and complex and can be suitable for short- to mid-range forecasting. As with all approaches, time series models range in complexity and how patterns in the historical data are carried into the future. An example of a very simple model is to assume that all future observations will be the same as the last recorded observation. Another simple model assumes that forecasted values will be equal to the average of all the values that have occurred in the past. More complex methods might involve identifying various sub-patterns in the historical time series and assuming more complicated relationships between past and future values. The structure of time series and patterns found within those are discussed in Chapters 6 and 7. For now it suffices to note that time series forecasting falls under formal approaches where data needs to be explicitly accounted for. The method is quantifiable and allows for reevaluation of both the data that was used and the exact model that was chosen. Since time series rely on historical observations, they naturally demand that a linguistic phenomenon of interest should be documented regularly so forecasts can be made. Thus, they might lead to the welcome side-effect of more thorough documentation of changes of interest. This is the case for the study of two complex prepositions in Icelandic in Chapter 8 and the study of changes in case marking with *hlakka til* ‘look forward to’ in Chapter 9.

Concluding this chapter, it should be noted that when formal models are used to make predictions, the forecast may often need to be supplemented with some form of judgmental predictions or comments. For instance, when making predictions on the proportion of individuals in a community exhibiting some linguistic feature or other, one might choose to restrict the prediction in such a way that it only allows for positive outcomes. This is because a statistical model might yield a prediction involving negative numbers. However, it is nonsensical to assume that a negative number of individuals could adopt a certain linguistic feature; it must be zero (no one in the community) or higher. There is never a situation in which a negative proportion of a population uses a particular linguistic feature. Therefore, even though forecasting is approached in a systematic way, using quantitative methods, there may be a need to supplement the forecast, so it fits with what is known about the topic and the world.

5 Measuring change over time

5.1 Units of time

Time is continuous and can only be discussed and referred to when broken up into units. The units can be of various lengths. When describing events occurring over a single *year* in the life of an individual, words like *day*, *week* or *month* may be used, all of which indicate a period of time. If, however, the aim is to report how long it takes to walk between two places within the same city or town, units such as *hours*, *minutes* and *seconds* are likely more appropriate. This suggests that the units that are used to measure time depend somewhat on the nature of what is being measured. Choosing the right unit of time for constructing regular time series containing language data is therefore important.

Time series vary considerably in the frequency of measurement. Some consist of observations made every nanosecond, others contain hourly, monthly, or annual data (Castle, Clements, Hendry, 2019:15). When statistical models relying on time series are used for forecasting, predictions are typically made one or more steps into the future based on historical data. The "steps" refers to the units of time the data is organized by. If the time series consists of daily data, one step-ahead forecast will predict the state of affairs one day into the future; if the data is weekly, the value for one or more weeks into the future can be predicted. Thus, the units of time that are chosen to measure the phenomenon under study may also influence how far into the future predictions can be made. If data is gathered and organized based on reference to generations of individuals, as is sometimes done within linguistics, predictions will also be made in terms of generations. As illustrated in Section 5.2, "generation" is not an optimal time scale for studying language change. This might be surprising given that changes are often considered to take place during acquisition which emphasizes how new generations construct language based on data from older generations (Chapter 3, Section 3.3). Discarding the notion of generation, the question arises which time units are relevant for measuring (and predicting) language change. The answer depends partially on the number of observations needed for a particular forecasting method, how quickly changes are perceived to occur, and what is a possible and realistic sampling frequency. In what follows, it is proposed here that standard measurements such as months, quarters, years and decades should be employed and not generations.

5.2 Meaningful and meaningless sampling frequencies

As illustrated in 5.1 above, certain units of time are more appropriate for measuring one thing than another. It makes equally little sense to measure the duration of a single lunch time in terms of weeks as it does measuring the time it takes to walk from home to work in nanoseconds. The scale of measurements must be in accordance with what is measured. Thus, although the propagation of language change takes place gradually, it is usually unfeasible to sample data at very short intervals such as every hour, every day or every week. It is even debatable whether measurement should be taken monthly or yearly (although this is discussed further below).[30] Certainly, no drastic language changes are expected to take place overnight or over the course of a few days, weeks or months.

[30] The chosen unit of time depends entirely on what exactly is being investigated. If the question is about tracking and predicting how the discourse of a particular topic unfolds online, one may want to look at everything from hourly data to yearly data.

When a contemporary language is studied, the exact time at which data is gathered is not always regarded as a significant factor. Data from two “similar” or “adjacent” periods is often treated as equally representative of the state of affairs within some time frame. If data was gathered in May 2020, there is not necessarily any reason to think that data gathered in November 2020 would have yielded significantly different results. Any difference between the two points in time might be considered to be due to variation in measurement rather than representing change. However, this way of thinking about time and change may be problematic.

When language change is not observed – or not expected to be observed – from one day to the next, it is not necessarily because of no change taking place. Rather, it may be related to how quickly or slowly language change is *perceived* to take place and how it can best be measured. This is best exemplified with a simple analogy to the growing of a plant. If one sits opposite a plant and constantly stares at it with the intention of watching it grow, one is likely to get bored fairly quickly as not much difference can be observed from one second (or minute) to the next. Even if the plant were to be measured in great detail every second, the difference in height from a second to second is so small that it is likely meaningless. However, if one steps away for a few hours or perhaps a few days, the amount that the plant has grown starts to be *noticeable* and *meaningful*.

Similar to how plants continue to grow steadily over the course of days and weeks, language is constantly changing as time moves forward. Depending on the frequency of the observations, not much difference may be observed from one point in time to the next. Estimating the situation of an ongoing change on a weekly basis may not give a meaningful insight when measurements at adjacent times are compared. Thus, the situation at week 2 may not be significantly different from that of week 1. Week 3 might give a similar picture as week 2 gave, and week 4 may not be very different from week 3. However, if week 4 and 1 are compared, a subtle difference – a change – might start to become noticeable.[31] This is presented visually in Figure 5.1.

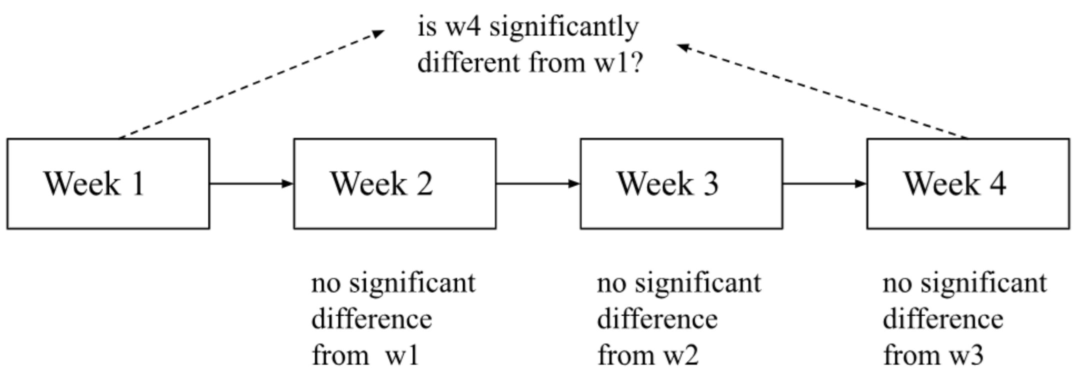


Figure 5.1. Adjacent weeks may be very similar to each other, not showing any significant difference in terms of change. However, when comparing measurements a couple of weeks apart a change might start becoming noticeable.

The fact that language is constantly changing (albeit slowly) necessarily raises the question of what constitutes a meaningful measurement of change over time.

Documenting a change in as much detail as possible would, of course, be ideal. However, there is a trade-off between how frequently data can theoretically be sampled and how much can be processed and analyzed within a given time period. The amount of data that can be gathered and analyzed varies somewhat depending on whether dealing with convenient or specialized language data (see discussion in Chapter 3). While the former may already exist in abundance and lend itself fairly easily to some form of

[31] If four weeks sound like too short of a time for a change to be noticeable, one might try thinking about it in terms of months or years.

periodization, it still needs to be gathered, analyzed and annotated. The latter, the specialized language data, must be actively created, starting with designing an experiment and recruiting participants. Although the same experiment can be run multiple different times over an extended period of time, there are some restrictions on how frequently it is feasible to do so. It takes time to find participants, run the experiment and process the results. Thus, each type of language data imposes restrictions on what is a possible and realistic sampling frequency. For studying language change and the propagation of change with the goal of forecasting in mind, it is necessary to settle on some frequency that is consistent, "makes sense", and is both doable and feasible.

When deciding on a sampling frequency for studying and measuring language change over time, it may be tempting to take into account how language is transmitted. The reason for doing so would be that transmission and acquisition of language is usually regarded as the primary 'source' of language change.[32] Since transmission of language typically occurs from one generation (parents) to another (children), sampling language data based on generations might sound like a good idea. Generation is a period of time in which changes are expected to be observed. Often, there is a noticeable difference between the language of grandparents and parents, and the language of parents and children.

Although measuring language change in terms of generations may seem ideal, it poses several problems. One such relates to *what kind* of time unit 'a generation' is and *how* something can be measured using this unit. To complicate matters, different definitions of the term *generation* exist. For instance, parents and children are typically thought to represent two distinct generations, fitting well with ideas of transmission of

[32] See for instance Hermann Paul (1886:29–31) who notes that child language may be conceived of as the primary source of many new linguistic variants that later can be found as a part of the standard language.

change. However, generation can also be used in terms of a social construct where it refers to "all the people who were born at about the same time" (OED), which may or may not match with parents making up one generation and children another. Figure 5.2 shows how the two definitions may apply to a single family.

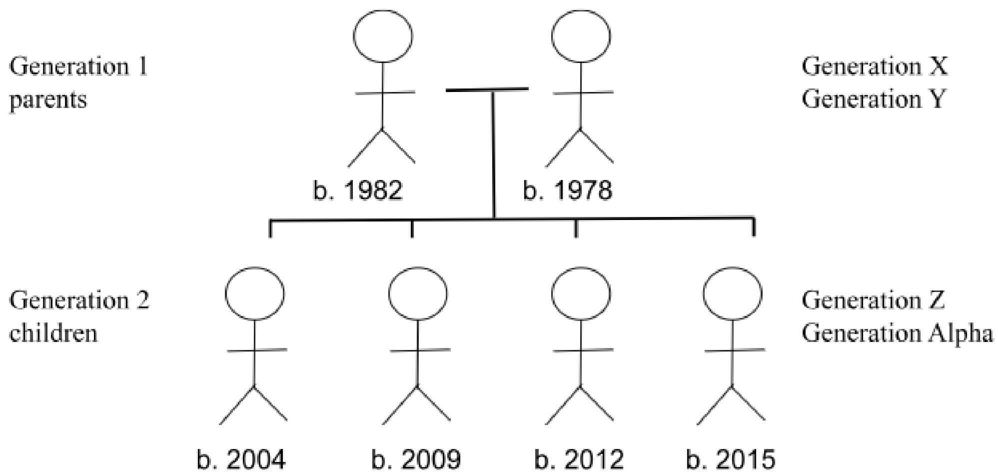


Figure 5.2. Generations can be understood in terms of parents and children, but also in terms of when someone was born and when they grew up.

Children with the same parents can be born in quite different time periods. Although siblings may be considered to belong to the same generation in terms of family relations, it is not guaranteed that they will count as a generation when viewed from the outside. In fact, they may have a vastly different experience of the world due to difference in age. They may even be placed into different conventional generational categories. This is illustrated in Figure 5.2, where the parents belong to Generation X and Y, and the children to Generation Z and Alpha. Note how the children born in 2004 and 2009, who are five years apart, are considered to belong to the same generation, Generation Z. However, the parents, being four years apart, are considered to be of different generations, Generation X and Generation Y.

With the goal of measuring language change in a consistent way over extended periods of time, the notion of generation is highly problematic. No matter whether it is viewed in terms of parent-child relationship or in terms of widely used generational categories, it introduces inconsistencies. Referring to grandparents, parents, and children by using generation 1, 2 and 3 means that there is no fixed link to how time is generally measured in years, how old is generation 1 for instance? Categories such as Generation X and Y are also not useful because changes do not necessarily take place in terms of these categories. A solution might be to view generations as fixed time periods. However, this is not as straightforward as it sounds. The gap between parents and children is not consistent, neither within nor across different cultures. Generational categories also do not have the same length, varying between 15 and 30 years. As a unit of measurement, *generation* is not a period of a consistent length, and is therefore not ideal for the study of change when the methods require sampling at regular intervals. Even if a fixed period of 15 years were to be used, there would still be issues. It is true that changes might be detected and noticeable when using a 15-year (or 10-year) interval. Unfortunately, it may also mean that a considerable amount of change may have taken place between two measurements, resulting in suboptimal documentation of the change.

Since a 15-year period is deemed to be too long for documenting change, the question is still what is a good sampling frequency. Moving away from the term generation, it is proposed here that yearly data should be used, possibly even quarterly or monthly data, depending on how many observations are needed for studying a relevant change and making predictions. The benefits of settling on monthly to yearly observations is that these are standardized time units that are used in other areas for measuring and predicting things. These time units also lead to a thorough documentation and abundance of data to work with, both for the sake of reference and for tracking variation and change. Yearly, quarterly, and monthly data can also be converted into other sampling frequencies if deemed appropriate, for instance bi-annual data or a series of observations every five years. The most obvious downside of gathering monthly to yearly data is that it takes a lot of time and effort. Additionally, one may run into issues of consistency with data sampling w.r.t. the type and nature of available data. Although this is always a problem for historical linguistics, it may prove to be extra challenging when observations are needed for every month or year.

5.2 Problems with consistencies in measurements

When studying a phenomenon over an extended period, it is important for individual observations to be comparable. For instance, when tracking outside temperature in a particular place (e.g., New Haven), one might go out every morning at the same time and note down the relevant measurement, using the same scale (Celsius or Fahrenheit). Doing this provides consistency in how things are measured, and it makes it possible to compare individual observations with the goal of track variation and change over time. In the case of language, making sure individual measurements are consistent and comparable can be quite challenging. Illustrating some of the problems that come up is useful as it can be informative about why certain decisions are made when choosing material when measuring changes over time.

Issues that arise when measuring language variations and change can be divided into at least three categories. These are: i) type of data used and whether each measurement is comparable, ii) the dating of the material used for tracking change, and iii) irregularities in sampling frequencies. A fourth category may be claimed to exist, dealing with how exactly measurements are taken and how they relate to the actual situation in the language. Since the nature of this issue is slightly different from the other three, it is treated in a separate section (Section 5.4 how exactly is propagation measured).

Starting with issues of the first type, it is most certainly the case that obtaining measurements that are truly comparable can be very challenging. Although potentially more prominent when measuring variation and change in older historical sources, this issue is also noticeable in studies based on more recent data. Thinking about the type of data used, it would be questionable to directly compare spoken data at one point in time to written data at another point in time while hypothesizing that the same things are being measured and that the measurements have a temporal relationship. Making sure the data is of the same type, e.g., always written data or always data from experiments or surveys, is very important although it does not on its own guarantee consistency.

Apparent time studies (e.g., Bowie 2005), set out to document language change by observing linguistic variants used by individuals of different ages. The language of children is compared to that of teens, adults and older people to establish a trajectory of a change. Information about the status of each of the groups is gained through similar means, i.e., by using questionnaires or by recording spoken language etc. The method assumes that the various age groups have something in common which makes their language comparable. An essential requirement is that they belong to the same language community, roughly defined. In some instances, it is important that they have similar kinds of backgrounds. In the light of individuals often having vastly different life experiences, the question arises whether two or more observations are indeed comparable. As an example, older individuals may have moved around more than younger individuals, causing them to pick up or change their language in different ways. Some age groups may also have larger social networks than others. Finally, even if all age groups are assumed to be comparable, it is difficult to be certain about to what extent the language of each group is truly comparable.

The question of whether individual observations are comparable or not also arises when using corpus data (see e.g. Hoffmann 2004:173). As recent studies on language change have shown, text types may play a role in whether and when new linguistic variants are found. Scientific writing is, for instance, stylistically very different from fiction and it may contain different types of linguistic constructions. It is therefore not appropriate to directly compare observations for one to the other. It would be similar to a meteorologist, interested in figuring out fluctuations in temperature at a particular location, comparing temperatures taken in two or more different places. The measurements end up not being fully comparable. Fortunately, compilers of corpora often attempt to balance the type and amount of material from each time period. Naturally, the balancing is restricted by the text types that are preserved from each period. This may be more noticeable the further back in time one goes as there is no way of controlling for which type of material happened to have survived to modern times.

Another issue with measuring change over time relates to the dating of individual observations. When working with texts preserved in manuscripts, a distinction is usually made between the text itself and the object it is preserved in. Problems with dating can arise in each of these areas as the same text can exist in multiple different manuscripts and which may be from different periods of time. While the manuscripts themselves have their origin at a certain point in time, the texts contained within the manuscripts are much more fluid in their temporal existence. When popular texts are written up repeatedly, some aspects of the text may reflect older stages of a language while some may be innovative. A text that is encountered in a manuscript may thus contain multiple layers of language, reflecting archaic usage and innovative variants. Take for instance the Prose Edda (also called Snorri's Edda), a textbook in poetry which contains stories of the Norse gods. The text is preserved in several manuscripts, dating from different time periods. The main manuscript and the oldest one is *Codex Regius* (Stofnun Árna Magnússonar í íslenskum fræðum, shelfmark GKS 2367 4to), which is dated to around 1300–1350. Another manuscript, *Codex Wormianus* (Den Arnamagnæanske samling, shelfmark AM 242 fol.), is dated to the mid to late 14th century, approximately between 1340 –1370. Since the same text is preserved in manuscripts from different time periods the question is should the text be dated to the time of the manuscripts, meaning that the two attestations of the texts are dated to different times? Or should it be dated to a hypothetical time of composure, which would be in the early 13th century, making it older than any of the preserved manuscripts? Choosing the latter can be problematic. In a corpus like The Icelandic Parsed Historical Corpus (IcePaHC Wallenberg et al. 2011; Rögnvaldsson et al. 2012) the dating of texts is sometimes based on when the text was likely originally written. The First Grammatical Treatise is thus claimed to be representative of 12th century language. However, text is only preserved in a single manuscript, *Codex Wormianus* which was written in the late 14th century. This leaves a gap of about two centuries between when the text was hypothetically composed and the time of the manuscript it is preserved in. It is not unlikely some aspects of the texts, including the language, changed when the text was copied. It may, therefore, be more appropriate to rely on the dating of the manuscript rather than the assumed date of the original text. If this is done, one should still note that the dating of individual manuscripts is not always accurate. In some cases, manuscripts cannot be dated to a particular time with absolute certainty and are instead assigned dating based on decades or centuries, similar to *Codex Regius* of the Prose Edda and *Codex Wormianus* mentioned above. Since manuscripts cannot always be dated precisely, it follows that language change, observed through manuscripts, cannot be tracked, or documented with absolute precision w.r.t. time.

Although discussing the dating of individual observations may seem more relevant for studies relying on older linguistic material, the problem also arises when using language data from more recent times. For instance, material found in online sources may have been edited at various points in time. This applies, in particular, to sources such as Wikipedia where the material is the result of collaborative work of multiple individuals over an extended period of time. In a similar fashion, news articles online may be edited more than once, although these edits tend to be within several hours or few days instead of stretching over multiple years. If the goal is to measure something on a quarterly or yearly basis, online newspaper articles might not be problematic. As for printed material, books and academic papers published in a particular year may not necessarily be written in the year of publication. The author may have worked on the text for an extended period of time, perhaps up to several years before the work became available. Taking into account who the author of the text is, what age they are etc. may complicate matters. While commenting on the dating of modern texts in this way may seem trivial, it is important when taking into consideration accurately measuring change like the time period and the type of data that is used.

The final issue mentioned here for measuring change over time is that of irregularity in sampling. As noted elsewhere (Chapter 6), language change is often not documented at regular time intervals. There are various reasons for this. Sometimes, restrictions are imposed on the sampling frequency by the type and amount of preserved material in the target language. For a language like Icelandic, a fair amount of written material exists from the twelfth century onwards, preserved in manuscripts and manuscript fragments (for an overview of Icelandic manuscripts and scribes see e.g., Gunnlaugsson 2005:245–264). Even though the quantity of texts and material may give a semi-holistic picture of the language at the time of writing, these materials are not evenly distributed over time. Due to how material culture is transmitted and preserved, there is more material from later times than earlier times. As for language data from more recent times, perhaps within the last 100–200 years or so, the situation might be slightly different, provided the language under discussion has an established tradition of writing or some other form of documentation. In these cases, language data tends to be available at more regular intervals. This includes both convenient E-language data in the form of written and recorded material and specially specialized language data in the form of grammaticality judgments, recordings, or experimental results (on the different types of data see Chapter 3, Section 3.2). However, problems with possible sampling frequencies can still occur.

Specialized language data does not exist without being systematically gathered or generated (see Chapter 3, Section 3.2). Usually, data on a single linguistic phenomenon is not gathered at regular intervals with the intention of tracking change. Rather, the goal is to answer theoretical questions about the existence of language in the mind and the internal grammar of speakers. Only rarely are follow-up studies carried out with the goal of tracking changes. When such studies are done, they are not necessarily carried out at regular time intervals. For instance, studies on changes in subject case marking with the predicate *hlakka til* ‘look forward to’ in Icelandic have been done several times over the last few decades (see further in Chapter 9, Section 9.3.2). The first major study was done in the 1980s (see Svavarsdóttir 1982). A comparable follow up study took place in the 2000s (Jónsson & Eythórsson 2003) and another one in 2006–7 (Thráinsson 2013, Thráinsson et al. 2013). While these studies offer valuable insight into the propagation of change, they do not meet the requirements that regular time series analysis calls for as they were not conducted at regular time intervals.

As for convenient E-language data, abundance of properly dated material is fairly easy to come by. However, the amount of data does not guarantee that the linguistic structures of interest are necessarily attested frequently enough for quantitative analysis. They may not even occur at regular intervals. Furthermore, a considerable part of naturally occurring language data is in the form of officially published writing, for instance various types of articles and novels. This type of material may not always reflect everyday language as spoken or used within the language community. This is because published material often goes through copy-editing and proofreading and is therefore likely to be corrected in such a way as to match some ideal language standard. Although the language standard can be close to everyday language, more often than not it more accurately reflects an older standard of the language under discussion, with a considerable gap being between the language standard and the modern language. With the advent of social media, more copious amounts of non-proofread and non-standard language is attested online and can be used for language research.This type of data, as well as published material, has allowed for more detailed and systematic studies on language w.r.t. variation and change over time, although sampling frequency may still occur.

The discussion above serves to illustrate how available language data (corpus material as well as other types of data) may affect possible sampling frequencies when measuring language change over time. It also points out that individual observations are not always truly comparable. When measuring language change over time there is no equivalent to the consistency in how temperature is measured for meteorological studies. Additional challenges, not discussed above, include how exactly language variation is measured at various points in time and how accurately the observations represent the situation at that point.

5.4 What exactly is being measured

Documenting the propagation of a change implicitly or explicitly assumes the existence of a population speaking a certain language. However, since it is usually impossible to include every single individual or every single utterance in the language community in documenting changes, subparts of the population or attested utterances are used to estimate the overall situation. The subparts can be thought of as windows into the state of affairs and ideally should be kept constant over time to make sure the measurements are comparable (see Section 5.3). Figure 5.3 visualizes this, showing a snapshot of a language at a particular time. The "I" represents the collection of all grammar of individuals that produce everything that is uttered, and “E” represents all utterances (see also Chapter 3, Section 3.3). The situation of a particular linguistic phenomenon in a language is assessed by looking at subparts of the language, i.e., through a “window” which might be in the form of convenient E-language data from a particular newspaper or an online social platform.

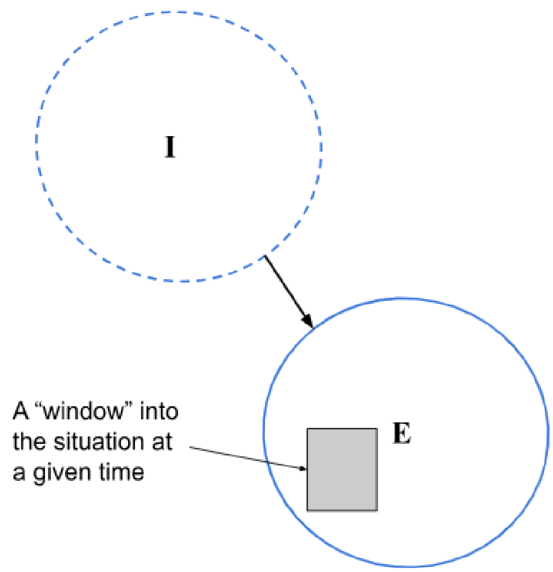


Figure 5.3. A snapshot of a language at a particular time with I referring to a collection of mental grammars that produce utterances (E) that are part of the language.

When sampling data with the intention of forecasting, the forecast will necessarily be limited by the “window” into the community. That is to say, predictions generated for future periods will be about the language as seen through a similar type of window in the future. If one chooses to use blog posts to measure change over time, a forecast for a future period will necessarily hypothesize about the status of the language in similar types of blog posts. Thus, the forecast will only say as much about the actual state of affairs in the whole population as the type of data used for the forecast does. Figure 5.4 shows the relationship between comparable "windows" into language at any given time.

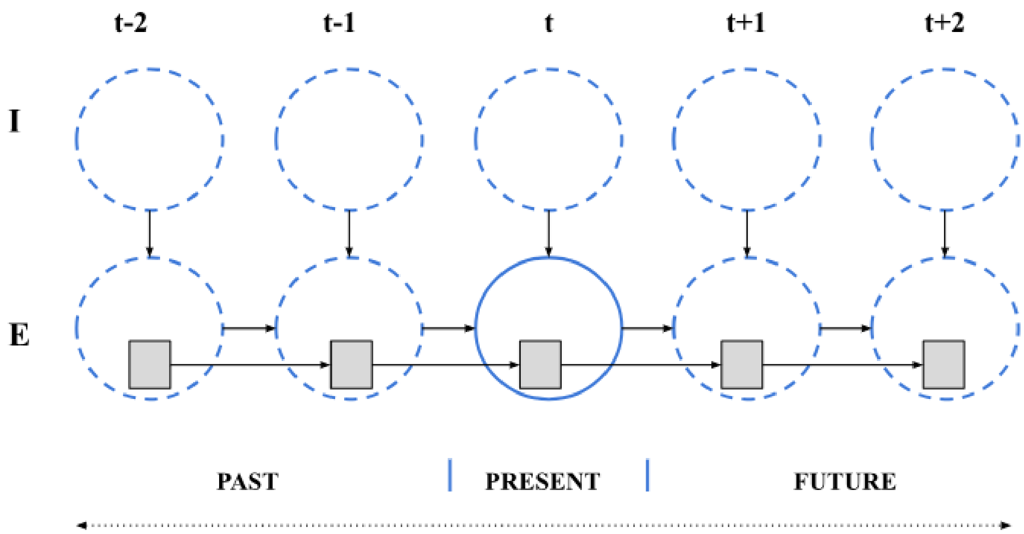


Figure 5.4. The gray area represents the "window" into the language at each time. Ideally, the window should remain constant over time to make sure that each observation is comparable. Any prediction about the future applies to a similar type of window as was used to observe the state at particular times in the past and the present.

When documenting changes through various "windows" into the language of the population there are a few ways in which the status of a change might be measured. Proportion of old and newer variants can be computed based on what individuals *can* and *do* say, or on attested utterances without referring to individuals.[33] Although the difference between these two may at first sight seem trivial, they yield different results. Figure 5.5 depicts five individuals, and their utterances of two linguistic variants, A and B. Variant A

[33] A further way of measuring the propagation of change is to compute the proportion of new and old variants over grammars of individuals, i.e., focus on what individuals can say (I-language) without accounting for what they do say (E-language). Measuring propagation in this way is potentially problematic as overt usage of language is a crucial factor in transmission, and the latter ultimately affects what the next generation will acquire (see also Chapter 3, Section 3.3).

represents an older linguistic feature while B stands for innovation. Measuring the propagation of variant B based on attested examples is simple enough; out of ten attestations, the new variant appears five times. In other words, variant B makes up 50% of all attested examples while variant B makes up the other 50%.

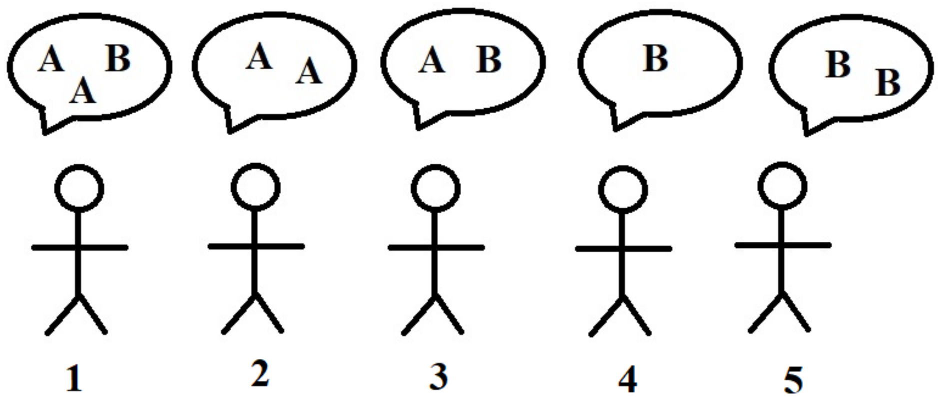


Figure 5.5. In a group of five individuals where an older linguistic variant is represented by *A* and a novel one by *B*, computing the propagation of the novel variant can be done in terms of attested utterances or in terms of individuals and what they can say. In this case variant B makes up 50% of attested examples.

If, however, the propagation is measured based on individuals and the variants they use rather than computed over attested examples, a different picture emerges.[34] Now, four out of five individuals, or 80% of the group in Figure 5.5 use the B variant. As for the A variant, three out of five individuals use it, or 60% of the group. Note that when added up the total percentage exceeds 100%. This is because the same individual can use more than one variant and is therefore counted twice. In the case that calculations are based on attested examples, the total percentage will always be exactly 100%. Measuring change in terms of

[34] The reason for computing propagation based on individuals rather than attested examples may be that individuals make up language communities and may have influence on other speakers to different degrees.

what speakers *can* do as opposed to measuring change in terms of observed examples can provide a different perspective on the same set of facts. Summarizing Figure 5.5, 80% of the population (if the five individuals are considered to make up a population) can use the novel variant B. Despite this, variant B only makes up 50% of the attested examples. Were the same set of people to be observed at a later time, the proportion of old and new variants might very well be different. The proportion of individuals who can use the new variant will not necessarily have changed. Figure 5.6 depicts the same individuals as Figure 5.5, but at a different point in time. While the number of individuals using variant A (3 individuals or 60%) and variant B (four individuals or 80%) is the same, variant B has increased and now makes up 70% of attested examples instead of 50% as earlier. If Figure 5.6 is sampled at a later time than Figure 5.5, one might say that the population is converging on using the newer variant B.

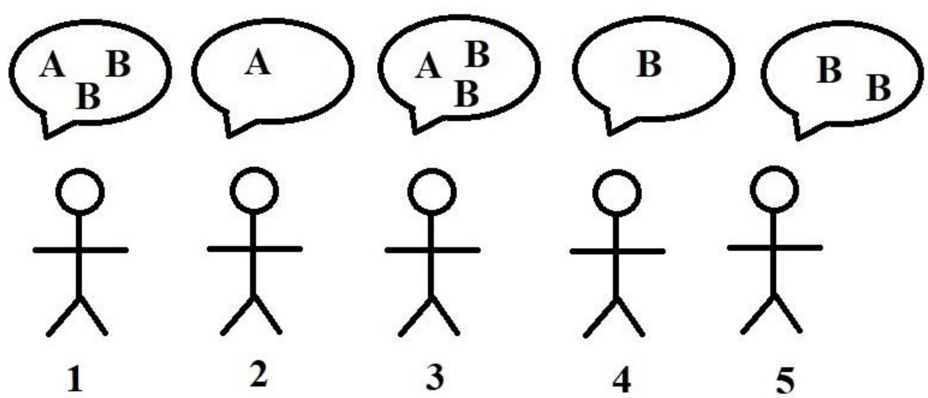


Figure 5.6. Same group of individuals as in Figure 5.5. The proportion of the innovative (B) vs. older (A) variant has changed without a change in the proportion of speakers who use the relevant variants. In this case variant B makes up 70% of attested examples.

When tracking change with forecasting in mind, the question necessarily arises which method of measuring change is more informative and useful for making accurate predictions. Since language forecasting has not been systematically practiced, this is hard to determine. Although sociolinguistic approaches to change often involve incorporating information on individuals which may play a role in the selection of certain linguistic variants, there is a tradition of simply measuring change based on attested examples when documenting change over longer periods. Choosing to measure change based on the number of attested examples within a given source (the "window" into the language community) has some benefits. For instance, it is fairly easy to compute and can be applied to a broad variety of data. Conveniently, it takes less time to obtain usable data for this kind of measurement, and one does not need to know how many individuals contributed to the sample at each time or who exactly contributed which variants. The individual is simply treated as an undefined part of the whole sample (see also Chapter 3, Section 3.3). I therefore suggest that measuring change based on attested examples within a given window into the language of a community is probably the best kind of measurement to experiment with for forecasting. I nevertheless acknowledge that, in the future, it may turn out that more attention should be paid to individuals and their contributions of utterances.

5.5 The working methodology

With the goal to generate a forecast for the propagation of language change using regular time series, it is essential to sample and measure change at regular intervals over some period of time. Accordingly, there is a need to decide on appropriate sampling frequency and the type of data that is used.

As already mentioned, the sampling frequency that is chosen for documenting change should be appropriate for the study of change while also taking into account the type of material available. Even though some changes may be expected to propagate more quickly than others, it may be beneficial to use the same frequency for all changes up to the extent to which it is possible. The reason for this is to make sure of consistency in how things are measured as well as emphasizing a thorough documentation of the propagation of the change of interest.

Although changes are often thought to occur during language acquisition, taking place over generations, measuring change in terms of "generations" is bound to lead to problems. Aside from being an ill-suited time unit for forecasting, there is a risk of missing out on observing the propagation of a change properly. Generations, defined either in terms of societal conventions, can span everything from 15 to 30 years and within that timespan some changes may have started and be (virtually) completed. For these reasons, more concrete and standardized measurements should be employed when measuring change over time, for instance *month*, *quarter* or *year*. These units have the benefits of being used for making predictions in other areas (weather, economics, spread of diseases, consumption, sales etc.) and models from those areas can be adapted and used on language data. Measuring change frequently also ensures that the resulting time series will have enough observations for time series analysis. If it turns out that it is more informative to only measure propagation every two to every four years, it is simple enough to convert monthly, quarterly and yearly time series into that format. Doing it the other way around, i.e., to go from bi-annual time series into quarterly time series, is more complicated.

When deciding on the frequency of a time series, it should be kept in mind that the time unit that is chosen, month, quarter, year, will serve as the basis for forecasting, i.e., one step ahead forecast will be based on the frequency of observations. Thus, there is a need to consider the balance between how frequently a change is measured, how many observations are needed for a time series analysis and how far into the future prediction should be made. Typically, the further ahead predictions are made, the more uncertain forecasts become. While yearly observations may be optimal, quarterly observation may be required in some cases, in particular if there is a need to obtain more observation for the time series.[35] To link this back to the discussion of short, medium, and long-range forecasts (see Chapter 4), one might assume that predictions about the situation of a language 30 years or more into the future belong to long-range forecasting. As for short- and medium-range forecasts, a working definition of these could be based around the idea of generations. Under this view, short-range forecasts would apply to the period from today until 15 years into the future and medium-range forecasts cover the 15–30 years (the time span of a generation) into the future. Of course, these estimations need to be tested and refined in connection to how accurate predictions for these periods turn out to be, and which methods are appropriate to use for each range. When relying on time series analysis and forecasting, the accuracy may depend somewhat on the frequency of the time series. The proposed periods for short-, medium-, and long-range forecasts are based on the assumption that annual (or bi-annual) data is used.

[35] If data spans ten years, yearly time series would consist of ten observations but quarterly series of 40. For time series analysis to be possible, a certain number of observations is needed.

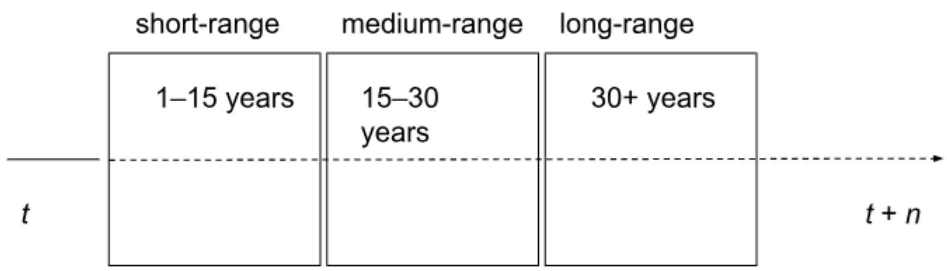


Figure 5.7. Short-, medium- and long-range forecasts covered different times in the future. It is hypothesized here that short-range forecasts might cover 1–15 years, medium-range 15–30 years and long range everything above 30 years. These definitions will likely need to be revised once language more experience has been gained in language forecasting.

Once the sampling frequency has been decided, it is appropriate to make explicit notes on exactly how changes are measured. As discussed in Section 5.4, this may be done in at least two ways, i.e., by measuring proportions of older and newer variants based on all attested examples from the sampling-window, or by taking into account individuals and their linguistic behavior. Since it is more straightforward to do the former than the latter, this is what is proposed to be done. There is a chance, however, that taking the linguistic behavior of individuals into consideration might be appropriate for language forecasting in the long run. This is left as a question for the future.

6 Patterns in language data over time

6.1 The future through extrapolation

Forecasting often relies on historical data to make predictions. Unusually, it is assumed that the phenomenon of interest exhibits some regularities over time which can be used to extrapolate from into the future (Castle, Clements & Hendry 2019:21).

Language change, both in the sense of innovation and propagation, shows regularities over time. These regularities are manifested in restrictions on which changes can occur at any given point (e.g., Wedel 2015) and how changes may propagate through a language community. Observed regularities can be useful in language forecasting as they give rise to expectations towards what *can* or *might* happen. They also hold a clue to what *cannot* or is *unlikely* to happen. Regularities observed in the propagation of change involve the emergence of various curves. One such curve is the S-curve (see further in Section 6.2) which is well documented and has even been claimed to be a feasible point of departure for language forecasting (e.g., Sanches-Stockhammer 2015, Nevalainen 2015; Van de Velde 2017).

While commonly observed curves in propagation of change can be useful in giving rise to expectations toward ongoing changes, these do not need to be taken as a priori assumptions regarding how changes will proceed. It is not necessary to assume that changes will unfold in a certain way when making predictions. Instead, emerging patterns can be studied through regular documentation and by comparing documented patterns to expectations towards what is likely to happen.

In a previous Chapter (see Chapter 5), it was argued that language forecasting could benefit from systematic gathering of data at regular intervals where the sampling frequency relies on units such as *month*, *quarter* or *year*. Gathering data at regular intervals gives rise to time series which can be useful in several different ways. First, individual observations can serve as points of reference for a situation at a given time. Second, the complete time series allows for studying various patterns in language over time, both patterns that have already been documented as well as patterns that have not gained much attention. Third, time series can serve as input for certain types of forecasting models, for instance models that allow for the emerging patterns to be approached in a fairly "neutral" way as they rely on taking historical patterns and extrapolating them into the future. Putting together expectations towards propagation of change and detailed documentation of language through time series, the contribution of this chapter can be summarized in two points, (6.1)–(6.2).

(6.1) There are regularities in language change that can be useful in forecasting. The regularities give rise to expectations towards what *can* or *might* happen.

(6.2) By systematically gathering language data at regular (small-scale) intervals patterns in language over time can be studied in detail. Time-series patterns can be compared to expectations (6.1) that previous knowledge has given rise to. They can also be used to extrapolate from into the future.

The Chapter is structured as follows. Section 6.2 illustrates some expectations toward how language change might propagate through a language community. Section 6.3 deals with the structure of time series, decompositions and which kinds of patterns may be observed in language data over time. Section 6.4 summarizes the main arguments and explains how time series data can be useful in both studying the propagation of change and in making predictions about the future. In short, studying language at regular short-time intervals may provide insight into the distinction between variation related to language use and variation tied to language change. Understanding these is important for language forecasting in general.

6.2 Observed patterns in language change

6.2.1 Expectations towards change

Language change is *not* random but exhibits both regularities and directionality.[36] Regularities can be manifested in various ways such as through common types of changes that occur within or across languages at different times, through regular correspondence sets that reflect systematic changes (Osthoff & Brugmann 1878:XIII on changes occurring without exception), and through the fact that language cannot change in just any way at any given time (e.g. Wedel 2015).

Many regularities in languages that are observed synchronically have been argued to be the results of common diachronic changes rather than inherent restrictions on language (e.g., Haspelmath 2019 who cites Greenberg 1969; Bybee 1988; Givón 1979 and

[36] The directionality is, of course, not intended in the sense that language consciously moves in a certain direction. Rather, when viewed at a certain level, multiple nuances and changes make it seem like there is a directionality (see also discussion in Andersen 1990).

Lehmann 2015; see also Blevins 2006; Anderson 2016). Anderson (2016:11), for instance, explicitly suggests that "combination of contingent historical developments and biases in the learning algorithm" are responsible for observed generalizations in phonology, and that generalizations in morphology and syntax are "indeed the product of diachronic change rather than synchronic constraint".

Aside from regularities being observed through types of changes that occur over and over again, they are also observed in changes that do not take place or are uncommon. Changes that have never occurred in any language have been termed impossible changes (Honeybone 2016). One such is /f/ > /θ/ which is claimed to be unattested (Honeybone 2016). Whether the absence of this change is due to conditions for the change being typologically rare or the change being "unnatural" in some way is not important in the current context. The main point is that since the change is not attested, it is generally not expected to occur. The opposite change, however, /θ/ > /f/, is reasonably well attested and might occur in a language with the phoneme /θ/.

Regularities in language change are also observed through grammaticalization, i.e., the conventionalization of words or string of words into a functional element (Hopper & Traugott 2003[1993]:4; see also Kurylowicz 1965/1975:52; Heine, Claudi & Hünnemeyer 1991:2). This type of change is extremely common and is responsible for creating items such as prepositions and indefinite articles. The opposite change, where a functional element becomes a lexical word (degrammaticalization), is by far less common (on the status of degrammaticalization see e.g., Norde 2009, 2012; see also Van de Velde & Norde 2016:10–11 and Narrog 2016:115 on exaptation in relation to various types of changes, including unexpected changes and (de)grammaticalization). Given a language at any point in time, the occurrence of grammaticalization would be understandable, even half-expected. Degrammaticalization, on the other hand, is typically not expected. These expectations are summarized in (6.3).

(6.3)	EXPECTED (common)	UNEXPECTED (uncommon or impossible)
	/θ/ > /f/	/f/ > /θ/
	grammaticalization	degrammaticalization

Another way in which regularities are manifested is through directionality in the propagation of change. Novel linguistic variants are usually expected to show some kind of directionality over a period of time in their propagation through a language community. They may either catch on and be used more frequently, diminish in use and disappear, or continue to exist alongside an older variant. The exact pattern of propagation is not always immediately noticeable. However, perceived direction may play a role in how changes are first noticed.

There are primarily two ways in which changes and directionality are initially encountered by individuals. First, through observing a variant that was not attested before or by noticing that a particular variant is no longer used; Second, by noticing an increase or decrease in the usage of a variant over time. The perception of a change is linked to how the situation at two or more points in time are compared. A variant that used to be uncommon (rarely encountered) at one point in time might become more common (encountered more frequently) at a later time. Awareness of age-grading may also play a role. Taken together, these may lead to an awareness of an ongoing trend where the increase of a new variant is understood in terms of directionality. When directionality has been established there is often no reason to think the course of the direction might be altered. In some cases, the directionality may be *perceived* to be in line with linear growth as in Figure 6.1. Note, however, that how individuals initially perceive change is not always in line with the true patterns emerging from documentation over a longer period of time. For an overview on linguistic work focusing on individuals' awareness of language change, see Kootstra & Muysken (2019).

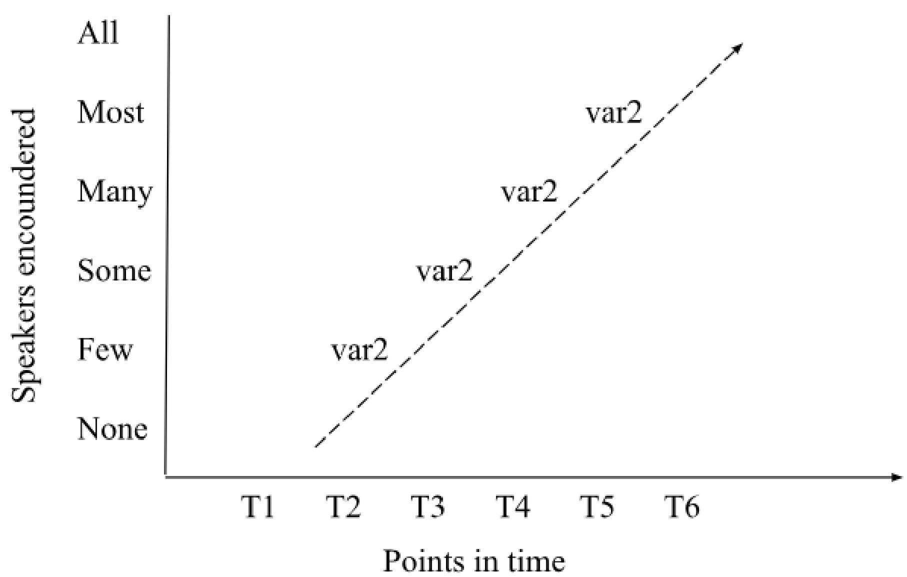


Figure 6.1. Perceived directionality. Individuals might understand directionality in an abstract way through indirectly comparing two or more points in time w.r.t how many encountered speakers use a particular feature. The perceived directionality might be in line with a straight linear growth.

Various factors contribute towards emergent directionality in the propagation of change. The transmission of linguistic variants plays a crucial role and so does age-grading and possible (expected) and impossible (unexpected) changes. Grammaticalization can be taken as an example. As already noted, lexical words (or string of words) may acquire the properties of a functional element through grammaticalization. The conventionalization of grammaticalized elements gives rise to the expectation that these elements should eventually be adopted by the whole language community. This is because grammaticalized elements can be transmitted in a regular way as well as being introduced through reinterpretation (grammaticalization) of variants already in the language. Furthermore, these items are not expected to degrammaticalize as degrammaticalization is not as common as grammaticalization. Although factors like these may partially account for directionality in the propagation of a change, they do not provide information on the exact trajectory a change may take. Naturally, there is a sense in which a change is expected to catch on over time, especially if increased usage has been perceived in the past. However, even in the case where a new variant is catching on, the adaptation of the variant by the language community is unlikely to be in the form of a straight linear growth. Instead, a more complex pattern might emerge, a pattern that might ultimately depend on the granularity and type of information available for studying the pattern.

6.2.2 Attested trends and curves

Although changes may initially be perceived to follow a straight linear growth, they will likely exhibit a more complex pattern when viewed systematically over longer periods of time. The most common pattern in propagation of a change is an S-curve which initially shows a slow uptake of a novel linguistic variant, followed by a rapid adoption by the language community and finally the propagation slows down as the change reaches completion, Figure 6.2.

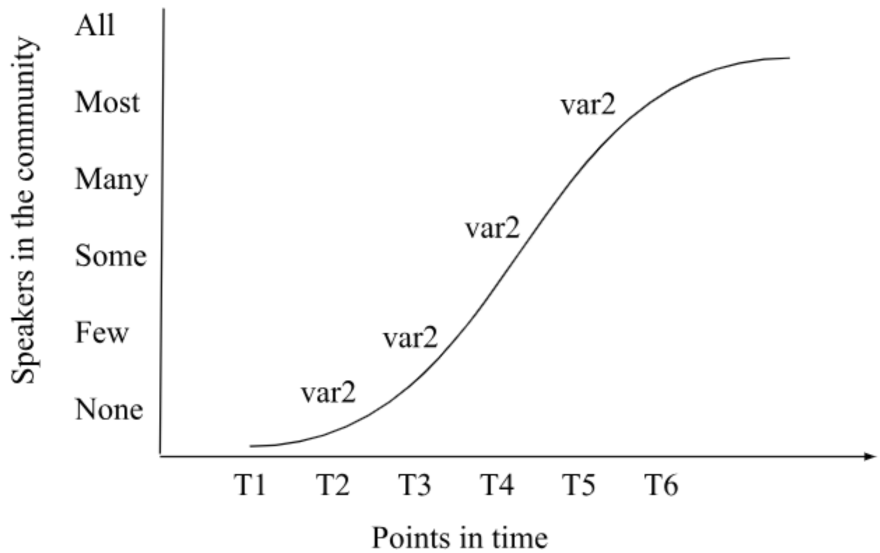


Figure 6.2. Documented directionality in the propagation of change tends to follow an S-curve.

The repeated occurrence and the importance of the S-curve in language change has been pointed out multiple times (Osgood and Seboek 1954; Bailey 1973; Kroch 1989; Pintzuk 1999; Croft 2000:183; Denison 2002; Pintzuk and Taylor 2006; Wallenberg 2009; Nevalainen 2015; Pintzuk, Taylor & Warner 2017:221; Nevalainen & Raumolin-Brunberg 2017:53–58). It has even been noted that an approximately S-shaped curve may show up on an individual level when lifespan changes are investigated (Raumolin-Brunberg 2005; Nevalainen & Raumolin-Brunberg 2012; Nevalainen 2015). As mentioned by Nevalainen & Raumolin-Brunberg (2017:79) the emergence of curves depends somewhat on how changes are analyzed, namely the material used to measure the change and whether changes are discussed in terms of binary choices (old vs. new variant) or not (on this problem see Denison 2002:59–62). Furthermore, the propagation of change does not always follow a perfectly shaped S-curve. Of the 14 changes Nevalainen & Raumolin-

Brunberg (2017:79) discuss, many do not “replicate the model in an unequivocal way” although they can still be related to it (see also Denison 2002:68 for S-curves in language change not being as uniform as sometimes claimed). Additionally, the rate of change, i.e., how quickly new variants propagate, does not appear to be consistent across the board for many changes; some are completed relatively quickly while others take a longer time to reach completion. As to when changes are considered completed or not, one might follow Nevalainen and Raumolin-Brunberg (2017:54–55) in mapping Labov’s (1994:67, 79–83) five stages of ongoing changes onto an S-shaped curve in the following way:

Incipient	below 15 per cent
New and vigorous	between 15 and 35 per cent
Mid-range	between 36 and 65 per cent
Nearing completion	between 66 and 85 per cent
Completed	over 85 per cent

Two stages, mid-range and nearing completion, can be linked to a part of an S-curve where rapid adaptation of a variant is observed. This fits with a perception of propagation that is ‘slow, slow, quick, quick, slow’ (Denison 2002:56). The reason for the emergence of this pattern may lie in the fact that variation may start among very few individuals, so the majority of the population never encounters it. Later, speakers may encounter the new variant to a greater extent, causing it to spread faster. Finally, uptake may slow down because “the number of speech events where the shift can occur diminishes” (Nevalainen & Raumolin-Brunberg 2017:53; Labov 1994:65–66).

Although S-curves (or approximation of S-curves) are commonly observed in the propagation of change, not all changes follow this trajectory. Occasionally, changes are not completed. They may show cyclic behavior where an alternation between increasing and decreasing in the novel variant is observed over time cf. Figure 6.3, or they may come to a halt, even be reverted. Discussing failed changes, Postma (2010:282–203) suggests that the these follow a bell-shaped curve, that they are related to successful S-curve patterns and can be mathematically derived from them. Although interesting and relevant for predictions that assume a certain trajectory, it is not discussed further here.

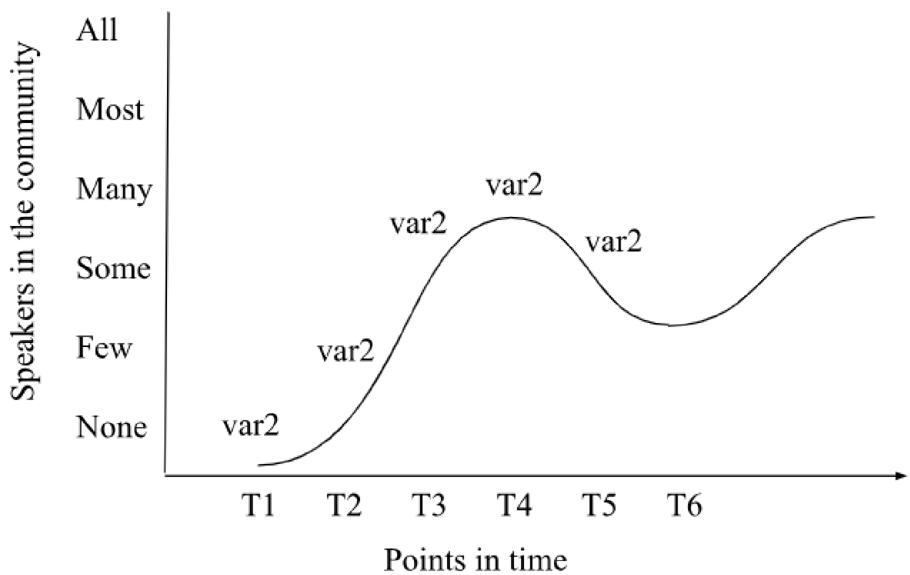


Figure 6.3. Sometimes changes may show a cyclic behavior where a repeated pattern of increase followed by decrease in the new variant is observed.

Changes documented over an extended period may be expected to show propagation in the form of an S-curve when successful.[37] However, when documented frequently over a short period, the question arises what kind of pattern may arise in addition to long-time propagation curves or stable variation. This is discussed in Section 6.3.

[37] The changes discussed by Nevalainen & Raumolin-Brunberg (2017) are documented over a period of almost three centuries, i.e., from the early 15h century (1410) to the late 17th century (ca. 1681).

Summarizing what was said above, it may be noted that language change, when considered overall, is *not* random. The non-randomness gives rise to certain expectations towards types of patterns that might emerge over an extended period of time. In case of propagation of a change, a novel variant that becomes prominent is typically expected to increase in use, i.e., unless there are special reasons to assume the opposite. Sometimes a newer variant will not catch on but continue to exist alongside an older variant in the form of stable variation. It goes without saying that it is not possible to know beforehand the exact pattern the propagation of a change will follow nor how quickly certain situations will be reached. However, with regular and systematic documentation of individual changes, e.g., through time series with fixed intervals, it is possible to use emerging patterns to predict what the future trajectory will be.

6.3 Patterns in language as seen through time series data

6.3.1 The structure of time series

A time series consists of individual data points (observations) that are sequentially gathered over some period of time (Box, Jenkins & Reinsel 2008:1; Castle, Clements, Hendry, 2019:14–15; Hyndman & Athanasopoulos 2021:17). The points may consist of aggregate data, such as how much beer was sold during a particular week, or they may represent a single measurement, e.g., the temperature in a particular place at a particular point in time. Time series may show various changes and fluctuations over time. Often, they can be viewed as being composed of sub-patterns such as trends, cycles and seasonality along with a remainder component which represents what is left when the other three have been accounted for.

The distinction between seasonality and cycles is important. Seasonality refers to fluctuations in the time-series that occur at regular, fixed-length intervals and can be linked to the time of day, week, month or year. For instance, more beer might be sold on Saturdays than other days of the week, or the mean temperature in Iceland might peak over the summer (June–July) and be lower all other months. Since seasonality is of a fixed length, it is fairly predictable and can be easily accounted for (for further information see Section 6.3.3). Cycles behave differently. Although they also consist of upwards and downwards changes over time, they are not of a fixed length and cannot be linked to the calendar year. In some ways, cycles are similar to trends which represent long-term upwards or downward change over time. Sometimes trend and cycle are lumped together into a single trend-cycle component (this is done in Hyndman & Athanasopoulos 2021:59; see also further in Section 6.3.2).

The decomposition of time-series involves identifying the key components (or sub-patterns) of the series. The value *y* of an observation at any given time *t*, can be explained by seasonality (S_t), trend-cycle (T_t) and a remainder component (R_t) at that time. The remaining component is typically expected to fluctuate around zero, reflecting random variation not captured by the other parts. The decomposition of a time-series can be either additive or multiplicative, as in (6.4). Which one is used depends on the magnitude of seasonal fluctuation and variation in the trend-cycle component. A multiplicative decomposition is appropriate when variation of the seasonal and trend-cycle component "appears to be proportional to the level of the time series" (Hyndman & Athanasopoulos 2021:64–65).

(6.4) Additive decomposition $y_t = S_t + T_t + R_t$

Multiplicative decomposition $y_t = S_t \times T_t \times R_t$

Although there is a tradition for documenting language change over extended periods of time, the observations are often few and far between, and do not necessarily occur at regular time-intervals. Constructing time series for language data where observations occur at fixed time-intervals can be quite challenging. For the dataset to be useful for time-series analysis and forecasting, it needs to include an adequate number of observations taken at an appropriate frequency. While some suggest 50–100 observations (Box, Jenkins, Reinsel & Ljung 2016:15), others maintain that there is no magic number (Hyndman & Athanasopoulos 2021:419).[38] To ensure individual observations are comparable they must be taken from a similar type of source over the whole period. This imposes restrictions on the data that can be used and where it can be obtained from. Further restrictions come from the frequency of the time series. If observations were to be recorded every ten years, with each datapoint containing aggregated 10-year information on the attestation of particular linguistic structure, one would need a minimum of 100 to 140 years of comparable data sources and measurements to construct a time series of 10 to 14 observations. If the frequency were to be lower, e.g., aggregated data for every year, one might only need data from a minimum of 10–14 years period. As argued in Chapter 5, a small time scale has many benefits (especially when it comes to obtaining information on ongoing language

[38] Hyndman & Athanasopoulos (2021:419) note that the "only theoretical limit is that we need more observations than there are parameters in our forecasting model. However, in practice, we usually need substantially more observations than that".

change in recent years) and it might provide valuable information on all sorts of fluctuations and short time changes in language over time.

A priori, one may have certain expectations towards which patterns are found in language data in light of regular time-series. Seasonality might, for instance, be expected to emerge when observing the use of certain lexical items. It would not be surprising to find that some vocabulary items tend to be used more frequently at certain times of the day than at other times. As for language change, seasonal fluctuation is generally expected to emerge as grammars are not hypothesized to change on time of day, week, month or year. Instead, trends and cycles might be observed and play an important role. These types of patterns are discussed further in the following two sections.

6.3.2 Trends and cycles

A trend is present when time-series exhibit an increase or decrease over an extended period of time. Trends may appear in combination with cyclic patterns that show repeated rises and falls in the time series. Although cyclic patterns can be somewhat regular, they differ from seasonal patterns in that they do not occur at fixed "seasonal" intervals (see Section 6.3.3).

Both cyclic patterns and trends may be observed in language data over time. In the case of language change, a new linguistic variant is generally expected to show an upwards trend (see Section 6.2), reflecting the fact that it is being transmitted and adopted by a larger and larger part of the population. Conversely, a variant that is disappearing from the language is expected to show a downwards trend over time.

In addition to trends, cyclicity can also be present in time series data that reflect the propagation of language change. Various factors, such as language purism and selection of variants, might contribute to the emergence of cycles. Sometimes changes are actively fought against but never fully eradicated, resulting in the new variant repeatedly increasing and decreasing over extended periods of time. Lexical elements or certain stylistic variation may also show this kind of pattern. They may become popular and widely used, only to fall out of use and later become popular again.

The introduction of the word *pabbabrandari* ‘dad joke’ into Icelandic shows a strong positive trend when viewed as being in competition with older words that refer to the same (or similar) concepts. The word, a compound made from the genitive singular indefinite of *pabbi* ‘dad’ and nominative singular indefinite *brandari* ‘joke’, is used to refer to silly jokes and puns often associated with dads. As Kristinsson (2023) notes, the word is seemingly a direct translation from English *dad joke* or *daddy joke*, of which the following is an example: *I'm afraid for the calendar; its days are numbered*.[39]

Pabbabrandari is fairly recent in the language and has gained much attention and popularity in the last few years, i.e., between c. 2016 and 2022. Older words that refer to the same concept include *fimmaurabrandari* ‘cheap joke (lit. five *aurar* joke)’, *aulabrandari* ‘silly joke, idiot joke’ and *orðaleikjabrandari* ‘word-game joke’. Although some might argue that *pabbabrandari* is more specialized than both *fimmaurabrandari* and *aulabrandari*, with dad jokes being a subset of the latter two (all dad jokes are silly, but not all silly jokes are dad jokes), the words can nevertheless be viewed to be in competition.

[39] At least three other variants of the calendar joke are attested on the internet: *I’m afraid of the calendar. Its days are numbered* (Twitter, Chris@HoopSpaces, Oct 12, 2023), *I’m afraid of the calendar because my days are numbered* (Twitter, niki !!!@starrydrms, January 30th, 2024) and *I fear for the calendar, its days are number* (Twitter, cee@xeeriuss, January 9th, 2023).

The introduction of *pabbabrandari* might even be regarded to represent lexical replacement. To demonstrate this, let us consider data from both Twitter and The Icelandic Gigaword Corpus.

A search on Twitter (Twitter API v2) via the R-package *academictwitteR* (Barrie & Ho 2021) targeted all forms, singular and plural, definite and indefinite, of the words *aulabrandari* and *pabbabrandari* between January 1st 2009 and December 31st 2021.[40] No results emerged for the year 2009 but both words were attested in the years that followed. The frequencies of the two words for each each year are shown in Table 6.1 and Figure 6.4. If the 12-year period is summarized, the proportion of *pabbabrandari* is about 61% and *aulabrandari* about 29%. However, when viewed on a yearly basis, it quickly becomes apparent that *pabbabrandari* takes over as the more common word in 2017 when it makes up about 60% of the examples. From 2017 onwards it continues to grow in usage. When observing the proportion of *pabbabrandari* for every year, a strong positive trend emerges. Figure 6.5 shows the proportion of *pabbabrandari* for every year from 2010 to 2021. A straight trend line with a confidence interval of 0.95 has been added to the plot using geom_smooth() from the ggplot2 package (Wickham 2016).

[40] At the time of the search, I overlooked the word *fimmaurabrandari*. When examining the data in 2023 and realizing this, I no longer had access to the Twitter API v2 and could consequently not add the word to the dataset.

Twitter: *aulabrandari* vs. *pabbabrandari*			
Year	***aulabrandari***	***pabbabrandari***	**Total**
2010	1 (100%)	0 (0%)	1
2011	5 (c. 83%)	1 (c. 17%)	6
2012	4 (100%)	0 (0%)	4
2013	3 (75%)	1 (25%)	4
2014	10 (c. 67%)	5 (c. 33%)	15
2015	22 (c. 51%)	21 (c. 49%)	43
2016	24 (c. 55%)	20 (c. 45%)	44
2017	17 (c. 40%)	26 (c. 60%)	43
2018	15 (c. 34%)	29 (c. 66%)	44
2019	11 (c. 31%)	24 (c. 69%)	35
2020	14 (25%)	42 (75%)	56
2021	11 (c. 8%)	123 (c. 93%)	134
TOTAL	**137 (c. 29%)**	**292 (c. 61%)**	**477 (100%)**

Table 6.1. Frequencies of the words *aulabandari* and *pabbabrandari* on Twitter over a 12-year period, from 2010 to 2021. If viewed as being in competition, the word *pabbabrandari* takes over in 2017 as the most common word for ‘silly jokes’.

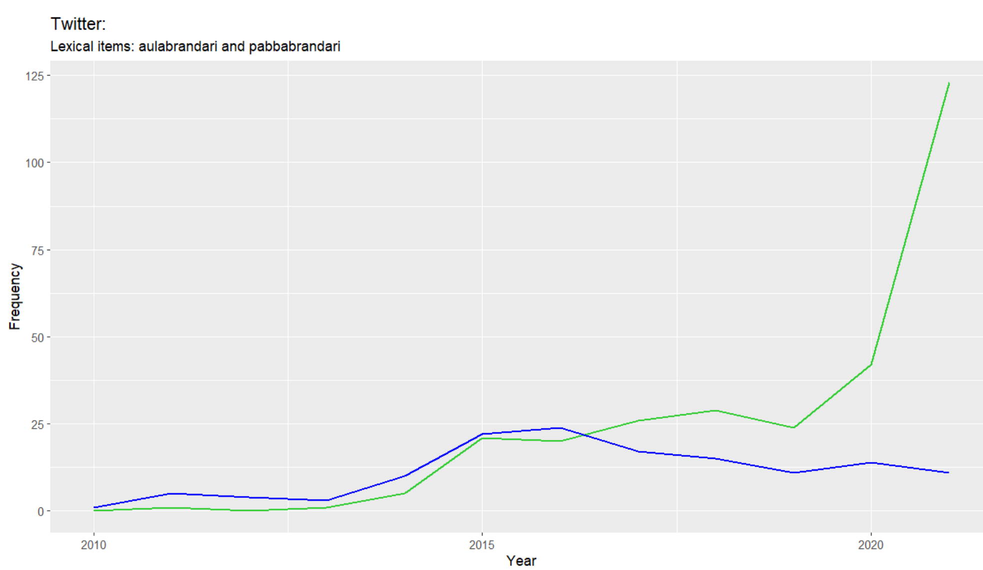


Figure 6.4. Data from Twitter shows an increased usage of *pabbabrandari* ‘dad joke’ starting around or shortly after the year 2015 and peaking after 2020. Note that an older word, *aulabrandari* ‘silly joke’ has a seemingly rather stable frequency from 2010 to 2022.

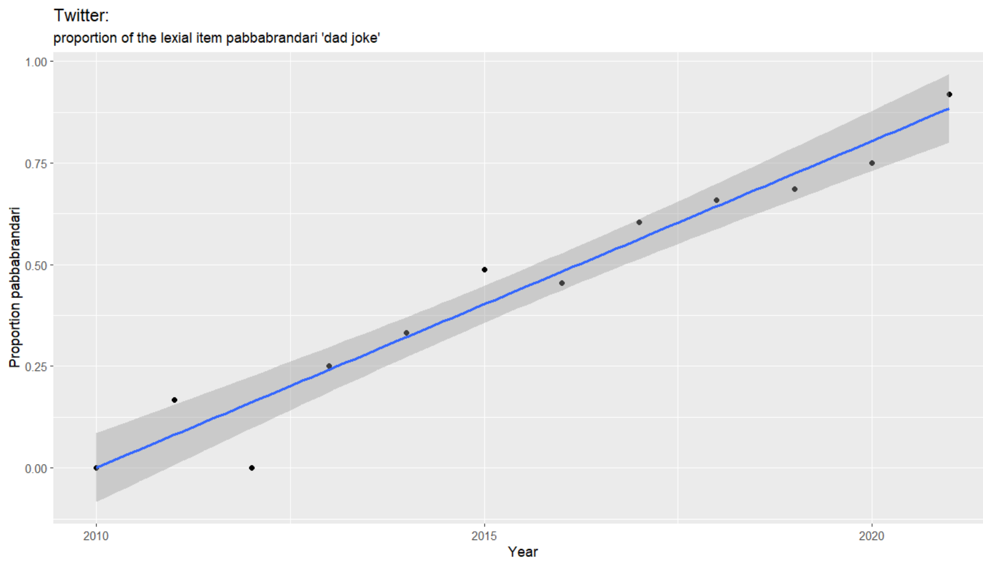


Figure 6.5. When considered to be in competition with *aulabrandari* 'silly joke', a strong positive trend is observed in the usage of the word *pabbabrandari* 'dad joke' from 2010 to 2022.

Data from the Icelandic Gigaword Corpus (IGC, rmh=2022 cf. Steingrímsson et al. 2018) provides additional insight into the use of *pabbabrandari*. The search in IGC targeted the lemma of the three lexical items *pabbabrandari*, *aulabrandari* and *fimmaurabrandari*. Of the 87 sub-sources available, 86 were used to observe the frequency of the relevant words over a 22 year period (from 2000 to 2021, both years included).[41] It should be noted that the sources in IGC contain multiple different text types, i.e, everything from copy edited newspapers to informal discussion threads on the internet (for further information on sources in IGC see Chapter 8 and 9). Given that copy edited material containing formal language is unlikely to quickly adopt certain kinds of innovation it may not be surprising that the innovative word, *pabbabrandari*, appears to have a slower uptake in IGC than on Twitter. Table 6.2 and Figure 6.6 summarize the number of examples of each lexical item

[41] In fact, all 87 sub-sources in IGC were initially selected but once data had been obtained, examples from Twitter were excluded resulting in only 86 of the sources being used. This was done because Twitter-data had already been obtained independently through Twitter API v2.

for every year. Note that *pabbabrandari* only makes up about 7% of the total examples over the whole period. However, when attestations are viewed on a yearly basis, we see that the earliest examples are from 2008 and 2009, with the word picking up in 2019 and 2020. In 2021 it exceeds both *aulabrandari* and *pabbabrandari* in raw frequency although it only makes up around 46% of the examples.

IGC: *aulabrandari* vs. *fimmaurabrandari* vs. *pabbabrandari*				
Year	***aulabrandari***	***fimmaurabrandari***	***pabbabrandari***	**Total**
2000	1 (c.14%)	6 (c. 86%)	(0 (0%)	7
2001	11 (c. 65%)	6 (c. 35%)	0 (0%)	17
2002	18 (c. 67%)	9 (c. 33%)	0 (0%)	27
2003	56 (c. 78%)	16 (c. 22%)	0 (0%)	72
2004	50 (c. 62%)	31 (c. 38%)	0 (0%)	81
2005	55 (c. 64%)	31 (c. 36%)	0 (0%)	86
2006	69 (c. 63%)	41 (c. 37%)	0 (0%)	110
2007	51 (c. 77%)	15 (c. 23%)	0 (0%)	66
2008	48 (c. 67%)	23 (c. 32%)	1 (c. 1%)	72
2009	37 (c. 63%)	20 (c. 34)	2 (c. 3%)	59
2010	29 (c. 55%)	24 (c. 45%)	0 (0%)	53
2011	17 (c. 41%)	24 (c. 59%)	0 (0%)	41
2012	15 (c. 35%)	28 (c. 65%)	0 (0%)	43
2013	8 (c. 24%)	26 (c. 76%)	0 (0%)	34
2014	16 (c. 36%)	28 (c. 62%)	1 (c. 2%)	45
2015	19 (c. 43%)	25 (c. 57%)	0 (0%)	44
2016	16 (c. 43%)	21 (c. 57%)	0 (0%)	37
2017	13 (c. 38%)	20 (c. 59%)	1 (c. 3%)	34
2018	15 (c. 19%)	60 (c. 78%)	2 (c. 3%)	77
2019	12 (c. 21,4%)	32 (c. 57%)	12 (c. 21,4%)	56
2020	6 (c. 14%)	26 (c. 60%)	11 (c. 26%)	43
2021	14 (c. 13%)	44 (c. 41%)	50 (c. 46%)	108
TOTAL	**576 (c. 48%)**	**556 (c. 46%)**	**80 (c. 7%)**	**1212 (100%)**

Table 6.2. Overview of the frequency of the three lexical items, *aulabrandari*, *fimmaurabrandari* and *pabbabrandari* in IGC from 2000 to 2021.

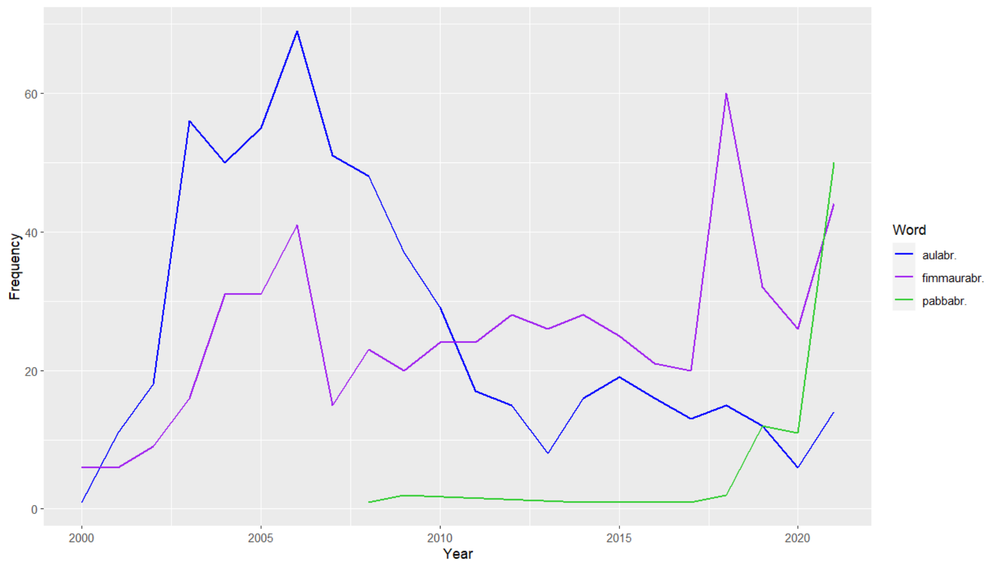


Figure 6.6. A summary of the raw frequency of *aulabrandari*, *fimmaurabrandari* and *pabbabrandari* as they appear in IGC between 2000 and 2021.

Each lexical item can be viewed proportionately to all attested examples every year, giving rise to three separate time series that show a trend-cycle pattern. Looking first at the proportions of *pabbabrandari* for every year, we can note that the lexical item is hardly attested in IGC until 2017, after which it suddenly shows an upwards trend into the 2020s. This is shown in Figure 6.7 where a LOESS smoothing line with a confidence interval of 0.95 has been added to the plot using geom_smooth() from the ggplot2 package in R.

Changes in the proportion of *fimmaurabrandari* 'cheap joke' look very different from that of *pabbabrandari*. As observable in Figure 6.7, the lexical item accounts for the majority of the examples in the year 2000 (out of seven attestations, *fimmaurabrandari* occurred six times). Until around 2007, a downward trend in the use of the item is observed, followed by an upward trend until around 2015. In or shortly after 2015 a second downward trend occurs, continuing into the 2020s. The "ups" and "downs" give rise to a cyclic pattern which is not linked to seasonality but occurs somewhat irregularly over longer periods of time. Interestingly, the initial downward trend in the use of *fimmaurabrandari* is mirrored by an upward trend in the use of *aulabrandari*, shown in Figure 6.9. The usage of *aulabrandari* peaks around 2005, after which it shows a continued downward trend into the 2020s.

Viewing the three time series as parts of a single story, it might be hypothesized that that up until ca. 2005 the words *fimmaurabrandari* and *aulabrandari* were in competition, with younger generations favoring *aulabrandari.* The mid 2010's witnessed increasing usage of a new lexical item *pabbabrandari* which was modeled after the English *dad joke*. Around the same time, *fimmaurabrandari* shows a mild come-back, possibly as a witness to older generations partaking in discussions around 'silly jokes' which are now typically referred to as *pabbabrandarar*.

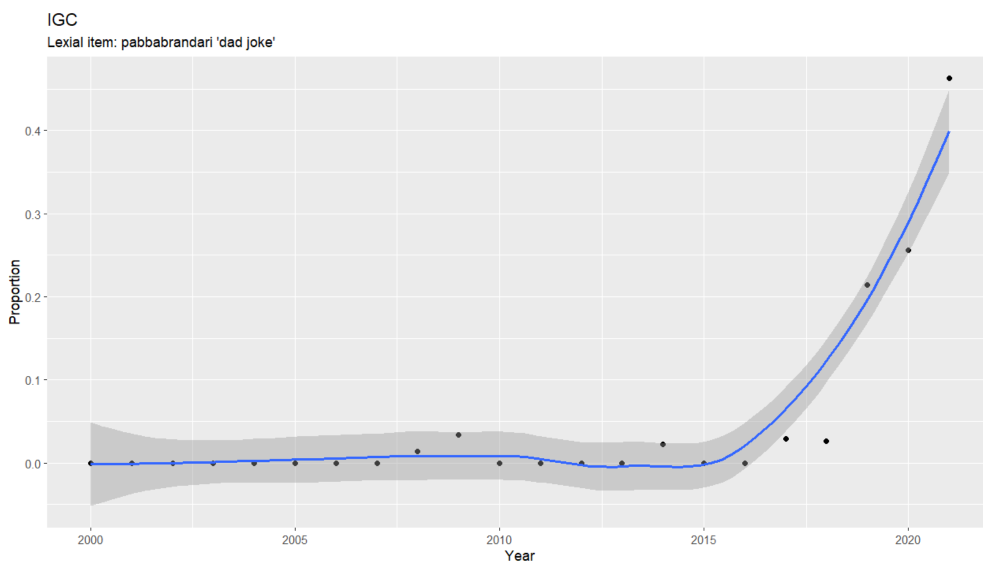


Figure 6.7. The lexical item *pabbabrandari* is hardly attested from the early 2000s (there are one or two examples here and there, cf. Figure 6.6. It is not until in and after 2019 that it starts picking up and showing an upwards trend into the 2020s. A LOESS smoothing line has been added to the plot to show the general trend of the usage of the word.

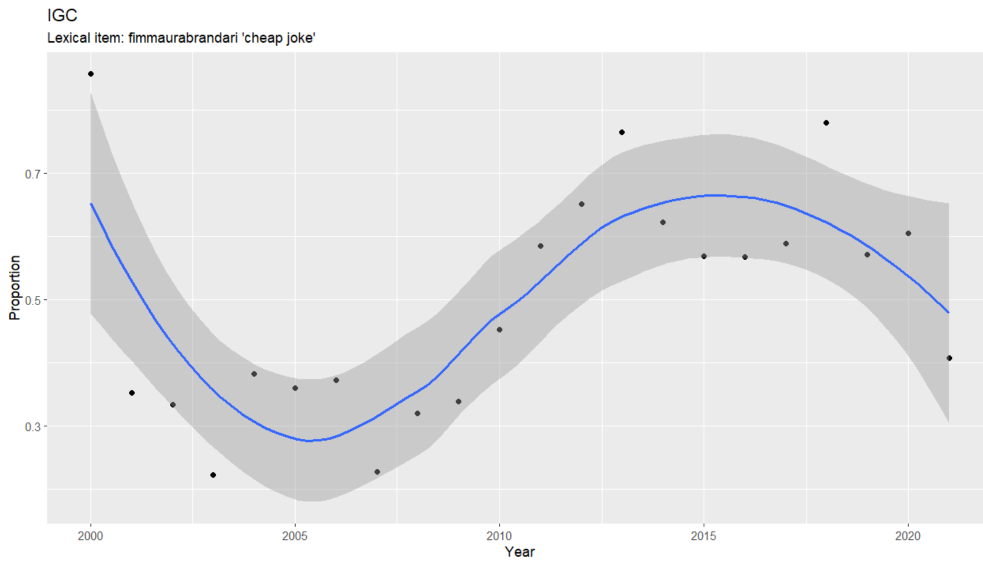


Figure 6.8. The lexical item *fimmaurabrandari* appears to show a somewhat cyclic pattern. From 2000 to 2005 there is a downwards trend, followed by an upwards trend to 2015 and then a second downwards trend into the 2020s. A LOESS smoothing line has been added to the plot to show the general trend in the usage of the word.

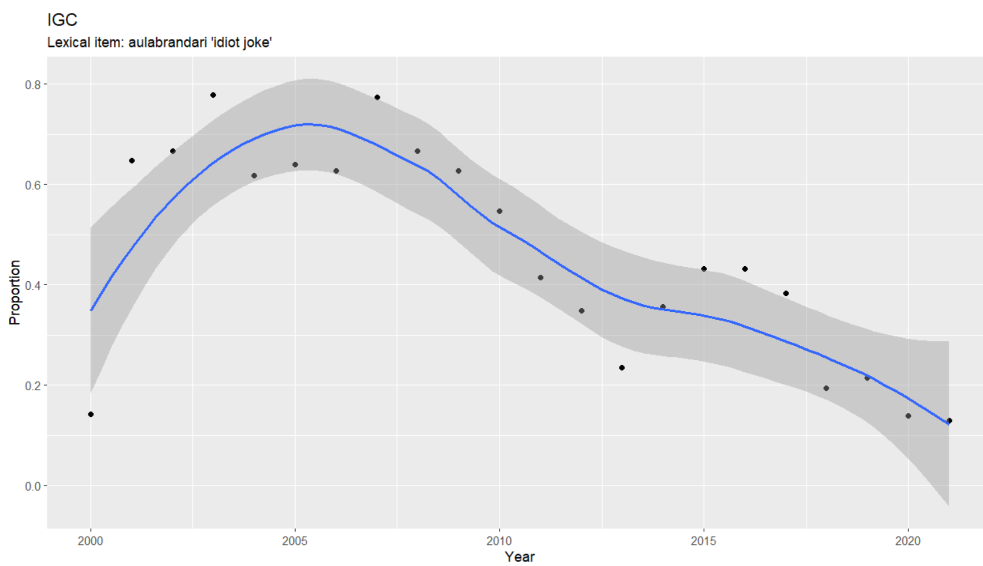


Figure 6.9. From 2000 to 2005 there appears to be an upwards trend in the use of the lexical item *aulabrandari*, followed by a downward trend that continues into the 2020s. A LOESS smoothing line has been added to the plot to show the general trend in usage of the word.

The question is whether the introduction and propagation of *pabbabrandari* ‘dad joke’ into Icelandic represents a true instance of lexical replacement or not. As already noted, some individuals maintain that the new word is not necessarily completely identical in meaning to *aulabrandari* and *fimmaurabrandari*. It has been suggested (e.g., Kroch 1989; Wallenberg 2019) that when two structures have a similar meaning, they may either compete for functional space or specialize. Based on this, one may want to keep the possibility open that *aulabrandari* and *fimmaurabrandari* are in direct competition which may result in lexical replacement. In fact, there are several reasons to lean towards this option.

Using the online library portal timarit.is it is possible to construct a chronology of the usage of the three words. The oldest word is *fimmaurabrandari*, being attested since 1948 (*Þjóðviljinn* 1948, April 27th). Note that the word itself is closely linked to the Icelandic currency system (*Icelandic króna*) where *eyrir* (pl. *aurar*) was a coin representing a hundredth of a *króna*. Naturally, the word *fimmaurabrandari* is only transparent to individuals who are familiar with *aurar*. Since the distribution of *aurar* stopped in the late 80's and early 90's (Numismatic collection of the Central Bank and National Museum of Iceland 2002) and the coins ceased being accepted as currency in the early 2000's, individuals born around and after that time may be considered unlikely to adopt the word *fimmaurabrandari*. Interestingly, this fits well with the introduction of *aulabrandari* which was first attested in 1976 (*Vísir* 1976, July 22nd).

As to the rise of the word *pabbabrandari* in the 2010's, increased usage of the word can be linked to how prominent the phenomenon has been in public discussions in recent years. The saliency is not just obvious in Iceland but also abroad, reflected in multiple online articles on the phenomenon. The quick adaptation of the word *pabbabrandari* may have spurred a slight increase in the usage of *fimmaurabrandari*. At least some individuals partaking in early discussions on 'dad jokes' can be assumed to have been between the ages of 45 and 60. If this is the case, they must have been born between 1955 and 1970, placing them in a time where *aurar* were still being issued and used. It would thus not be surprising if they were more familiar with the lexical item *fimmaurabrandari* than the newer word *pabbabrandari*. The diminishing use of *aulabrandari* remains to be explained but might be linked to English influence.

The example taken above, i.e., the documentation of the usage of three lexical items over a 12 (Twitter) to 22 (IGC) year period, serves the purpose of pointing out trends and cycles in time series data. Ultimately, trends and cycles (or a trend-cycle component, when the two are treated together) are the patterns that matter the most in language change. They capture both the general direction of changes as well as some "fluctuations" that may be relevant. Interestingly, it is possible to observe changes in directionality in the usage of some lexical items in less than a few years. Note that this is a shorter period than a generation, i.e., when 'generation' is taken to refer to a period of 15–30 years. The short-time changes highlight the importance of observing language use rather than just acquisition. They also suggest that it might be informative to track and document language change over relatively short periods of time.

6.3.3 Seasonality

Time series sometimes show regular variation that can be linked to seasonality, for instance the time of day, week, month, or year. Seasonal variations always occur at fixed time intervals and are repetitive and predictable (Hyndman & Athanasopoulos 2021:59); they are different from cyclic patterns discussed in Section 6.3.4 which do not occur at fixed time intervals.

Time series that record phenomena like temperature or sales figures frequently show seasonality. In these cases, the variation is linked to the time of day or the time of year. Figure 6.10 shows the monthly mean temperatures recorded at Stórhöfði between the years 2000 and 2015. Note that lower temperatures systematically occur around the beginning and end of each year, while higher temperatures are recorded during the summer months.

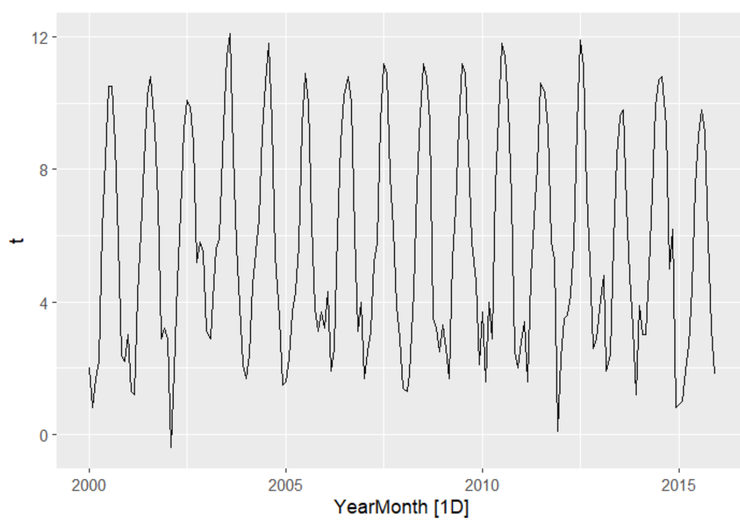


Figure 6.10. Monthly mean temperature (given in °C) at Stórhöfði, recorded from 2000 to 2015 (data obtained from the Icelandic Meteorological Office https://vedur.is/). Observe the seasonal fluctuations.

Seasonal fluctuations may occur for various reasons all of which can be linked (directly or indirectly) to the time of day or the time of year. Air temperature, for instance, is determined by multiple factors, including incoming solar radiation which in turn is affected by the rotation of the earth. Sales figures reflect human behavior which is often influenced by the weather and events in the real world. The question here is whether seasonality should be expected in language data.

In order for language data to show seasonal patterns, regular and consistent fluctuations must be present. It is not enough for the series to show time related variation once or twice. Rather, the variation must be regular and occur at fixed time intervals. In this context, it is useful to distinguish between grammar and language use. It is, for instance, hard to picture a scenario where the mental grammar of an individual is subject to seasonal variation. An individual who always rejects certain syntactic structures in the early morning while finding those same structures consistently grammatical in the afternoon simply sounds fantastical. However, the use of certain lexical items may – and indeed does – show seasonality. Below are some examples ranging from daily to yearly seasonality.

Yearly seasonality emerges quite expectedly in data reflecting language use. Many cultural phenomena such as festivals, travels, and school related events are directly linked to the calendar year. It is only natural that topics related to these events should be more frequently discussed leading up to the time of the events until shortly afterwards. For instance, people may talk more about Christmas in November, December and at the beginning of January than they do in May or June. Words such as *jólaskraut* 'Christmas ornament', *jólapakki* 'Christmas present' and *jólaandi* 'Christmas spirit' are thus likely to correlate with the Christmas period. Figure 6.11 shows the raw frequency of the word *jólapakki* 'Christmas present' every month from January 1998 to December 2019. The data was obtained from IGC.[42]

[42] The search in IGC targeted the lemma *jólapakki* which means that all forms of the word, nominative, accusative, dative and genitive both singular and plural were included.

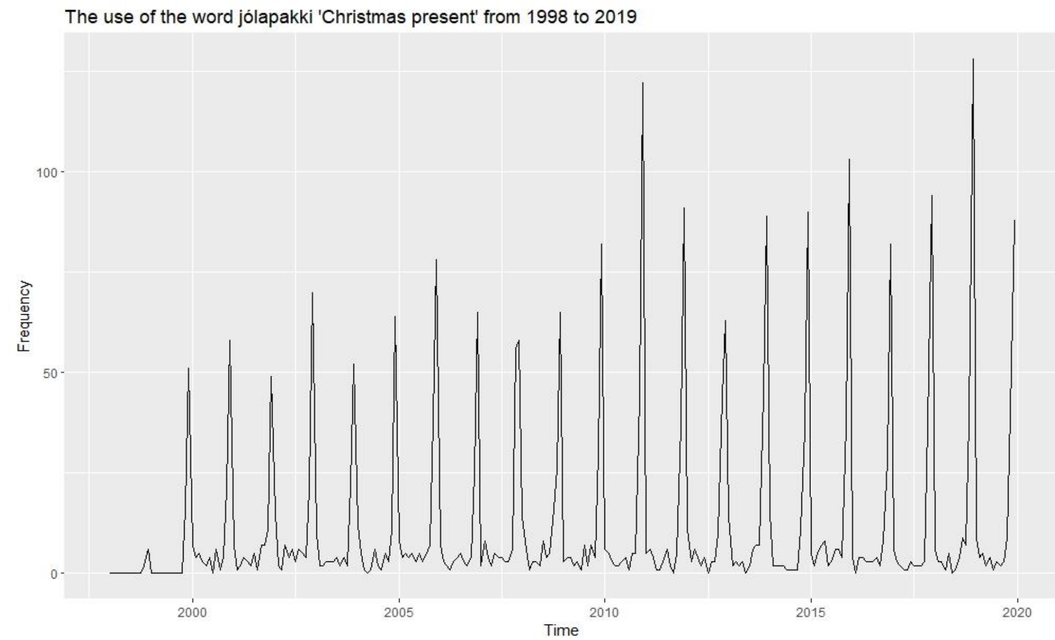


Figure 6.11. Monthly frequency of the word *jólapakki* ‘Christmas present’ from January 1998 to December 2019.

Notice the regularities in the pattern in Figure 6.11. This is due to the word *jólapakki* mostly appearing in written material between November and January. During the other nine months (February to October) the word occurs rarely or not at all. Although it may seem like the overall trend is for the word to increase in usage from year to year (compare the end of year 1999 with roughly 50 occurrences of *jólapakki* to the end of year 2018 with over 125 occurrences), this might simply be due to the corpus having more material in later years.

Much like in the case of the word *jólapakki* ‘christmas present’ which occurs more frequently in the time shortly before and after Christmas, the use of words related to monthly events might correlate with the occurrence of those events. Examples of phenomena that might be expected to give rise to monthly seasonality in Icelandic language data include monthly salaries, bills, horoscopes and the publication of some periodicals, cf. (6.5).

(6.5) *fullt tungl* ‘full moon’, *leiga* ‘rent’, *mánaðarleg tímarit* ‘monthly periodicals’, *reikningar* ‘bills’, *útborgað* ‘paying of salaries’, *útborgunardagur* ‘payday’.

In an attempt to establish monthly seasonality in language data a search was made on Twitter for some of the items in (6.5) for the period 2020–2022. Unfortunately, no seasonal patterns were detected. This need not disprove the existence of monthly seasonality in language data, but could simply be due to there not being enough data for the seasonality to emerge. The results for *útborgunardagur* ‘payday’ might be in favor of this assumption. While the word occurs very rarely on Twitter, the few instances of it are clustered around the end and beginning of each month.

Unlike monthly seasonality, weekly seasonality is readily detected in Icelandic language data on Twitter. Examples are provided featuring two lexical items: *föstudagur* ‘Friday’ and *fössari* ‘Friday (slang)’. Although both words refer to Friday they are generally not considered to be in direct competition. The word *föstudagur* is the regular, unmarked lexical item while *fössari* is a recent slang, typically used when individuals want to capture the feeling of there really being a Friday (Kristinsson 2023).

For *föstudagur*, only information on the nominative singular form (*föstudagur*) was gathered. For *fössari*, both the nominative singular (*fössari*) as well as accusative, dative and genitive singular (*fössara*) were obtained. Note that the data is not scaled or transformed in any way. The results simply show the raw frequency of occurrence.

Figure 6.12. Shows the raw frequency of the word *fössari* on Twitter from January 2022 to December 2022. Although not immediately obvious, the use of the word spikes roughly every seven days, suggesting weekly seasonality. This becomes clearer once each week of 2022 is plotted on top of each other on the same figure, showing the frequency one day at a time as in the seasonal plot in Figure 6.13.

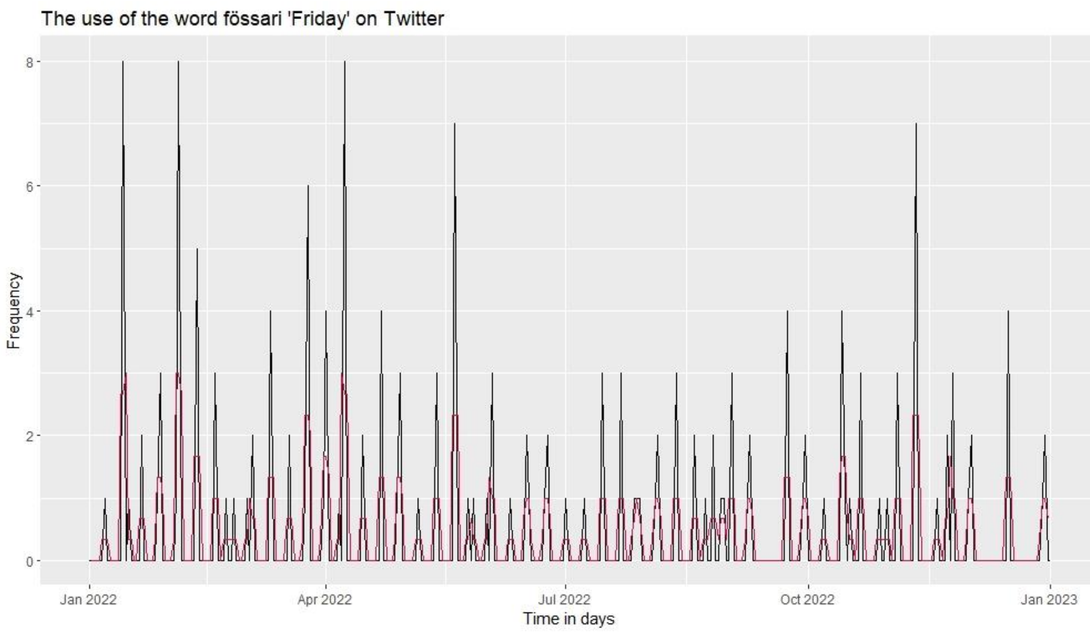


Figure 6.12. The word *fössari*, slang for ‘Friday’, on Twitter in 2022. The raw frequency is plotted in black; a moving average of order 3 is plotted in pink, highlighting regularities in the pattern.

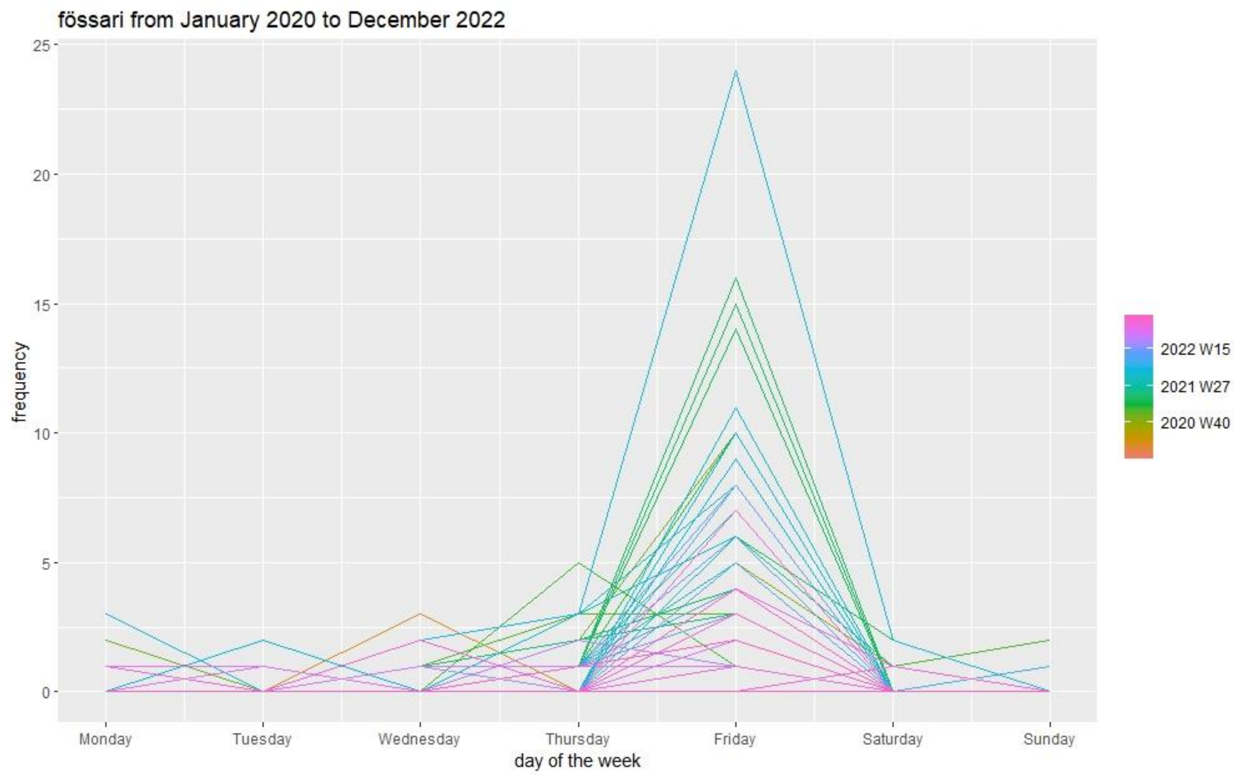


Figure 6.13. The frequency of the word *fössari* 'Friday (slang)' on Twitter plotted for every day of the week from January 2020 to December 2022.

It might be difficult to conceive of daily seasonality arising in language data. For such a phenomenon to emerge, a pattern must show regularities that correlate with the same time every day. Interestingly, the use of the word *kvöld* 'evening' does exactly this. A search on Twitter targeted all indefinite singular forms of the word, i.e., *kvöld* (nominative/accusative), *kvöldi* (dative) and *kvölds* (genitive). Results for the period January 2020 to December 2022 yielded a total of 46.855 examples. These are shown on a hourly time series graph in Figure 6.14.

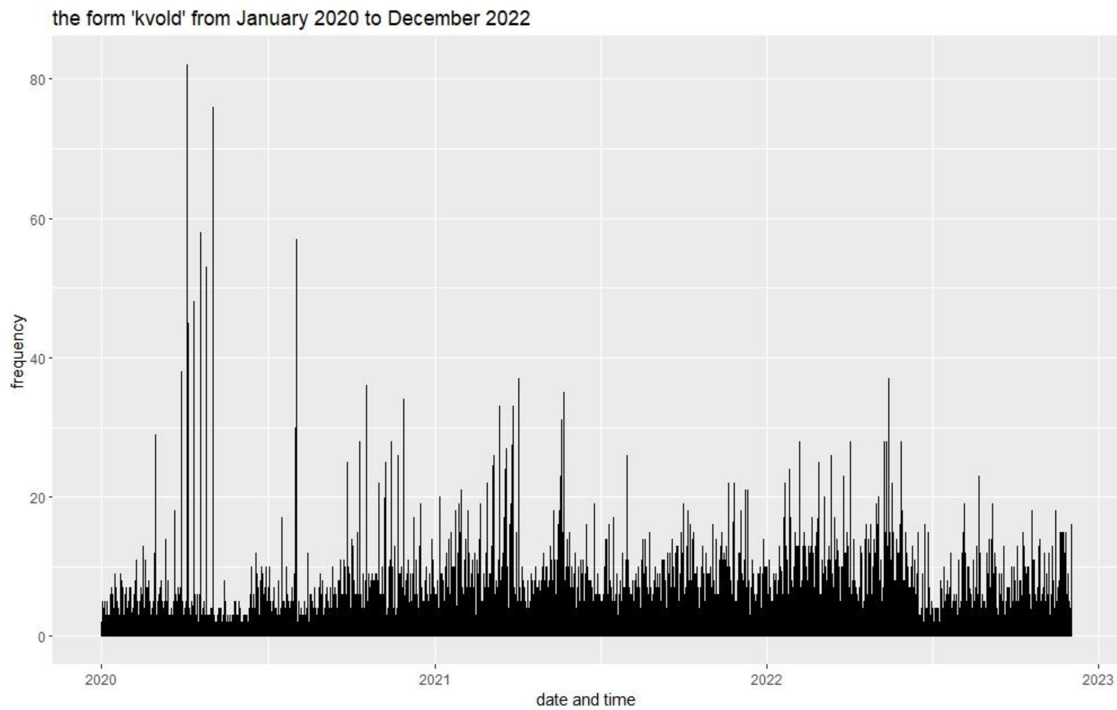


Figure 6.14. Hourly frequency of the indefinite singular forms of *kvöld* ‘evening’ in all cases (nominative, accusative, dative, genitive) over the period January 2020 to December 2022.

When each week of the three year period (2020–2022) is plotted independently, hour-by-hour, a seasonal pattern emerges, shown in 6.15. Here the frequency of the word is the highest in the early afternoon of each day and is followed by a sharp drop in use over nights. Also note the presence of weekly seasonality where a spike in usage of *kvöld* occurs repeatedly on Saturdays.

.

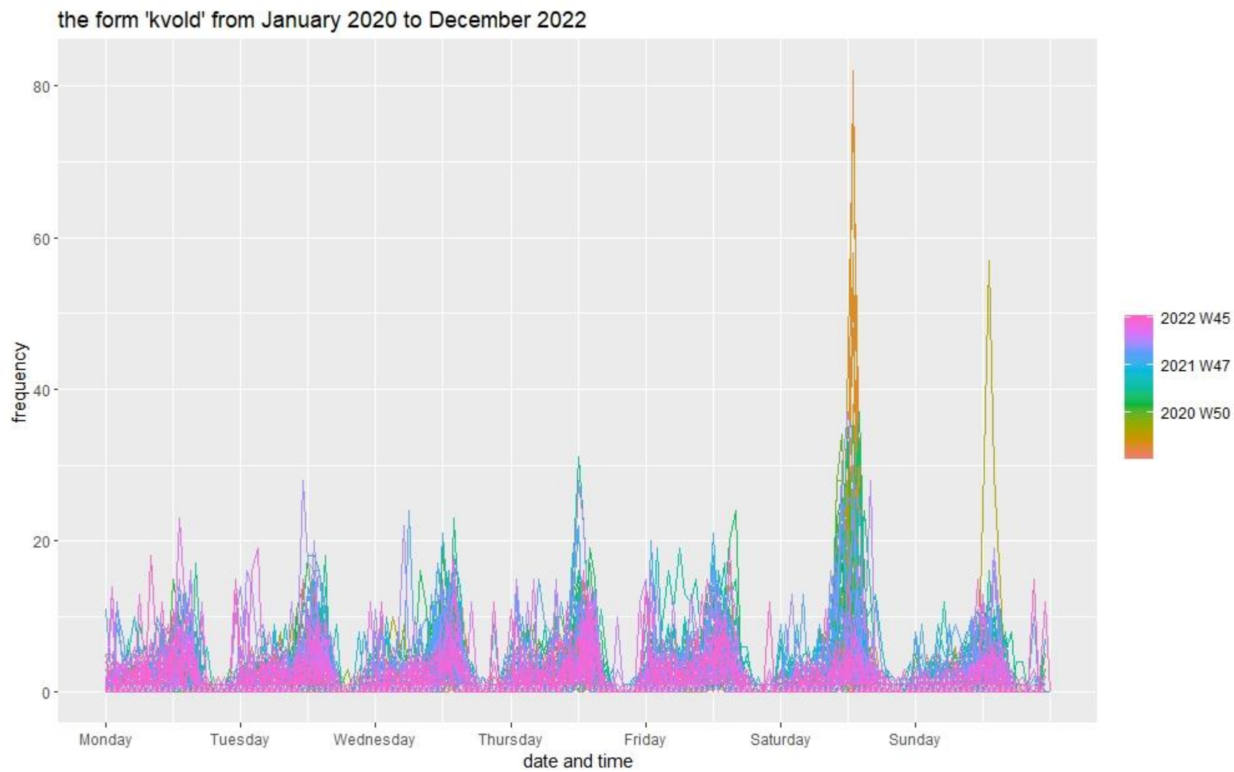


Figure 6.15. The frequency of the word *kvöld* 'evening' on Twitter from January 2020 to December 2022. Each week (Monday to Sunday) is plotted hour-by-hour on top of each other.

When the same data points are plotted hour-by-hour over the period of a single day the daily seasonality becomes more noticeable, Figure 6.16. While there is certainly some variation over the course of the day, a peak in use is nevertheless detected from shortly before noon into the early afternoon. This suggests that individuals on Twitter use the indefinite singular of the word *kvöld* 'evening' in all cases more during a 5-hour period between 10:00 to 15:00 than at any other time of the day.

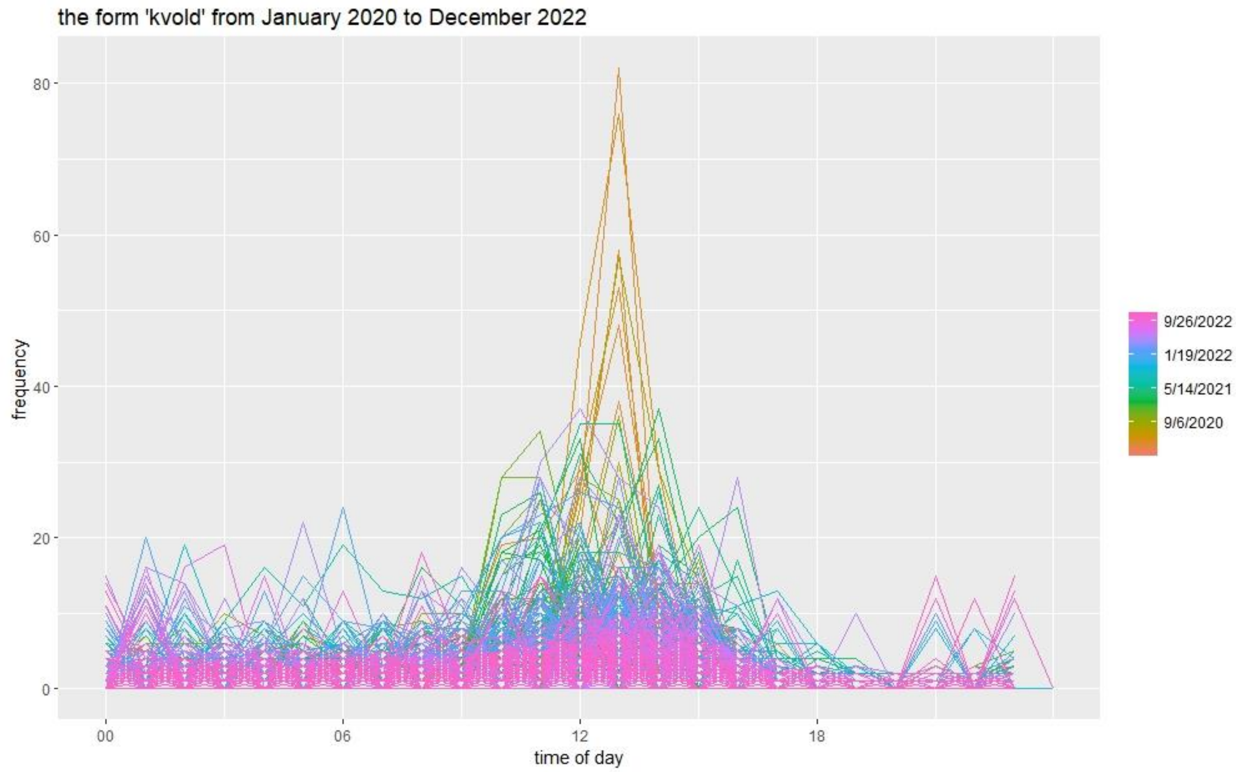


Figure 6.16. The frequency of the word *kvöld* 'evening' on Twitter from January 2020 to December 2022. Each day is plotted hour-by-hour on top of each other.

All the examples above (the use of the words *jólapakki* 'Christmas present', *föstudagur* 'Friday', *fössari* 'Friday (slang)' and *kvöld* 'evening') show that everything from daily to yearly seasonality can be detected in data reflecting language use. As for sub-daily seasonality, the existence of such a pattern seems not applicable. First of all, access to data of such granularity is needed to figure out if hourly patterns exist. Second, it seems highly unlikely that a pattern in language use would repeat itself every hour. Should such a pattern exist it would be quite surprising.

Note that the seasonal patterns discussed in this subsection all apply to the use of certain lexical items and ***not*** to grammatical phenomena. In other words, they do not cover things such as variation in grammaticality judgments, lexical replacement, case marking with certain predicates, changes in functional categories or any other such phenomenon.

Although seasonality may not turn out to be an essential factor in language change, it is important to be aware of such patterns in the data. When measuring language change over time, there are chances that seasonality might emerge as a result of language use. As already noted (Chapter 3), language use plays a crucial role in transmission of linguistic variants and therefore it is important for the study of language change.

6.4 Extrapolating from emerging patterns

Already established patterns of propagation give a valuable insight into how various types of changes diffuse through a community of speakers. The most common pattern found in propagation of change is an S-curve (or an approximation of an S-curve) which has been documented for numerous changes and is typically observed over an extended period of time, i.e., over decades to hundreds of years (see Section 6.2). However, not all changes follow an S-curve. Occasionally changes do not catch on but continue to exist in the form of stable variation or show some form of cyclicity when innovative variants go through periods of increase and decrease. In some cases, changes may even be reverted and disappear. For these reasons it may not always be ideal to simply assume a certain development a priori. It may prove useful to study emerging patterns on a smaller time scale and use those to extrapolate from into the future.

In Chapter 5, it was argued that changes should be documented at regular short-time intervals, such as over months, quarters, and years. Tracking change in this fashion gives rise to regular time series which may provide new insight into propagation of change and language data in general. As discussed in Section 6.3, individual observations in time series (y_t) can be regarded to be composed of sub-patterns or component parts. These components are trend-cycle component, seasonal component, and a remainder component. The trend-cycle component is no doubt the most important one for language change as it can show general trajectories of propagation over time. While a seasonal component is not expected to emerge in data on language change,[43] it does show up in language data that represents language use. For instance, words like *jólapakki* 'Christmas present' can be shown to mostly occur in printed material during the months leading up to Christmas and in early January. At other times of the year the word is either absent or occurs very rarely, thus showing yearly seasonality. Weekly seasonality can be detected in words such as *föstudagur* 'Friday' and *fössari* 'Friday (slang)' where data from Twitter shows that these tend to (unsurprisingly) occur most frequently on Fridays. Even daily seasonality may be found in data representing language use. As an example, the word *kvöld* 'evening' mostly shows up in the second half of the day. Needless to say, the emergence of seasonality in online data is due to the behavior of individuals and not to how language works. There is nothing inherent about language that calls for the use of *föstudagur* 'Friday' more on Fridays than any other days of the week. It is simply the case that Fridays appear to be culturally important; it is the last day of the typical work week, occurring right before the weekend. It is possible that words such as *miðvikudagur* 'Wednesday' could also show

[43] Although seasonality will unlikely show up in data documenting language change, certain types of time-series decompositions will nevertheless hypothesize a seasonal component in case a time series consists of weekly, quarterly, or monthly observations. However, the seasonal component is so small that it does not find its way into time series modeling and forecasting (see further in discussion of the time series in Chapter 8 and 9).

seasonality, but this is less certain as people (at least in Iceland) may have less reasons to openly talk about Wednesdays consistently on a particular day of the week.

Turning back to language change, time series data gathered at regular intervals (monthly, quarterly, yearly) may prove useful when studying propagation of change. Not only is it possible to learn more about both language use and language change from gathering this type of data, but it also offers a somewhat neutral way of approaching change over time. This is because they do not assume that changes will propagate in any particular way, but rather allow for patterns (and changes in patterns) to unfold over time. By documenting changes at relatively short intervals, one may learn more about patterns that have already been established (such as S-curves, cycles, stable variation and failed changes) and something of lesser known patterns such as seasonality and how much randomness may appear in data over time. This may lead to a better understanding of the distinction between variation in language usage over time and long-time trends in language change over time. It may also provide information on speed of changes, how quickly the direction of propagation may change and how language use affects propagation of change. In short, time series data may provide a new perspective on the propagation of language change through the creation of a new type of data. Such data is also compatible with forecasting methods relying on regular time series and can thus be used to make predictions about changes in the future. Using time series for forecasting can lead to a better understanding of the phenomenon of study and of the type of data needed to generate better forecasts. Finally, such data is also useful in evaluating forecasts and the status of changes at a particular period of time.

Studying patterns in time series data and using them to produce forecasts does not mean that commonly observed patterns in language change have no role in the forecasting process. On the contrary, they can be used indirectly to evaluate forecasts in such a way that the output of a forecasting model is compared to expectation based on what has previously been noted on the trajectory of change. In other words, they can function as a way of "checking" if the forecasts match what is previously known about language change. Thus, one may ask questions such as whether emerging patterns are in accordance with expectations towards propagation of change (e.g. do they show an emerging S-curve) or, if not, *why* unexpected patterns appear and *what role* they may play in language transmission, language use and language change.

7 Preliminaries

7.1 The forecasting situation

7.1.1 The purpose of the forecasting

Forecasts are usually generated with a specific purpose and a particular audience in mind. The purpose and the audience are usually intrinsically linked as the goal of forecasts is typically to assist users with decision-making on some level (e.g., Hoff 1983:3; Makridakis, Wheelwright & Hyndman 2018:2–3; Castle, Clements & Hendry 2019:50; Hyndman & Athanasopoulos 2021:14–15). As discussed in Chapter 2, this is not always the case for language forecasting. While language forecasting can be useful for language planning and revitalization, it was argued to also be a method to study language change.

The forecasts presented in the following chapters (see Chapter 8, Sections 8.5.2–8.5.4 for *á bak við* and *við hliðina á*, and Chapter 9, Section 9.5.2–9.5.4 for *hlakka til*), along with description of methodologies and data (Chapter 8, Section 8.4 and Chapter 9, Section 9.4), are directed towards linguists who are interested in language change, the propagation of change, and predictions made about a future state of a language. The purpose of the forecasts is to test commonly used forecasting methods on available language data in an attempt to predict the state of affairs in the future. The methods are based on the general assumption that the past contains information on what the future has in store. The models that are used rely on regular time series and may incorporate exponential smoothing or autoregression (see 7.2.4). The hope is that the results will contribute towards better understanding of language forecasting, especially regarding what type of data can be used for these purposes, what factors need to be taken into consideration and how far into the future forecasts can be made. As already discussed (Chapter 6), studies on language change have given rise to expectations towards how changes come about and how they might propagate through a language community. Documentation of variation and change through regular time series allows for (re)evaluation of these expectations. Additionally, by generating predictions about the future, expectations can be further evaluated, and new ones generated.

7.1.2 The context of the forecasting

No forecast is made in a vacuum or without a context. The forecasts presented in the following chapters presuppose several factors relevant for their interpretation. The most important of these is the continued transmission of the Icelandic language and the relevant linguistic structures to new generations of native speakers. If this condition is not met the forecast will fail, not because the methods or the data were inappropriate, but because an unexpected event, not factored into the forecast, occurred. It can be useful to distinguish between failures due to unexpected events and failure based on bad data or inappropriate modeling, as the latter is typically more informative about the forecasting procedure including data gathering and annotation. To illustrate the difference further, one might imagine an analogy to the copying of manuscripts. Each time a text in a manuscript is copied, changes (including errors) may be introduced. Some errors might be predictable. When a manuscript has been copied, observed errors can be compared to predicted errors. Any method and data used to predict the errors can then be evaluated in the light of what ended up happening. In the case the manuscript was not copied at all, none of the predictions will be born out. Since the reason for the predictions not being born out can be traced to the manuscript not being copied at all, there is no way of evaluating any of the forecasting methods or the data used to generate the prediction. The predictions are based on copying taking place. Similarly, predictions regarding the propagation of particular structures within a given language presuppose a continuous transmission of that language.

For the case studies presented here (see Chapters 8 and 9) it is not enough for the language to be transmitted, it must also continue to be *used*. If speakers suddenly opt to avoid the relevant structures altogether, naturally occurring E-language data will not contain the structure under investigation and predictions cannot be checked against attested observations. Thus, unless otherwise specified, forecasts based on E-language data tacitly assume continued transmission and use of particular structures.[44]

Other factors that are presupposed and matter for the interpretation of the forecasts include the assumption that the language data that is used to generate a forecast is, in some way or another, representative of the language community, and that emerging patterns in the time-series are not random or meaningless. Unlike the previous two factors, i.e., that the language must continue to be transmitted and used, the meaningfulness of the data can be tied to how variation and change is measured. Language data sampled at very short intervals may show seasonal- or trend-cycle patterns that do not necessarily contain obvious signals relevant for language change. Conversely, data gathered at very long intervals may miss out on important information. The amount of data and frequency of a construction of interest may also play a role in whether a meaningful signal emerges from historical data. There are, of course, numerous ways in which variation can be measured

[44] An exception from this is when a forecast explicitly accounts for diminishing use or disappearance of a particular language or a structure within a language.

(see Chapter 5). For the case studies presented here, documentation is based on the proportion of the innovative variant(s) versus other possible variants at a given time.

To sum up, the factors in (7.1) are tacitly assumed to hold for the forecasts produced in the following chapters. Note that the first two can be thought of as social factors that represent necessary conditions for the forecast to be interpreted. If (7.1-i) and (7.1-ii) are not met, the forecast will necessarily fail.

(7.1) i. The language continues to be transmitted and used (sociolinguistic factor)
ii. Relevant structures continue to be transmitted and used (sociolinguistic factor)
iii. The data is representative of the situation (nature of data)
iv. Emerging patterns are meaningful (data and methodology)

The Factors in (7.1-iii) and (7.1-iv) are of a different nature than those in (7.1-i)–(7.1.ii), being linked to the quality of the data and the general forecasting methodology. If the factors in (7.1-iii) and (7.1-iv) do not hold, the forecast can still be evaluated and may lead to improvement in these areas.

One might ask how likely the assumptions in (7.1-i)–(7.1-ii) are to hold for a language like Icelandic that is spoken by less than 400,000 people. There is no straightforward answer. Scholarly literature (Hilmarsson-Dunn & Kristinsson 2010) and recent discussion on social media such as Facebook and in newspaper articles point towards the status of Icelandic not being as strong as it used to be. For instance, concerns have been raised about Icelandic youth not understanding various vocabulary items (for a discussion see Rögnvaldsson 2023, October 18th), some of which might be claimed to be quite "normal but perhaps a bit old fashioned" (Sæmundsson 2022, December 31st). Related to this, recent results from the Programme for International Student Assessment (PISA) suggest that only 53% of Icelandic boys aged 15 have basic reading comprehension skills in Icelandic (Sigurjónsdóttir 2023, December 5th) so it does perhaps not come as a surprise that vocabulary may affected. Furthermore, it has been pointed out that the domain of language use might be shrinking for Icelandic (Rögnvaldsson 2020). English is becoming more prominent in advertisements of all sorts, and at least one company has recently chosen to change the name of one of their products from Icelandic to something that sounds more international, i.e., from *Toppur* to *Bon Aqua* (Online news article 2023, June 30th). Rögnvaldsson, who has been very forward in discussing the current situation of Icelandic, has noted that the use of foreign languages has increased substantially in the last few years. He notes, echoing Rask's 1813 foreboding prediction about the future of Icelandic, that "if everything continues as it is, it is not inconceivable, even likely, that English will have succeeded Icelandic as the main language of communication in the country around middle of this century" (Rögnvaldsson 2023, May 12th).[45] Despite these concerns, data gathering, annotating and forecasting in this dissertation is done under the assumption that Icelandic will continue to be transmitted and used in the next two decades, at least.

Currently, no limits have been established on the length of the forecasting horizon for the propagation of language change. In Chapter 5, it was suggested that time series data might be used for short-range language forecasts, potentially reaching into the mid-range. Short-range forecasts were defined as being shorter than 15 years and mid-range forecasts

[45] Translation mine. Original text is as follows: "En ef svo fer fram sem horfir er alls ekki óhugsandi, og jafnvel líklegt, að um miðja þessa öld muni enska hafa tekið við af íslensku sem aðalsamskiptamálið í landinu."

were considered to apply to everything between 15 to 30 years. Beyond 30 years was claimed to belong to long-range forecasting. These definitions will no doubt need to be revised in coming years, depending on the accuracy of times series forecasting with language data.

Forecasts based on time series are made one or more steps ahead, where each step is equal to the time interval between individual observations in the series. In case of a yearly series, one step ahead prediction provides values for one year into the future. In case of a quarterly series, the one step ahead prediction reaches one quarter into the future. Forecasts tend to become more uncertain the further ahead predictions are made. Thus, forecasting 15 periods into the future, which is within the claimed short-range of language forecasting, can give rise to inaccuracies. How (in)accurate the short-range language forecast might depend on whether the time series consists of monthly, quarterly or yearly observations.[46] The frequency of observations in the time series matters for how far predictions can be made (Castle, Clements & Hendry 2019:15). Forecasts made 15 steps ahead may cover different lengths of time depending on the frequency of observation. For yearly time series, 15 steps equal 15 years. For quarterly series, it amounts to little less than four years. For monthly time series, 15 steps equal one year and three months. If predictions are to be made regarding the situation of a particular linguistic phenomenon in the next 20 years, a 20 step ahead prediction for yearly series would be needed, 80 step ahead prediction for quarterly series and 240 step ahead prediction for monthly series. It stands to reason that anything above 15–30 steps ahead might result in a poor forecast accuracy, i.e., for the predictions furthest in the future.

[46] Inaccuracies will also depend on whether “raw” time series are used to make predictions or whether some form of smoothing or extraction of trend is done.

When preparing to make predictions about the future, a time series is usually split into a training set and a test set. The training set is used to choose and fit a model while the test set is used to evaluate how well the model does at forecasting. According to convention, the test set should typically include about 20% of the time series and "should ideally be at least as large as the maximum forecast horizon required" (Hyndman & Athanasopoulos 2021:135). Both of these general rules of thumbs are flouted in the present study. The training period used here is about 10% of the series and predictions are consciously made for anything between 15 to 30 steps into the future. There are several reasons for doing this. First, the time series used here already contain relatively few observations that cover a short time span (from ca. 11 to 19 years) and setting aside observations risks making the series too short to be used for time series analysis and forecasting. Second, setting aside large amounts of data from an already short time series means recent trends might not be accounted for in model selection. Third, language forecasting is in its infancy, and it is still unknown how many steps into the future it makes sense to project language data. Fourth, there is nothing to lose by projecting further into the future than just three or four steps. Rather, information might be gained about where (how many steps into the future) things start to go horribly wrong.

7.2 Forecasting methods

7.2.1 General background, workflow, and model choice

The methods used here fall under non-explanatory statistical approaches and they rely on historical data in the form of regular time series (e.g. Makridakis & Wheelwright 1978:14–16). Typically, time series consist of sequential observations equally spaced in time (e.g., Box, Jenkins & Reinsel 2008:1).[47] Importantly, more recent data values are thought to be dependent on previous values. Thus, since time series analysis and forecasting deals with picking out patterns in the series and extrapolating them into the future, there is no such thing as one model fits all. Models are selected based on the properties of the time series of interest. However, which family of models is to be used, i.e., which general approach one adopts, can be chosen beforehand.

For the case studies and forecasts in Chapters 8 and 9, a tidy workflow was followed, as recommended by Hyndman & Athanasopoulos (2021:105). This involves i) preparing the data, ii) visualizing the data, iii) specifying a model, iv) fitting the model to training data, v) evaluating the outcome and, finally, vi) forecasting. The majority of these steps were done using R (R Core team 2021) and RStudio (Posit team 2023) with the package fpp3 (Hyndman 2023) which attaches a number of additional R packages relevant for forecasting and visualization, e.g., fable (O'Hara-Wild, Hyndman & Wang 2021) and ggplot2 (Wickham 2016). Descriptions of models and evaluations of fit and forecasts are adopted from Hyndman & Athanasopooulos (2021).

Data preparation was, no doubt, the most time-consuming task of the six steps in the workflow. It was done between January 2021 and June 2023 and involved extracting data from the Icelandic Gigaword Corpus (rmh=2019, rmh=2022 cf. Steingrímsson et al. 2018) as well as from Twitter (https://twitter.com/), cleaning the data, annotating it and making sure it was in the correct format for time series analysis, i.e., it had to contain one observation for each time period with the observations being equally spaced in time. Observations were based on an aggregated number of examples for each time period and

[47] Any series that contains observations ordered sequentially in time is considered a time series (e.g., Box, Jenkins & Reinsel 2008:1). Regular time series contain observations equally spaced in time.

documented the proportion of a novel variant. As an example, if a structure were to be attested 200 times in Q1 of a particular year, having an innovative feature 50 times, the observation for Q1 of that year would be 50/200 or 0.25 (25%). A more detailed description of the data gathering and annotation can be found in Chapter 8, Section 8.4 and Chapter 9, Section 9.4. Visualization of important aspects of the overall data is also provided there.

As already noted, a time series is typically split into a training set and a test set before forecasts are made. Values fitted to training data are referred to as fitted values. Forecasted values are those that predict the test data and any future period. The fit of a model to the training data can be evaluated based on residuals, i.e, the difference between observed values and fitted values (7.2). For evaluating how well a model does at predicting new observations, forecast errors (7.3) are used.

(7.2) residuals

$$e_t = y_t - \hat{y}_t$$

(7.3) forecast errors

$$e_{T+h} = y_{T+h} - \hat{y}_{T+h|T}$$

Further information on how models and forecasts are evaluated are provided in section 7.2.5. Figure 7.1 summarizes the difference between training data, test data, fitted values, forecasted values, residuals and forecast errors.

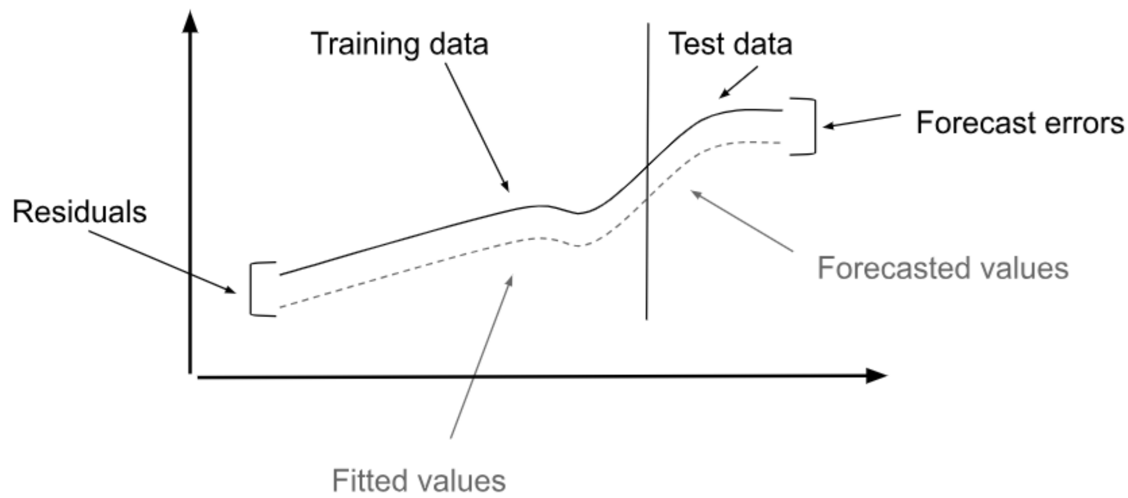


Figure 7.1. In preparation of forecasting, a time series is split into training and test data. A model is fitted to the training data (fitted values) and used to generate a forecast for the period of the test data.

Before a model is fitted to the training data, there is a choice of using the raw time series or transforming or changing the series in some way. For instance, moving averages might be used to smooth out the data and eliminate noise that is considered unimportant. The higher the order of the moving average part, the smoother the data. While moving averages can be useful to smooth data, they tend to somewhat shorten the time series.[48] Since the time series used in the present study are already relatively short, a choice was made to use the raw series. In some cases, a log transformation was used in the forecasting to ensure predictions would stay on a positive scale (Hyndman & Athanasopoulos 2021:411–413) as it makes little sense to obtain negative predictions when the values concern proportion of

[48] I make a distinction here between using moving averages to smooth data before forecasting (data preparation) and using moving averages as a part of forecasting. Forecasting methods that rely on time series decomposition may use moving averages to estimate the trend-cycle component (Hyndman & Athanasopoulos 2021:69–75). Other forecasting methods, e.g., ARIMA, may use moving averages of error terms as a part of predicting a future value (Hyndman & Athanasopoulos 2021:276–277).

examples featuring an innovative variant. When doing this, forecasts are computed on a transformed series and then automatically back-transformed to the original scale.

As already mentioned (Chapter 4), various methods exist for making forecasts, ranging from informal guessing to complex formal forecasting systems (see discussion in Castle et al., 2019:22; Hyndman & Athanasopoulos, 2018:12-14). The methods used here rely on regular time series and focus on extrapolating some aspects of the historical data into the future. The methods are non-explanatory in that they only take into account already-observed values of the phenomenon of interest and do not incorporate other factors that might affect propagation of change. The assumption is, of course, that the future will in some way be similar to the past, i.e., similar to today, yesterday and so on.

Three to four simple models were fitted to the training data and used to predict observations for the test period. For time series with yearly observations, three basic models were fitted, a Naïve model, a Mean model, and a Drift model. For time series with quarterly observations, a Seasonal naïve model was also used. For explanations of these, see Section 7.2.3. The simple models served as benchmark forecasting methods, i.e, they were used to compare more complex models to. Only if the complex models performed better at predicting new values than the simple models, were they used to generate forecasts.[49] Explanations of how model fit, point forecasts and forecast distributions are evaluated are found in Sections 7.2.5 and 7.2.6, respectively.

Once simple models had been fitted to the training data and used to predict the test data, more complex models were tested. These were innovations state space models for exponential smoothing (ETS for Error, Trend, Seasonal) and Autoregressive Integrated

[49] In at least one instance, a more complex method was used to generate a forecast for future time periods when a simple model provided a more accurate prediction for the test period.

Moving Average models (ARIMA) that rely on lagged-time values of the phenomenon of interest. Both approaches involve studying patterns in the time-series and extrapolating them into the future. Exponential smoothing methods rely on weighted averages of past observations with the weight decreasing exponentially the older the observations are. ARIMA models focus on autocorrelation using time-lagged values of the phenomenon to be forecasted (e.g., Hyndman & Athanasopoulos, 2018; Makridakis et al., 1978). In general, ARIMA models as well as exponential smoothing methods are widely used for forecasting and are appropriate for most, if not all, types of time-series data. They can and have been used on everything from economics and spread of Covid 19 to behavior of people, for instance when it comes to beer consumption. ETS and ARIMA models are explained further in Section 7.2.4.

7.2.2 Time series decomposition and other features

As described earlier (Chapter 6), time series can be decomposed into a trend, cycle, and seasonal component. Sometimes the trend and cycle are lumped together into a trend-cycle component or simply trend (this is done in Hyndman & Athanasopoulos 2021). Thus, a value y at a given time t in a time series, can be explained by seasonality (S_t), trend (T_t) and a remainder component (R_t) at the relevant point in time. The remainder component is simply what is left over when trend and seasonality have been removed. The decomposition can be either additive or multiplicative, see (7.4), and which one is used depends on the magnitude of seasonal fluctuation and variation concerning trend and cycle.

(7.4) Additive decomposition $y_t = S_t + T_t + R_t$

Multiplicative decomposition $y_t = S_t \times T_t \times R_t$

Many forecasting methods can take trend, seasonal and cyclic patterns into account when producing forecasts. This includes exponential smoothing discussed in section 7.2.4. Instead of classical decomposition, Seasonal and Trend decomposition using Loess (STL) is used to show the decomposition of the relevant time series in Chapters 8 and 9. This is also the decomposition used for generating forecasts that only take into account the trend-cycle component of the time series. The STL decomposition (Cleveland et al. 1990) is convenient as it allows for the smoothness of the trend-cycle to be controlled. For the case studies in Chapters 8 and 9, the trend window was set to 13 to obtain a very smooth trend-cycle component. Consequently, the remainder component, which was considered to mostly represent "noise" in measurements, is larger than when using a smaller trend-cycle window. STL also allows for the seasonal component to either change over time or be set to a fixed window. Since seasonal variation is hypothesized to not play an important role in language change over time, the window was set to "periodic" to ensure it would stay the same for all observations.

In addition to showing the STL decomposition of individual series in Chapter 8 and 9, some general features of the series are also noted. The strength of a trend F_T is based on STL decomposition and given in a value between 0 and 1. It is calculated as in (7.5) where T refers to the smoothed trend component and R to the remainder component. When the data has a strong trend, the variation of the remainder component (R_t) should be smaller than the seasonally adjusted data (T_t + R_t), but for data with little trend these should be roughly the same.

(7.5) $$F_T = max\left(0,1 - \frac{Var(R_t)}{Var(T_t+R_t)}\right)$$

(Hyndman & Athanasopoulos 2021:92)

A correlogram displaying autocorrelation coefficients is used to show autocorrelation in the time series. An autocorrelation coefficient r_k at lag k gives the relationship between observation y_t and lagged value y_{t-k} (7.6).

(7.6) $$r_k = \frac{\sum_{t=k+1}^{T}(y_t - \bar{y}_t)(y_{t-k} - \bar{y})}{\sum_{t=1}^{T}(y_t - \bar{y})^2}$$

(Hyndman & Athanasopoulos 2021:52)

For determining whether a time series can be regarded as white noise or not, a Ljung-Box test can be used to evaluate whether autocorrelation is within acceptable limits or not (see Section 7.2.5).

For determining whether a series is stationary, it is sometimes enough to visualize the series and observe whether there is an obvious positive or negative trend. When the level of a series changes over time, e.g., the series shows a negative or positive trend, the series is non-stationary. Otherwise, it is stationary. A Kwiatkowski–Phillips–Schmidt–Shin test (KPSS, (Kwiatkowski et al. 1992) can be used to determine whether a series is stationary or not.

Features of time series are obtained by using a relevant function in R. For more detailed discussion on how features are calculated and interpreted, see Hyndman & Athanasopoulos (2021).

7.2.3 Some simple methods: Naïve, Seasonal naïve, Mean and Drift

Simple forecasting methods are often used as benchmark methods to compare more complex forecasting models. If a more complex model is to be used, it needs to perform better than the simple forecasting methods, for instance by having better residuals, returning better point forecasts and better forecasting distribution (see sections 7.2.5 and 7.2.6). Four simple methods are considered here. These are a Naïve method, a Seasonal naïve method, Average method, and Drift method (see Hyndman & Athanasopoulos 2021:110–112). The first one, the Naïve method, assumes that all future observations will be the same as the last recorded observation in the time series. Thus, if the last recorded observation shows that an innovative variant is attested 65% of the time, all future periods will be hypothesized to also have the innovative variant 65% of the time. A variant of the Naïve model takes seasonality into account and assumes that future values are equal to the last observed value of the same period. With quarterly data, this means that the first quarter of all upcoming years will have the same value as the first quarter of the last recorded year, the second quarter of all upcoming years will have the same value as the second quarter of the last recorded year and so on. The equations for the Naïve and Seasonal naïve models are provided in (7.7). As earlier (section 7.2.1), observed values are denoted by y and forecasted values by $\hat{y}$. The number of steps into the future are given by h, the last observed time is T. The seasonal period m (which in case of quarterly time series is 4) and k is (h-1)/m (see Hyndman & Athanasopoulos 2021:110–111).

(7.7) Naïve: $\hat{y}_{T-h|T} = y_T$

Seasonal naïve: $\hat{y}_{T-h|T} = y_{T+h-m(k+1)}$

A third simple method, the Average or Mean model, assumes that all forecasted values in the future will be equal to the mean of the whole series. This is represented in (7.8) where every observation in the series from y_1 to y_T is summed up and divided by the number of observations in the series.

(7.8) Mean: $\hat{y}_{T-h|T} = (y_1 + \ldots y_T)/\,T$

Finally, the Drift method assumes that changes over time are equal to the average change in the historical data, cf. (7.9).

(7.9) Drift: $\hat{y}_{T+h|T} = y_T + \frac{h}{T-1}\sum_{t=2}^{T}(y_t - y_{t-1}) = y_T + h(\frac{y_T - y_1}{T-1})$

It sometimes happens that simple forecasting methods give more accurate predictions than complex models. For instance, if a series has a lot of variation without any noticeable trend or seasonality, a Mean model may end up producing a forecast closer to observed values than a more complex model that takes into account autoregression. Essentially, the future remains unknown until it arrives and any model is simply an estimation of what might happen.

7.2.4 More complex models: ETS and ARIMA

Models used for forecasting in Chapter 8 and 9 that are more complex than those introduced in Section 7.2.3 come from two families of commonly used forecasting methods. These are i) innovations state space models that rely on exponential smoothing and factor in trend and seasonality (referred to as ETS for error, trend, and season), and ii) ARIMA (autoregressive integrated moving average) models that deal with autocorrelation in the data (Hyndman and Athanasopoulos 2021:265). Both approaches seek to project some aspect of the historical data into the future. These are now explained in turn.

Forecasting methods involving exponential smoothing (Brown, 1959; Holt, 1957; Winters, 1960) do not incorporate external explanatory values into forecasting.[50] Rather, they attempt to capture general patterns in the data by relying on past values. Although exponential smoothing methods come in a few different flavors, they all rely on weighted averages of previous observations that decrease exponentially the further back in time the observations are. Thus, future values are considered to be composed of varying weights of previous values. If the smoothing parameter α, having a value between $0 \leq \alpha \leq 1$, controls the weight of previous observations, the smoothing formula can be written as in (7.10) where $\hat{y}$ is a forecasted value, y is an observed value and T is the time of the last observation in the time series.

[50] By *external explanatory values*, I mean real-world factors that might affect the phenomenon of interest. In the case of language forecasting this might be the prestige of different linguistic variants, normative pressure, number of second language speakers and so on. The "explanatory" factors in exponential smoothing (if they can be referred to as explanatory) come from previous observations of the phenomenon of interest.

(7.10) $\hat{y}_{T+1|T} = \alpha y_T + \alpha(1-\alpha) y_{T-1} + \alpha(1-\alpha)^2 y_{T-2} + \cdots$

(from Hyndman and Athanasopoulos 2021:229)

As noted earlier (Section 7.2.2) observations in a time series can be decomposed into a trend, seasonal and a remainder component. Forecasting models based on exponential smoothing range from being quite simple to more complex, depending on how the trend and seasonal factor are accounted for in the model, and "the way in which these enter the smoothing method" (Hyndman and Athanasopoulos 2021:228). For instance, simple exponential smoothing involves only a forecasting equation and a single smoothing equation. More complex methods also make use of equations that deal with linear trends and/or seasonality (overview in e.g., Hyndman & Athanasopoulos, 2018). For ETS models (error, trend, season) with prediction intervals, error terms are also included. Like trend and seasonality, these can be either additive (A) or multiplicative (M). When conveying which type of exponential smoothing model is used, it is convenient to indicate how errors, trend and seasonality are accounted for. An ETS(M,A,N) model has multiplicative errors (M), additive trend (A) and no (N) seasonality. This particular model is commonly referred to as Holt's linear method with multiplicative errors.

Equations for ETS forecasting models with additive and multiplicative errors are shown in Table 7.1. Values at time *t* are given by y_t. Note that the value y_t always consists of (estimated) smoothed value (or level) ℓ_t at time *t* and estimated ε_t error at time *t*. For methods with trend, b_t is the estimated trend at time *t* and β is the smoothing parameter for the trend. For methods with seasonality, s_t is the estimated seasonality at time *t* and γ is the smoothing parameter for the seasonality. The damping parameter for models with dampening, is given by ϕ.

ADDITIVE ERROR MODELS

Trend	Seasonal N	Seasonal A	Seasonal M
N	$y_t = \ell_{t-1} + \varepsilon_t$ $\ell_t = \ell_{t-1} + \alpha\varepsilon_t$	$y_t = \ell_{t-1} + s_{t-m} + \varepsilon_t$ $\ell_t = \ell_{t-1} + \alpha\varepsilon_t$ $s_t = s_{t-m} + \gamma\varepsilon_t$	$y_t = \ell_{t-1}s_{t-m} + \varepsilon_t$ $\ell_t = \ell_{t-1} + \alpha\varepsilon_t/s_{t-m}$ $s_t = s_{t-m} + \gamma\varepsilon_t/\ell_{t-1}$
A	$y_t = \ell_{t-1} + b_{t-1} + \varepsilon_t$ $\ell_t = \ell_{t-1} + b_{t-1} + \alpha\varepsilon_t$ $b_t = b_{t-1} + \beta\varepsilon_t$	$y_t = \ell_{t-1} + b_{t-1} + s_{t-m} + \varepsilon_t$ $\ell_t = \ell_{t-1} + b_{t-1} + \alpha\varepsilon_t$ $b_t = b_{t-1} + \beta\varepsilon_t$ $s_t = s_{t-m} + \gamma\varepsilon_t$	$y_t = (\ell_{t-1} + b_{t-1})s_{t-m} + \varepsilon_t$ $\ell_t = \ell_{t-1} + b_{t-1} + \alpha\varepsilon_t/s_{t-m}$ $b_t = b_{t-1} + \beta\varepsilon_t/s_{t-m}$ $s_t = s_{t-m} + \gamma\varepsilon_t/(\ell_{t-1} + b_{t-1})$
A_d	$y_t = \ell_{t-1} + \phi b_{t-1} + \varepsilon_t$ $\ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha\varepsilon_t$ $b_t = \phi b_{t-1} + \beta\varepsilon_t$	$y_t = \ell_{t-1} + \phi b_{t-1} + s_{t-m} + \varepsilon_t$ $\ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha\varepsilon_t$ $b_t = \phi b_{t-1} + \beta\varepsilon_t$ $s_t = s_{t-m} + \gamma\varepsilon_t$	$y_t = (\ell_{t-1} + \phi b_{t-1})s_{t-m} + \varepsilon_t$ $\ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha\varepsilon_t/s_{t-m}$ $b_t = \phi b_{t-1} + \beta\varepsilon_t/s_{t-m}$ $s_t = s_{t-m} + \gamma\varepsilon_t/(\ell_{t-1} + \phi b_{t-1})$

MULTIPLICATIVE ERROR MODELS

Trend	Seasonal N	Seasonal A	Seasonal M
N	$y_t = \ell_{t-1}(1 + \varepsilon_t)$ $\ell_t = \ell_{t-1}(1 + \alpha\varepsilon_t)$	$y_t = (\ell_{t-1} + s_{t-m})(1 + \varepsilon_t)$ $\ell_t = \ell_{t-1} + \alpha(\ell_{t-1} + s_{t-m})\varepsilon_t$ $s_t = s_{t-m} + \gamma(\ell_{t-1} + s_{t-m})\varepsilon_t$	$y_t = \ell_{t-1}s_{t-m}(1 + \varepsilon_t)$ $\ell_t = \ell_{t-1}(1 + \alpha\varepsilon_t)$ $s_t = s_{t-m}(1 + \gamma\varepsilon_t)$
A	$y_t = (\ell_{t-1} + b_{t-1})(1 + \varepsilon_t)$ $\ell_t = (\ell_{t-1} + b_{t-1})(1 + \alpha\varepsilon_t)$ $b_t = b_{t-1} + \beta(\ell_{t-1} + b_{t-1})\varepsilon_t$	$y_t = (\ell_{t-1} + b_{t-1} + s_{t-m})(1 + \varepsilon_t)$ $\ell_t = \ell_{t-1} + b_{t-1} + \alpha(\ell_{t-1} + b_{t-1} + s_{t-m})\varepsilon_t$ $b_t = b_{t-1} + \beta(\ell_{t-1} + b_{t-1} + s_{t-m})\varepsilon_t$ $s_t = s_{t-m} + \gamma(\ell_{t-1} + b_{t-1} + s_{t-m})\varepsilon_t$	$y_t = (\ell_{t-1} + b_{t-1})s_{t-m}(1 + \varepsilon_t)$ $\ell_t = (\ell_{t-1} + b_{t-1})(1 + \alpha\varepsilon_t)$ $b_t = b_{t-1} + \beta(\ell_{t-1} + b_{t-1})\varepsilon_t$ $s_t = s_{t-m}(1 + \gamma\varepsilon_t)$
A_d	$y_t = (\ell_{t-1} + \phi b_{t-1})(1 + \varepsilon_t)$ $\ell_t = (\ell_{t-1} + \phi b_{t-1})(1 + \alpha\varepsilon_t)$ $b_t = \phi b_{t-1} + \beta(\ell_{t-1} + \phi b_{t-1})\varepsilon_t$	$y_t = (\ell_{t-1} + \phi b_{t-1} + s_{t-m})(1 + \varepsilon_t)$ $\ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha(\ell_{t-1} + \phi b_{t-1} + s_{t-m})\varepsilon_t$ $b_t = \phi b_{t-1} + \beta(\ell_{t-1} + \phi b_{t-1} + s_{t-m})\varepsilon_t$ $s_t = s_{t-m} + \gamma(\ell_{t-1} + \phi b_{t-1} + s_{t-m})\varepsilon_t$	$y_t = (\ell_{t-1} + \phi b_{t-1})s_{t-m}(1 + \varepsilon_t)$ $\ell_t = (\ell_{t-1} + \phi b_{t-1})(1 + \alpha\varepsilon_t)$ $b_t = \phi b_{t-1} + \beta(\ell_{t-1} + \phi b_{t-1})\varepsilon_t$ $s_t = s_{t-m}(1 + \gamma\varepsilon_t)$

Table 7.1. Equations for ETS models. *A* stands for additive, *M* for multiplicative and *N* for none. A lowercase *d* in A_d refers to dampening (from Hyndman & Athanasopoulos 2021:253).

Since seasonality is not hypothesized to play a role in language change, language forecasting will likely only rely on models with no seasonality (N). In fact, the quarterly data in Chapters 8 and 9 had such a small seasonal component that it never found its way into any of the ETS or ARIMA models that were used.

An appropriate ETS model can be selected based on minimizing Akaike's Information Criterion (7.11), the corrected Akaike's Information Criterion for small sample bias (7.12) or the Bayesian Information Criterion (7.13). For the case studies in Chapter 8 and 9, the selection of ETS models was done automatically in R with the function ETS(). The function chooses a model by minimizing AICc (cf. Hyndman & Athanasopoulos 2021:255).

(7.11) Akaike's Information Criterion

$$\text{AIC} = -2\log(L)+2k$$

(7.12) AIC corrected for small sample bias

$$\text{AIC}_c = \text{AIC} + \frac{2k(k+1)}{T-k-1}$$

(7.13) Bayesian Information Criterion

$$\text{BIC} = \text{AIC} + k[\log(T) - 2]$$

Aside from choosing an appropriate model, the smoothing parameters along with the initial value of the level ℓ need to be estimated. This is done by minimizing the sum of squared errors (7.14) where the errors are the residuals from the model fit, i.e., residuals *e* at time *t* equals the difference between observed and fitted value: $e_t = y_t - \hat{y}_{t|t-1}$.

(7.15) Sum of squared errors

$$\text{SSE} = \sum_{t=1}^{T} (y_t - y_{t|t-1|})^2 = \sum_{t=1}^{T} e_t^2$$

As already noted, EST models assume the future is in some ways similar to the past. They are designed to capture patterns in the time series, such as trend, seasonality and error, and extrapolate them into the future. Due to the nature of the approach there may be a wide error range, resulting in the forecasting distribution becoming relatively large. However, this is not necessarily a problem as point forecasts are produced along with prediction intervals and the whole forecast is interpreted in the context of the relevant language change (see 7.2.7).

Unlike ETS models, autoregressive integrated moving average (ARIMA) models rely on autocorrelation in the data to obtain future values. Essentially, the models involve studying the relationship between values of the same variable at different times and using the information to extrapolate into the future. ARIMA models combine differencing with autoregression and moving averages of error terms. The differencing simply refers to calculating the difference between consecutive observations in the time series and it serves the purpose of making the series stationary. Sometimes a series needs to be differenced more than once to make it stationary.[51] The autoregression is similar to regular regression, except time-lagged values of the relevant phenomenon are used as predictors (for an overview see e.g. Hyndman & Athanasopoulos 2018; Makridakis et al. 1997). Finally, the moving average part relies on past forecast errors. In a way, the models in the ARIMA family can be said to be self-projecting (Hoff 1983:9).

ARIMA models can be either seasonal or non-seasonal. As noted above, seasonality is hypothesized to not play an important role in propagation of language change. For a non-

[51] While non-stationary series show changes in properties over time, stationary series have "statistical properties that do not depend on the time at which the series is observed" (Hyndman & Athanasopoulos 2021:265). The term *integration* (the I in the ARIMA acronym) refers to "the reverse of differencing" (Hyndman & Athanasopoulos 2021:278).

seasonal ARIMA (p, d, q) model, p represents the order of autoregressive part (i.e., how many past variables are used), d the degree of first differencing needed to make the time series stationary and q the order of the moving average part. The formula for a full ARIMA model with p^{th} order of autoregression and q^{th} order of the moving average part is given in (7.16). The y'_t stands for the differenced series at time t, c is a constant, ϕ is the autoregressive parameter, θ is the moving average parameter, and ε_t is the error term at time t.

(7.16) $y'_t = c + \phi_1 y'_{t-1} + \cdots + \phi_p y'_{t-p} + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t$

(from Hyndman & Athanasopoulos, 2018:278)

Just like with ETS models, appropriate ARIMA models are selected depending on patterns in the relevant time series. A Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test (Kwiatkowski et al. 1992) can be used to determine whether a series is stationary or not and this can help determine the order of differencing needed. The order of p and q can often be determined by looking at the autocorrelation and partial autocorrelation in the series. It is also possible to rely on AIC, AICc and BIC to select between ARIMA models of different complexity. Hyndman & Athanasopoulos (2021:285) recommend using AICc which is as in (7.17) where p and q refer to order of the autoregressive part and the order of moving average part. If there is no constant in the model (c = 0) then k = 0, otherwise k = 1 (Hyndman & Athanasopoulos 2021:285).

(7.17) $\text{AICc} = \text{AIC} + \frac{2(p+q+k+1)(p+q+k+2)}{T-p-q-k-2}$

The selection of an ARIMA model along with the value of relevant parameters can be done automatically in R. The function ARIMA() from the fable package (O'Hara-Wild, Hyndman & Wang 2021) uses a version of the Hyndman-Khandakar algorithm (see Hyndman & Athanasopoulos, 2018:244–247) to automatically select an appropriate model for the relevant time series. The algorithm involves several steps which are described in Hyndman & Athanasopoulos (2021:286). These include, determining the number of differencing needed to make the series stationary, selecting the values of p and q by minimizing AICc, fitting four initial models and choosing the model with the lowest AICc, testing closely related models and selecting the one with the lowest AIC. The ARIMA() function estimates the parameters ϕ and/or θ by relying on maximum likelihood estimation. According to Hyndman & Athanasopoulos (2021:284) this is similar to least squares estimates.

The ARIMA() function is extremely useful for selecting an appropriate ARIMA model and was used for the case studies in Chapter 8 and 9. In a few instances, an attempt was made to find a better ARIMA model manually by going through the appropriate steps and selecting values for p, d, and q. However, the best fitting models turned out to be those selected by the automatic function that relied on the Hyndman-Khandakar algorithm.

Since the processes behind the emerging patterns in the time series are not incorporated into the models discussed above, the models do not explain the reasons for why particular patterns emerge. This does not mean that causes of certain patterns in time-series datasets is not known. In fact, many processes that affect language change and diffusion have been studied, e.g., transmission of language, language contact, language policy and prestige. However, the interactions between these factors are complex and causation is often difficult to determine. Furthermore, these factors may be hard to quantify and incorporate into time-series forecasting.

7.2.5 Evaluating model fit and point forecasts

Models within a particular model family, e.g., ARIMA or ETS, are chosen by minimizing AIC, AICc and/or BIC. Innovation residuals can then be used to evaluate the fit of a model to the training data. When a transformation has been used, the residuals are on a transformed scale. In the case of a log transformation where $w_t = log(y_t)$ the innovation residuals are equal to the difference between an observed value of the training data on a transformed scale and forecasted value on a transformed scale, i.e, $w_t - \hat{w}_t$. When no transformation has been used, the innovation residuals are the same as regular residuals, i.e, the difference between observed value in the training data and forecasted value: $y_t - \hat{y}_t$ (cf. Hyndman & Athanasopoulos 2021:115–116). According to Hyndman & Athanasopoulos (2021:117) innovation residuals should have the properties listed in (7.18), where (7.18-1)–(7.18-2) are considered essential and (7.18-3)–(7.18-4) desirable.

(7.18) Innovation residuals should

1. Be uncorrelated
2. Have zero mean
3. Have constant variance
4. Be normally distributed

(cf. Hyndman & Athanasopoulos 2021:117)

To test whether innovation residuals are uncorrelated, it is possible to use a portmanteau test such as the Ljung-Box test based on (7.19) where ℓ is the maximum lag considered, r_k is the autocorrelation for lag *k*, and *T* is the number of observations.

(7.19) $$Q^* = T(T+2)\sum_{k=1}^{\ell}(T-k)^{-1}r_k^2$$

Ljung-Box statistics and a corresponding *p*-value for innovation residuals of models used in Chapter 8 and 9, were obtained by requesting the ljung_box statistics with the function feat() from the feasts (O'Hara-Wild, Hyndman & Wang 2021) package included in fpp3 (Hyndman 2023). If the p-value was larger than 0.05, the series was considered a white noise series (cf. Hyndman & Athanasopoulos 2021:96, 121).

Once a model has been selected and a forecast generated for the test data, a point forecast can be evaluated based on forecast error, i.,e, the difference between forecasted values and observed values in the test data. Evaluation can be done in several ways, for instance by observing *mean absolute errors* (7.20) and *root mean squared errors* (7.21).

(7.20) Mean absolute error

$$\text{MAE} = \text{mean}(|e_t|)$$

(7.21) Root mean squared error

$$\text{RMSE} = \sqrt{\;}\text{mean}(|e^2_t|)$$

Further ways of evaluating point forecasts include relying on percentage errors (7.22), or rather the *mean absolute percentage error* (7.23). These are unit free and can be used to compare forecasts for different data sets. However, there is a slight catch to using these when values in the time series are zero (or very close to zero). In these cases MAPE is given as infinite or undefined (Hyndman & Athanasopoulos 2021:137).

(7.22) Percentage errors

$p_t = 100e_t/y_t$

(7.23) Mean absolute percentage error

MAPE = mean($|p_t|$)

Finally, *mean absolute scaled errors* and *root mean squared errors* can be used to evaluate forecast performance, i.e, how close forecasted values are to observed values. These were proposed by Hyndman and Koehler (2006) and scale the forecasting errors based on mean absolute errors of the training data when a simple forecast method, such as a Naïve method, is used.

(7.24) Scaled errors for non-seasonal time series using a Naïve forecast

$$q_j = \frac{e_j}{\frac{1}{T-1}\sum_{t=2}^{T}|y_t - y_{t-1}|}$$

(7.25) Mean scaled errors

MASE = mean($|q_j|$)

(7.26) Root mean squared scaled errors

RMSSE = $\sqrt{mean(q_j^2)}$

In Chapter 8 and 9, MAE, RMSE, MAPE, MASE and RMSSE are provided for point forecasts of the test period for various models. These help evaluate how close to observed values the forecasted values are. A lower score is associated with more accurate prediction. Sometimes, the different methods of evaluation (MAE, RMSE, MAPE, MASE and RMSSE) favor a different model.

7.2.6 Evaluating distributional forecasts

Forecast distribution is evaluated slightly differently than point forecasts. In these cases, the prediction intervals are taken into consideration. Prediction intervals can be written as in (7.27), where c represents a multiplier of a relevant percentage and $\hat{\sigma}_h$ is the standard deviation of the forecast distribution h-step ahead. For 80% forecast interval, c is 1.28 and for 95% intervals it is 1.96 (Hyndman & Athanasopoulos 2021:123), cf. (7.28).

(7.27) Prediction interval

$$\hat{y}_{T+h|T} \pm c\hat{\sigma}_h$$

(7.28) 80% prediction interval $\hat{y}_{T+h|T} \pm 1.28\hat{\sigma}_h$

95% prediction interval $\hat{y}_{T+h|T} \pm 1.96\hat{\sigma}_h$

(Hyndman & Athanasopoulos 2021:123)

One step ahead predictions intervals are estimated based on standard deviation of residuals, (7.29).

(7.29) Standard deviation of residuals

$$\hat{\sigma} = \sqrt{\frac{1}{T-K}\sum_{t=1}^{T} e_t^2}$$

(Hyndman & Athanasopoulos 2021:124)

A Quantile Score and Winkler Score (Winkler 1972) can be used to evaluate forecast distribution and intervals. For the case studies in Chapter 8 and 9, these were obtained automatically by using accuracy() with the function quantile_score() and winkler_score(). The Quantile Score, $Q_{p,t}$ is based on evaluation of quantile forecast *f* with probability *p* at time *t*, written as $f_{p,t}$ in (7.30).

(7.30) Quantile Score

$$Q_{p,t} = \begin{cases} 2(1-p)\left(f_{p,t} - y_t\right), & \text{if } y_t < f_{p,t} \\ 2(y_t - f_P, t), & \text{if } y_t \geq f_{p,t} \end{cases}$$

The Winkler Score takes into consideration the length of the prediction interval and the penalty assigned when an observed value is outside of the interval. For the Quantile and

Winkler score, a lower value is associated with a "better estimate of the quantile" and "narrow intervals" (Hyndman & Athanasopoulos 2021: 142–143).

A further way to evaluate forecast distribution is to take into account the whole forecast distribution and calculate Continuous Ranked Probability Score, or CRPS (Gneiting and Katzfuss 2014). A lower value is associated with better forecasts.

Finally, scale-free skill scores based on CRPS can be used. These evaluate how well a new model does w.r.t. a given benchmark method. When requesting skill scores using the accuracy() function, the automatic benchmark method is the Naïve method for non-seasonal time series and Seasonal naïve method for seasonal series (Hyndman & Athanasopoulos 2021:145). An example of how skill scores are calculated based on a Naïve model and some other model (*model2*) is provided in (7.31).

(7.31) skill score $= \frac{CRPS_{Naïve} - CRPS_{model2}}{CRPS_{Naïve}}$

If a skill score for a model is positive, it is considered an improvement over the benchmark model. If a skill score is negative, the relevant model performs worse than the benchmark method. A model that obtained a skill score of 0.12 when compared to a Naïve method is considered to perform 12% better than the Naïve model.

In Chapters 8 and 9, the Quantile Score, Winkler Score, CRPS and skill score are used to compare the forecast distribution of the relevant forecasting models. Note though that according to Hyndman & Athanasopoulos, reliable error measures can only be calculated provided the test set is large enough. Since the test set in Chapters 8 and 9 is only about 10% of the whole time series, containing either two or four observations, evaluations of forecasting performance of particular models should be taken with a grain of salt.

7.2.7 Interpreting forecasts

The general philosophy adopted here is that predictions make little sense out of context. Even though forecasts may be generated according to well-founded procedures and return predictions that feel reliable or look good, they still need to be contextualized. For instance, if one encounters a forecast claiming that the temperature outside will be 10 °C tomorrow, the prediction is not very informative unless one also knows the location and the time of year for the prediction. It is also advantageous to know what the temperature was for the last few days, since how much the temperature fluctuates can influence how the temperature is perceived. Other types of information might also matter, such as whether one is located at the place of the forecast or whether one is traveling there from a different climate or not. For reasons such as these, I have tried as much as possible to contextualize the language forecasting done here by providing a thorough background on the changes under investigation, commenting on expected direction of propagation, describing the data in a useful way, noting features of the time series, and by discussing the relevant forecast and forecast intervals. Note that forecast intervals provide information about how certain point predictions are and therefore "...point forecasts can be of almost no value without the accompanying prediction intervals" (Hyndman & Athanasopoulos 2021:124, see also Castle, Clements & Hendry 2019:2).

When contextualizing a forecast, a commentary might sometimes provide a kind of narrative or a story about the potential future developments. Such narratives are sometimes referred to as *forediction* (Castle, Clements & Hendry 2019:191). Of course, one has to be careful about biases in such narratives as they may include comments about what one *wishes* to see happening (perhaps due to a belief in a certain theory or unspecified internalized biases) in mixture with what models are suggesting. In this context, it is important to remember that forecasts represent an attempt to predict events that have not occurred. They are not statements about what will happen, but an hypothesis about what could or might happen. Even with good performance on previously unobserved data, it is not guaranteed that a forecasting method will provide accurate predictions for an unobserved future.

7.3 Potential issues

Predicting the future is not an easy task. The task is certainly not made easier by language forecasting being a relatively new field within linguistics. As already discussed (see Chapters 3 and 5), language data has generally not been systematically gathered with forecasting in mind. This applies to convenient E-language data as well as data generated by means of specifically designed experiments. For these reasons, any forecast is necessarily constrained by resources and quality of language data that is available. Various possible problems have been pointed out and discussed in previous chapters. In the present context, it is worth summarizing three factors that relate to the quality of the language data used for forecasting. These are presented in (7.32) and pertain to i) how language variation and change is measured, ii) whether the measurements accurately reflect the situation in the language community and iii) how often observations were made, i.e., the sampling frequency and length of the time series. The three factors have the potential to affect both the forecasting problem and the interpretation of individual forecasts.

(7.32) i. How is the variation and/or change measured?

Proportion of new variant at a given time

Proportion of individuals using new variant at a given time

Proportion of individuals capable of using new vs. old variant at a given time

ii. Is the measurement representative of the situation in the language community?

Nature of the language data the forecasting is based on

Specialized language data

Convenient E-language data

Nature of the source of the data?

What is the language style and/or register

Is the type of the data consistent over the sampling period

Does the data reflect the actual situation in the language community?

Consider normative pressure and/or prescriptivism

Consider a potential gap between what speakers know vs. what they use in the relevant language settings

iii. How often are observations made?

Sampling frequency

Length of time series

Decisions made for the case studies in Chapter 8 and 9 are as follows. Convenient E-language data (written language) was used to gather information regarding the status of the relevant linguistic variation. Measurements were in the form of the proportion of innovative variants in a given corpus over a given period. The style or register of the language covered what is hypothesized to be informal and semi-formal language. Material from sources thought to contain formal language was not used for forecasting (on the categorization of informal semi-formal and formal see Chapter 8 Section 8.4.3 and Chapter 9, Section 9.4.3). The register-type of the data was consistent for the whole time period taken into consideration. Since the data is in the form of written language there are good chances that normative pressure and prescriptivism have an effect on the proportion of innovative and traditional variants in the sources. This is especially relevant for case marking of subjects with the predicate *hlakka til* 'look forward to' (Chapter 9), which has gained much attention in both scholarly literature and within the school system. Ever since innovative case marking was first noticed, there have been consistent attempts to eradicate it (see discussion and references in Chapter 9).

Data obtained from the Icelandic Gigaword corpus was projected into yearly time series. This is made possible by the corpus containing a relatively consistent amount of informal and semi-formal data over the past twenty years or so. Data from Twitter, on the other hand, was projected into quarterly time series. The choice of using quarterly series was mainly for the purpose of increasing the number of observations and thereby lengthening the time series. In theory, it would have been possible to use monthly time series for the Twitter data to increase the number of observations further. However, taking into consideration how many years ahead prediction for the future were to be made, this would have been impractical. The goal was to provide forecasts for at least 3 to 5 years into the future. Projecting 15 steps ahead using monthly time series would only yield predictions that cover a time period of one year and three months. If one were to project 5 years into the future using a monthly series, 60 steps ahead forecast would be required. As already pointed out (see discussion in Section 7.1.2) the more steps ahead predictions are made, the more uncertain a forecast becomes. Thus, a choice was made to use quarterly series for Twitter data and yearly series for data from the Icelandic Gigaword corpus. The time series used for forecasting in Chapters 8 and 9 are relatively short, or between 19 (= 19 years) and 44 (= 11 years) observations.[52] When divided up into training and test data, the test period is only 2 to 4 observations. Although the short test period (used to evaluate models) might be somewhat problematic, this was done so that more information from the time series could be used when fitting a model. In general, all decisions made regarding the data and the time series were aimed towards striking a balance between what is required by the methodology, e.g., that observations need to be equally spaced in time and be comparable, and what kind of data can reasonably be obtained for the variation and change under investigation. Hopefully, the case studies presented in the following two chapters can be used to evaluate what could be improved in language forecasting and, more generally, contribute towards better understanding of language forecasting.

[52] Having 44 observations may be considered enough (or even plenty) under certain circumstances. However, one should keep in mind that 44 quarterly observations amount to 11 years which is not a long time period for observing syntactic change.

8 Grammaticalization of complex prepositions

8.1 Introduction

In Modern Icelandic (1540 – present), the complex prepositions *við hliðina á* ‘next to, beside (lit. by the side of)’ and *á bak við* ‘behind (lit. at back of)’ frequently appear in a simplified form in the written language, *hliðiná* ‘next to’ and *bakvið* ‘behind’ respectively. Some naturally occurring examples are given in (8.1) and (8.2). Note that the a-examples are fully comparable with the b-examples w.r.t. the context the prepositions occur in, their meaning (conveying spatial position) and the time of occurrence (there is less than a year between the a- and the b-examples).

(8.1) a.

sá	*fjórði*	*sat*	***við***	***hliðina***	***á***	*mér*	*í*
that	fourth	sat	by	side-ACC	on	me-DAT	in

salnum.
the.hall

‘...the fourth one sat next to me in the hall.’

(Twitter, Haukur Bragason@HaukurBragason, Jan 31, 2016)

b.

fólkið	*sem*	*sat*	***hliðiná***	*mér*	*í*	*eymundsson…*
people	that	sat	side.of	me-DAT	in	Eymundsson

‘The people that sat next to me in Eymundsson…’

(Twitter, Helga Dögg@DoooHelga, Jan 8, 2016)

(8.2) a.

Er	*myndatökunaðurinn* [sic]	*að*	*fela*	*sig*	***á***	***bak***
is	cameraman	to	hide	self	on	behind
við	*lýsistunnu?*					
by	lýsi.barrel					

‘Is the cameraman hiding behind a barrel of lýsi?’

(Twitter, Sindri Geir@sindrigeir, Feb 27, 2022)

b.

Alvöru	*karlmenn*	*fela*	*sig*	*greinilega*	***bakvið***
real	men	hide	self	clearly	behind
nafnlausa	*accounta…*				
nameless	accounts				

‘Clearly, real men hide behind nameless accounts…’

(Twitter, Tómas Ingi@tomasingiad, Jan 24, 2023)

Since written language tends to be conservative in many aspects, it can preserve forms and structures which may not fully reflect everyday language. This applies to the forms *við hliðina á* ‘beside’ in (8.1a) and *á bak við* ‘behind’ (8.2a) which indicate a structure consisting of a preposition followed by a noun and a second preposition. These types of structures are sometimes referred to as complex prepositions, phrasal prepositions, compound prepositions, or simply PNP-constructions (e.g., Quirk & Mulholland 1964; Seppänen et al. 1994; Hoffmann 2004; Vincent 2020; Stefanowitsch et al. 2020). In what follows, they are presented as in (8.3) where $\mathbf{P_1}$ refers to the first preposition in the complex structure, $\mathbf{N_1}$ to the following noun (*hliðina* 'side' or *bak* 'back') and $\mathbf{P_2}$ to the second preposition; the notation (+N_2) is used to indicate the nominal argument that follows the complex preposition and whose case is assigned by $\mathbf{P_2}$.

(8.3) $\mathbf{P_1 N_1 P_2}$ (+N_2)

The innovatively-written forms in (8.1b) and in (8.2b) can be taken to reflect a change which involves the grammaticalization of a noun and a following preposition into a single preposition (suggested by Friðjónsson 2004, 2007; Rögnvaldsson 2021), *hliðiná* and *bakvið*. As such, the sequence in (8.3) is replaced by (8.4).

(8.4) **P** (+N_2)

The grammaticalization of a string of elements into a single preposition is widely attested in the world's languages and is regarded as one of the most common grammaticalization paths that have been observed (e.g., Lehmann 1991:501). This type of change has never been systematically studied for Icelandic and documentation of the change from (8.3) to (8.4) in context of the complex prepositions *á bak við* and við *hliðina á* is lacking. The present study remedies that by offering an insight into the diachrony of *á bak við* and *við hliðina á*. It is claimed that in Modern Icelandic the two strings are understood as a single unit with the function of a preposition. It is furthermore suggested that the absence of $\mathbf{P_1}$ in examples like those in (8.1b) and (8.2b) is due to phonological erosion, a process arguably linked to grammaticalization; this part of the change is still ongoing in Modern Icelandic.

Aside from providing a general overview of variation in the complex preposition *á bak við* and *við hliðina á* in Modern Icelandic, the present study offers a novel type of documentation of the propagation of variants lacking $\mathbf{P_1}$. The data comes in the form of regular time series where each observation is based on examples attested in written sources and contains information about the proportion of variants with and without $\mathbf{P_1}$. Examples were obtained from two sources, the Icelandic Giga Word corpus and Twitter. Time series constructed based on data from IGC covered the period 2000 – 2021, although only the years 2003-2021 were used for time series analysis and forecasting. Time series based on material from Twitter contained examples from Q1 2009 to Q4 2022, although only Q1 2012 to Q4 2022 were used. The reason for not using the whole series was that early observations are based on relatively few examples and may introduce noise into the series. A second reason for only using part of the series is to make the time series comparable to the ones presented in Chapter 9. The novelty of the study does not only lie in general documentation of the complex prepositions and the documentation through regular time series, but also in containing predictions about propagation of variants lacking $\mathbf{P_1}$. Each time series that was taken into consideration was split into a training and test set. Initially three to four models were fitted to the training set (a Naïve model, a Seasonal naïve model, a Mean model, and a Drift model) and used to predict observation in the test set. Next, slightly more complex models were fitted to the training set and used to predict attested observations. The more complex models involved methods such as exponential smoothing (ETS models) and autoregression (ARIMA) models. In some instances, the simple models generated more accurate point predictions and forecast distribution for the test set than the more complex models. Forecasts for future periods were generated using models that were deemed appropriate based on the time series and how well models did at predicting test data.

The main results of the study are as follows. Data from Twitter generally showed a higher proportion of examples lacking **P**$_1$ than data from IGC. Additionally, the preposition *á bak við* showed a higher proportion of variants lacking **P**$_1$ than *við hliðina á*, suggesting that *á bak við* is further along in the trajectory of grammaticalization than *við hliðina á*. As to individual sources, regular time series based on data from Twitter show a relatively stable variation over an 11-year period. Forecasting models, relying on aspects of the historical data, suggest a relatively unchanging future. Time series based on IGC data behave slightly differently, showing decrease over time in forms lacking **P**$_1$. This applies especially to complex preposition *á bak við*, which shows a very consistent decrease over time in the use of variants lacking **P**$_1$. This is not in line with expectations based on grammaticalization and phonological erosion, but it could be explained in terms of the type of material behind the IGC data. Part of the IGC data comes from online news media. Another explanation is that retention or restoration of **P**$_1$ might be linked to Icelandic generally favoring multi-word prepositions (Berthele et al. 2015). This is further discussed in Section 8.6.

The structure of the chapter is as follows. Section 8.2 provides a general background on prepositions in Icelandic (8.2.1) and discusses expectations towards direction of change in light of grammaticalization (Section 8.2.2). Section 8.3 focuses on variation and change in the complex prepositions *á bak við* and *við hliðina á* in Icelandic. Variation found in Modern Icelandic is introduced in Section 8.3.1. Arguments in favor of the strings already being grammaticalized in the modern language are provided in Section

8.3.2 (Variation in writing and usage as an indicator of grammaticalization) and Section 8.3.3 (Syntactic (in)flexibility as an indicator of grammaticalization). Documentation of the phrases at older stages can be found in Section 8.3.4 with a summary in 8.3.5. Section 8.3.6 contextualizes the grammaticalization of *á bak við* and *við hliðina á* further and summarizes expectation for the propagation of variants lacking $\mathbf{P_1}$. In Section 8.4, data used for forecasting is described, e.g., in terms of how it was obtained, how annotations were made and how regular time series were constructed. Section 8.5 focuses on forecasting. Four time series were taken into consideration, two for each data source. Descriptions of the four time series are found in Section 8.5.1 and fitting of models and forecasts in Section 8.5.2–8.5.4. Results from the study are discussed in Section 8.6.

8.2 On prepositions and directions of change

8.2.1 Prepositions in Icelandic

Prepositions in Icelandic assign an oblique case, either accusative, dative or genitive, to their NP complement. They furthermore take part in conveying information about location in time or space or movement along a path (for an overview of characteristics of prepositions in Icelandic see Thráinsson 2005:109,113–120, 122; for a comprehensive discussion on selected prepositions and prepositional phrases see Kress 1982:469–505; Friðjónsson 1988, 2005; Berthele et al. 2015).[53] Some simple prepositions, consisting of a single word, are shown in (8.5). Assuming prepositions are heads of prepositional phrases,

[53] Prepositions in Icelandic are often defined in terms of their morphological and syntactic behavior. Thus Thráinsson (2005:109) notes (transl. mine): "Prepositions are function words (uninflected words) which take a complement. The complement is usually a noun phrase whose case (oblique case) is assigned by the preposition".

a simplified structure of the PPs in (8.5b) might be represented as in (8.5), using (8.5a) as an example.

(8.5) a. *Ég* *lagði* *bókina* [$_{PP}$ ***á*** *borðið*].

I-NOM laid book.the-ACC on table.the-ACC

'I laid the book on the table'

b. *Ég* *svaf* [$_{PP}$ ***í*** *tjaldi*].

I slept in tent-DAT

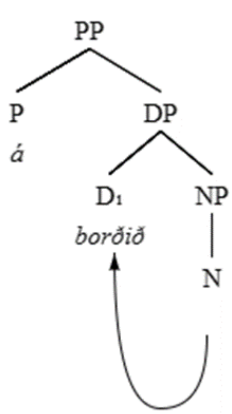
'I slept in a tent.'

c. *Þessi* *gjöf* er [$_{PP}$ ***til*** *þín*].

This present is to you-GEN

'This present is for you.'

(8.6)



In addition to single-word prepositions, Icelandic also has prepositions that contain more than one element, often a combination of an adverb and a preposition.[54] In cases where an adverb precedes a preposition, the preposition can be regarded as the head of the phrase and the preceding adverbial may add information on orientation in space (Friðjónsson 1988:19–36; Berthele, Whelpton, Næss & Duijff 2015:88).

(8.7)	*Skálin*	*stóð*	***ofan***	***á***	*borðinu.*
	bowl.the	stood	on.top	on	table.the-DAT

‘The bowl stood on the top of the table.’ (Berthele et al. 2015:88, example (11))

In cases where the adverb follows the prepositions, what counts as head of the phrase is less clear. Berthele et al. (2015:89) provide the example in (8.8), noting that *austan* has the form of an adverbial and *fyrir*, a preposition, is responsible for the accusative case on *húsið* ‘house’.

(8.8)	*Tveir*	*kettir*	*læddust*	***fyrir***	***austan***	*húsið.*
	two	cats	snuck	at	eastwards	house.the.ACC

‘Two cats snuck around in the area to the east of the house.’

(Berthele et al. 2015:89, example (20))

[54] The main distinction between adverbs and prepositions is that adverbs typically do not take a complement or assign case, while prepositions typically have a complement and always assign case. The distinction is not always this straightforward as some adverbs, especially particles (a subcategory of adverbs) may sometimes behave eerily similar to prepositions. On the distinction between these categories see Thráinsson 1979:25ff, 2005:122, 2007:139–140.

Combinations of adverbials and prepositions like in (8.7) and (8.8) are referred to as complex prepositions by Berthele et al. (2015:87). The term complex preposition has also been used to cover a string of a preposition, a noun and a second preposition. Alternatively, such strings may be referred to as phrasal prepositions, compound prepositions or PNP-constructions (e.g., Quirk & Mulholland 1964:65, Vincent 2020). While PNP-sequences are quite common in Modern Icelandic (for a list of many such see Kress 1982:469–505), their internal syntactic structure may differ. Compare, for instance, (8.9) and (8.10) which do not appear to have the same status; *í sambandi við* 'in touch with' conveys a different type of information than *á bak við* 'behind' and *við hliðina á* 'next to'. Only the phrases in (8.10) appear to function as 'complete prepositions'.

(8.9)	*Ég*	*verð*	***í***	***sambandi***	***við***	*þig.*
	I	be	in	touch-DAT	with	you-ACC
	'I'll be in touch with you.'					

(8.10)	a.	*Kötturinn*	*er*	***á***	***bak***	***við***	*sófann.*	
		cat.the	is	on	back	by	couch	
		'The cat is behind the couch.'						
	b.	*Köturinn*	*er*	***við***	***hliðina***		***á***	*sófanum.*
		cat.the	is	by	side.the		of	couch
		'The cat is next to the couch.'						

The structure of (8.10), including whether it really consists of a preposition, a noun and a second preposition, is discussed further below (see Section 8.3.1–8.3.3). For now, it should be noted that single-word prepositions are probably the most prototypical type of prepositions in Icelandic. More complex prepositions are, however, also frequently used, especially for certain spatial descriptions. When compared to other Germanic languages such as Frisian, Swiss German, Standard High German and Norwegian, Icelandic has been shown to use complex expressions for spatial descriptions more than the other languages (Berthele et al. 2015). This observation was made on the basis of structures containing an adverb and a preposition, not a sequence of preposition, noun and a second preposition.

8.2.2 Direction of change viewed through grammaticalization

Grammaticalization is usually defined in terms of a content word assuming "the grammatical characteristics of a function word" (Hopper & Traugott 2003[1993]:4; see also Kurylowicz 1965/1975:52; Heine, Claudi & Hünnemeyer 1991:2). As such, the emergence of prepositions from nouns or other content words is a typical case of grammaticalization, prepositions often being regarded as functional elements (although see Déchain 2005 for the view of prepositions being a borderline lexical category).

The emergence of a preposition from a noun, or a combination of noun and other elements, is well attested in the world's languages and has been claimed to be one of the most common grammaticalization paths found (Lehmann 1991:501; on this type of change see also Heine 1995; Heine & Kuteva 2002:271–2; Hoffmann 2004; van Gelderen 2011:182–187). Examples of prepositions with such origin are English *beside* from the Old English phrase *be sidan* 'by the side$_{\text{DAT.SG}}$ of' (Hopper & Traugott 2003:110; van Gelderen

2011:183–187), German *wegen* 'because of' from the noun *Wegen* 'ways$_{\text{DAT.PLUR}}$' (Hopper & Traugott 2003:110), and French *chez* 'with' from Latin *casa* 'house' (Vincent 1999, Longobardi 2001). A further example is French *avant* 'before' from Latin *ab ante* 'from before' which involves the univerbation of two prepositions (Vincent 1999:1133, 2020).

For Icelandic, some prepositions found in Old Icelandic (also used in the modern language) developed from nouns already in the prehistory of the language (Magnússon 1989 s.v.; Faarlund 2004:107). Thus, the preposition *gegn* 'against, opposite', which assigns dative case to its complement, is related to the noun *gagn* 'advantage' (8.11a) and the preposition *hjá* 'at, by, with' is derived from Germanic **hiwa*, meaning 'household, family' (8.11b).

(8.11) a. *gegn* 'towards, *gegnt* 'against, opposite', related to *gagn* 'advantage'
b. *hjá* 'at, by, with' derived from Germanic **hiwa* 'housold, family'
c. *til* 'to' from Germanic **tila* 'goal' (cf. German *Ziel* 'goal')

The examples mentioned above are all instances of single word prepositions developing out of a multi word structure or a single noun. Such changes do not take place in one go but may consist of various subprocesses. Complex prepositions in the form of PNP-structures (as mentioned in Section 8.2.1 above) may bear witness to this. Although seemingly consisting of multiple elements, these may sometimes function more like a single preposition. As Quirk & Mulholland (1964:64) note "there is considerable variation in the degree of interdependence between the four elements of such sequences".[55] In other

[55] By four elements, Quirk & Mulholland include the N complement of the second preposition, i.e., the sequence $\mathbf{P_1N_1P_2N_2}$.

words, some PNP-sequences may be fairly fixed, even unchangeable, while others can be modified in one or more ways (I return to this in Section 8.3.3). Sequences that are unchangeable, may in time come to be represented by a single element in much the same way as the prepositions mentioned above.

In a short discussion of the complex preposition *á bak við* 'behind', Friðjónsson (2007:28) notes that uses of the form *bak við* (sometimes written <*bakvið*>) involves at least two changes, i.e,. i) *bak* 'back' acquiring a new meaning and ii) the initial preposition **P**$_1$ á being omitted.[56]

> Í orðasambandinu á bak við e-ð hefur stofnorðið bak misst eigin merkingu og fengið það sem kalla má hlutverksmerkingu. Þetta má m.a. sjá af því að fs. á er oft felld brott og þá stendur einungis bak við e-ð.
>
> In the collocation *á bak við smth*, the head noun has lost its lexical meaning and turned into a function word. This can, for instance, be observed when the preposition *á* is omitted, resulting in the form *bak við smth*.

Under this view, grammaticalization (here the formation of a single-word preposition from a string containing more than one element) appears to be simultaneously a single change and multiple subprocesses. Newmeyer (1998:226), has in fact proposed that "there is no such thing as grammaticalization", but rather grammaticalization is the result of interactions of three types of changes, i.e., downgrading analysis (which might be more generally referred to as reanalysis), semantic change and phonetic reduction (on this view see also Janda & Gil 1980; Janda 2001). This is represented visually in Figure 8.1 (from

[56] Faarlund (2004:109) claims that when prepositions and nouns followed by a noun in the dative are used in Old Icelandic, the first noun has already been grammaticalized and is a part of the preposition. He provides two arguments, namely that nouns typically do not assign dative case to other nouns and that the meaning of these is already semantically bleached. One of the examples he provides is *á bak* + DAT 'behind (lit. on back)'.

Newmayer 1998:260), where downgrading analysis, semantic change, and phonetic reduction are shown to exist (and be able to occur) independently of grammaticalization.

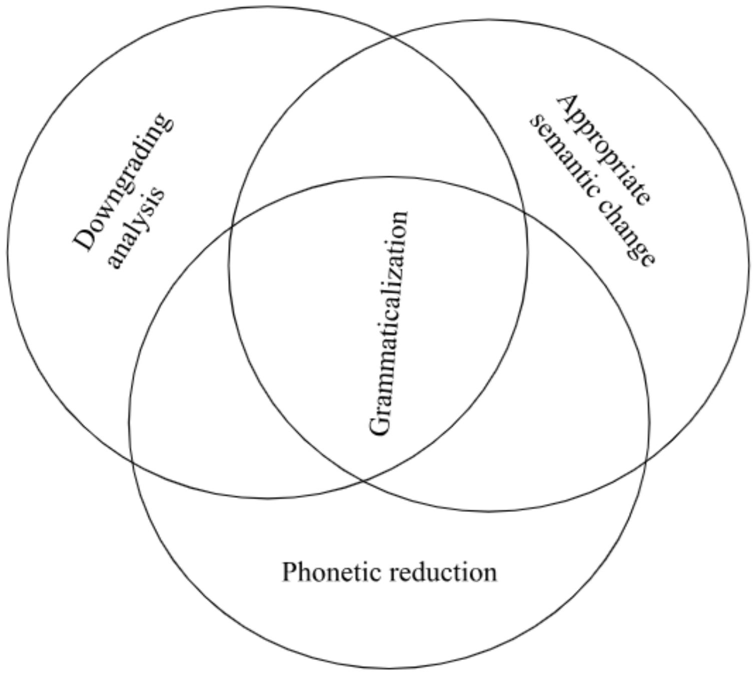


Figure 8.1. According to Newmeyer (1998:260), grammaticalization does not exist as an independent phenomenon. Rather, it is an epiphenomenon, emerging at the intersection of downgrading analysis, semantic change, and phonetic reduction.

If semantic change is taken to be a part of grammaticalization, one might expect to observe some form of semantic change when complex prepositions are grammaticalized, for instance through metaphorical use of relevant elements or through semantic bleaching of some sort. An important indicator is when speakers stop making a connection between two items (a grammaticalized element and a lexical item) even though diachrony indicates a common origin. Hoffmann (2004:180) describes this in terms of decategorialization:

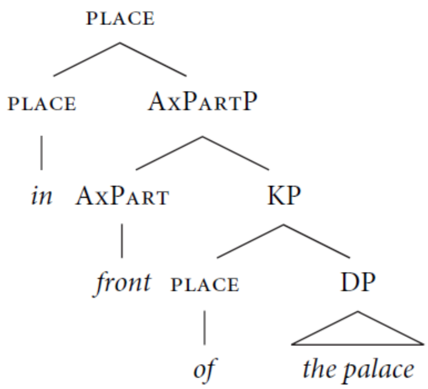
> ... the grammaticalization of complex prepositions manifests itself in a number of ways. [...] the nominal element of the construction over time loses the features that define its categorial status as a noun. For example, in the complex prepositional use of *in view of*, *view* cannot occur in the plural or with a determiner, nor can it be premodified by an adjective. The noun *view* has thus undergone the process of **decategorialization**.

As soon as there has been a dissociation with an element in a complex preposition and a diachronically related lexical item, downgrading analysis (or reanalysis) can be argued to already have taken place. Since syntactic structure depends (at least partially) on categorization of elements and categorization is (at least partially) related to meaning, it seems impossible to make a clear distinction between semantics and syntax.

Approaches that deal with the conceptual decomposition of prepositions (e.g., Jackendoff 1973; see Asbury et al. 2008 for discussion) tend to reflect the close relationship between syntax and semantics. Under these type of accounts, prepositional phrases may contain multiple functional projections, realized by a string of words (or a single word), conveying a direction, path or place (e.g., den Dikken 2010, Svenonius 2004, 2006, 2008; cf. discussion in Asbury et al. 2008:9). In line with Svenonius (2004, 2006), a prepositional phrase conveying location consist of a placement projection (PLACE), a type of coordinating projection (AxPartP), case assigning element (under KP) and a DP (a lexical complement) that represents the ground to which something is coordinated.[57] Such a structure (adapted from Asbury et al. 2008:9 which base it on Svenonius 2004, 2006), showing the string *in front of the place*, is provided in (8.12).

[57] For a more formalized (and detailed) discussion of this type of decomposition of prepositional phrases and labeling, see e.g., Svenonius (2004, 2006, 2008) and Roy & Svenonius (2009). For a discussion on the difference between figure and ground in prepositional phrases (the ground is always the complement) see Svenonius (2004:15) and citations there, e.g., Talmy (1978, 2000).

(8.12)

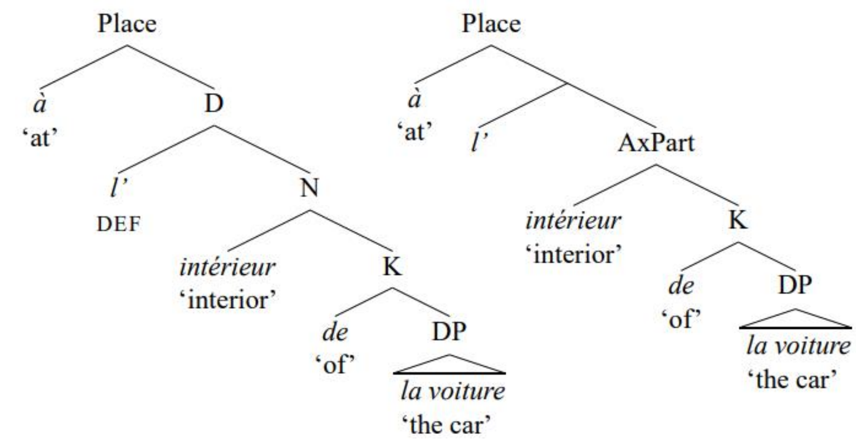


In (8.12), it is immediately noticeable that *front* in AxPart resembles the lexical noun *front*. This is no coincidence as elements in complex prepositional phrases often trace their origin to nouns although they do not behave as such synchronically; they can, for instance, not be pluralized nor be modified with an adjectives (Svenonius 2006:50, 56; Hoffmann 2004; see also Quirk & Mulholland 1965 for fourteen diagnostics of the behavior of $P_1N_1P_2$ structures). In short, “N is undeniably an important source for AxParts diachronically” (Svenonius 2006:74) and it turns out that certain types of nouns, including those referencing body parts like ‘back’ and ‘head’, tend to be a source for prepositions (Heine 1995:123, 125).[58]

As an example of the emergence of an AxPart element, Roy & Svenonius (2009) argue that, *l’intérieur* ‘the interior’ in the French prepositional phrase *à l’intérieur de* ‘at the interior of’ was originally a DP, with *intérieur* being the head of N. Thus, the change is that from (8.13a) to (8.13b) which apparently amounts to some form of relabeling, i.e., a former NP (or DP) is rebranded as AxPart.

[58] The data Heine’s observation is based on comes from various African and Oceanic languages.

(8.13)



(Roy & Svenonius 2009:4)

Like previous accounts on changes in complex prepositions (e.g., Hoffmann 2004:180), the transition from (8.13a) and (8.13b) involves a semantic change (a "decategorization" of a noun) occurring simultaneously with a change in the underlying structure of a prepositional phrase. For this step, the surface string may stay the same. The change is compatible with some definitions of (syntactic) reanalysis whereby only an alteration of the underlying structure needs to occur (see in particular Langacker 1977:58 and Campbell 2020:279 who assume reanalysis does not have to affect surface strings).

Interestingly, Roy & Svenonius (2009:4) note that they are "unsure about the structural status of the vestigial definite article" which leads them to leave it unlabeled in the tree showing the grammaticalized structure in (8.13). One solution might be to assume that the D (not the N) is the source of the AxPart. This would be in line with the work of Longobardi (2001) and van Gelderen (2011) who argue that the grammaticalization of a noun as a preposition may take place via a D-head. According to this, an N (with a locative feature) undergoes a movement from N to D before being incorporated into P (van Gelderen 2011:183). This would mean that N does not directly become AxPart, but rather that D does.

The dissociation of an element from a lexical category, along with structural reanalysis (a part of grammaticalization), raises questions about how surface strings are interpreted. For instance, might multi word strings still be understood as a single element? Such a view is arguably compatible with Lehmann's (2002) ideas of lexicalization.[59] Here, a string which was previously decomposed in a systematic way is now understood as a single unit. If this idea is aligned (or contrasted) with the work of Svenonius (2006, 2008) and Roy & Svenonius (2009), it brings into question the association of surface elements with underlying projections such as AxPart and K. Perhaps the change is not as simple as a former N becoming an AxPart. Rather, a dissociation from N may cause a whole string (simultaneously) being linked to multiple projections at once. Although interesting, this line of thought is not pursued further here. It is left for future exploration.

To wrap up the discussion on grammaticalization and direction of change in prepositions, a few things may be noted. First, prepositions often develop from nouns or from multi-word strings with various elements. Such a change may consist of various subprocesses, namely downgrading analysis (or reanalysis), semantic change and phonological reduction. The exact relationship between these three is not fully clear, but one may expect the grammaticalized element (or a part of it) to be is somewhat "detached" from the meaning of a cognate lexical item, that it no longer behaves like the source item,

[59] Lehmann argues for lexicalization and grammaticalization being two separate phenomena. This claim is only understandable in light of the lexicon being concerned with "those signs which are formed irregularly, and which are handled holistically", while grammar (hence the term grammaticalization) is "concerned with those signs which are formed regularly and which are handled analytically" (Lehmann 2002:1).

and that it (eventually) shows some signs of phonological reduction. In this way, grammaticalization generates expectations towards development of prepositions (see also Heine 1995 for grammaticalization of this type allowing for language evolution to be predicted up to a certain extent). Keeping this in mind when viewing the $\mathbf{P_1N_1P_2}$ strings *á bak við* 'behind' and *við hliðina á* 'next to, beside' in Icelandic, a change to *bakvið* and *hliðiná* is unsurprising, even expected. The dropping of $\mathbf{P_1}$ and the univerbation of $\mathbf{N_1}$ and $\mathbf{P_2}$ might be viewed as signs of phonological reduction. The characteristics of the strings in Modern Icelandic and variation in use of the two complex prepositions are discussed further in Section 8.3.

8.3 The complex prepositions *við hliðina á* 'next to' and *á bak við* 'behind'

8.3.1 Noticing variation in Modern Icelandic

Modern Icelandic exhibited variation in the use of the complex prepositions *á bak við* 'behind' and *við hliðina á* 'by the side of, next to', such that the initial preposition (*á* or *við*) is sometimes omitted. Some attested examples of the complex preposition *á bak við* are provided in (8.14).

(8.14)	a.	*…að*	*fólk*	*viti*	*hver*	*er*	***á***	***bak***	***við***	*hann*.
		that	people	know	who	is	on	back	by	him

'... that people know who is behind it [propaganda].'

(Twitter, Kristján@tyggjo, Nov 26, 2015)

b. *Veit* *ekki* *hvað* *er* ***bak*** ***við*** *hana*.

Know not what is behind by her

‘I don’t know what’s behind her.’

(Twitter, Gunnar Dofri@gunnardofri, Jan 9, 2018)

In addition to the absence of the initial preposition *á* in (8.14b), the subparts *bak* and *við* are occasionally written in a single form (8.15).

(8.15) a. *…hvað er* ***á*** ***bakvið*** *þessar tölur* *takk.*

what is on behind those numbers thanks

‘...what’s behind those numbers, thank you.’

(Twitter, Kristín Halla@KristinHallaL, Jan 9, 2023)

b. *…hvað er* ***bakvið*** *ykkur*.

what is behind you

‘...what's behind you.’

(Twitter, Eiríkur Kristjánsson@Eirikur_Gauti, Jan 25, 2017)

The kind of variation in (8.14)–(8.15) is also found with the complex preposition *við hliðina á*, i.e., the forms (*við*) *hliðina á* and (*við*) *hliðiná* are both attested.

To facilitate discussion of the complex prepositions, they will be referred to in term of their full forms, *á bak við* and *við hliðna á*, unless highlighting written variants which are indicated by the use of angle brackets such as <*á bakvið*> and <*bakvið*>. The presence and absence of the initial prepositions (*á* and *við*) is sometimes denoted with brackets as in (8.16).

(8.16) **Location in space**: (*á*) *bak við* ‘behind’, (*við*) *hliðina á* ‘next to, by the side of’

Although variation in the use of *á bak við* and *við hliðna á* is briefly mentioned in discussions on Modern Icelandic (see Friðjónsson 2004, 2007 and Rögnvaldsson 2021 who both note the direction of change to be *á bak við* > *bakvið* and *við hliðina á* > *hliðiná*), it has never been systematically investigated or documented. The present study remedies that.

In addition to the prepositions listed in (8.16) there is also variation in the use of several other prepositional phrases which can appear either with or without an initial prepositional element, *á* ‘on’ or *í* ‘in’ (Kress 1982:188, 191, 193; Friðjónsson 1988:35–36). Some of these are listed in (8.17), where they are grouped based on their primary meaning, i.e., whether they convey location in time or space, or a movement along a path.[60]

(8.17) **Ordered location in time and/or space**: (*á*) *eftir* ‘after, behind’, (*á*) *undan* ‘before, in front of’

Location in space: (*á*) *meðal* ‘among’, (*á*) *milli* ‘between’, (*á*) *móti* ‘opposite’, (*í*) *milli* ‘between’

Location in time: (*á*) *meðan* (að) ‘while’

Path: (*í*) *gegnum* ‘through’, (*í*) *kringum* ‘around, circa’

[60] The goal is not to give a complete overview of variation in prepositions in Icelandic, only to illustrate that such variation exists.

It seems appropriate to contrast the complex prepositions in (8.16) with the prepositional phrases in (8.17) as they have many things in common: i) They are multi-word prepositions where the initial element is frequently omitted, ii) Similar to some of the prepositions in (8.17), *við hliðina á* and *á bak við* denote location in space, both literally and figuratively, iii) Variation in the use of these preposition have existed for a long time (over 100 years, see Section 8.3.4–8.3.5), and iv) the complex and simplified forms are often used interchangeably (see discussion below).

Despite some similarities, the complex prepositions *á bak við* 'behind' and *við hliðina á* 'beside' appear different from the prepositions in (8.17) in at least one prominent way. They seem to be more transparent. The form *bak* can be identified as the accusative singular indefinite form of the neuter noun *bak* 'back' and *hliðina* is recognizable as the accusative singular definite form of the feminine noun *hlið* 'side'. Both of these nouns are widely used in their basic literal meaning (8.18) as well as in idiomatic structures with a metaphorical meaning (8.19).

(8.18)	a.	[*É*]*g*	*er*	*með*	*sítt*	*hár*	*niður*	*á*	***bak***
		I	am	with	long	hair	down	to	back

'I have long hair that reaches down to the back' (IGC, Bland.is, 2004)

b.	… *að*	*ég*	*keyrði*	*í*	***hliðina***	*á*	*öðrum*	*bíl*
	that	I	drove	into	side	of	another	car

'... that I drove into the side of another car' (IGC, Bland.is, 2004)

(8.19) a. *[Þ]egar maður er að drulla upp á* ***bak*** *í öllum*

when one is at poop up on back in al

verkefnum

tasks

'When one is performing extremely badly in all tasks'

(Twitter, Berglind Festival@ergblind, 2017)

b. *Twitter fór á* ***hliðina*** *í gær eftir að ...*

Twitter went on the-side yesterday after that

'Everyone on Twitter went crazy yesterday after …'

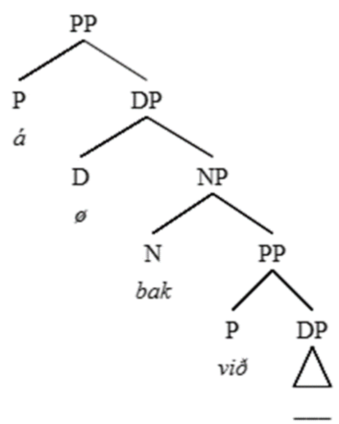
(nutiminn.is, 4. September 2018)

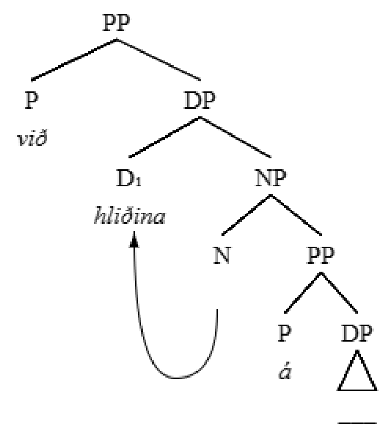
Taking into consideration that the nouns *bak* 'back' and *hlið* 'side' can be related to the relevant elements in the complex prepositions in (8.16), the full forms *við hliðina á* and *á bak við* can be viewed a sequence of a preposition, a noun and a second preposition. In line with earlier work on complex prepositions (e.g., Quirk & Mulholland 1964), this may be represented as in (8.20) where (+N_2) refers to the complement of the second preposition, $\mathbf{P_2}$.

(8.20) $\mathbf{P_1N_1P_2}$ (+N_2)

The presence of two prepositions, i.e., $\mathbf{P_1}$ and $\mathbf{P_2}$, naturally suggests the presence of two prepositional phrases. Given that $\mathbf{P_2}$ (+N_2) conveys information on the location of $\mathbf{N_1}$ rather than an additional (coordinated) location, one might assume a structure of the type in (8.22).

(8.21) a. b.





In what follows, I argue that the structure in Modern Icelandic is ***not*** the one presented in (8.21), but rather that the $\mathbf{P_1N_1P_2}$ string has been lexicalized – or grammaticalized – into a single element. The structure might thus be more similar to that of the prepositions in (8.17) than true $\mathbf{P_1N_1P_2}$ sequences that consist of two prepositional phrases. However, in order to keep the discussion as clear as possible, I will continue to refer to the individual parts of *á bak við* and v*ið hliðina á* in terms of elements in a $\mathbf{P_1N_1P_2}$ sequence.

8.3.2 Variation in writing and usage as an indicator of grammaticalization

Considerable variation is found in the writing of the $\mathbf{N_1P_2}$ elements of the complex prepositions *á bak við* and *við hliðina á*. The simplified forms <*bakvið*> and <*hliðiná*> occasionally appear in printed material, although they are more commonly found online, such as in blog posts and on social media (see examples (8.1b)–(8.2b) above). Other variants are also found on the internet and in informal writing. Variations I have come across in the writing of **N$_1$P$_2$** are listed in (8.22). Angle brackets are used specifically for signaling that these are written variants.

(8.22) Variation in the writing of **N$_1$P$_2$**

Behind: <*bak við*>, <*bakvið*>

Next to, beside:[61] <*hliðina á*>, <hliðin á>, <*hliðinaá*>, <*hliðin'á*>, <*hliðiná*>, <*hliðn'á*>, <*hliðná*>

The written forms in (8.22) do not indicate a significant variation in pronunciation of the **N$_1$P$_2$** sequence. Being subject to the pressures of normalized spelling, writing is often conservative. The traditional forms <*bak við*> and <*hliðina á*> are thus not necessarily informative on how the **N$_1$P$_2$** sequence is perceived by individuals. Innovatively written forms, on the other hand, may provide hints about (ongoing) changes.

As observed in (8.22), **N$_1$** and **P$_2$** are sometimes collapsed into a single form. The lack of space between **N$_1$** and **P$_2$** in <*bakvið*> and <*hliðinaá*> suggests a perception of them as single elements. The same can be said about the forms <*hliðiná*>, <*hliðin'á*> and <*hliðn'á*>, where the apostrophe indicates a conscious omission of the final -*a* in *hliðina*. The omission of <*i*> in the forms <*hliðn'á*> and <*hliðná*> can be viewed as a characteristic of casual speech or taken to signal phonological reduction, a part of grammaticalization (see discussion in Section 8.2.2).

[61] The form <*hliðina*> can also be found. However, this is a difficult form to deal with in relation to the complex preposition *við hliðina á* as searching for only <*hliðina*> returns multiple examples of the lexical noun *hlið* 'side' as well as instances where *hliðina* can arguably be analyzed as an adverb, i.e., having no complement.

If innovatively written forms are taken to indicate grammaticalization, the appearance of the traditional variants <*hliðina á*> and <*bak við*> are potentially deceptive. Their use might either signal the presence of two distinct elements (*bak* and *við, hliðna* and *á*) or be the result of normalized spelling. An individual writing <*bak við*> or <*hliðina á*> might still perceive the $\mathbf{N_1P_2}$ sequence as a single element.[62] In fact, the same individual may sometimes use more than one written variant. As an example, (8.23) shows <*bak við*> occurring in the same sentence as <*bakvið*>.

(8.23)	*… hlyti*	*þá*	*að*	*vera*	***bak***	***við***	*Vesturberg*	*en*	*þetta*
	must	then	to	be	back	by	Vesturber	but	this
	var	***bakvið***		*Lönguvitleysu*			*sýndist*	*mér*.	
	was	behind		Langavitleysa			appeared	me	

'...must be behind Vesturberg, but this appeared to me to be behind Langavitleysa.'

The presence or absence of an initial $\mathbf{P_1}$ (*á* or *við*) is slightly different in nature from the variation in written forms of $\mathbf{N_1P_2}$. While the latter relates to a phonological string being perceived as one or two words, the former can be linked to the presence or absence of a phonological material. Unless individuals are conscious of an initial $\mathbf{P_1}$ (or if they have been explicitly told that there is such an element), they may fail to indicate it in writing. In some cases, individuals may go back and forth on including and omitting $\mathbf{P_1}$. Observe (8.24) which shows the prepositions (*í*) *gegnum* appearing in Bændablaðið in 2023; the examples are from the same article in the paper.

[62] The matter of whether something constitutes one word or two words is the subject of a section in the official writing rules (*Ritreglur*, accessed January 2024; see also Sigtryggsson 2022).

(8.24) a. *…að* *loðnan* *fór* ***í*** ***gegnum*** *hnífasett…*

that the.capelin went in through knife.set

'... that the capelin went through a set of knives…'

(*Bændablaðið* 2023(3):16)

b. *Loðnan* *fór* ***gegnum*** *þrengingu…*

the.capelin went through narrowing

'The capelin went through a narrowing…' (*Bændablaðið* 2023(3):16)

Similar to (8.24), the presence or absence of **P_1** in *á bak við* and *við hliðina á* do not appear to be tied to a specific meaning or a particular use of the prepositions. For instance, the examples in (8.25), which are both taken from Morgunblaðið, show a comparable use of the forms <*á bak við*> and <*bak við*>.

(8.25) a. … *hefur* *myndast* *mikil* *saga* ***á*** ***bak*** ***við*** *þennan bíl.*

has formed great story on back by this car

'... a great story has formed behind this car.' (*Morgunblaðið* 2003(216-B):7)

b. *Hin* *óvenjulega* *saga* ***bak*** ***við*** *bíómyndina* *Salt.*

the unusual story back by the.film Salt

'The unusual story behind the film Salt.'(*Morgunblaðið* 2003(10-B):14)

The complex prepositions *á bak við* and *við hliðina á* both exist in a range of literal and metaphorical meanings. Some examples are provided below, starting with *á bak við*. Although I have chosen to use the full $\mathbf{P_1N_1P_2}$ sequence in the examples, forms lacking $\mathbf{P_1}$ are also grammatical with these meanings for speakers who are generally able to omit $\mathbf{P_1}$. In (8.26a) the preposition conveys a location behind someone's physical back, while in (8.26b and c) it refers to a more abstract "back-side" location.

(8.26)	a.	*Litli*	*drengurinn*	*skýldi*	*sér*	***á***	***bak***	***við***	*móður*
		little	boy	hid	self-DAT	on	back	with	mother
		sína.							
		his							

‘The little boy hid behind his mother.’

b.	*Þvoðu*	*þér*	***á***	***bak***	***við***	*eyrun*!
	wash	yourself-DAT	on	back	with	the.ears

‘Wash behind your ears!’

c.	*Þaðan*	*hefði*	*maðurinn*	*ekið*	***á***	***bak***	***við***	
	Therefrom	had	the.man	driven	on	back	with	
	Arion	*banka*	*og*	*hlaupið*	*í burtu*			
	Arion	bank	and	run	away			

‘From there, the man had driven behind Arion Bank…’

The examples in (8.27) show a metaphorical use of the preposition.

(8.27) a. *Hún fór* ***á*** ***bak*** ***við*** *foreldra sína.*

she went on back with parents hers

'She went behind her parents' back.'

Metaphorical: 'She was dishonest with her parents.'

b. *Hafðu þetta* ***á*** ***bak*** ***við*** *eyrað.*

have this on back with ear

'Keep this behind your ear.'

Metaphorical: 'Don't forget this.'

c. *Hver stendur* ***á*** ***bak*** ***við*** *þetta?*

who stands on back with this

'Who is behind this?'

Metaphorical: 'Who is responsible for this?'

In addition to metaphorical uses as in (8.27), the preposition occurs in some highly idiomatic expressions. Two such are (*á*) *bak við luktar dyr* 'in secrecy (lit. behind closed doors)' and (*á*) *bak við tjöldin* 'without anyone seeing (lit. behind the curtains)'. These can also occur both with and without the initial $\mathbf{P_1}$.

The complex preposition *við hliðina á* 'next to, beside' is mostly used to describe a location next to something that has a physical side, e.g., an individual (8.28a) or building (8.28b).

(8.28)	a.	*Konan*	*sat*	***við***	***hliðina***	***á***	*mér.*
		the.woman	sat	by	side	of	me
		'The woman sat next to me.'					
	b.	*Ísbúðin*	*er*	***við***	***hliðina***	***á*** *sundlauginni.*	
		the.ice.cream.store	is	by	side	of the.swimming.pool	
		'The ice cream store is next to the swimming pool.'					

The phrase can also be used in the metaphorical meaning 'compared to'. Example (8.29a) shows the written variant <*við hliðina á*> and (8.29b) the simplified form <*hliðiná*>.

(8.29)	a.	*auðmýkt*	*og*	*kærleikur*	*eru*	*einskis virði*	***við***	***hliðina***
		Humility	and	love	is	nothing worth	by	side
		á *allri*	*þeirri*	*gleði..*				
		of all	that	joy				
		'Humility and love are worth nothing compared to all that happiness…'						

b. *ég*	*er*	*dvergur*	***hliðiná***	*þér*	*bara*	*svona*
I	am	dwarf	side.of	you	just	so
svo	*þú*	*vitir*				
so	you	know				

‘I am a dwarf compared to you, just so you know…’

The usage in (8.29) is innovative and is likely influenced by English *next to* as in (8.30).

(8.30) ***Next to** him, I am nothing* (Addicted to Love 1997)

As mentioned earlier (Section 8.2.2) semantic change, potentially in the form of non-literal uses of a word, might be indicative of grammaticalization. If this is the case, various senses of the complex prepositions *á bak við* and *við hliðina á* could be claimed to provide cues about when the grammaticalization started. However, many of the non-literal senses are not new in the language and are already found in the uses of the nouns *bak* ‘bak’ and *hlið* ‘side’, not just in the complex prepositions. They can, therefore, not be used as indicators for the starting point of the grammaticalization.[63] However, some of these uses might be taken to indicate the dissociation of *hliðina* and *bak* from relevant lexical items (*hlið* ‘side’, *bak* ‘back).[64]

---

[63] Some non-literal uses of the prepositions *bakvið* and *hliðiná* appear to be relatively new in the language. This is the case for *við hliðina á* in the sense ‘compared to’. However, it can be argued that when this usage emerged, the grammaticalization of *við hliðina á* had already started.

[64] Dissociation from lexical items differs from a simple metaphorical use of those items. In the case of metaphors, speakers are usually conscious (on some level) about where the individual items come from, and this can affect the behavior of metaphorical expressions.

Since the *á bak við* and *við hliðina á* can be used in various different senses, both literal and metaphorical, one might expect correlation between the various meanings and whether the prepositions are used with an initial $\mathbf{P_1}$ or not. While this might be the case for some speakers, it is certainly not universal. In fact, I have only been able to verify the existence of speakers that (i) that never accept the omission of $\mathbf{P_1}$ in *á bak við* and *við hliðina á*,[65] and (ii) speakers who fluctuate between including and omitting $\mathbf{P_1}$, perceiving no difference between when one or the other should be used. As of yet, I have not come across individuals who only accept variants where $\mathbf{P_1}$ has been omitted.

In addition to variation in the presence and absence of $\mathbf{P_1}$, there are also a few instances where $\mathbf{P_1}$ is something other than *á* or *við* (see also Rögnvaldsson 2021a). In (8.31) the preposition *fyrir* 'for' appears instead of the expected *á* in *á bak við*.

(8.31)	*Það*	*voru*	*þrír*	*unglingstrákar*	*að*	*labba*	***fyrir***	***bakvið*** *mig…*
	there	were	three	teenage.boys	INF	walk	for	behind me

'Three teenage boys were walking behind me…'

(Twitter, Alex@alexirisar, Apr 28, 2023)

Rögnvaldsson (2021a) notes that *á bak við* expresses essentially the same meaning as *fyrir aftan* 'behind' and *aftan við* 'behind', although the three phases cannot always be used interchangeably. He furthermore suggests that the innovative variant *fyrir bakvið*, as in (8.31), may be under the influence of *fyrir aftan* and *aftan við*. On this view, *fyrir bakvið*

[65] Interestingly, this speaker is not consistent in their use of other prepositional phrases like (*í*) *gegnum* and (*í*) *kringum* and appears to be able to use them interchangeably, i.e., without any concrete semantic difference. However, this needs to be investigated further.

might be regarded as a multiple source construction in the sense of Van de Velde, De Smet and Ghesquiére (2013).

Just like **P**$_1$ in *á bak við* is occasionally something other than á, **P**$_1$ in *við hliðina á* is occasionally something other than *við*. In (8.32), **P**$_1$ is *fyrir* ‘for’.

(8.32)	…til	að	leika	***fyrir***	***hliðina***	***á***	félaga	sínum	hjá	Hearts.
	to	at	play	for	the.side	of	partner	his	at	Hearts

‘...to play by the side of his partner at Hearts.’

In addition to the occasional example of **P**$_1$ being other than expected, there are a handful of instances where **P**$_2$ in *við hliðina á* is something other than *á*. In (8.33) the preposition *hjá* ‘by’ is used.

(8.33)	*Minn*	*páfagaukur*	*er*	***hliðina***	***hjá***	*hennar*	*búri*
	mine	parrot	is	the.side	by	her	cage

‘My parrot is next to her cage…’

It is possible that variation in **P**$_1$ and **P**$_2$ occurs due to phonological reduction in the complex prepositions. On this view, speakers may perceive a presence of a **P**$_1$ or **P**$_2$ element without being able to recover the intended phonological form. Consequently, they insert **P**-forms that they deem to be appropriate. An alternative view is that forms with unexpected **P**$_1$ or **P**$_2$ elements are in fact multiple source constructions (Van de Velde, De Smet and Ghesquiére 2013), resulting from the blending of two or more source constructions.

Finally, further evidence for the prepositional phrases not always providing strong enough signal for their structure, occasional examples from social media indicate that speakers conflate *á bak við* with two other common phrases, namely *að baki e-u*(*m*) 'behind something/someone's back' and *á bak e-u*(*m*) 'on someone's back'. Unlike *á bak við*, which takes an accusative complement, *að baki e-u*(*m*) 'behind something/someone's back' and *á bak e-u*(*m*) take a dative complement. Two examples of conflated forms are provided in (8.34a) with expected (or corrected) forms shown in (8.35).

(8.34) a. *Mig langar að sjá tölfræðina* ***á*** ***bak*** *þessari*
Me wants to see statistics on back this-DAT
staðhæfingu
claim-DAT
'I want to see the statistics behind this claim.'

b. *Hugmyndafræðin* ***bak*** *þessari aðferð er…*
Ideology back this-ACC method-DAT is
'The ideology behind this method is…'

(8.35) a. *Mig langar að sjá tölfræðina sem liggur* ***að*** ***baki***
Me wants to see statistics.the which lies at back
þessari staðhæfingu
this-DAT claim-DAT
'I want to see the statistics behind this claim.'

b. *Hugmyndafræðin* ***að*** ***baki*** *þessari* *aðferð…*

Ideology.the at back this-DAT method-DAT

'The ideology behind this method is…'

In some cases, the form *bak* is found as a preposition. In fact, this usage occurs already in Old Icelandic as in (8.36) and has survived into the modern language.

(8.36) *fara* *síðan* *norðr* ***bak*** *jólum*

go then north back christmas

'...then they go north after Christmas.' (ONP, *Stu¹R440ˣ 2311*)

Although examples of the type in (8.34a)–(8.34b) and (8.36) do occur in the modern language, they are not very common.

Summarizing the points discussed here, it may be claimed that innovatively written variants of $\mathbf{N_1P_2}$ in *á bak við* and *við hliðia á* suggest an interpretation of $\mathbf{N_1P_2}$ as a single element. As for the presence and absence of $\mathbf{P_1}$, it seems that phonological cues for its presence may not be very strong.[66] Individuals who allow $\mathbf{P_1}$ to be omitted have no obvious

[66] While a reduced strength of $\mathbf{P_1}$ may be mostly attributed to the process of grammaticalization, it is plausible that general intonational rules of Icelandic also play a role. Árnason (1994–1995:109; 2011:286–287) has proposed a hierarchy of strength relations between word classes where prepositions are considered "weaker" than nouns and verbs, but stronger than personal pronouns, (i).

(i) noun > verb > preposition > personal pronoun (Árnason 2011:287)

Árnason (2011:286) notes that it can happen that "a word from a stronger class attracts phrasal stress away from a following word of a weaker class". If the whole $\mathbf{P_1N_1P_2}$ sequence is regarded as a single phrase, it is possible that $\mathbf{N_1}$ might attract stress away from other close elements, contributing towards a phonological reduction.

rule as to when it is present or absent. Instead, they may use the two forms (with and without $\mathbf{P_1}$) seemingly interchangeably. Further evidence for the phonological signal being rather weak involves the substitution of an expected $\mathbf{P_1}$ with a different preposition. In these cases, speakers may be able to infer the presence of an initial $\mathbf{P_1}$ but fail to pick out the "correct" form, causing them to insert a different preposition. Thus, evidence from written material (and judgments from two speakers) can be taken to suggest that the sequence ($\mathbf{P_1}$)$\mathbf{N_1P_2}$ is viewed as a single unit. Consequently, the dropping of $\mathbf{P_1}$ may be regarded as a form of phonological reduction.

8.3.3 Syntactic (in)flexibility as an indicator of grammaticalization

Some arguments for the grammaticalization of *á bak við* and *við hliðina á* into a single **P** element have already been provided in Subsection 8.3.2. In this subsection, further evidence, from syntactic behavior, is provided.

Quirk & Mulholland (1964:65) note fourteen parameters that may be used for determining behavioral properties of $\mathbf{P_1N_1P_2}$ sequences, assuming that such sequences fall on a continuum of being more grammatical or less grammatical (see also Svenonius 2006:50, 56; Hoffmann 2004 for behavioral properties of such sequences). By more and less grammatical they mean that some structures are rather inflexible and appear to function similar to single-word prepositions while others show a behavior that indicates compositionality. The properties used to determine how (in)flexible $\mathbf{P_1N_1P_2}$ sequences are (Quirk & Mulholland 1964:65) are as follows: i) $\mathbf{P_1}$ can be replaced, ii) $\mathbf{P_1}$ can be deleted showing that $\mathbf{N_1}$ is head of a Nom. Grp., iii) $\mathbf{N_1}$ has definite article, iv) $\mathbf{N_1}$ is concrete, v) $\mathbf{N_1}$ can take different deictics, vi) $\mathbf{N_1}$ can be premodified by adjective, vii) $\mathbf{N_1}$ can show number contrast in the $\mathbf{P_1N_1P_2N_2}$ sequence, viii) $\mathbf{N_1}$ is used as a member of a lexical set, ix) $\mathbf{P_2N_2}$ can be deleted, x) with P2N2 deleted, $\mathbf{N_1}$ can show number contrast, xi) $\mathbf{P_2N_2}$ can be replaced by genitive pronoun premodifying $\mathbf{N_1}$, xii) $\mathbf{P_2N_2}$ can be replaced by a demonstrative (that, such) premodifying $\mathbf{N_1}$, xiii) There is no transformation relation between $\mathbf{N_1}$ $\mathbf{N_2}$ and a VC structure respectively, and xiv) $\mathbf{N_2}$ must be concrete.

Unfortunately, most of the parametric tests proposed by Quirk & Mulholland cannot be applied straightforwardly to the Icelandic prepositions *á bak við* and *við hliðna á*. They are specifically put together for English and yield either ungrammatical structure in Icelandic, provide older or newer variants of the prepositions or result in strengthened connections to the lexical elements *bak* 'bak' and *hlið* 'side'. Despite this, three properties may be noted here.

First, adding a premodifying adjective to N_1 (*bak* and *hlið*) yields an ungrammatical sentence or radically alters the meaning. This is shown in (8.37) with the object *bolti* 'ball' which does not have a physical 'back' or a natural 'side' to it.

(8.37)	a.	**Ég*	*setti töskuna*	***(á) gult/slétt***	***bak***	***við***	*boltannann.*
		I put	the.bag	on yellow/straight	back	by	ball.the

b.	*Ég*	*setti*	*töskuna*	*(***við***)	***gulu/sléttu***	***hliðina***
	I	put	the.bags	by	yellow/straight	the.side
	á	boltanum				
	of	ball.the				

In (8.37b), the outcome is only grammatical provided the element *við* ‘with’ is present. In this case, the noun *hlið* ‘side’ must be understood as ‘surface’, leaving room for an interpretation where the ‘bag’ is placed next to the ‘yellow’ or ‘straight, smooth’ surface of the ball.[67] This is different from conveying that the ‘bag’ is placed (anywhere) next to the ball. Also, note how the presence and absence of $\mathbf{P_1}$ does not affect the outcome in (8.37a); it is always ungrammatical.

Second, modifying the number of $\mathbf{N_1}$ gives an ungrammatical structure for both *á bak við* and *við hliðina á*, irrespective of whether $\mathbf{P_1}$ is present or not (8.38).

(8.38)	a.	**Ég*	*setti*	*töskurnar*	(***á***)	***bökum***	***við***	*boltana*.
		I	put	the.bags	on	the.backs	by	the.balls
	b.	**Ég*	*setti*	*töskurnar*	(***við***)	***hliðarnar***	***á***	*boltunum*.
		I	put	the.bags	by	the.sides	of	the.balls

Unlike in (8.37b), *hlið* ‘side’ in (8.38b) can no longer be perceived as the ‘surface’ of a ball, likely because balls do not have multiple surfaces.

The third property tested for here involves the (im)possibility of modifying the definiteness of N_1. While Quirk & Mulholland (1964:65) specifically focus on whether N_1 can take a definite article or not, I have simply turned this property into whether or not the definiteness can be modified. As already explained (see Section 8.3.1), N_1 in *á bak við* has an indefinite form while N_1 in *við hliðina á* has a definite form.

[67] It might be the case that the surface of the ball is yellow or smooth in some areas and not others. Placing a bag next to the smooth (part of the) surface is thus a possible reading.

(8.39)	a.	**Taskan*	*er*	(***á***)	***bakinu***	***við***	*boltana*.
		the.bag	is	on	the.back	by	balls.the

	b.	**Taskan*	*er*	(***við***)	***hlið***	***á***	*boltanum*.	
		the.bag	is	by	side	of	the.balls	

Although (8.39b) is marked as ungrammatical, few instances of such use with P_1 *við* can be found. In those cases, the complement appears to be understood as having a physical 'side'. This is shown in (8.40).[68]

(8.40)	… *að*	*vera*	***við***	***hlið***	***á***	*vinsælum*	*veitingastað*
	to	be	by	side	of	popular	restaurant

'... to be next to a popular restaurant' (*Víkurfréttir* 2020(44):10)

If behavioral properties such as those tested in (8.37)–(8.39) provide information on how grammaticalized $\mathbf{P_1N_1P_2}$ sequences are, we may assume that *á bak við* is slightly further along in the grammaticalization than *við hliðina á,* simply because the latter is marginally more modifiable. However, it may also be called into question whether *á bak við* and *við hliðina á* are true $\mathbf{P_1N_1P_2}$ sequences. The examples in (8.37)–(8.39) above, testing modification of $\mathbf{N_1}$, suggest that the form of $\mathbf{N_1}$ is quite unchangeable. To further

[68] Wood (p.c.) suggested that examples like (8.40), with the indefinite form *hlið*, might be under the influence of the phrase *hlið við hlið* 'side by side'. This appears plausible and would mean that *við hlið á* should essentially be regarded as a multiple source construction (cf. Van de Velde, De Smet & Ghesquiére 2013).

demonstrate the inflexibility of the full sequences it is useful to compare them to similar sequences.

Many sequences apparently consisting of $\mathbf{P_1N_1P_2}$ can be found in Icelandic in daily use (for a list of such sequences see Kress 1982), some are listed in (8.41).

(8.41) Some $\mathbf{P_1N_1P_2}$ sequences in Icelandic

á bakinu á e-um 'on someone's back'

á hillunni í skápnum 'on the shelf in the cupboard'

á hliðinni á e-u 'on the side of something'

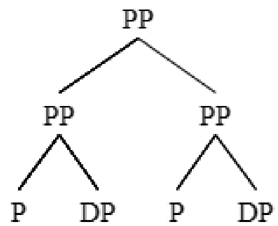
í kappi við tímann 'competing with time'

í skuld við e-n 'in someone's debt'

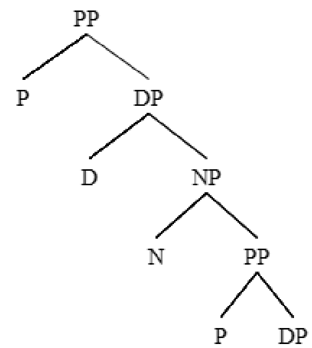
um borð í skipi 'on board a ship'

The structure of the sequences in (8.41) are not identical. For instance, *bak* 'back' is a physical part of an individual and *hilla* 'shelf' can be positioned inside a cupboard. However *kapp* 'competition' is not a part of time nor is *skuld* 'debt' a part of an individual. Reflecting various types of relationships between the different nouns, the syntactic structure of the sequences is likely not the same. Some may have two adjoined PPs, for instance as in (8.42a), while some may include a PP as a complement of a noun inside another PP (8.42b). Other structures, such as a PP adjoined to DP or NP, might also be possible for some $\mathbf{P_1N_1P_2}$ sequences.

(8.42) a.



b.



For the purpose of showing how the complex prepositions *á bak við* and *við hliðina á* behave differently from other structures with two full PPs, examples have been chosen that resemble how the complex prepositions may originally have been constructed. These are shown in (8.43) and include structures where the second PP is a complement of the lexical nouns *hlið* 'side' (8.43a) and *bak* 'back' (8.43b).

(8.43) a. *Það* *er* *ör* [$_{PP}$ ***á*** [$_{DP}$ ***bakinu*** [$_{PP}$ ***á*** *honum*]]].

there is scar on the.back on him

'There is a scar on his back.'

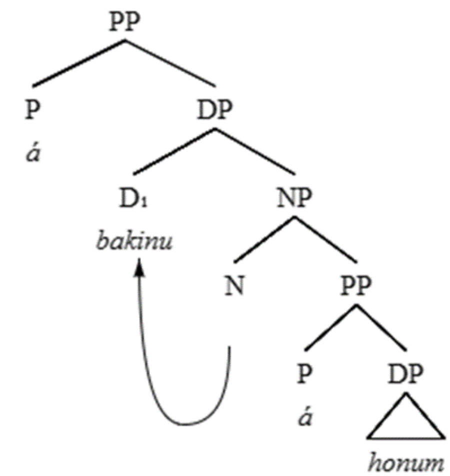
b. *Það* *er* *rispa* [$_{PP}$ ***á*** [$_{DP}$ ***hliðinni*** [$_{PP}$ ***á*** *bílnum*]]].

there is scratch on the.side on the.car

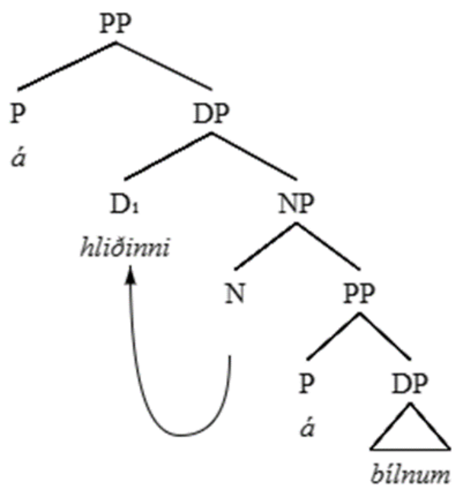
'There is a scratch on the side of the car.'

(8.44)

a.



b.



Even though the second PP is considered to be the complement of the noun in the first PP, the reverse order of (8.43) can be found. This is shown in (8.45), where the PPs *á bakinu* and *á hliðinni* appear to the right of *á honum* 'on him' and *á bílnum* 'on the car'. In some ways, this is reminiscent of elements that have been dislocated to the right and provide more specific information on a phrase appearing earlier (on Right Dislocation see Thráinssson 2007:363, 367–368).[69] It might also be the case that the underlying structure of the examples in (8.45) is different from that of the examples in (8.44).

---

[69] Observe for instance the example in (i) where the dislocated phrase (*hann*) *Alfreð* identifies *hann* 'he' who is mentioned earlier:

(i)	*Hann*	*er*	*langbestur*,	(*hann*)	*Alfreð*.
	he	is	long-best	(he)	Alfreð

'He is by far the best, Alfred.' (from Thráinsson 2007:636, ex (7.55))

(8.45) a. *Það* *er* *ör* ***á*** ***honum***, ***á*** ***bakinu***

there is scar on him on the.back

'There is a scar on him, on the back.'

b. *Það* *er* *rispa* ***á*** ***bílnum***, ***á*** ***hliðinni***.

there is scratch on the.car on the.side

'There is a scratch on the car, on the side.'

Returning to the more neutral word order in (8.43), we may note that the second PP can be topicalized to create a focus-type structure as in (8.46).[70]

(8.46) a. [$_{\text{PP}}$ ***Á*** ***bílnum***]$_{\text{i}}$ *hefur* *aldrei* *verið* *rispa* [$_{\text{PP}}$ *á*

on the.car has never been scratch on

[$_{\text{DP}}$ *hliðinni* [$_{\text{PP}}$ __]$_{\text{i}}$]].

on the.side

'On the car, there has never been a scratch on the side.'

---

[70] Harðarson (2017:204) marks extractions like in (i), which appears similar to (8.46), as ungrammatical. I disagree with this judgment. For me (and two individuals I consulted with), (i) is grammatical. My only comment is that it is slightly pragmatically odd to use *hverju* (dative singular neuter of the interrogative pronoun *hver* 'who') as it suggests an operation is being made on something inanimate. Changing the form to the dative singular masculine (*hverjum*) or feminine (*hverri*) eliminates this oddness.

(i) *[*Á* *hverju*]$_{\text{i}}$ *gerðu* [*þeir*] [*tíðar* *aðgerðir* __$_{\text{i}}$]?

On what performed they frequent operations

'What did they frequently operate on?'

b. [PP ***Á*** ***honum***]i *hefur aldrei verið ör* [PP *á* [DP *bakinu* [PP __]i] .

on him has never been scar on the.back

'On him, there has never been a scar on the back.'

Unlike in (8.46), topicalization of the second PP is not possible in the context of the complex prepositions *við hliðna á* and *á bak við.* This is shown in (8.47) where the result is ungrammatical.[71]

(8.47) a. *[PP ***Á*** ***bakaríinu***]*i* *er sundlaugin* [PP *við hliðina* __*i*].

on the.bakery is swimming.pool by the.side

b. *[PP ***Við húsið***]*i* *hefur aldrei verið brunnur* [PP *á bak* __*i*].

by the.house has never been well on back

The ungrammaticality of (8.47) suggests that the structure of the complex prepositions *á bak við* and *við hliðina á* may not be the same as of the $\mathbf{P_1N_1P_2}$ strings in (8.43).

Asymmetry in behavior of the complex prepositions in *á bak við* and *við hliðina á* and the structures in (8.43) is found in question-answer pairs. Observe how questions about the location of 'a scar' (8.48) or 'a scratch' (8.49) can be answered with a single prepositional phrase ($\mathbf{P_1N_1}$), but not with a $\mathbf{P_1N_1P_2}$ string.

[71] If *við hliðina á bakaríinu* and *á bak við húsið* in (8.47) consist of a single P projection, one might expect it should be possible to topicalize the DP and leave the preposition in situ. While topicalizing only the DP in (8.47a) and (8.47b) yields an ungrammatical structure, it does not necessarily speak against an analysis of a single P. As noted by Thráinsson (2007:345–346), Icelandic allows preposition stranding in some types of prepositional phrases whereas pied piping appears obligatory in others. In short, "the restrictions on preposition stranding in Icelandic remain to be studied in detail" (Thráinsson 2007:345).

(8.48)	Q:	*Hvar*	*á*	*honum*	*er*	*ör?*
		where	on	him	is	scar
		‘Where does he have a scar?’				
	A:	*Á*	*bakinu.*			
		on	the.back			
		‘On the back’				
	A’:	**Á*	*bakinu*		*á.*	
		on	the.back		of	

(8.49)	Q:	*Hvar*	*á*	*bílnum*	*er*	*rispa?*
		where	on	the.car	is	scratch
	A:	*Á*	*hliðinni.*			
		on	the.side			
	A:	**Á*	*hliðinni*		*á.*	
		on	the.side		of	

Contrary to (8.48)–(8.49), questions involving the location *á bak við* ‘behind’ cannot be answered with a sequence of $\mathbf{P_1N_1}$ but only with $(\mathbf{P_1})\mathbf{N_1P_2}$. This is presumably due to the $(\mathbf{P_1})\mathbf{N_1P_2}$ sequence being understood as a single functional element. Observe the example in (8.50).

(8.50) Q: *Hvar* *við* *húsið* *er* *brunnurinn*?

where by the.house is the.well

A: **Á bak*

A': (*Á*) *bak við*

Answers to questions involving the location *við hliðina á* 'next to' behave differently from both (8.48)–(8.49) and (8.50). Note the question is already pragmatically strange as it indicates that *skúrinn* 'the shed' is physically 'on' the house, a description that does not fit building practices in Iceland. In any case, neither a $\mathbf{P_1N_1}$ sequence nor a $\mathbf{(P_1)N_1P_2}$ sequence can be used (8.51).

(8.51) Q: *Hvar* *á* *húsinu* *er* *skúrinn?*

where on the.house is the.shed

A: **Við hliðina*

A' #(*Við*) *hliðina á*

The reason for the infelicity of A' in (8.51) is likely not due to the structure of the $\mathbf{(P_1)N_1P_2}$ sequence per se. Instead, it may be linked to semantic or pragmatic oddness. As already stated, the question in (8.51) signals that *hvar* 'where' must be located physically 'on the house' rather than 'next to' it. The location is then contradicted with the use of (*við*) *hliðina á* in the reply which cannot convey anything but 'next to' or 'by'. Thus, the question-answer test is, unfortunately, not applicable for the complex preposition *við hliðina á*.

Another piece of evidence in favor of the sequence $\mathbf{N_1P_2}$ (*bak við* and *hliðina á*) being a single element in the modern language involves instances of reduplication of $\mathbf{P_2}$. Such examples occasionally appear in informal language on the internet. In (8.52) the preposition *við* 'by' appears immediately after the written variant <*bakvið*>.

(8.52) a. ***bakvið við*** *spilaranna*

behind by the.player

'Behind the player.'

In (8.53), two examples of a reduplicated $\mathbf{P_2}$ in *við hliðina á* are provided, one with *á* occurring after the conservative written wariant <*við hliðina á*> (8.53a) and another with a more innovative written variant <*við hliðiná*> (8.53b).

(8.53) a.

…af	*hverju*	*við*	*erum*	*að*	*vinna*	***við***	***hliðina***
why		we	are	at	working	by	side
á,	***á***	*fólki*	*sem…*				
of	of	people	that				

'... why we are working beside people that…'

b.

…hann	*býr*	***við***	***hliðiná***	***á***	*mér*	*nánast*
he	lives	by	side.of	of	me	almost

'...he lives, so to speak, next to me'

For a comparison, repetition of $\mathbf{P_2}$ is not possible when $\mathbf{N_1}$ is a lexical noun taking a prepositional complement (8.54).

(8.54)	a.	**Örið*	*er*	***á***	***bakinu***	***á***	***á***	*honum*.
		the.skar	is	on	the.back	on	on	him
	b.	**Rispan*	*er*	***á***	***hliðinni***	***á***	***á***	*bílnum*.
		the.scratch	is	on	the.side	of	of	the.car

Reduplication as in (8.53) and (8.54) can be taken to be an indicator of grammaticalization as it shows that the two (reduplicated) elements likely serve a different function in the structure. In fact, these may be viewed in terms of divergence (Hopper 1991:24–25) which is a process resulting in two or more elements "having a common etymology, but diverging functionally." A further instance of divergence involves the development of a copula from a demonstrative pronoun. This is a common change in the language of the words (Heine & Kuteva 2002:108–109, Lohndal 2009) and may result in a surface string such as the one in (8.55), where the newly emerged copula can occur alongside the traditional demonstrative.

(8.55) Sranan CE

da	*somma* ***da***	*wan*	*boen*	*somma*.
that	person is	a	good	person

'That's a good person.' (Arends 1986:107, cited from Kuteva, et al. 2019:136)

Another example of reduplication, albeit of a slightly different kind, involves a reinterpretation of an infinitival marker as a complementizer. Some predicates that take a verbal complement with an infinitival marker *að* 'to', for instance *ætla* 'intend, be going to', sometimes appear with two instances of *að,* cf. (8.56). In cases like these, one *að* can be analyzed as a complementizer and the other as an infinitival marker.[72] On a similar type of "doubling" of the infinitival marker in Norwegian, see Faarlund (2015).

(8.56) *Og hann ætl-ar* ***að*** ***að*** *verða fullgildur meðlimur…*
and he intends COMP INF become fully.qualified member
'And he intends to become a fully qualified member…'
(https://www.visir.is/g/20071518417d)

Concluding this section, a few things may be noted about the complex prepositions *á bak við* 'behind' and *við hliðina á* 'next to, beside' in Modern Icelandic. First, there is considerable variation in the use of these prepositions, not all of which has been done justice here. The conservative variants <*við hliðina á*> and <*á bak við*> suggest a structure containing two prepositional phrases, one embedded within the other [$_{PP}$ [$_{DP}$ [$_{PP}$]]].

[72] A second type of "duplicated" infinitival marker also exists in Icelandic, shown in (i). This type of structure is also found in Old Icelandic (Benediktsson 1976; Faarlund 2004:137; Thráinsson 2005:489) and has been linked to the rise of the infinitival progressive (*vera að + infinitive*) in Icelandic (Jóhannsdóttir 2011:15).

(i) *Hún var* ***að*** ***að*** *mála allan daginn.*
she was at at painting all the.day
'She was painting all day.'

The doubling in (i) appears to differ from the one in (8.56) in that it contains a particle (or a preposition) *að* which together with the verb *vera* (cf. *vera* ***að***) describes how someone is 'at doing something'. The second *að* in (i) is an infinitive marker (Thráinsson 2005:489). In (8.56) no one is 'at doing something'. Rather, the first *að* (in *ætlar* ***að***) appears to function as a complementizer, with the second *að* being an infinitival marker. Doubling as in (8.56) is attested in Modern Icelandic with verbs such as *fara* 'go, begin', *hugsa* 'think', *langa* 'want', segja 'say', *vilja* 'want' and *ætla* 'intend'.

However, these written variants are somewhat deceptive. The appearance of contracted form <*hliðiná*> and <*bakvið*> indicate that the $\mathbf{N_1P_2}$ sequence is taken to constitute a single element. Behavior in selected syntactic tests (topicalization and question-answer pairs) show that *á bak við* and *við hliðina á* are not as flexible as structures where the second PP is a complement of a noun in a higher PP. Although inflexibility might, in theory, simply be a coincidence, is taken here to indicate conventionalization of the string as a single unit. Seeing that the writing of $\mathbf{N_1P_2}$ as one word <*hliðiná*> <*bakvið*> or two words <*við hliðina á*> and <*á bak við*> does not affect the inflexibility of the structure, it is concluded that the univerbation of $\mathbf{N_1P_2}$ has already taken place in Modern Icelandic. Furthermore, the variation between presence and absence of $\mathbf{P_1}$, without any noticeable change in meaning or usage, indicates that the whole $\mathbf{P_1N_1P_2}$ sequences may serve a single function, namely that of a preposition. In other words, we are dealing with the grammaticalization of $\mathbf{P_1N_1P_2}$ and the presence or absence of $\mathbf{P_1}$ is here taken to be related to phonological reduction.

8.3.4 Documenting older stages

To establish a long-time diachronic development of the complex prepositional phrases *á bak við* 'behind' and *við hliðina á* 'next to, beside', it is necessary to consider older stages of Icelandic.[73] For this purpose *Íslenskt Textasafn* (e. *Icelandic Text collection*, ÍT), *Ordbog over det norrøne prosasprog* (e. *Dictionary of Old Norse Prose*, ONP for short), and *Ritmálssafn Orðabókar Háskólans* (ROH) were consulted. *The Icelandic Text collection* (ÍT) and *Dictionary of Old Norse Prose* (ONP) both cover the period of Old Icelandic,

[73] It is outside the scope of the present study to provide a complete and exhaustive documentation of the historical development of the two complex prepositions, *á bak við* 'behind' and *við hliðina á* 'besides'. The following is a rough sketch.

from ca. the 12th century to 1640 (examples from younger manuscripts are included). ÍT consists of normalized searchable texts, including the Icelandic family sagas as well as some other medieval Icelandic texts. ONP contains a number of attested examples from Old Icelandic (and Old Norse). *Ritmálssafn Orðabókar Háskólans* (ROH) contains attested examples from 1540 to the early 20th century.

Starting with *á bak við* 'behind', no example of the full $\mathbf{P_1N_1P_2}$ sequence was found in ONP under the headword *bak*. However, a sequence of a preposition and the noun *bak*, followed by a noun in the dative case, was found to convey the meaning 'behind' (8.57).

(8.57) $\mathbf{P_1N_1}$ $(+N_2)_{\mathrm{DAT}}$

Two examples of the $\mathbf{P_1N_1}$ sequence *á bak* are provided in (8.58). In (8.58a), the sequence refers to a location 'behind the chair'; In (8.58b) it involves a movement to 'behind the tent'. The manuscript in which the examples were found along with an approximate date is provided in brackets.

(8.58)	a.	***a***	***bac***	*stolinom*	*stoð*	*buandamvgrinn…*
		on	back	chair.DAT	stood	crowd

'The crowd stood behind the chair...'

(ONP, ÓH 165 l.12, Holm Perg 2 4to, c 1250 – 1300)

b.	*Eptir*	*þat*	*gengr*	*hann*	*út*	*ok*	***á***	***bak***	*tjaldinu*
	after	that	walks	he	out	and	on	back	tent.DAT

‘After that, he goes out and behind the tent.’

(ONP, OrvM 176 l.9, AM 344 a 4to c. 1350 – 1400)

Given that ONP covers the period of Old Icelandic fairly extensively, it comes as a surprise that the $\mathbf{P_1N_1P_2}$ sequence *á bak við* is not found there. Since the oldest form of the complex preposition has previously been noted to be *á bak við* (see Friðjónsson 2004(125):52), the absence of the full $\mathbf{P_1N_1P_2}$ sequence from ONP raises questions about its age.

ONP is not the only source for Old Icelandic. Old Icelandic material is also found in ÍT where a search targeting the sequence *á bak við* yields three results of the complex preposition. All of these occur in one and the same Icelandic family saga, Heiðarvíga saga. An example is provided in (8.59).

(8.59)	*Gestur stendur* ***á***	***bak***	***við***	*tjaldið*
	Gestur stands on	back	with	the.tent.ACC

‘Gestur stands behind the tent. (ÍT, Heiðarvíga saga, Ch. 10)

For those familiar with Heiðarvíga saga, the presence of *á bak við* in the saga and the absence of the sequence in ONP immediately raises suspicion. Even though Heiðarvíga saga can be dated to the second half of the 13th century (Finley 2003), its preservation is problematic as the saga only exists in a single manuscript. Unfortunately, the first half of the manuscript was destroyed in the Copenhagen Fire of 1728, resulting in Jón Ólafsson of

Grunnavík (b. 1705, d. 1779) recounting the beginning of the saga in writing to prevent the content from being completely lost (Ólason 2006:111). Remarkably, the sequence *á bak við* only occurs early on in the saga, in chapters 9, 10, and 12, i.e., in the part retold by Jón. The examples can therefore not be taken to represent Old Icelandic, but must be attributed to Jón Ólafsson himself and taken to reflect 18th century language. Further examples of *á bak við* are found in ÍT and ROH in material from the 18th century onwards, suggesting the emergence of the sequence around or shortly before that time.[74] An example from the mid-18th century is provided in (8.60).

(8.60)	*...sem*	*hann*	*var*	*að ganta*	*um*	*oftlega*
	which	he	was	to joke	about	often
	á	***bak***	***við***	*mig*.		
	on	back	with	me-ACC		

'... which he often joked about behind my back.'

(ROH, *Sjálfsævisaga síra Þorsteins Péturssonar*, mid-18th c.)

Taken together, the data from ONP, ÍT and ROH suggest that *á bak við* $(+N)_{ACC}$ replaced an earlier structure *á bak* $(+N)_{DAT}$ somewhere around the 18th century, or in "Middle" Icelandic, (8.61).

[74] Since Jón Ólafsson was born in 1705 and he used the full form <*á bak við*>, it is likely that the full $\mathbf{P_1N_1P_2}$ $(+N_2)_{ACC}$ sequence had surfaced already in the late 17th century.

(8.61) The complex preposition *á bak við* 'behind'

P_1N_1 $(+N_2)_{DAT}$	replaced by	**$P_1N_1P_2$** $(+N_2)_{ACC}$
Old Icelandic		c. 18th century onwards

As to the omission of the initial **P_1**, the sequence *bak við* (without the initial *á*) is attested only once in Old Icelandic according to ONP. The example, which is provided in (8.62), comes from Lazaruss saga (e. *The story of Lazarus*) from a manuscript, Reykjahólabók (Holm Perg 3 fol), dating from the mid-16th century, c. 1530–1540 (ONP s.v.).

(8.62)

gieck	*hann*	*[...] j*	*eitt*	*lithed*	*skoth*	***bak***	***vid***	*hvrdina*	*aa*	*kierkivne*
walked	he	into	one	small	corner	back	by	door	of	the.church

'... he went [...] into a small corner behind the church door'

(ONP, *LazReyk 17621*)

The example in (8.62) appears to be the oldest example of **N_1P_2** $(+N_2)_{ACC}$.[75] Viewing it as an early attestation of the change *á bak við* > *bak við* is somewhat problematic since the full sequence **$P_1N_1P_2$** $(+N_2)_{ACC}$ only appears about two centuries later, or in the 18th century (cf. examples above).[76] Additionally, the text in (8.62) is a translation that shows many "un-Icelandic" characteristics (Reykjahólabók I 1969:xxxix-xl; Axelsdóttir 2002:124–

[75] Friðjónsson (2004) states that the oldest examples of the change *á bak við* > *bak* is from the first half of the 16th century. The example from The story of Lazarus could be what he has in mind.

[76] Alternatively, the example in (8.62) might be viewed as representing the change *bak* > *bak við*; the form *bak* appearing as a preposition early on in the language.

125). The scribe (and the translator) is thought to be Björn Þorleifsson who spent a part of his life in Bergen, Norway and thus, the language might be under Norwegian influence. Leaving (8.62) aside, the sequence *bak við* is next found in texts from the 18th century (8.63). Note that one of the examples (8.63b) is from Jón Ólafson who was already mentioned above in relation to an early attestation of the full form *á bak við*.

(8.63)	a.	*Kviðan [Rígsþula] er*	*til*	***bak***	***við***	*Eddu*	*þá,*	*sem…*
		poem Rígsþula is	to	behind	by	Edda	that	which
		‘The poem Rígsþula exists behind the Edda that…’						

(ÍT, PVídSkýr, early 18th c.)

b.	*… í*	*húsinu*	***bak***	***við***	*heitann*	*kakalofn*.
	in	the.house	back	with	hot	furnace
	‘… in the house behind a hot furnace.’ (ÍT, Nikulás Klím)					

However, it was not until the 19th century and early 20th century that *bak við* became frequent.[77] An early 19th century example is provided in (8.64). Note the metaphorical use of the expression.

(8.64)	*Viljir*	*þú*	*heldur*	*að*	*eg*	*hafi*	***bak***	***við***	*eyrað…*
	want	you	rather	that	I	have	back	with	ear
	‘If you rather want that I keep in mind …’ (ÍT, BiskGörð: early 19th century)								

[77] The form *bak við* is found in two a popular songs and a poem: *Sigling* ‘Sailing’ (Song: Friðrik Bjarnason (1880-1962), Text: Örn Arnarson), *Lagið um það sem er bannað* ‘The song about that which is not allowed’ (song and text: Sveinbjörn I. Baldvinsson), and *Ljáðu mér vængi* ‘Lend me wings’ (text: Hulda, Unnur Benediktsdóttir Bjarklind).

Finally, the written variant <*bakvið*> appears in material from the mid 19th century onwards. An example from a 19th century translation of One Thousand and One Nights is proved in (8.65).

(8.65) ***Bakvið*** *fortjaldið* *muntu* *finna* *gullna* *hurð…*

behind the.curtain will.you find golden door

'Behind the curtain you will find a golden door…' (ÍT, Þús1, mid-19th century)

Turning to the complex preposition *við hliðina á* 'beside, next to', no examples were found in Old Icelandic through a search in ONP and ÍT. The oldest examples where the noun *hlið* 'side' appears in a configuration conveying a location in space dates from the 18th century. In example (8.66a), the **P_1N_1** sequence takes a complement in the dative case (*honum* 'him') which has been topicalized. In (8.66b), the complement is in the genitive case.

(8.66) a. *<u>honum</u>* *hvervetna /* ***við*** ***hlið*** *fylgin.*

him-DAT everywhere by side follow

'Following everywhere by his side' (ÍT, MiltPar, late 18th-century)

b. *Vil* *eg* *hvílast* ***við*** ***hlið*** *<u>hans</u>* *og* *kýs…*

want I rest by side him-GEN and choose

'I want to rest by his side and choose…' (ÍT, Draupn, late 19th century)

The alternation between dative and genitive is not surprising. A special use of the dative, *dativus sympatheticus* (sometimes nicknamed 'dative of body parts'), is frequently found in older Icelandic, especially in relation to inalienable possessions like body parts (Skard 1951; Bjarnadóttir 1989). In examples such as in (8.66), the dative appears when coordinating a position in relation to a human being. The genitive, on the other hand, can either be used in relation to a human or a non-human.

If sequences as in (8.66) are considered to be the origin of the complex preposition *við hliðina á*, the initial status might be represented as in (8.67).

(8.67) $\mathbf{P_1N_1}$ $(+N_2)_{DAT/GEN}$

Earliest instances of the full $\mathbf{P_1N_1P_2}$ $(+N_2)_{ACC}$ sequence, *við hliðina á*, appear in material from the 19th century. Two examples are given in (8.68).

(8.68) a.	*var*	*hann*	*jarðaðr með*	*mikilli*	*viðhöfn*		*í*	*stórhertugans*
	was	he	buried with	great	ceremony		in	grand.duke
	grafkapellu		***við hliðina***	***á***	*vini*	*sínum*		
	mausoleum		by the.side	of	friend	his		

'He was buried with a grand ceremony in the mausoleum of the grand duke, by the side of his friend.' (*Skírnir* 1832:113)

b.	*og*	*spurði*	*so*	*franskann*	*mann,*	*sem*	*sat*	***við***
	and	asked	then	french	man	who	sat	by
	hliðina		***á***	*mér…*				
	the-side		of	me				

'... and asked then a frenchman who sat next to me…'(ÍT, bref_GTogBP)

The shift from a $\mathbf{P_1N_1}$ sequence taking a dative or genitive complement to a $\mathbf{P_1N_1P_2}$ with an accusative complement thus seems to have occurred between the 18th and 19th century, (8.69).

(8.69) The complex preposition *við hliðina á* 'next to, beside'

$\mathbf{P_1N_1}$ $(+N_2)_{DAT/GEN}$	replace by	$\mathbf{P_1N_1P_2}$ $(+N_2)_{DAT}$
18th century		19th century

Finally, examples where the initial $\mathbf{P_1}$ has been omitted are found from the late 20th century.

8.3.5 Summarizing two major changes

Summarizing the documentation of both complex prepositions, *á bak við* and *við hliðina á*, the following trajectories of the sequence of elements might be suggested, (8.70)–(8.71). Note that I have chosen to omit the 16th century outlier of <*bak við*> that appeared in a translation of The story of Lazarus (see example (8.62) bove). Consequently, the oldest example of the $\mathbf{N_1P_2}$ sequence lacking the initial $\mathbf{P_1}$ is from the 18th century.

(8.70) Development of the sequence of elements in the complex preposition *á bak við*

Sequence I	Sequence II	Sequence III	Sequence IV
$\mathbf{P_1N_1}$ $(+N_2)_{\mathrm{DAT}}$	$\mathbf{P_1N_1P_2}$ $(+N_2)_{\mathrm{ACC}}$	$\mathbf{N_1P_2}$ $(+N_2)_{\mathrm{ACC}}$	$\mathbf{N_1P_2}$ $(+N_2)_{\mathrm{ACC}}$
á bak	***á bak við***	***bak við***	***bakvið***
Old Icelandic	c. 18th century	c. 18th century	19th century

(8.71) Development of the sequence of elements in the complex preposition *við hliðina á*

Sequence I	Sequence II	Sequence III	Sequence IV
$\mathbf{P_1N_1}$ $(+N_2)_{\mathrm{DAT/GEN}}$	$\mathbf{P_1N_1P_2}$ $(+N_2)_{\mathrm{ACC}}$	$\mathbf{N_1P_2}$ $(+N_2)_{\mathrm{ACC}}$	$\mathbf{N_1P_2}$ $(+N_2)_{\mathrm{ACC}}$
við hlið	***við hliðina á***	***hliðina á***	***hliðiná***
c. 18th century	19th century	20th century	20th century

Viewing (8.70) and (8.71), it is immediately noticeable that *á bak við* (8.70) and *við hliðina á* (8.71) exhibit similar development despite the timeline for the sequences being different. Supposing that the sequence in III-IV involve grammaticalized elements where *bak við* and *hliðina á* no longer consist of a noun and a preposition ($\mathbf{N_1P_2}$) but rather represent a single preposition (**P**), sometimes appearing in the written language as <*bakvið*> and <*hliðiná*>, we might claim that the diachronic development consists of two major changes. The first change involves a replacement of $\mathbf{N_2}$ in a dative (or genitive) case with a prepositional phrase. The head of the new prepositional phrase ($\mathbf{P_2}$) assigns accusative or dative to N_2.

The second change is the grammaticalization of the sequence $\mathbf{N_1P_2}$ as single **P**; The omission of $\mathbf{P_1}$ can is here regarded as a part of this process and subsumed under phonological reduction.

(8.71) Major changes in the complex prepositions *á bak við* and *við hliðina á*

Major Change I:	$\mathbf{P_1N_1}$ $(+N_2)_{DAT/GEN}$	replaced by	$\mathbf{P_1N_1P_2}$ $(+N_2)_{ACC}$
Major Change II:	$\mathbf{P_1N_1P_2}$ $(+N_2)_{ACC}$	>	**P** $(+N)_{ACC}$

Two things may be noted about the first major change in (8.72). Firstly, the replacement of a dative (or genitive) complement with a prepositional phrase ($\mathbf{P_2}$ $(+N_2)_{ACC}$) occurs in the 18th and 19th centuries, during a time at which Icelandic was in contact with the mainland Scandinavian languages, especially Danish. Since some of the earlier examples of (*á*) *bak við* are from individuals that spent parts of their lives abroad (Jón Ólafsson, for instance, lived in Copenhagen for many years), it is tempting to think that language contact may have played a role in the change. Secondly, the change is reminiscent of ongoing changes in other domains of the language. Thráinsson (2007:94) among others has noted that genitives of possessive tend to be replaced by a prepositional phrase (see also Sigurðsson 2006, 2017b; Árnason 2011:4; Pfaff 2023). This can occur when conveying possession of non-alienable body parts (8.73) or physical objects (8.74), or when noting a close relationship of two items (8.75).

(8.73) a. *Puttinn* *minn*.

the.finger my

b. *Puttinn* *á* *mér*.

the.finger on me

‘My finger.’

(8.74) a. *Tölvan* *mín.*

the.computer my

b. *Tölvan* *hjá* *mér*.

the.computer by me

‘My computer.’

(8.75) a. *Þak* *hússinss*.

roof the.house-GEN

b. *Þakið* *á* *húsinu*.

roof of the.house

‘The roof of the house.’

The use of the b-examples in (8.73)–(8.75) as well as major change I in (8.72) may be taken to reflect changes operating on the typological level of the language, i.e., causing a shift from an inflectional to non-inflectional language through the replacement of an inflected element in an idiosyncratic case (typically genitive but also dative) by a more "transparent" prepositional phrase with non-idiosyncratic case assignment. Although a replacement of a genitive with a prepositional phrase does not eliminate the use of case per se, it shrinks the domain in which the genitive is used and may thus contribute towards the dissolution of the case system. Such a shift may naturally occur in a step-by-step fashion, affecting one subdomain of the language first and then another. Note, for instance that (8.73)–(8.75) exclusively affects possessive pronouns, while (8.75) affects nouns.

Taken together, major change II in (8.72) is a piece in the puzzle of the slow dissolution of the inflection system. In fact, Icelandic is not unique when it comes to changes of this type. A replacement of a noun in the genitive case by a prepositional phrase is also observed in the history of English. Pseudo-partitives have, for instance, gained interest in recent years (Grestenberger 2015).

Without going into further detail about changes that cause a typological shift like this, it is interesting to note that replacing an oblique case with a prepositional phrase is commonly observed within the diachrony of Indo-European languages (Hewson & Bubenik 2006). Modern Icelandic may thus simply be following the route of many other Indo-European languages in this respect, albeit somewhat slowly.

Turning to major change II in (8.72), we observe that it is a shorthand for grammaticalization of **(P$_1$)N$_1$P$_2$** as one preposition, **P**. Grammaticalization is generally viewed as a change of lexical item into a functional item (Kurylowicz 1965/1975:52; Heine, Claudi & Hünnemeyer 1991:2; Hopper & Traugott 2003) and the emergence of new prepositions "through the combination of a (mostly relational) noun with either a preposition or a case suffix is one of the most common grammaticalization process in the world" (Lehmann 1991:501, cited via. Hoffmann 2004:172). The grammaticalization has arguably already occurred in Modern Icelandic as the phrases are fairly fixed and the elements *hliðina* and *bak* are not directly associated with the lexical nouns *hlið* 'side' and *bak* 'back'. Phonological erosion in the form of leaving out the initial P_1 may be argued to be ongoing as not all speakers accept forms lacking P_1 (see also Section 8.3.2 and Section 8.3.3 above).

8.3.6 Expectations

The type of diachronic process argued to be observed in the grammaticalization of *á bak við* and *við hliðina á* can occur multiple times within the same language, affecting different lexical elements at different time periods. Within the history of Icelandic at least two prepositions have already gone through a parallel development at earlier times. These are (*í*) *gegnum* 'through' and (*í*) *kringum* 'around', shown in (8.76).

(8.76)	a.	(*í*) *gegnum* (< *í gegn um*)	'through'
	b.	(*í*) *kringum* (< *í kring um*)	'around, circa'

Although the elements in (8.76) were already mentioned in Section 8.3.1 in relation to the presence and absence of $\mathbf{P_1}$ (cf. Kress 1982:188; Friðjónsson 1988:35–36) it is worth pointing out here that the *gegnum* and *kringum* are likely originally derived from nouns followed by the preposition *um* 'about' (Magnússon 1989 s.v.; Faarlund 2004:107).[78]

A further example from Modern Icelandic where a noun has been (or is in the process of being) grammaticalized, involves the phrase *á borð við* (lit. 'on table with') which is used to convey the meaning 'like, such as' (8.77).

(8.77)	*smitsjúkdóma*	***á***	***borð***	***við***	*HIV*	*og*	*lifrarbólgu*	*C*
	diseases	on	table	with	HIV	and	hepatitis	C

'... diseases such as HIV and hepatitis C'

(https://www.raudikrossinn.is/verkefni/innanlandsverkefni/heilbrigdi-og-velferd/skadaminnkun/)

Although the element *borð* in the sequence in (8.77) is easily identified as coming from the lexical noun *borð* 'table, desk', two items have little or no connections in the modern language.

Another close parallel to grammaticalization of *við hliðina á* and *á bak við* is development and variation in the temporal conjunction (*um*) *leið og* (lit. 'on path also') meaning 'as soon as, at the same time as' (8.78).

[78] Magnússon (1989 s.v.) is careful in his wording, claiming that the prefix *gagn-* 'opposite, against' has an unclear origin but may be derived from a noun or an adjective. As for *kringum*, he notes that it appears to be from the accusative of the noun *kringur* 'circle' followed by a preposition *um*.

(8.78)	*Lömb*	*sett*	*út*	***um***	***leið***	***og***	*hægt*	*er.*
	lambs	put	out	about	path	and	possible	is

'Lambs are sent outdoors as soon as possible.'

(https://www.bbl.is/frettir/lomb-sett-ut-um-leid-og-haegt-er)

The element *leið* in (8.79) is originally from the noun *leið* 'path'. The dissociation of **N_1** *leið* in in a structure meaning 'as soon as' likely occurred already in Old Icelandic. Example (8.79a) is ambiguous between showing the temporal conjunction 'soon as' and the lexical 'via the path that'.

(8.79)	a.	*Þat*	*var*	*þeira*	*ráð*	*bræðra,*	*at*	*þeir*	*kómu*
		it	was	their	plan	brothers	that	they	came
		báðir	*til*	*Staðar,*		***um***	***leið***	***er***	*Kálfr*
		both	to	Staður		through	path	which	Kálfr
		reið	*norðr*	*yfir*	*heiði*				
		rode	north	over	heath				

Possibility 1: 'The brothers' plan was such that they arrived to Staður via the path which Kálfur rode north over the heath.'

Possibility 2: 'The brothers' plan was such that they arrived to Staður at the same time as Kálfur rode north over the heath.'(ONP, *Stu^{II}R11127^{x} 3098*) b.

Þegar	*vm*	*leið*	*hæggr*	*hann*	*til*	*drekans*
immediately	about	path	hews	he	to	the.dragon

'Immediately at the same time, he swings [the sword] towards the dragon.'

(ONP, Þiðr 198 l.11, Holm perg 4 fol, c. 1275-1300)

The examples in (8.76)–(8.79) consist of dissociation of lexical nouns with an element in prepositions and adverbial.

The directionality observed in changes in the complex prepositions *á bak við* and *við hliðina á* is that from a complex form to a more simplified one. If grammaticalization is taken to consist of several subprocesses, the changes can be decomposed into disassociation of $\mathbf{N_1}$ with the nouns *hlið* 'side' and *bak* 'back', conventionalization of the sequence $\mathbf{P_1N_1P_2}$, and a reanalysis of $\mathbf{(P_1)N_1P_2}$ as a single prepositional phrase. Phonological reduction may be reflected in both the "merger" of $\mathbf{N_1P_2}$ into a single element (occasionally written <*hliðiná*> *a*nd <*bakvið*>) and the cliticization or omission of $\mathbf{P_1}$. Unfortunately, standardized spelling, which demands the sequence $\mathbf{N_1P_2}$ to be written in two words (<*hliðina á*> and <*bak við*>), makes it difficult to measure when speakers regard $\mathbf{N_1P_2}$ as a single element and when not. Of course, the greater the dissociation between the $\mathbf{N_1}$ and the relevant lexical item (*hlið* 'side' or *bak* 'back'), the likelier speakers may be to write it as a single word. Pressure from prescriptivism may still work in favor of keeping the traditional written form alive for longer. Arguably, however, the univerbation has already taken place as indicated by the inflexibility of the structures in the modern language (see Section 8.3.2 and Section 8.3.3 above). The omission of $\mathbf{P_1}$ from the $\mathbf{P_1N_1P_2}$ sequence may be viewed as phonological reduction and it is the only aspect of the change that is still ongoing. The presence vs absence of $\mathbf{P_1}$ is easily measured in corpus material and this is used for measuring the ongoing change in following sections, 8.4 and 8.5.[79]

Although the grammaticalization path seems straightforward (*við hliðina á* > *hliðiná*; *á bak við* > *bakvið*), the change is slow and is seemingly countered by other pressures within the language system. Changes in the use of the prepositions *á eftir* and *á undan* exhibit opposite directionality. According to Friðjónsson (2005, 2009) these originally lacked the initial $\mathbf{P_1}$ which was first introduced the 14th century (*á eftir*, Friðjónsson 2005:23–24) and in the 17th century (*á undan,* Friðjónsson 2005:19), with *á eftir* not being properly established until the 19th century. Importantly, according to Friðjónsson, the introduction of a $\mathbf{P_1}$ element was slow and somewhat unsystematic, reflecting instability in the whole prepositional system. In Modern Icelandic, it appears that forms containing $\mathbf{P_1}$ have been generalized, with forms without $\mathbf{P_1}$ being used in special circumstances. In this context, Friðjónsson (2009:68) notes that *á undan* is associated with ordered location in time and space, while *undan* (without $\mathbf{P_1}$) is used when denoting direction of movement or a reason for something (Friðjónsson 2009:68). Given the complex history of *á undan* and *á eftir*, it may be concluded that changes in a few lexical items do not occur in isolation, but must be viewed in the light of the whole grammatical system. While a process such as grammaticalization, involving semantic change, reanalysis and phonological reduction, might cause complex structures to be simplified, counterprocesses or pressures (either from the grammatical system itself or from the language society) may cause complexity to be retained or reintroduced. In the case of the

[79] It is possible to quantify the occurrences of various written variants of $\mathbf{N_1P_2}$, e.g. <*bakvið*> and <*hliðiná*>. However, as I have argued above (Sections 8.3.2–8.3.3) this is a less reliable measurement than keeping track of the presence or absence of $\mathbf{P_1}$.

complex prepositions *við hliðna á* and *á bak við*, the simplified variants, lacking **P_1**, are expected to win out.

8.4 Data annotation and description

8.4.1 Data source

Quantitative data for documenting variation in Modern Icelandic and generating predictions for the use of the two complex prepositions under investigation, i.e., *á bak við* and *við hliðina á*, come from two different sources: Twitter (Twitter API v2) via the R-package *academictwitteR* (Barrie & Ho 2021) and The Icelandic Gigaword corpus (IGC). These sources contain material from different registers and cover slightly different periods of time.

Data from Twitter is taken to represent semi-informal, non-proofread language. Tweets can be posted by anyone with a Twitter account, and they do not go through a formal review or approval before being posted online.[80] However, due to language purism and prescriptivism being widespread in Iceland, individuals who tweet may try to conform somewhat to preconceived ideas of the standardized language and general rules of punctuation. It is mostly for these reasons that Twitter data is not taken to reflect fully informal language.

When obtaining data from Twitter, the target language was set to Icelandic and retweets were excluded from the query. A search was made for various forms of the N_1P_2

[80] Although anyone and everyone can sign up for a Twitter account, not everyone does. This results in data from Twitter being subject to self-selection bias such that only tweets from individuals who have chosen to be active users of Twitter figure into the dataset. The assumption here is nevertheless that this type of data remains relatively consistent over time and that it is informative of language use of some part of the population.

part of each of the two prepositions, see Table 8.1. Although Twitter contains data from 2006 onwards, Icelandic tweets are relatively few until around 2009. Due to this, data prior to 2009 is not well suited for a time-series analysis. As a result, the Twitter data used here only covers the period from 2009 to 2022, both years included. An example of a query made on the 27th of January 2023 is provided in (8.80). As can be seen, the target string in this case was "bak við" within Icelandic (lang:is). The search covered the period from January 1st 2022 to January 1st 2023 with retweets being excluded.[81]

(8.80) "bak við" -is:retweet lang:is

2022-01-01T00:00:01Z

2023-01-01T00:00:00Z

The second source of data, namely The Icelandic Gigaword Corpus, contains about 2,429 million running words (IGC-2022) from several different text types, ranging from online forum threads to printed newspaper articles. Most of the material dates from after 2000. Due to material before 2000 being irregularly spaced in time only data from 2000 to 2022 (rmh=2022) was used for the present study. Search in IGC targeted (unannotated) $N_1(P_2)$ word forms of both complex prepositions, see Table 8.1.

As for individual sources within IGC, care was taken to only include material with a timestamp containing information on the year of creation.[82] This resulted in some sources

[81] Data for the period January 1st, 2009, to December 31st, 2021, had been abstracted previously.

[82] Originally, the intention was to only obtain data that included information on year and month of creation. However, some of the timestamps turned out to have a placeholder "00" for both month and day such that only the year (and not month) of creation could be established.

being excluded, for instance published books and academic journals which may not have been created the year they were published.

Although the latest version of the Icelandic Gigaword Corpus (IGC-2022) contains data from Twitter, this source was not used since Twitter data had already been obtained separately. A list of the type of material obtained from IGC is presented in (8.81).

(8.81) **IGC-2022 SOURCES**:

News media: television and radio news, printed newspapers, online newspapers

Social media: blog posts, online threads

Special material: various local news media, specialized news material (agriculture and fish news), sports news, tabloids, cultural material

SOURCE	**FORMS**	
	Next to'	**Behind'**
Twitter	*hliðina á*	*bak við*
(2009–2022)	*hliðin'á*	*bakvið*
	hliðn'á	
	hliðiná	
IGC	*hliðina*	*bak*
(2000–2021)	*hliðiná*	*bak við*
	hliðna	
	hliðná	
	hliðinaá	

Table 8.1. A list of forms searched for on Twitter and IGC.

8.4.2 Data annotation

Data from Twitter and the Icelandic Gigaword corpus was carefully annotated using Microsoft Excel (version 2308, Microsoft Corporation 2022). The annotation was done semi-automatically and by hand. The semi-automatic method involved searching for particular strings using the IF function and annotating relevant examples in batches. In many cases, the semi-automatic annotation was manually verified. All non-prepositional uses of the forms *hliðin*(*a*) 'side' and *bak* 'back' were marked specially and excluded from further analysis.[83]

The annotation provided information on whether P_1 (*á* or *við*) was present or not in the relevant example, and whether N_1P_2 was written in a compact form (*bakvið* and *hliðiná*) or with a space between N_1 and P_2 (*bak við* and *hliðina á*). The second type of annotation does not figure into the analysis. In a handful of instances, examples from IGC did not contain enough context to determine whether the example was an instance of a preposition or not, or whether it included P_1 or not. When this occurred, the example was annotated with the string UNID (for ***unid****entifiable*) and excluded from further analysis. Other examples (very few) that were deemed unclear or strange were marked ATH and also excluded analysis. Labels of sub-corpora in IGC were normalized and, for small local news media sources no distinction was made between printed publication and internet sources (these were collapsed into a single label). For larger news media, the distinction between printed material and online material was maintained as the former is known to have gone through proofreading while the latter typically has not.

[83] These include examples such as *í hliðina á* 'into the side of' and *hina hliðina á* 'the other side of'.

During annotation of material from Twitter and IGC, it turned out that some examples appeared multiple times. It is unclear why this was the case. The problem was more prominent with IGC than Twitter. For the latter, the same account may have tweeted identical content at slightly different times.[84] For IGC, the database appears to sometimes contain multiple instances of the same text. This can potentially be attributed to blog posts and online news articles sometimes being updated several times over the course of a few hours or few days. The edits may add material or correct previously written text. If all versions are included in IGC, it is only natural that duplicated examples will appear. Duplicate examples were removed automatically in Microsoft Excel by using the Remove Duplicates function. Decision regarding whether something was considered a duplicate or not was based on information on where the data was from (which source within the IGC) as well as whether the examples were identical or not.[85]

8.4.3 General overview of data

After the data had been cleaned-up and annotated, the number of examples of each of the complex prepositions, *bak við* and *við hliðina á*, in IGC and Twitter were as in Table 8.2. The label NOPE refers to occurrences without $\mathbf{P_1}$.

[84] Duplicated examples from Twitter were not retweets. Reetweets were already excluded in the initial search and in the rare case that retweets (labeled RT) found their way into the results these were removed. Duplicated examples that contained two different forms of either of the prepositions were not removed but rather annotated once for each example.

[85] Data from IGC is formated in such a way that the search target (cf. Table 8.1) appears in a separate column from the left and right context of the target. Examples were considered identical provided the left context, the target and right context were deemed identical by Excel.

Source	*á bak við* 'behind'			*við hliðina á* 'next to'		
	NOPE	**P₁ *á***	**Total**	**NOPE**	**P₁ *við***	**Total**
Twitter (2009–2022)	5,443 (c. 34%)	10,364 (c. 66%)	15,807 (100%)	958 (c. 12%)	6,792 (c. 88%)	7,750 (100%)
IGC (2000–2021)	27,669 (c. 23%)	92,205 (c. 77%)	119,874 (100%)	2,621 (c. 7%)	35,494 (c. 93%)	38,115 (100%)

Table 8.2. Distribution of examples of *á bak við* 'behind' and *við hliðina á* 'next to' with and without $\mathbf{P_1}$ in Twitter and IGC.

As shown in Table 8.2, the total number of examples of *á bak við* from Twitter over a 14-year period (2009–2022, both years included) was 15,807. Of these, 5,443 (c. 34%) lacked the initial $\mathbf{P_1}$ *á*. For a comparison, the total number of examples from IGC over a 22-year period (2000-2021, both years included) was 119,874 with 27,669 (c. 23%) occurring without $\mathbf{P_1}$ *á*. The proportion of examples lacking $\mathbf{P_1}$ appears to be slightly lower in IGC than in Twitter. This might be due to IGC stretching over a longer period or due to data from there being from different registers. I return to this below.

The total number of examples of *við hliðina á* from Twitter was 7,749, with 957 (c. 12%) lacking the initial $\mathbf{P_1}$ *við*. For a comparison, the total number of examples from IGC was 38,115. and 2,621 of these (c. 7%) did not have $\mathbf{P_1}$ *við*. Like in the case of *á bak við*, the proportion of examples lacking $\mathbf{P_1}$ appears slightly lower in IGC than in data from Twitter.

Comparing the two complex prepositions, it is immediately noticeable that *á bak við* has a higher proportion of examples without initial $\mathbf{P_1}$ than *við hliðina á* in both Twitter and IGC. As noted in Sections 8.3.4–8.3.5, changes in the complex prepositions *á bak við* started earlier than changes in *við hliðina á*. The difference in proportion of examples lacking **P1** between the two might therefore be attributed to *á bak við* being further along on the grammaticalization route than *við hliðina á*.

While Twitter can be regarded as a uniform data source (all examples come from tweets), this is not the case for IGC. The sub-sources within IGC vary in nature and are not necessarily homogeneous. For instance, material from a printed newspaper like *Morgunblaðið* may have different characteristics from texts that appear in online forums or blog posts. Figures 8.2 and 8.3 show the distribution of *á bak við* and *við hliðina á*, with and without the initial **P1**, in individual sub-sources within IGC. Note how the lack of **P1** in *við hliðina á* is mostly found within certain sub-sources. This does not appear to be the case for **P1** *á* in *á bak við*.

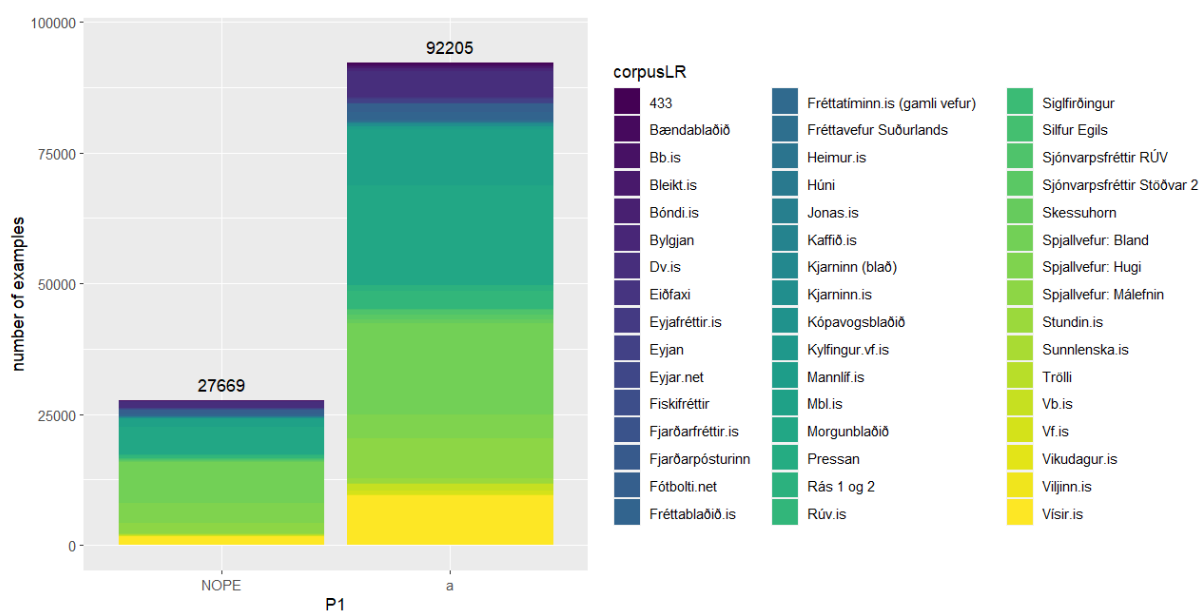


Figure 8.2. Distribution of *á bak við* with and without **P1** *á* in various sources within IGC.

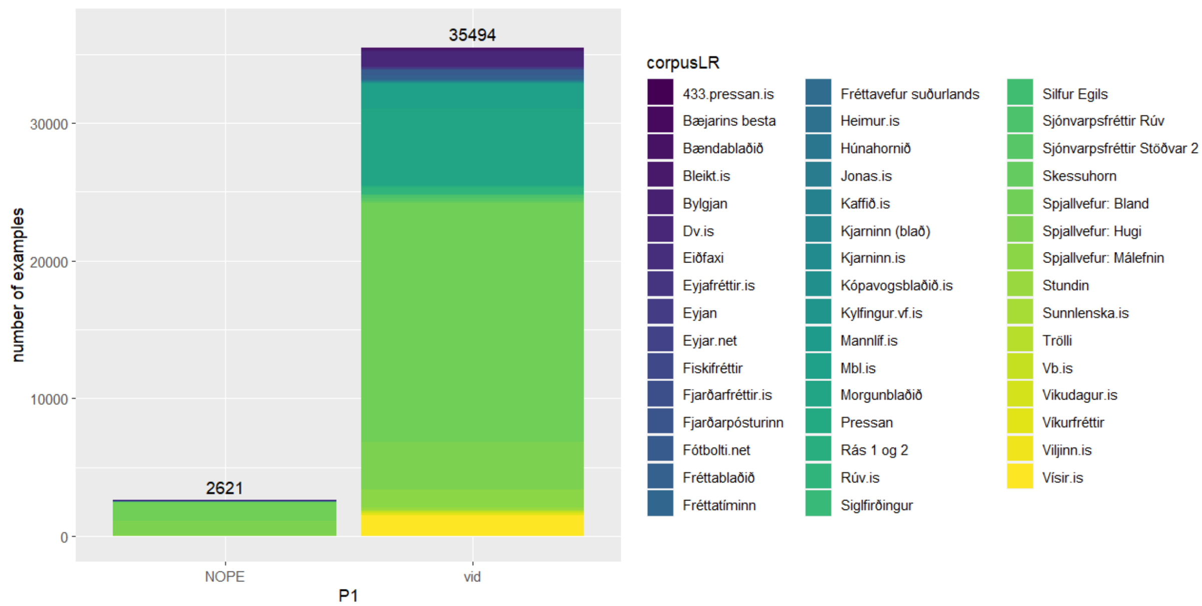


Figure 8.3. Distribution of *við hliðina á* with and without **P1** *við* in various sources within IGC.

Visualization of the distribution of variants with and without **P1** in individual sources within IGC can also be done as in Figures 8.4 and 8.5. In addition to indicating how many examples have **P1** and how many do not, these figures give an idea of the amount of material contributed by different sources. Setting aside the fact that *á bak við* has overall a lot more attestation than *við hliðina á*, we may note that seven sources consistently make up most of the data in each case. These are: Dv.is, Mbl.is, Morgunblaðið, Spjallvefur: Bland, Spjallvefur: Hugi, Spjallvefur: Málefni, and Vísir.is.

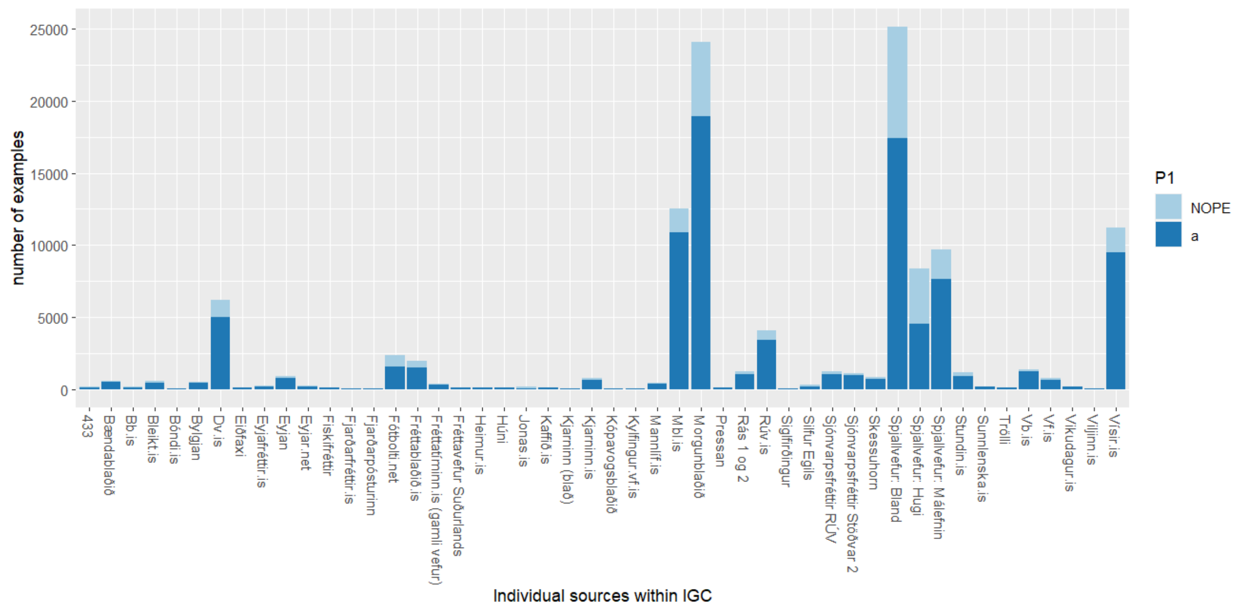


Figure 8.4. **P₁** in the complex preposition *á bak við* ‘behind’ in various sources in IGC. The label NOPE stands for lack of **P₁**.

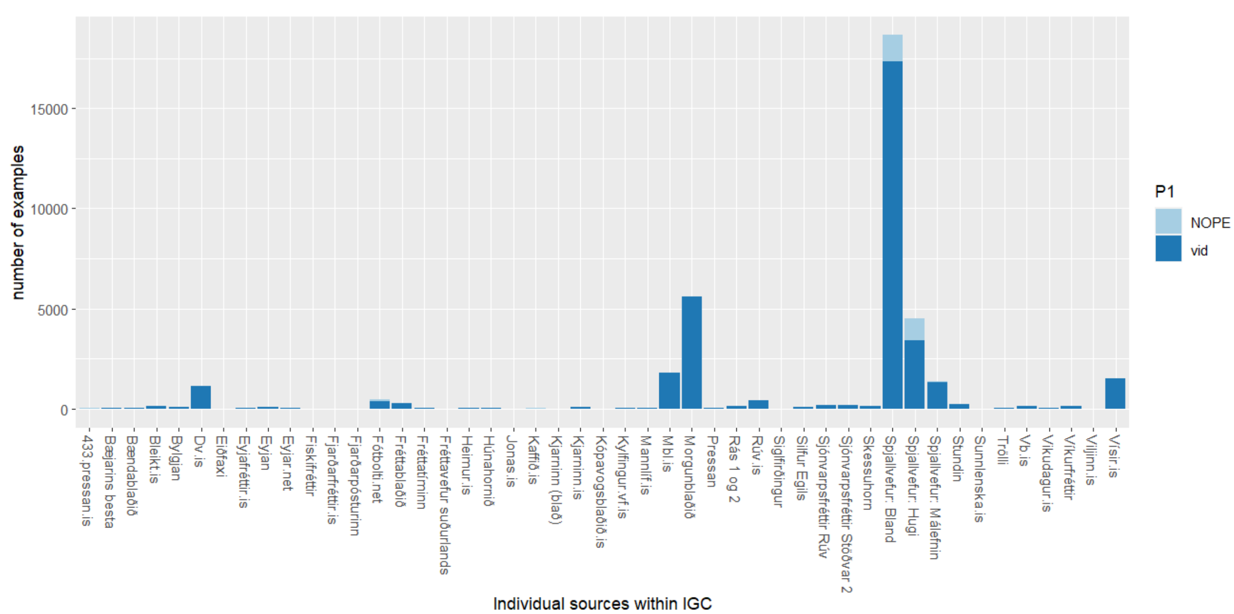


Figure 8.5. **P₁** in the complex preposition *við hliðina á* ‘next to’ in various sources in IGC. The label NOPE stands for lack of **P₁**.

The seven sources that make up the bulk of the attestation of *á bak við* and *við hliðina á* are not identical in nature. In fact, they can be taken to represent three different register types. Morgunblaðið, a daily newspaper that began its publications in 1913, is the most prestigious of these sources. The language appearing there tends to be highly standardized, reflecting the fact that it is quite formal in nature, is targeted at the general population of Iceland and goes through proofreading before publication.

The online counterpart of Morgunblaðið, Mbl.is, is also targeted at the general population. The use of "good" language is considered appropriate. Unlike Morgunblaðið, news articles published on Mbl.is do not go through copy-editing. The same applies to material on Vísir.is, an online newspaper, and Dv.is, an online tabloid. Together, Mbl.is, Vísir.is and Dv.is may be taken to represent semi-formal language, i.e., they are not as formal as Morgunblaðið, but also not informal.

The last group consists of three online discussion forums (Icelandic nom. sg. *spjallvefur*). These are labeled Spjallvefur Bland, Spjallvefur Hugi and Spjallvefur Málefnin. Messages appearing in threads on the forums are not proofread and are only targeted at individuals participating in individual discussion threads. The register of the language can be regarded as informal.

Viewing Figure 8.5. in light of formality-type (formal, semi-formal and informal), suggests that the absence of **P_1** in *við hliðina á* might be linked to formality. Material from Spjallvefur Bland and Spjallvefur Hugi are the two sources in IGC with the most attestations of forms without the initial **P_1**. Note that although this also appears to be the case with *á bak við* in Figure 8.4, it is harder to ascertain by just looking at the figure.

In an attempt to capture the observations that the lack of **P_1** in the complex prepositions may be tied to language registers, individual sources in IGC were labeled either as representing formal, semi-formal or informal registers. Formality-labeling was based on i) whether the material was likely to be proofread or not, ii) how the material was mediated, i.e., via printed publication, radio, or online, and iii) which group of people the material was targeted at, i.e., the general population or just a few individuals. The labeling of individual sources within the IGC is shown in (8.82).[86]

(8.82) The annotation of three types of registers

Form: Formal language, printed or broadcasted material. The use of "standard" and "good" language is considered very important. The material is likely proofread.

Sources: *Bændablaðið, Eiðfaxi, Fjarðarpósturinn, Fréttablaðið, Fréttatíminn, Kjarninn (blað), Kópavogsblaðið, Morgunblaðið, Rúv.is, Siglfirðingur, Sjónvarpsfréttir RÚV, Sjónvarpsfréttir Stöðvar 2, Stundin*

SForm: Semi-formal language, mostly online material. The material is typically not proofread although the use of "good" language is considered appropriate.

Sources: *Bæjarins besta, Bb.is, Bleikt.is, Bóndi.is, Bylgjan, Dv.is, Eiðfaxi, Eyjafréttir.is, Eyjan, Eyjar.net, Fiskifréttir, Fjarðarfréttir.is, Fjarðarpósturinn, Fótbolti.net, Fréttablaðið.is, Fréttatíminn,*

[86] In (8.82) *Fréttatíminn* appears to show up twice, once in formal material and once in semi-formal material. The one listed under formal material is a printed newspaper (published and distributed between 2010 and 2017) while the other (under semi-formal material) is an online new media page.

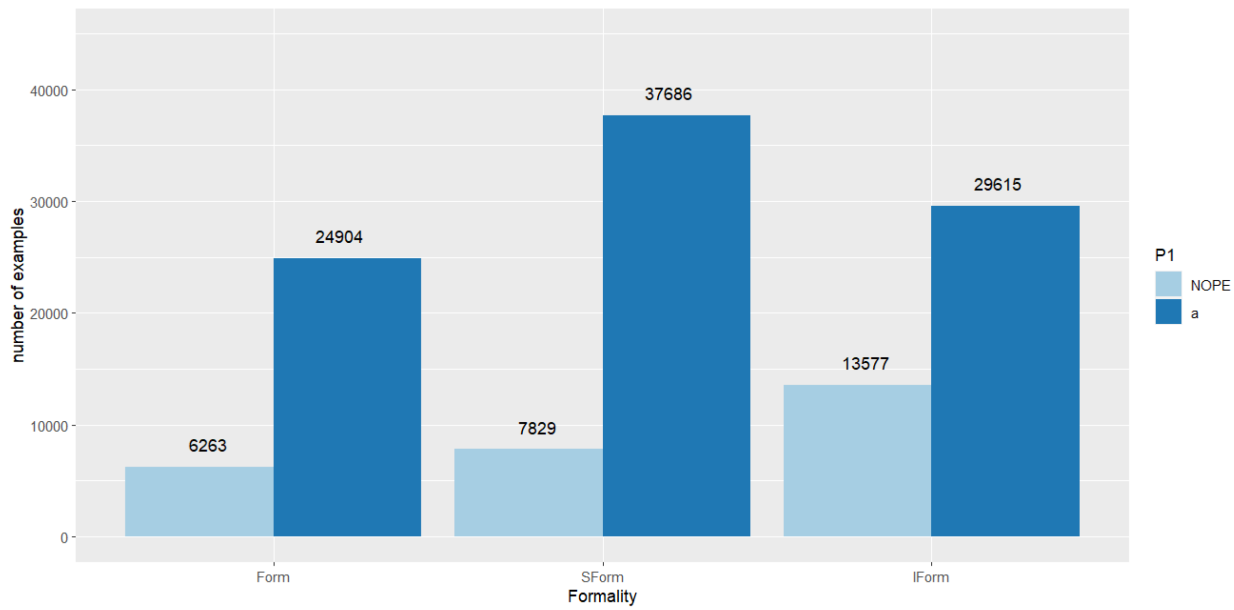
Fréttatíminn.is (gamli vefur), Fréttavefur Suðurlands, Heimur.is, Húnahornið, Húni, Jonas.is, Kaffið.is, Kjarninn.is, Kópavogsblaðið.is, Kylfingur.vf.is, Mannlíf.is, Mbl.is, Pressan, Rás 1 og 2, Siglfirðingur, Silfur Egils, Skessuhorn, Stundin.is, Sunnlenska.is, Trölli, Vb.is, Vf.is, Vikudagur.is, Víkurfréttir, Viljinn.is, Vísir.is, 433.pressan.is

IForm: Informal language, online material. The material is not proofread and conforming to a "standard" or "proper" language is typically not considered crucial as the material is not necessarily intended to be read by everyone.

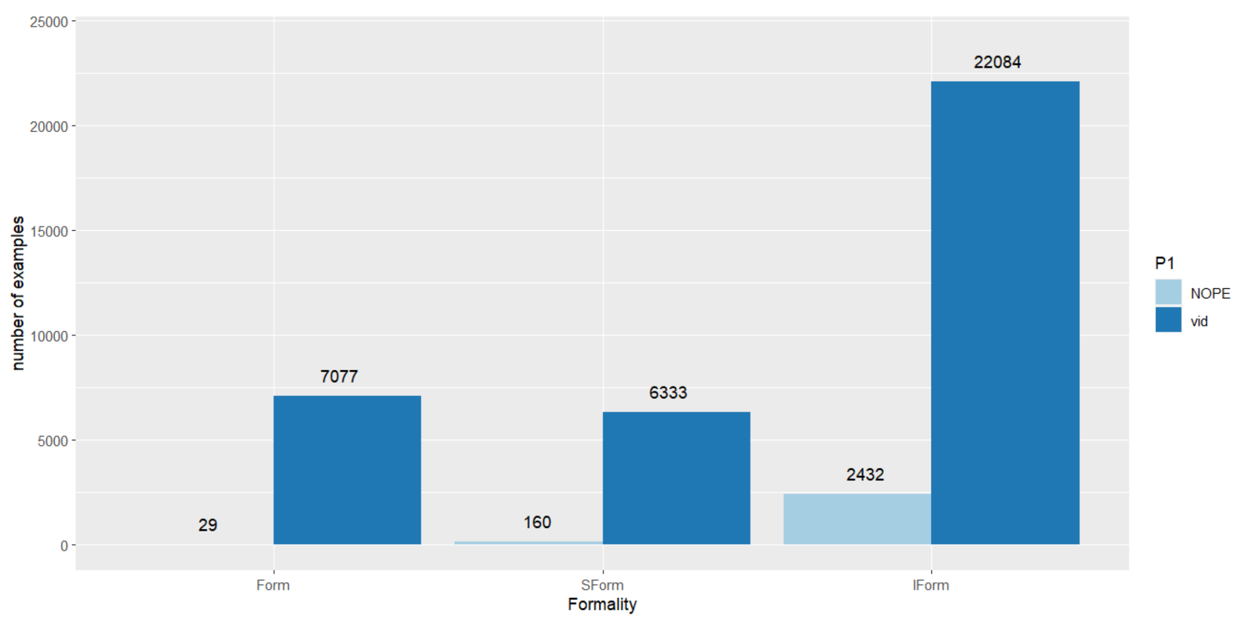
Sources: *Spjallvefur Bland, Spjallvefur Hugi.is, Spjallvefur Málefnin*

The division into formal, semi-formal informal and informal language is not only rooted in how the examples of the two complex prepositions appear in IGC, but also in how language change might be expected to surface in different sources. As already noted, language purism is prominent in Iceland and prescriptivist attitudes towards Icelandic are very common. Innovative and non-prestigious forms may be expected to be either avoided or replaced, especially in formal registers or in material that is meant to be directed at a broader audience. This may leave venues where speakers interact on a more casual level, perhaps in small groups, as the likeliest places for change and non-prestigious variants to surface.

Figures 8.6 and 8.7 show the distribution of examples with and without $\mathbf{P_1}$ in IGC in material labeled formal, semi-formal and informal. Table 8.3 summarizes the number of examples and the proportion of each type of form.



Figures 8.6. Distribution of forms of the complex prepositions *á bak við* with and without **P1** in Formal, Semi-Formal and Informal material in IGC.



Figures 8.7. Distribution of forms of the complex prepositions *við hliðina á* with and without **P1** in Formal, Semi-Formal and Informal material in IGC.

Formality	*á bak við* 'behind'			*við hliðina á* 'next to'		
	NOPE	**P_1 *á***	**Total**	**NOPE**	**P_1 *við***	**Total**
Formal	6,263 (c. 20%)	24,904 (c. 80%)	31,167 (100%)	29 (c. 0.4%)	7,077 (c. 99.6%)	7,106 (100%)
Semi-formal	7,829 (c. 17%)	37,686 (c. 83%)	45,515 (100%)	160 (c. 2%)	6,333 (c. 98%)	6,493 (100%)
Informal	13,577 (c. 31%)	29,615 (c. 69%)	43,192 (100%)	2,432 (c. 10%)	22,084 (c. 90%)	24,516 (100%)

Table 8.3. Distribution of examples of *á bak við* 'behind' and *við hliðina á* 'next to' with and without **P_1** in formal, semi-formal and informal sources in IGC. Informal sources have the highest proportion of examples without initial **P_1**.

As can be observed from Table 8.3, the occurrence of forms without **P_1** varies depending on formality. Most noticeable is the difference between formal and informal register in *við hliðina á*. In formal sources, the proportion of forms without P_1 is right around 0.4% (29 examples). In informal sources, it is around 10% (2432 examples).

In the case of *á bak við*, variants lacking P_1 appear in all types of registers. Although there does not seem to be a considerable difference between register types, forms without P1 are still the most common in informal material, where they make up around 31% (13,577 examples) of attested examples.

Of the three different types of formality, the informal category in IGC resembles data found on Twitter the most, i.e., if we just observe the proportion of examples without **P_1**. Examples of *á bak við* lacking the initial **P_1** constitute around 34% of examples on Twitter and about 31% of examples in informal sources in IGC. Examples of *við hliðina á* without initial **P_1** make up 12% of the Twitter data and 10% of the informal sources in IGC.

8.4.4 Projecting data into regular time series

In the discussion above, data from IGC and Twitter was treated holistically as coming from a single uniform period (Twitter 2009–2022, IGC 2000–2009). These can be converted into time series. Unfortunately, not all the examples from IGC had a proper timestamp containing information on the month of publication in addition to the year. Due to this, IGC data is only converted into yearly time series. The Twitter data, on the other hand, has more detailed information on time of appearance and can be converted into quarterly time series. Figure 8.8 shows how the complete data was projected into yearly and quarterly series.

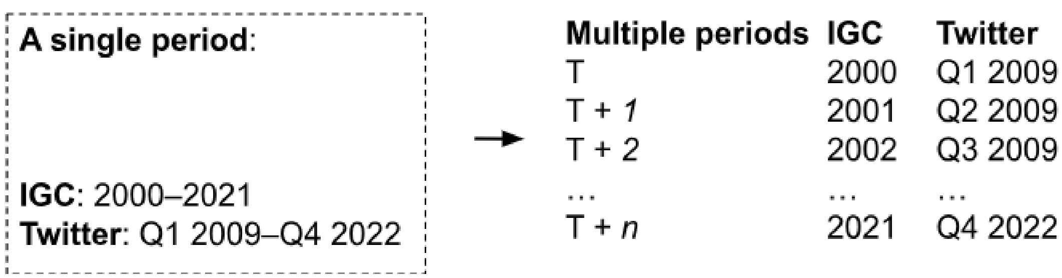


Figure 8.8. Taking data from a period of 14–22 years and creating yearly (IGC) and quarterly (Twitter) time series.

Figure 8.9 to Figure 8.16 summarize distribution of variants with and without **P_1** over time. In the case of data from IGC, the distribution is shown for each year. For data from Twitter data, the distribution is shown for each quarter of the year.

It is worth making a special note about data from Twitter, namely that tweets have a certain character limit which might affect whether individuals choose a shortened form (without **P_1**) of the complex prepositions or not. Originally, the character limit was 140, but is now 280 (Burgess & Baym 2020:34), with a paid premium version of the platform (X Premium) allowing for up to 25,000 characters (https://help.twitter.com/en/using-x/x-premium). The 280-character limit was introduced in fall 2017 (Rosen 2017). If the extended character limit had a direct effect on which form of the complex prepositions were chosen, one might expect the Twitter time series to show an altered "behavior" after fall 2017. However, this does not appear to be the case. Any trend seemingly observable in Figures 8.21–8.24 extends to prior to 2017. This fact does not exclude the possibility of tweet-length generally affecting which form of the complex preposition is chosen. The potential correlation between tweet length and choice of form for each of the complex prepositions was not tested in the present study.

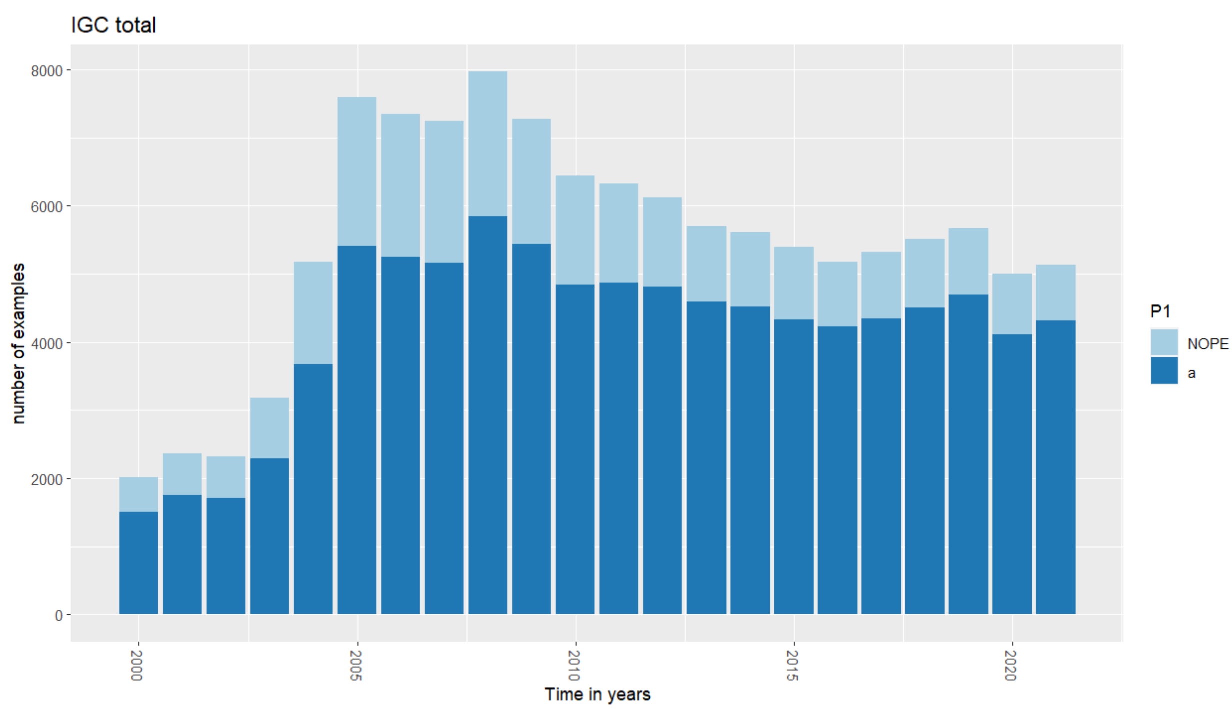


Figure 8.9. Distribution of presence and absence of **P1** in the complex preposition *á bak við* over time. Most of the examples each year have the **P1** *á*.

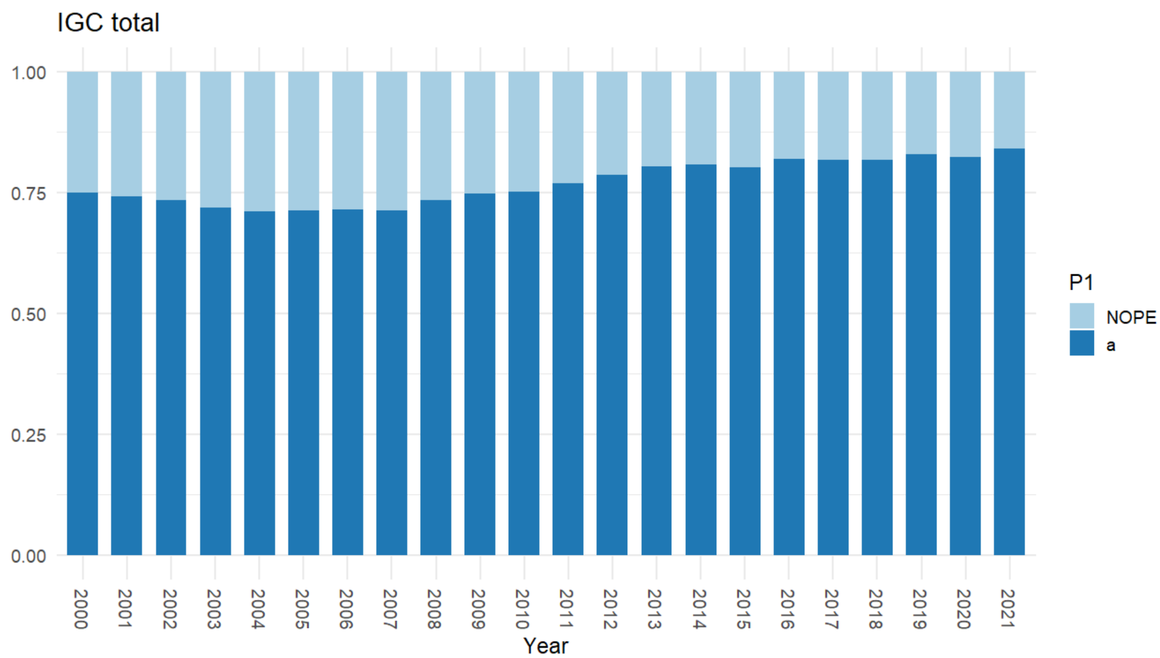


Figure 8.10. Proportion of examples with and without **P₁** in the complex preposition *á bak við* over time. Most of the examples each year have the **P₁** *á*.

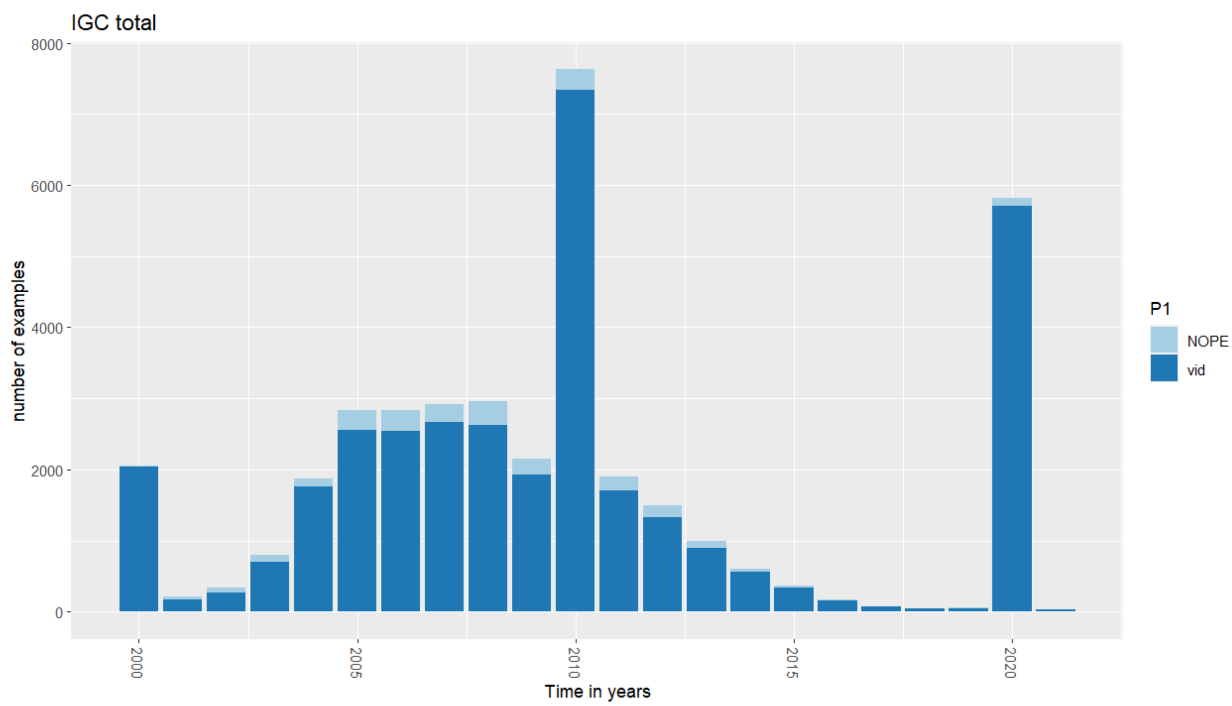


Figure 8.11. Distribution of presence and absence of **P₁** in the complex preposition *við hliðina á* over time. Most of the examples each year have the **P₁** *við*.

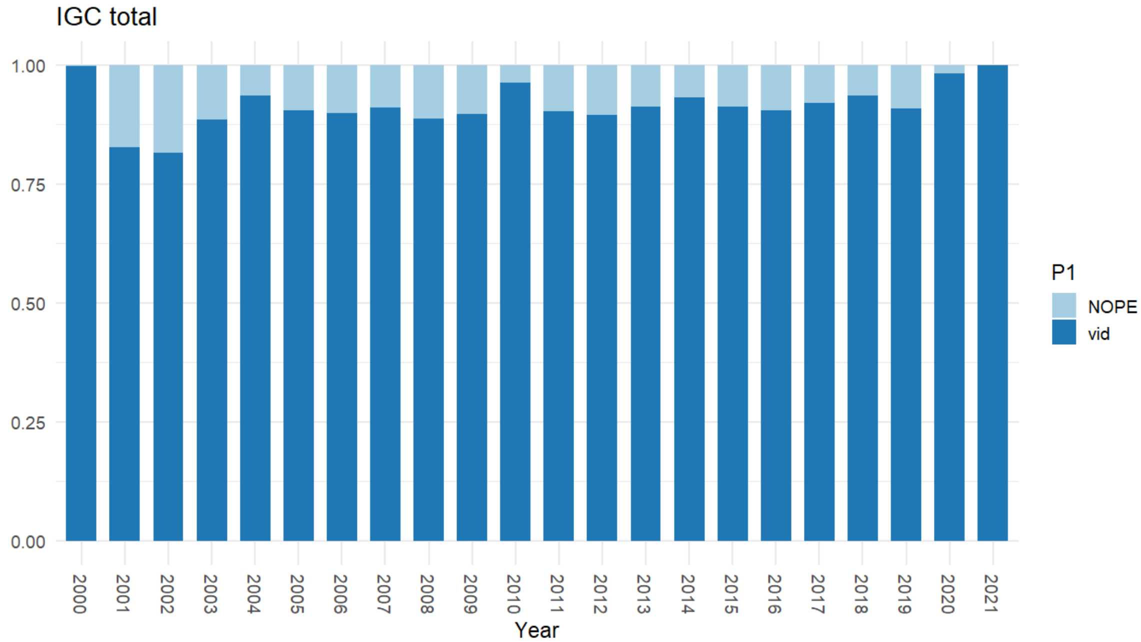


Figure 8.12. Proportion of examples with and without **P1** in the complex preposition *við hliðina á* over time. Most of the examples each year have the **P1** *við*.

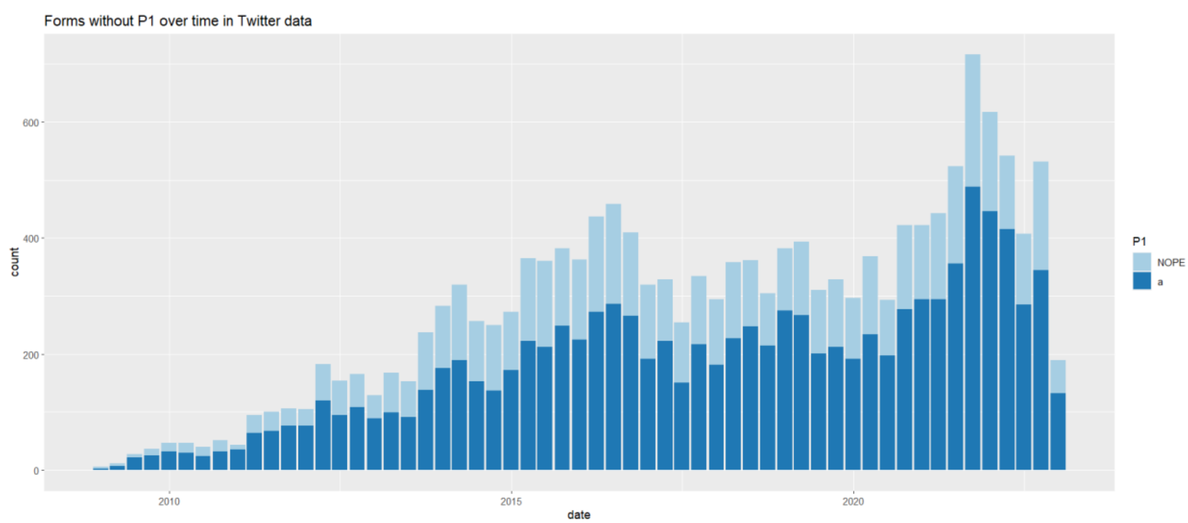


Figure 8.13. A quarterly overview of the data from Twitter, showing the number of examples of *á bak við* 'behind' with and without **P1** *á*.

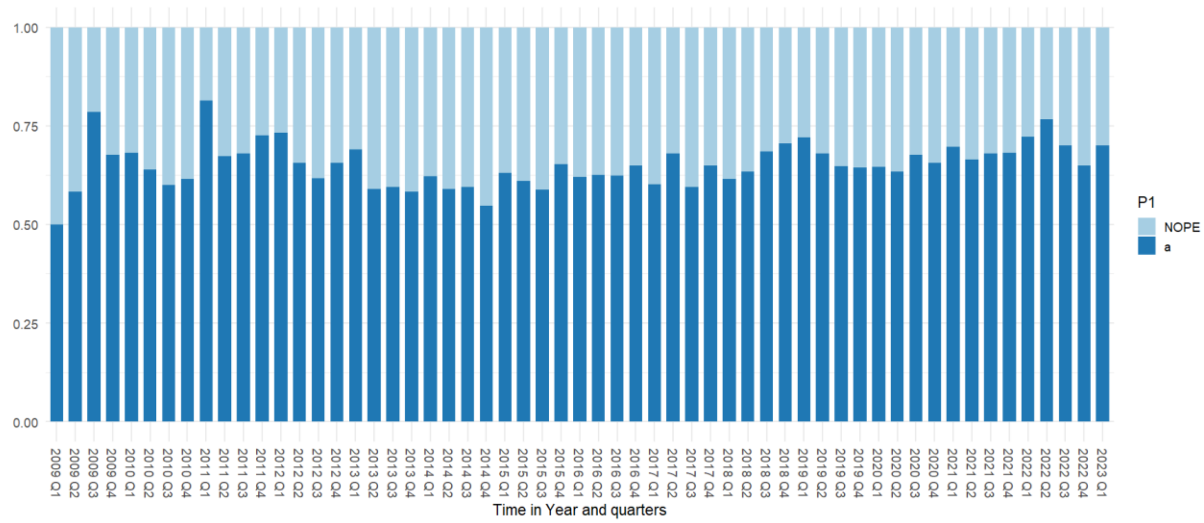


Figure 8.14. A quarterly overview of the data from Twitter, showing the proportion of examples of *á bak við* 'behind' with and without **P1** *á*.

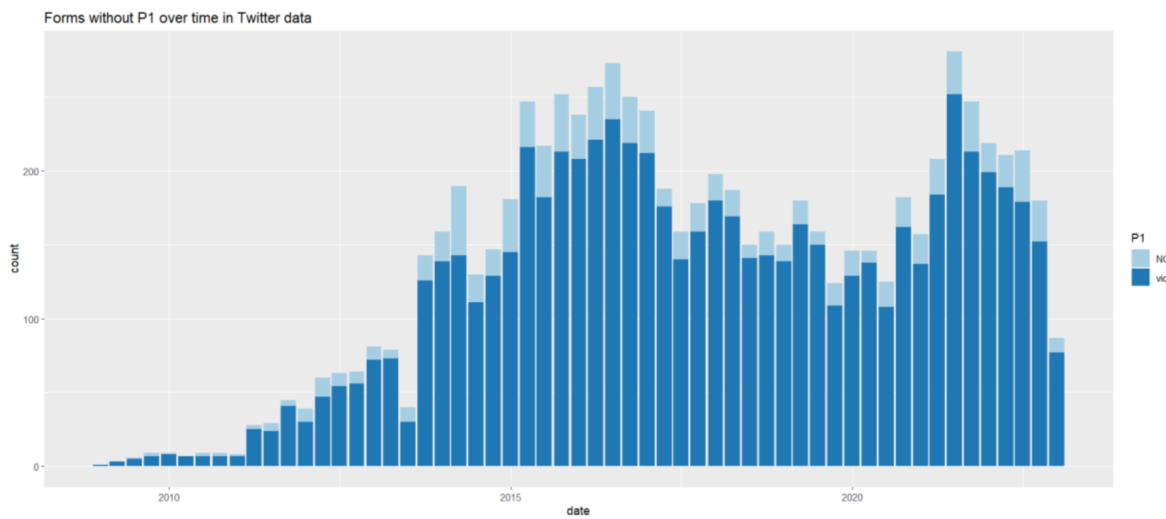


Figure 8.15. A quarterly overview of the data from Twitter, showing the number of examples of *við hliðina á* 'next to' with and without **P1** *við*.

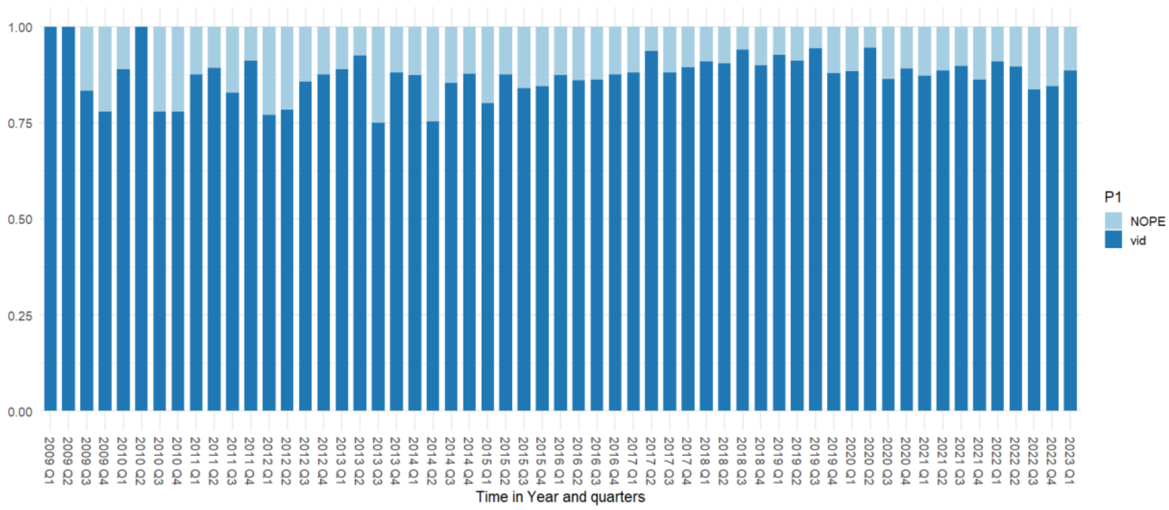


Figure 8.16. A quarterly overview of the data from Twitter, showing the proportion of examples of *við hliðina á* ‘next to’ with and without $\mathbf{P_1}$ *við*.

By glancing at the figures in Figure 8.9– Figure 8.16, a few things can be observed. First, we may note that there are generally more examples of the complex preposition *á bak við* than *við hliðina á* in both IGC and on Twitter. Second, there is a higher proportion of forms lacking $\mathbf{P_1}$ in *á bak við* than in *við hliðina á*. Third, by looking at figures showing the proportion of forms with and without $\mathbf{P_1}$ it may appear like forms without $\mathbf{P_1}$ become less common over time. This is unexpected in light of the discussion in Sections 8.2.2 and 8.3.6. Properties of the time series are discussed further in Section 8.5.

8.5 Time series analysis and forecasting

8.5.1 The time series

Having described the data obtained from both IGC and Twitter in Section 8.4, it is appropriate to take a closer look at the time series. Four time series are taken into consideration, two for each complex preposition. For both *á bak við* and *við hliðina á,* data from semi-formal and informal sources in IGC is projected into a yearly time series stretching from 1999 to 2021, both years included. This gives a total of 22 observations sequenced in time for each preposition. Data from Twitter is projected into a quarterly time series, starting in Q1 2009 and ending in Q4 2022, giving a total of 56 observations sequenced in time. The four time series are shown in Table 8.6. The time series show the proportion of examples without $\mathbf{P_1}$ for each time period. Measurements for each time period are based on aggregated numbers over that period. No external factors are taken into consideration. The focus is strictly on how the proportion of forms without $\mathbf{P_1}$ changes (or stays stable) over time within a given data source.

Overview of data for the two complex prepositions						
Source	**Period covered**	**Observations**	**Series**	**N examples for complex preposition**		
				without P1	*with P1*	*total*
IGC	1999–2021	22 yearly	við hliðina á	2,592	28,417	31,009
Twitter	Q1 2009–Q4 2022	56 quarterly	við hliðina á	958	6,792	7,750
IGC	1999–2021	22 yearly	á bak við	21,406	67,301	88,707
Twitter	Q1 2009–Q4 2022	56 quarterly	á bak við	5,443	10,364	15,807

Table 8.4. A summary of the time series, the number of observations and number of examples behind each series. As an example, the IGC time series on *við hliðina á* contains a total of 31,009 examples which make up 22 yearly observations.

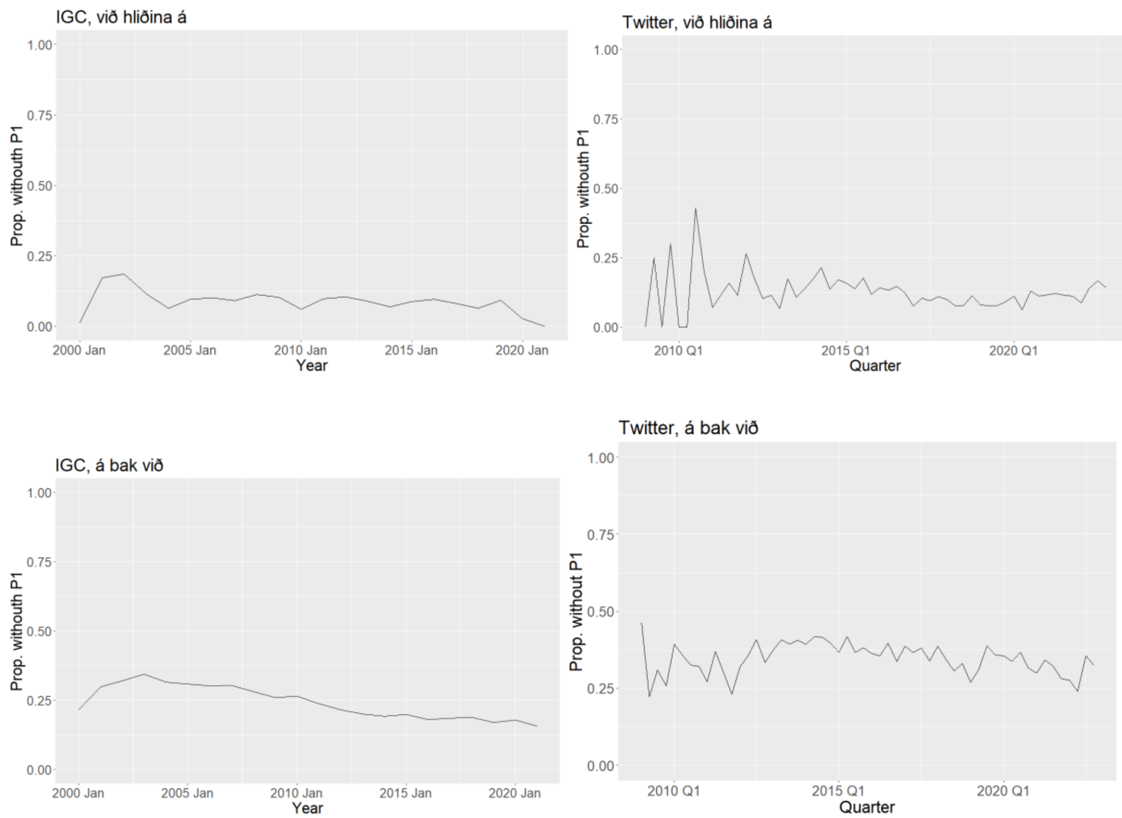


Figure 8.17. Four time series, showing the proportion of examples lacking **P₁** for the complex prepositions *við hliðina á* and *á bak við* in material from IGC and Twitter. The y-axis shows the proportion between 0 (0%) and 1 (100%).

When glancing at the series in Figure 8.17, the initial few observations in each series may appear slightly different from the rest of the series, i.e., they might be interpreted as containing more "noise" than the rest of the series. Recalling that the number of examples behind the first few observations are not as many as for the rest of the series, this is perhaps not surprising. To minimize the effect of early observation, these were not taken into consideration for analysis, model fitting or forecasting. In other words, the initial few observations were removed. This was done in the following way. Examples from Twitter are relatively few until around 2012 so Q1 2009 – Q4 2011 were removed from the series, leaving 44 observations stretching from Q1 2012 to Q4 2022. For IGC, the years 1999 –

2002 were removed, leaving 19 observations stretching from 2003 to 2021. A second reason for removing the initial observation is to make the series cover the same period as the time series in Chapter 9. The shortened time series, henceforth referred to as the complete time series, are depicted in Figure 8.18. A summary of the series is provided in Table 8.5.

Overview of data for the two complex prepositions						
Source	**Period covered**	**Observations**	**Series**	**N examples for complex preposition**		
				without P1	*with P1*	*total*
IGC	2003–2021	19 yearly	við hliðina á	2,488	27,548	30,036
Twitter	Q1 2012–Q4 2022	44 quarterly	við hliðina á	934	6,663	7,597
IGC	2003–2021	19 yearly	á bak við	20,722	65,658	86,380
Twitter	Q1 2012–Q4 2022	44 quarterly	á bak við	5,238	9,906	15,144

Table 8.5. Overview of the shortened time series (henceforth the complete time series), the number of observations and number of examples behind each series. For instance, the IGC time series containing information on *við hliðina á* consists of 30,036 examples that make up 19 observations.

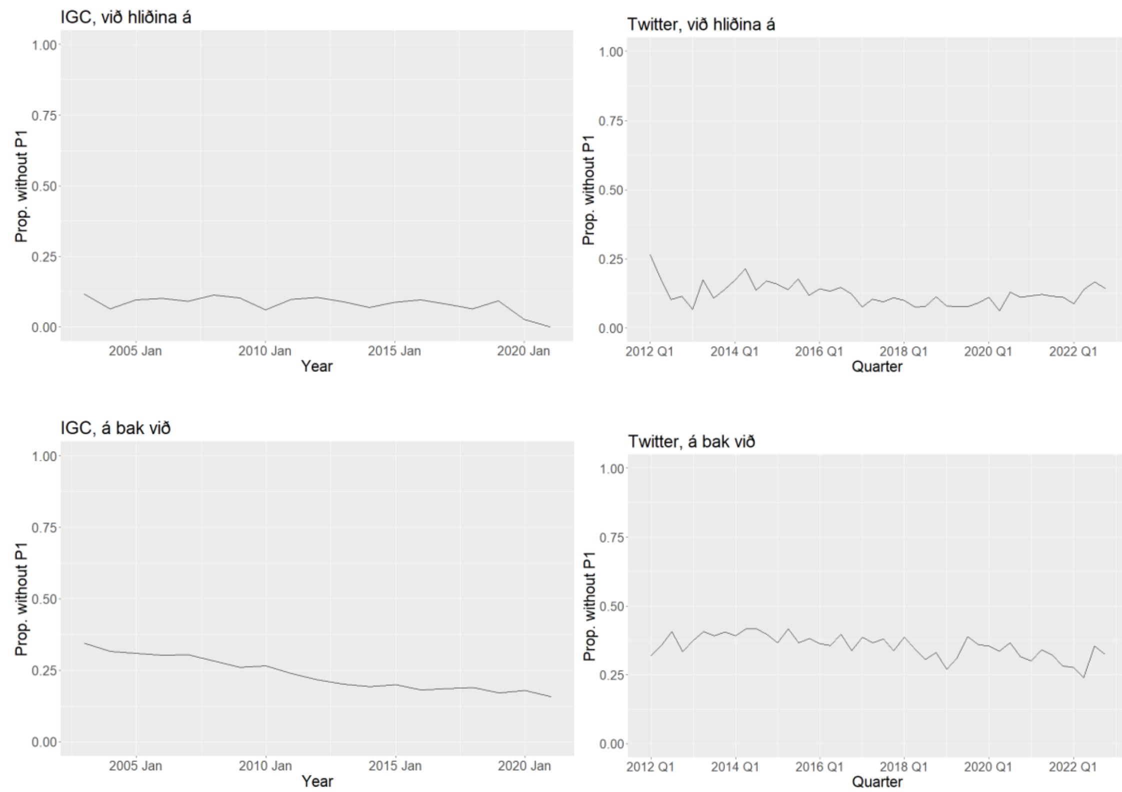


Figure 8.18. The four time series once the initial observations have been removed. Data from Twitter now consists of 44 observations for each series and data from IGC 19 observations. Again, the y-axis shows the proportion between 0 (0%) and 1 (100%).

In general, the overall proportion of examples without initial $\mathbf{P_1}$ varies depending on whether the data is from Twitter or IGC, and whether the preposition is *við hliðina á* or *á bak við*. The proportion of examples lacking $\mathbf{P_1}$ is generally lower for the complex preposition *við hliðina á* than *á bak við*. In IGC, examples lacking $\mathbf{P_1}$ make up about 11.6% of attested examples in 2003. The proportion remains relatively stable over time until 2020 when it drops to 2.6% and ends in 0% in 2021. Either the last two observations should be taken with a grain of salt (possibly considered as outliers), or there is a negative trend for dropping $\mathbf{P_1}$. In Section 8.5.2, the last two observations are handled like any other observation in the series. The proportion of examples of *við hliðina á* without $\mathbf{P_1}$ in Twitter is slightly higher than in IGC. In Q1 2012, the proportion is 26.4%. This series is also relatively stable over time, although an ever so slight negative trend might be claimed to be observable until 2020 when one observation (Q2 2020) shows the proportion to be 6.2%. By the end of 2022, the proportion of examples without $\mathbf{P_1}$ is ca. 14%.

The level of the series containing information on the proportion of examples of *á bak við* without initial $\mathbf{P_1}$ is higher than that for *við hliðina á*. In IGC, a clear and consistent negative trend is observed, with the proportion of examples lacking $\mathbf{P_1}$ being 34% in 2003, but around 15.7% in 2021. Data from Twitter shows that $\mathbf{P_1}$ is more frequently dropped in this type of material than in IGC-type material. The proportion of examples lacking initial $\mathbf{P_1}$ in Q1 2012 is around 32% and it is similar in Q4 2022.

The difference in the level of series containing information on *við hliðina á* and *á bak við* can be argued to be expected as grammaticalization of *á bak við* appears to have started earlier than for *við hliðina á* (see Section 8.3.4). However, a decline in dropping of P1 (as seems to be observed for at least *á bak við* in IGC) is not fully expected. The general direction of change as witnessed by the time series and expectations rising from forecasting are further discussed in Section 8.3.6.

The four time series were split into training and test sets. Under normal circumstances, 80% of the observations in a time series is used to train a model and 20% to test how well the model performs (Hyndman & Athanasopoulos 2021:135). However, since the time series dealt with here are already on the shorter side, setting this amount of observations aside for testing models will result in models being fitted to fewer observations further back in time. Additionally, since the goal is to produce a forecast for future time periods, it may be beneficial to include as much information as possible in the training set. For data from Twitter, the training set contained 40 observations (= 10 years) or c. 91% of the time series, and the test set contained 4 observations (= 1 year) or c. 0.9% of the time series. For IGC, the training set contained 17 observations (= 17 years) or c. 89% of the time series, and the test set contained 2 observations (= 2 years) or c. 11% of the time series. A summary of the training period and test period for each time series is shown in Table 8.6.

Source	**Series**	**Training period**		**Test period**		**Whole series**	
		Time	*N observations*	*Time*	*N observations*	*Time*	*N observations*
IGC	*við hliðina á*	2003–2019	17 (c. 89%)	2020–2021	2 (c. 11%)	2003–2021	19 (100%)
IGC	*á bak við*	2003–2019	17 (c. 89%)	2020–2021	2 (c. 11%)	2003–2021	19(100%)
Twitter	*við hliðina á*	Q1 2012– Q4 2021	40 (c. 91%)	Q1 2022–Q4 2022	4 (c. 9%)	Q1 2012– Q4 2022	44 (100%)
Twitter	*á bak við*	Q1 2012– Q4 2021	40 (c. 91%)	Q1 2022–Q4 2022	4 (c. 9%)	Q1 2012– Q4 2022	44 (100%)

Table 8.6. A summary of the training period and test period for each time series. The test set makes up about 10% of the whole series.

The procedure that was followed for each of the time series used to forecast was as discussed in Chapter 7, Section 7.2.1, namely to specify which model was to be used (define a model), fit the model to the training data (train the model, estimate parameters), and check the performance of the model (evaluate) by looking at the residuals and (when applicable) forecast errors. Finally, using a model where the residuals are within acceptable limits a forecast was produced.

Forecasts for future periods, i.e., periods after 2021 for IGC data and periods after Q4 2022 for the Twitter data were done in two ways, i.e., i) by fitting the same type of model used for the training data to the whole dataset and then projecting into the future, and ii) by fitting a model to the trend of the decomposed series and projecting into the future. Having provided a general overview of the four time series, each of them will now be discussed in turn including information on their properties, the models fitted and the results.

8.5.2 IGC *við hliðina á*

The IGC time series containing information about the proportion of examples of *við hliðina á* lacking $\mathbf{P}_1$ consists of 19 (yearly) observations from 2003 to 2021. The proportion of examples lacking $\mathbf{P}_1$ is generally very low, or less than 12% (the mean is ca. 8%). Since the time series is made up of yearly observations there is no seasonality. The series does not show any strong trends, although it drops to zero in 2021 when the observation is made up of 27 examples, all of which contain $\mathbf{P}_1$ *við*. As suggested by the autocorrelation function (ACF), cf. Figure 8.19, as well as the Ljung-Box statistics (statistics 4.51, p-value 0.921), the series is indistinguishable from white noise series. An STL decomposition (see Chapter 7, Section 7.2.2) of the series, where each observation is assumed to consist of a trend component and a remainder component (cf. Figure 8.19), shows a negative trend from around 2015 to 2021. The trend strength is 0.6325553, the linearity -0.07958893 and the curvature -0.04996428. The decomposition also shows a very small amount of noise, which lies somewhere between -0.02 and 0.01. Note that the gray boxes in each of the panels in Figure 8.19 are there to show the scale; they are of the same "size".

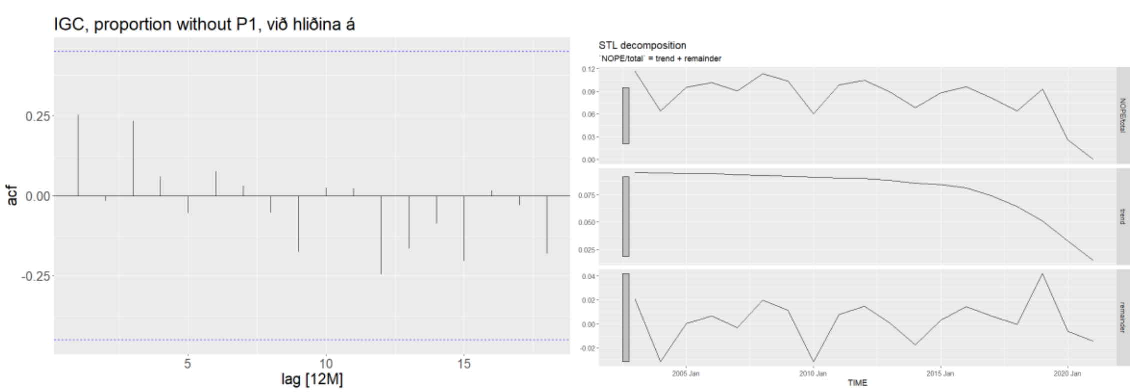


Figure 8.19. ACF and STL of the IGC series for the proportion of examples lacking $\mathbf{P}_1$ *við* in examples of *við hliðina á*. The series has a clear negative trend from around 2015 to 2021. The remainder component is very small, between 0.01 and -0.02.

Before choosing and fighting forecasting models, the series was divided into a training and test set. As noted earlier (Section 8.5.1), the training set contained 17 observations from 2003–2019, and the test set 2 observations from 2020–2021.

Three simple models were fitted to the training data, i.e., i) a Naïve model which assumes that all future observations will be the same as the last recorded observation in the series, ii) a Mean model which assumes that all future values will be equal to the mean of the whole series, and iii) a Drift model where changes over time are assumed to be average change in the historical data (see further in Chapter 7). Point predictions for each of the models for the test period are shown in Figure 8.20 along with the whole series. The three models give very similar results and appear near indistinguishable in Figure 8.20, with the proportion of examples without $\mathbf{P_1}$ being around 9%. Observed values for the period are 2.6% and 0%. A summary of the fit of each of the models to the training data as well as how accurate the point forecasts are, is provided in Table 8.7. Based on the RMSE, MAE, MASE and RMSSE (Table 8.7), the Drift method appears to perform marginally better than the other two models.

IGC, *við hliðina á*							
Model	source	type	RMSE	MAE	MAPE	MASE	RMSSE
Mean	IGC	Training	0.0166	0.0134	16.6	0.62	0.654
Naïve	IGC	Training	0.0253	0.0215	27.1	1	1
Drift	IGC	Training	0.0253	0.0215	26.8	1	0.998
Drift	IGC	Test	0.0783	0.0773	Inf	3.59	3.09
Mean	IGC	Test	0.0777	0.0766	Inf	3.56	3.07
Naïve	IGC	Test	0.0806	0.0795	Inf	3.69	3.18

Table 8.7. A comparison of the accuracy of the fit and the point forecast of each of the three simple models. Since the series goes to zero, MAPE does not return usable information.

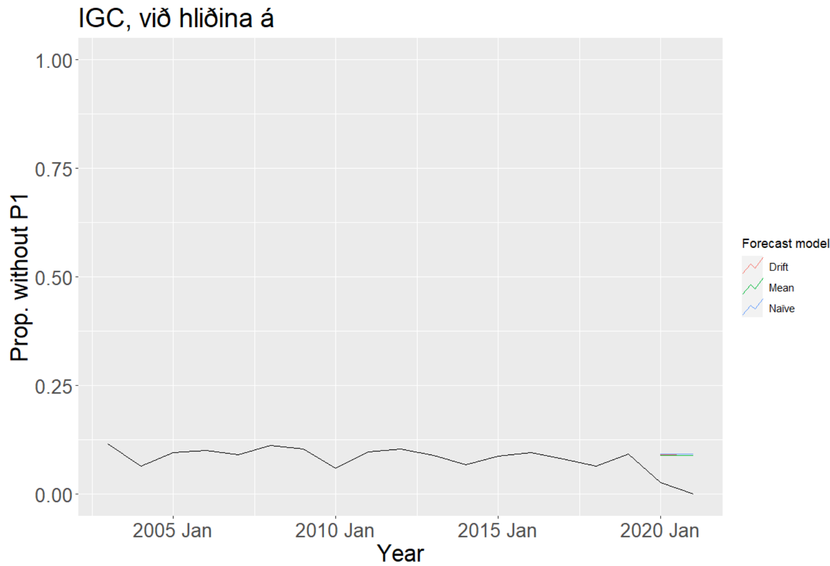


Figure 8.20. Three simple forecasting methods used to predict the data in the test set, i.e., the two-year period from 2020–2021. Although predictions from the three models appear near indistinguishable, the Drift model is marginally more accurate than the other models when taking into consideration RMSE, MAE, MASE and RMSSE (Table 8.7).

In addition to the three simple forecasting models mentioned above, an ETS and an ARIMA model were also fitted to the training data. Seeing that the time series goes down to zero in 2021, a log transformation was used to keep point predictions and prediction intervals within a positive range. The type of ETS and ARIMA model along with initial states was determined using the ETS() and ARIMA() functions from the fable package. While the ETS() function selects a model by minimizing the AICc (Hyndman & Athanasopoulos 2021:254), the ARIMA() function relies on the Hyndman-Khandakar algorithm for automatic ARIMA modeling, taking into account how often the series needs to be differenced to making it stationary and relying on AICc (Hyndman & Athanasopoulos 2021:286).

An automatic selection of an ETS model resulted in simple exponential smoothing with additive errors, ETS(A,N,N). The smoothing parameter was alpha = 0.0001000007 and the initial state l = -2.430083. The residuals of the model were deemed acceptable and plausibly interpreted as white noise (Ljung-Box statistics = 6.21 with p-value 0.798). An automatic selection of ARIMA resulted in ARIMA(0,0,0) with mean which is essentially a white noise model. The coefficient constant was -2.4301 (s.e. 0.0479). The residuals of this model were also deemed acceptable and plausibly interpreted as white noise (Ljung-Box statistics = 6.21 with p-value 0.798). Point forecasts with prediction intervals are shown in Figure 8.21. Even though different types of models are being used, the predictions for the training period are very similar. In both instances, the models suggest that the proportion of examples without **P₁** *við* should be just below 9%.

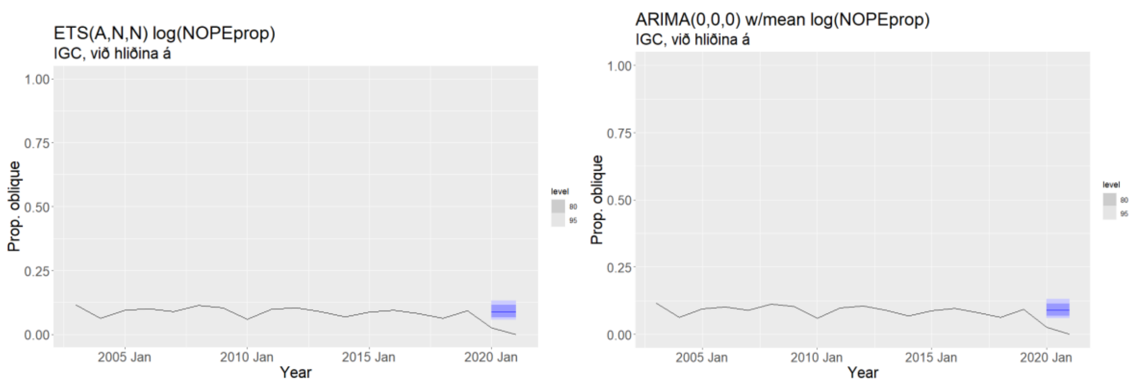


Figure 8.21. The ETS(A,N,N) model and ARIMA(0,0,0) mean produce similar forecasts for the test period.

A summary of the accuracy of the point predictions of the two models, i.e, the ETS(A,N,N) and ARIMA(0,0,0) model, is provided in Table 8.8. Note that the models give virtually identical results in accuracy of point predictions. An evaluation of the forecast distribution using quantile scores, Winkler scores, Continuously Ranked Probability Scores and Scale-free comparison using skill scores, suggests that both of the models perform slightly worse than the Naïve method and the Drift model. Skill scores (based on crps) suggest that the

Drift model is about 4.5% better than the Naïve method. This indicates that a simple forecasting method might be most appropriate for forecasting, although keeping in mind that the evaluation is based on a very small (2 steps ahead) test set.

IGC, *við hliðina á*							
Model	source	type	RMSE	MAE	MAPE	MASE	RMSSE
ETS(A,N,N) log(NOPEprop)	IGC	Training	0.0166	0.0138	16.7	0.639	0.657
ARIMA(0,0,0) w/mean log(NOPEprop)	IGC	Training	0.0166	0.0138	16.7	0.639	0.657
ETS(A,N,N) log(NOPEprop)	IGC	Test	0.078	0.0769	Inf	3.57	3.08
ARIMA(0,0,0) w/mean log(NOPEprop)	IGC	Test	0.0779	0.0767	Inf	3.56	3.07

Table 8.8. A comparison of the accuracy of the fit and the point forecast of the ETS(A,N,N) and ARIMA(0,0,0) model. The two models give virtually identical results.

IGC, *við hliðina á*						
Model	source	type	qs probs = 0.1	winkler level 80	crps	skill
Drift	IGC	Test	0.0651	0.444	0.0594	0.0474
Mean	IGC	Test	0.0973	0.586	0.0667	-0.0697
Naïve	IGC	Test	0.0726	0.481	0.0623	0
ARIMA(0,0,0) w/mean log(NOPEprop)	IGC	Test	0.0985	0.594	0.0665	-0.0672
ETS(A,N,N) log(NOPEprop)	IGC	Test	0.0974	0.589	0.0663	-0.0635

Table 8.9. Based on quantile, Winkler, crps and skill scores, it appears that the Drift method provides the best forecast distribution for the test period 2020–2021.

Since both the ETS(A,N,N) and ARIMA(0,0,0) performed worse than two of the benchmark models, these are not considered further. The Drift model appeared to generate the best forecast for the test period so this method was used on the decomposed trend of the series (STL decomposition with a trend window = 13). Predictions were generated a forecast 30 steps ahead into 2051. The results, shown in Figure 8.22, suggest that omission of $\mathbf{P_1}$ will be at minimum in the coming years.

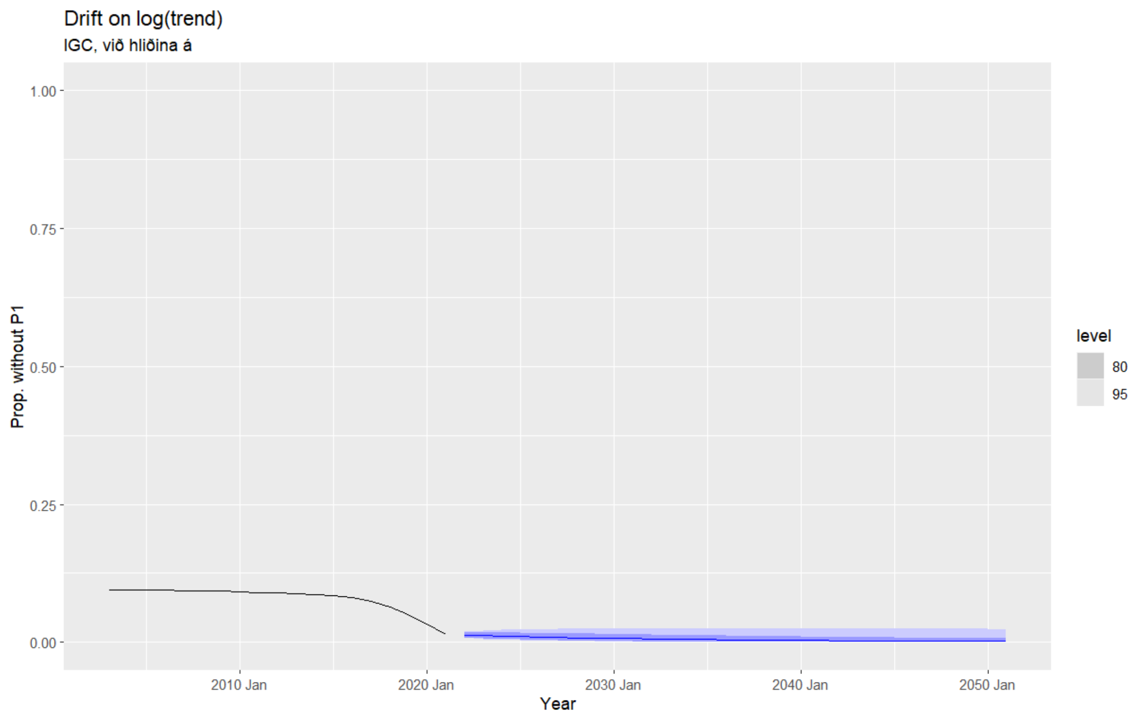


Figure 8.22. A Drift model used on the trend component of the series to generate a forecast 30 steps into the future.

The general picture that emerges from studying *við hliðina á* in IGC is that the proportion of examples lacking **P**$_1$ *við* is less than 12%, appearing mostly stable over time although the proportion drops to zero in 2021. No model was able to accurately predict the test period, possibly due to the sudden change in the series. Of the models that were tested, the Drift model appeared to be the best (although still quite poor), based on accuracy of point predictions and forecast distribution. Using the Drift model on the trend of the series, the proportion of examples without **P**$_1$ is expected to be close to zero in the near future. It should be noted that the last observation of the series, which was zero, might be treated as an outlier, in which case the proportion of examples without **P**$_1$ might be expected to be higher than zero but still lower than 10%. The expectations generated from the forecasts are discussed further in Section 8.6.

8.5.3 Twitter *við hliðina á*

The Twitter time series containing information about the proportion of examples of *við hliðina á* lacking **P_1** consists of 44 (quarterly) observations from Q1 2012 to Q4 2022. The proportion of examples without **P_1** appears to be relatively stable over time, hovering around 13% (the mean of the series is ca 12.4%) although occasional observations go higher (ca. 26%) or lower (ca 6%). As suggested by the autocorrelation function (ACF) which has one spike well outside of the significant level and according to Ljung-Box statistics (statistics 6.87, p-value 0.00874) the series is not simply white noise. A KPSS test (statistics 0.525, p-value 0.0361) suggests the series is not stationary and would need to be differenced once to make it so. An STL decomposition of the series, where each observation is assumed to consist of a trend component and a remainder component (cf. Figure 8.23), shows changes in the trend as time passes. There is a slight positive trend, followed by a negative trend and a second positive trend. The trend strength is 0.7327042, the linearity -0.113984 and the curvature 0.1027369. The decomposition also shows extremely small seasonal fluctuations, or between -0.0050 and 0.0025. The remainder component is quite large, between -0.05 and 0.15. Note that the gray boxes in each of the panels in Figure 8.23 are there to show the scale; they are of the same “size” and the larger the box the smaller the relevant component.

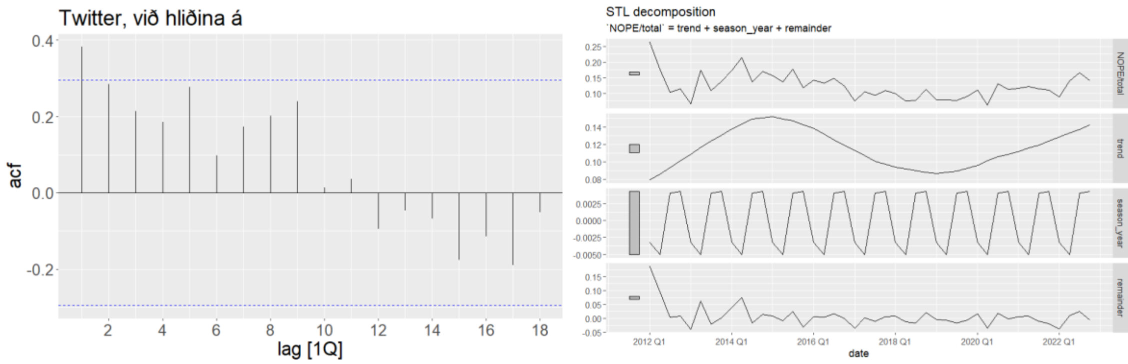


Figure 8.23. ACF and STL decomposition of the Twitter series for the proportion of examples lacking **P$_1$** in the complex prepositions *við hliðina á*. The series has an interesting trend component, quite a lot of noise (remainder) and extremely small seasonality.

The series was divided into a training and test set. As noted earlier (Section 8.5.1), the training set contained 40 observations from Q1 2012 to Q4 2021, and the test set 4 observations from Q1 – Q4 2022.

Four simple models were fitted to the Twitter first person training data: i) a Naïve model which assumes that all future observations will be the same as the last recorded observation in the series, ii) a Seasonally naïve model which assumes each quarter in the future will be the same as last reworded quarter of the same type, iii) a Mean model which assumes that all future values will be equal to the mean of the whole series, and iv) a Drift model where changes over time are assumed to be average change in the historical data (see further in Chapter 7, Section 7.2.3). The point-predictions of each of the models for the test period are shown in Figure 8.24 along with the whole series. The model relying on the mean of the training data appears to give the most accurate point prediction, claiming examples without **P$_1$** should make up just above 12% of attested examples in Q1 2022 – Q4 2022 when the observed values lie between ca. 9% and 17%, depending on the quarter. A summary of the fit of each of the models to the training data as well as accuracy of the point forecasts is provided in Table 8.10.

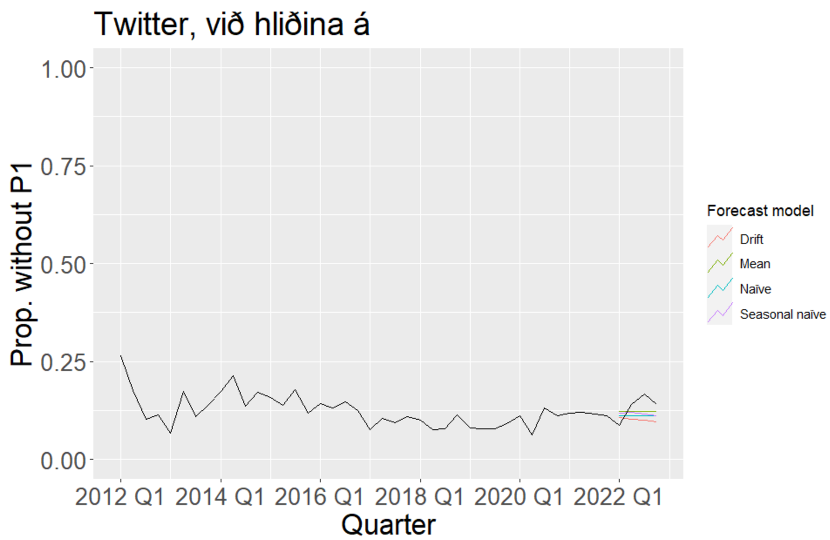


Figure 8.24. Four different models were fitted to the training data of the Twitter time series containing information on proportion of examples of við *hliðina á* lacking **P1**. These models were: a Drift model, a Mean model, a Naïve model, and a Seasonal naïve model. The fitted model was used to forecast Q1 – Q4 2022.

Twitter, *við hliðina á*								
Model	source	type	RMSE	MAE	MAPE	MASE	RMSSE
Mean	Twitter	Training	0.0412	0.0312	27.4	0.962	0.846
Naïve	Twitter	Training	0.0406	0.0311	26.8	0.961	0.832
Seasonal naïve	Twitter	Training	0.0488	0.0324	31.2	1	1
Drift	Twitter	Training	0.0404	0.0311	26.4	0.96	0.828
Drift	Twitter	Test	0.0457	0.0423	30.2	1.31	0.938
Mean	Twitter	Test	0.0305	0.0283	22.7	0.875	0.626
Naïve	Twitter	Test	0.0366	0.0345	25.4	1.06	0.752
Seasonal naïve	Twitter	Test	0.0346	0.0325	24.7	1	0.709

Table 8.10 Report on the fit of the simple models to the training data and the accuracy of the point predictions for the training period Q1 – Q4 2022.

The four simple models above serve as a benchmark when fitting and choosing other models. Both ETS and ARIMA models were fitted to the training data in the hope that they would perform better than the simple models above. The type of ETS and ARIMA model along with initial states was determined using the ETS() and ARIMA() functions from the fable package (O'Hara-Wild, Hyndman & Wang 2021). While the ETS() function selects a model by minimizing the AICc (Hyndman & Athanasopoulos 2021:254), the ARIMA() function relies on the Hyndman-Khandakar algorithm for automatic ARIMA modeling (Hyndman & Athanasopoulos 2021:286).

An automatic selection of an ETS model resulted in ETS(M,N,N) which is a model with multiplicative errors, no seasonality and no trend. The smoothing parameters were alpha = 0.3221618 , with the initial state l = 0.2000737. The residuals of the model are within desirable limits (mean of innovation residuals is -0.03414046) and can be interpreted as white noise (Ljung-Box statistics = 5.17 with p-value = 0.879, lag = 10).

The automatic selection of ARIMA resulted in (0,1,1) coefficients ma1 = -0.5567 (s.e. .e. 0.2298). The residuals from this model are also within desirable limits (mean of the innovation residuals is -0.007172366) and the series (Ljung-Box statistics = 8.20 with p-value 0.609, lag = 10). Note that neither of these models pick up on the extremely small seasonal component found in the time series. Point forecasts with prediction intervals are shown in Figure 8.25 and a summary of the accuracy of the fit of the models and accuracy of point forecasts are provided in Table 8.11. The ARIMA(0,1,1) appears to generate a minimally more accurate point forecast than the ETS(M,N,N), but still worse than the Mean model above.

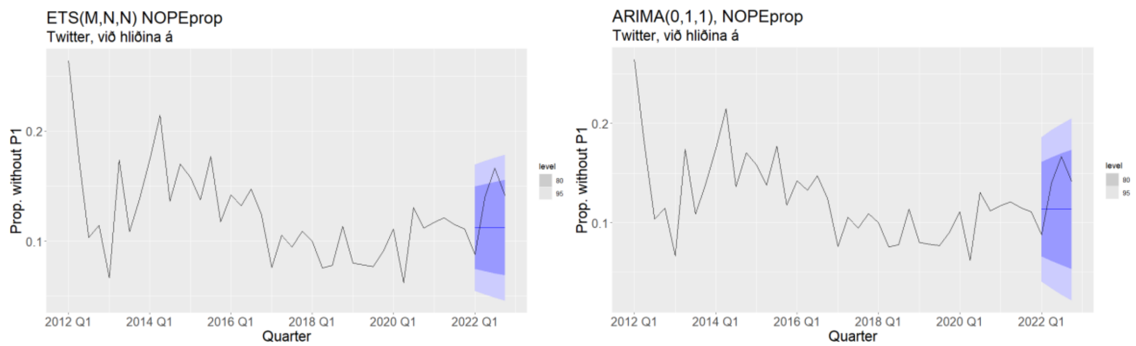


Figure 8.25. An ETS(M,N,N) and ARIMA(0,1,1) were fitted to the training data. The models were then used to predict the test data, i.e., Q1 2022 to Q4 2022.

Twitter, *við hliðina á*							
Model	source	type	RMSE	MAE	MAPE	MASE	RMSSE
ETS(M,N,N) NOPEprop	Twitter	Training	0.0366	0.0277	24.9	0.854	0.75
ARIMA(0,1,1) NOPEprop	Twitter	Training	0.0362	0.0264	23.8	0.814	0.742
ETS(M,N,N) NOPEprop	Twitter	Test	0.0359	0.0338	25.2	1.04	0.736
ARIMA(0,1,1) NOPEprop	Twitter	Test	0.0352	0.0333	24.9	1.03	0.722

Table 8.11. Report on the accuracy of the point forecasts for Q1 – Q4 2022 generated with ETS(M,N,N) and ARIMA(0,1,1). The ARIMA model appears to be minimally more accurate.

The point forecasts of the ETS(M,N,N) and ARIMA(0,1,1) are minimally different. The former predicts that the proportion of examples without $\mathbf{P_1}$ from Q1 – Q4 2022 to be 11.23224% and the latter predicts 11.34912%. Observed values lie between ca. 8.8% and 16.7% which is fully within the 80% confidence interval of the ARIMA(0,1,1) model, but not of the ETS(M,N,N) model.

An evaluation of the forecasts distribution of all the methods discussed above using quantile scores, Winkler scores, Continuously Ranked Probability Scores and Scale-free comparison using skill scores, suggests that the ETS(M,N,N) performs ca. 4% worse than the Seasonal naïve method when skill scores (based on crps) are considered. The

ARIMA(0,1,1) with drift is only estimated to be around 0.3% better. The Mean model appears to be almost 12% better than the Seasonal naïve model, although keeping in mind that the evaluation is based on a very small (4 steps ahead) test set.

Twitter, *við hliðina á*						
Model	source	type	qs probs = 0.1	winkler level 80	crps	skill
Drift	Twitter	Test	0.023	0.165	0.0264	-0.279
Mean	Twitter	Test	0.013	0.108	0.0181	0.119
Naïve	Twitter	Test	0.0206	0.16	0.0225	-0.0942
Seasonal naïve	Twitter	Test	0.0161	0.125	0.0206	0
ETS(M,N,N) NOPEprop	Twitter	Test	0.0124	0.113	0.0214	-0.0379
ARIMA(0,1,1) NOPEprop	Twitter	Test	0.0149	0.108	0.0205	0.00311

Table 8.12. Based on quantile, Winkler, crps and skill scores, it looks like the Mean model has the best forecast distribution for the test period Q1 2022 – Q4 2022.

The Mean model was fitted to the whole series and used to predict 20 steps into the future. The results suggest that for the next several years, the proportion of examples without $\mathbf{P_1}$ will be around 12.4%.

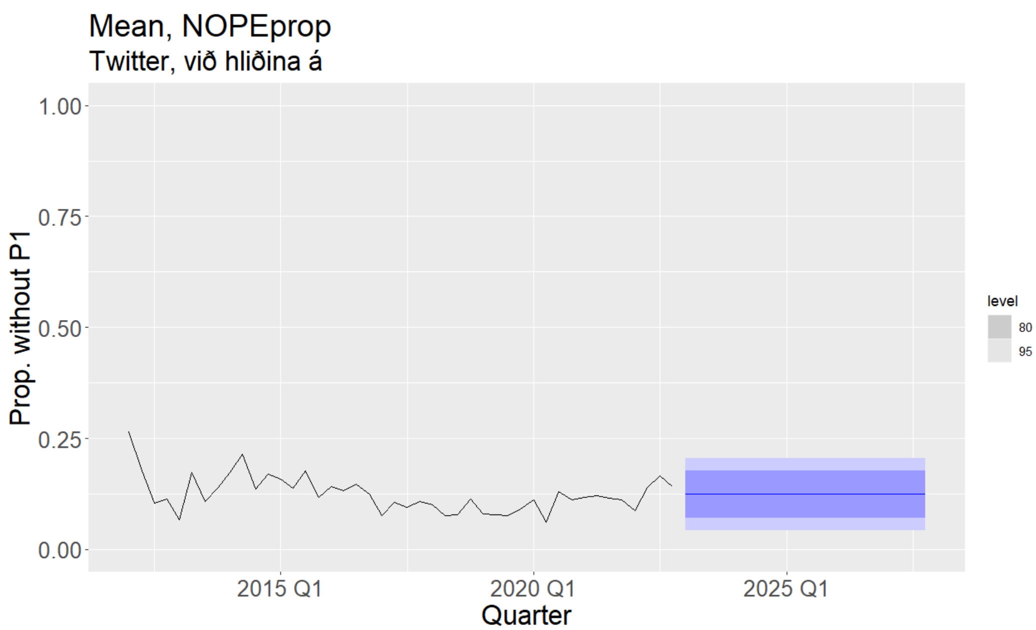


Figure 8.26. A 20-step ahead forecast using a Mean model predicts that the proportion of examples without $\mathbf{P_1}$ will be around 12.4% for the next few years.

As noted above, an STL decomposition of the time series suggested that the remainder component was quite large. If the remainder component is claimed to not hold meaningful information about the proportion of examples without $\mathbf{P_1}$ over time, it might be omitted from the series along with the seasonal component. The remaining trend shows a pattern that appears to tend towards cyclic and the series is now stationary. The mean is ca. 11.6% so if a Mean model were to be used, all future periods would be hypothesized to hover around that mark. In Figure 8.27 an ARIMA model was selected automatically for the trend component and used to predict 20 steps into the future. The automatic selection resulted in a seasonal ARIMA(2,0,2)(0,0,1). The point forecast suggests a slight increase in examples without $\mathbf{P_1}$ over time, or from 14.7% in Q1 2023 to ca. 17.7% in Q4 2027.

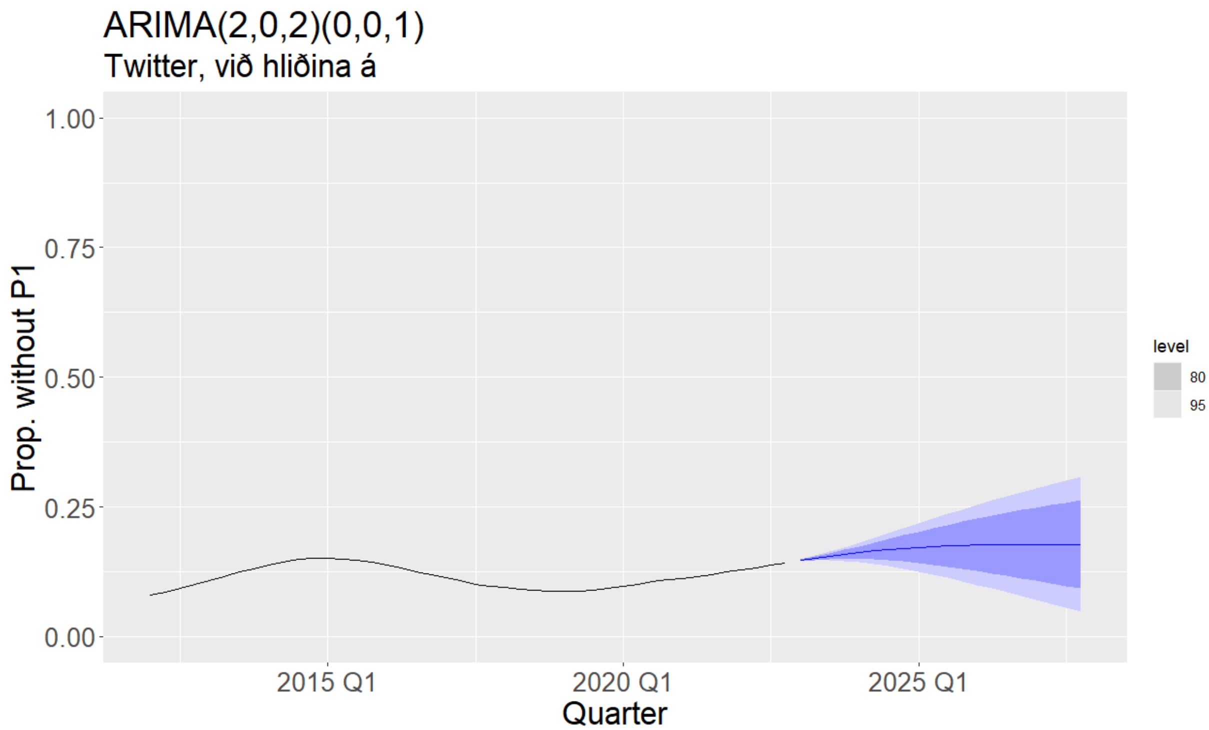


Figure 8.27. A seasonal ARIMA(2,0,2)(0,0,1) model fitted to the trend of the whole series and used to predict the proportion of examples without $\mathbf{P_1}$ in the complex preposition v*ið hliðina á* from Q1 2023 to Q4 2027.

The general picture that emerges from studying the proportion of examples without $\mathbf{P_1}$ in the complex preposition við hliðina á in data from Twitter, is that the proportion stays relatively stable over time if one ignores the large amount of noise in the data. In Q1 2012 examples without $\mathbf{P_1}$ make up about 26.4% of the attested examples on Twitter and in Q4 2022 they make up about 14% with the periods in between fluctuating between these numbers. The mean of the series is just around 12.4%.

In addition to four simple models (Naïve, Seasonal naïve, Drift and Mean model), an ETS(M,N,N) and ARIMA(0,1,1) were fitted to the training data. Both of these produced worse forecasts for the test data than some of the simple models. The Mean model gave the best point forecast and forecast distribution for the training data, even though the residuals of the model were not perfect. The Mean model suggests that examples without **P_1** will make up around 12.4% of attested examples in all future perdios. The forecast intervals suggest they might drop below 10% or go close to 20%. If the remainder component and the very small seasonal component are removed from the raw series, an automatically selected seasonal ARIMA(2,0,2)(0,0,1) applied to the trend component, predicts a slight increase in of examples without **P_1**, or from 14.7% in Q1 2023 to ca. 17.7% in Q4 2027. Note, however, that the prediction intervals after 2025 become larger. Expectations generated from the forecasts in this section are discussed further in Section 8.6.

8.5.4 IGC *á bak við*

The IGC time series containing information about the proportion of examples of *á bak við* lacking **P_1** consists of 19 (yearly) observations from 2003 to 2021. The proportion of examples lacking **P_1** decreases over time, being ca. 35% in 2003 and 16% in 2021. The mean of the series is just below 24%. Since the time series is made up of yearly observations there is no seasonality. The series is non-stationary and would have to be differenced once to make it so. As suggested by the autocorrelation function (ACF), cf. Figure 8.28, as well as the Ljung-Box statistics (statistics 51.3, p-value 0.000000153), the series is not a white noise series. An STL decomposition (see Chapter 7, Section 7.2.2) of the series (trend window = 13), where each observation is assumed to consist of a trend component and a remainder component (cf. Figure 8.28), shows a clear negative trend. The trend strength is 0.9799812, the linearity -0.2455351 and the curvature 0.03981829. The decomposition also shows a very small amount of noise, which lies somewhere between -

0.010 and 0.015. Note that the gray boxes in each of the panels in Figure 8.28 are there to show the scale; they are of the same "size".

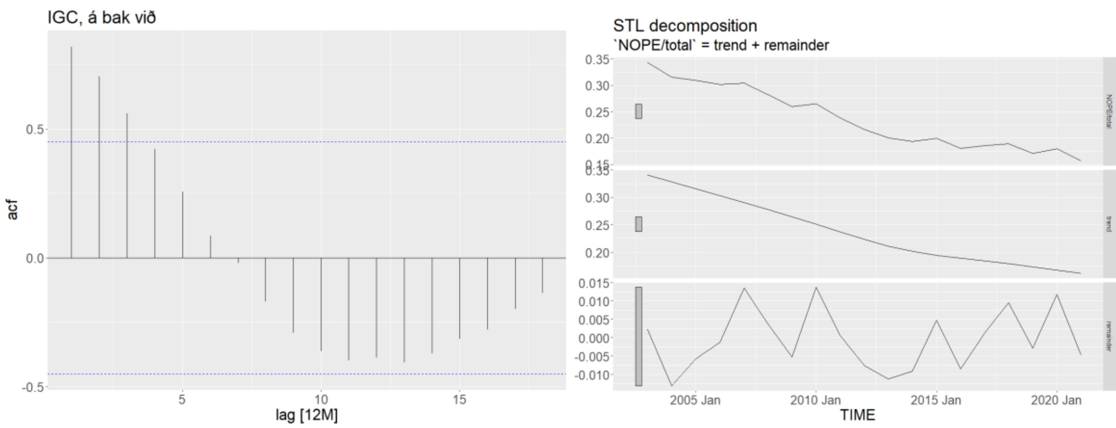


Figure 8.28. ACF and STL decomposition of the Twitter series for the proportion of examples lacking **P₁** in the complex prepositions *á bak við*. The series shows a clear negative trend and a small remainder component, which might be interpreted as noise.

Before choosing and fighting forecasting models, the series was divided into a training and test set. As noted earlier (Section 8.5.1), the training set contained 17 observations from 2003 – 2019, and the test set 2 observations from 2020 – 2021.

Three simple models were fitted to the training data, i.e., i) a Naïve model which assumes that all future observations will be the same as the last recorded observation in the series, ii) a Mean model which assumes that all future values will be equal to the mean of the whole series, and iii) a Drift model where changes over time are assumed to be average change in the historical data (see further in Chapter 7). Point predictions for each of the models for the test period are shown in Figure 8.29 along with the whole series. The Mean model gives quite different results from the other two, with the proportion of examples without **P₁** being predicted to be 24.4%. The Naïve model suggests a slightly lower proportion, or 17%. The Drift model predicts the proportion in 2020 should be just around

16% and in 2021 just below 15%. Observed values for the period are 17.9% and 15.7%. A summary of the fit of each of the models to the training data as well as how accurate the point forecasts are, is provided in Table 8.13. The Naïve model appears to give the best predictions of the three.

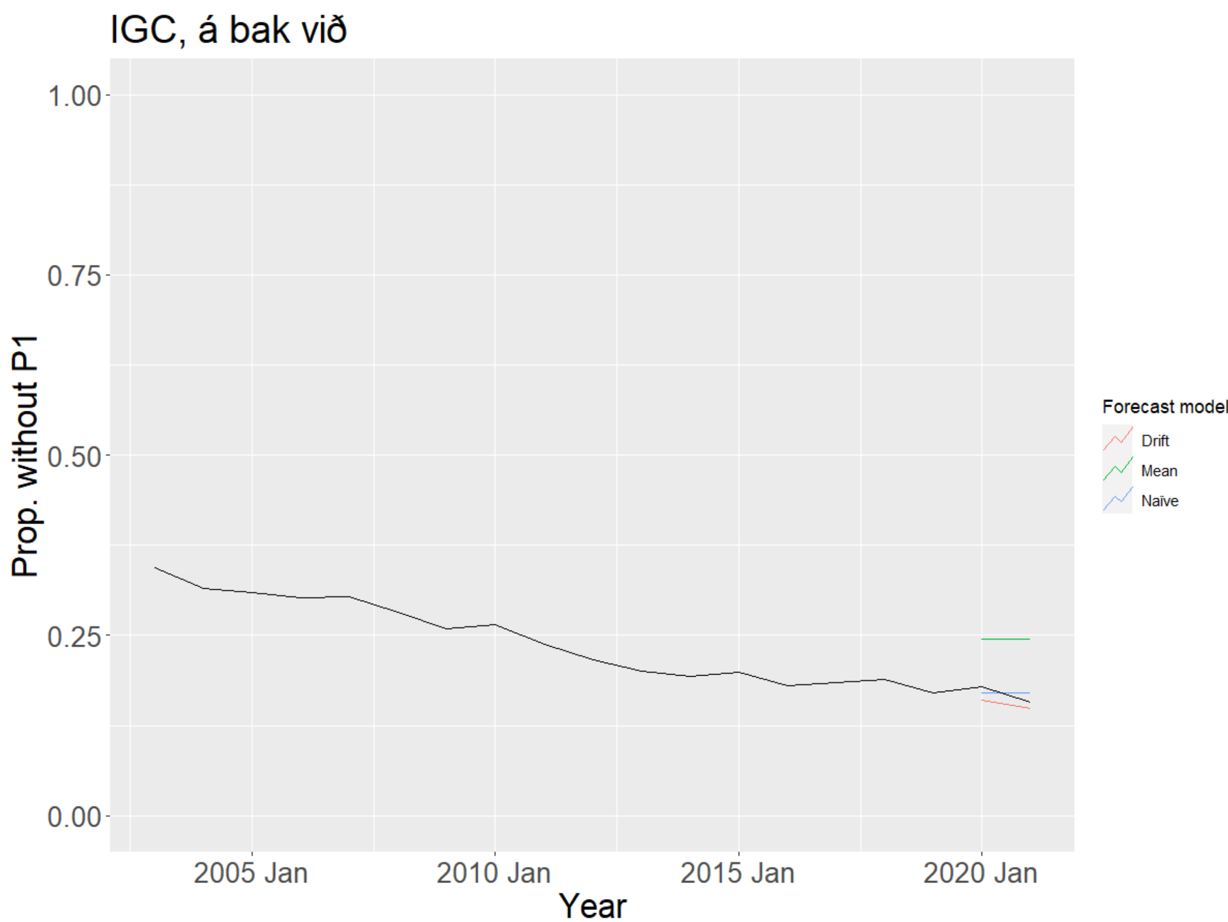


Figure 8.29. Three different models were fitted to the training data of the IGC time series containing information on proportion of examples of *á bak við* lacking $\mathbf{P_1}$. These models were: a Drift model, a Mean model and a Naïve. The fitted model was used to forecast 2020 and 2022.

IGC, *á bak við*							
Model	source	type	RMSE	MAE	MAPE	MASE	RMSSE
Mean	IGC	Training	0.0551	0.0502	21.5	3.68	3.39
Naïve	IGC	Training	0.0162	0.0136	5.93	1	1
Drift	IGC	Training	0.0121	0.0111	4.83	0.812	0.747
Drift	IGC	Test	0.0148	0.0136	7.89	0.997	0.913
Mean	IGC	Test	0.0769	0.0761	45.9	5.57	4.74
Naïve	IGC	Test	0.0116	0.0113	6.83	0.826	0.712

Table 8.13. Report on the fit of the simple models to the training data and the accuracy of the point predictions for the training period 2020 – 2021.

The three simple models above served as benchmark methods when fitting and choosing other models. Both ETS and ARIMA models were fitted to the training data in the hope that they would perform better than the more simple models. The type of ETS and ARIMA model along with initial states was determined using the ETS() and ARIMA() functions from the fable package (O'Hara-Wild, Hyndman & Wang 2021). While the ETS() function selects a model by minimizing the AICc (Hyndman & Athanasopoulos 2021:254), the ARIMA() function relies on the Hyndman-Khandakar algorithm for automatic ARIMA modeling (Hyndman & Athanasopoulos 2021:286).

Since the series has a negative trend, it seems reasonable to use log-transformation to ensure both point forecasts and prediction intervals stay positive. An automatic selection of an ETS model resulted in ETS(A,A,N) which is a model with additive errors, additive trend and no seasonality. The smoothing parameters were alpha = 0.6376826 and beta = 0.0001000014, with the initial states l = -1.036822 and b = -0.04240781. The residuals of the model are within acceptable limits although one lag (lag 6) was slightly outside the significance level. Otherwise, the residuals (mean of the innovation residuals is 5.199569e-

05) are normally distributed and can be interpreted as white noise (Ljung-Box statistics = 12.2 with p-value = 0.272).

The automatic selection of ARIMA resulted in ARIMA(0,1,0) with drift, which is essentially a random walk model (constant -0.0437 (s.e. 0.0127)). The residuals from this model can be regarded as white noise (mean of the innovation residuals is -6.029307e-05, Ljung-Box statistics = 9.02 with p-value 0.531, lag = 10) although residuals have a slight positive right-sided tail in their distribution which might slightly skew prediction intervals. Like the ETS model, lag 6 is very close to the significance level. An attempt was made to find a better ARIMA model, but the ARIMA(0,1,0) w/drift was the one with lowest AICc score.

Point forecasts with prediction intervals are shown in Figure 8.30 and a summary of the fit of the models and accuracy of point forecasts are provided in Table 8.14. The ETS(A,A,N) appears to generate a minimally more accurate point forecast than the ARIMA(0,1,0) with drift.

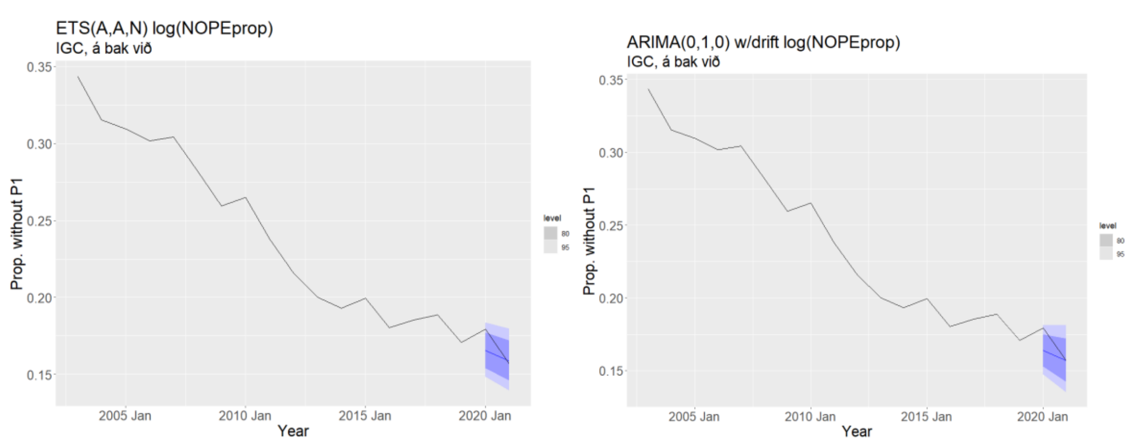


Figure 8.30. An ETS(A,A,N) and ARIMA(0,1,0) with drift were fitted to the training data. The models were then used to predict the test data, i.e., 2021 – 2022.

IGC, ***á bak við***							
Model	source	type	RMSE	MAE	MAPE	MASE	RMSSE
ETS(A,A,N) log(NOPE)prop	IGC	Training	0.0107	0.00961	4.17	0.704	0.66
ARIMA(0,1,0) w/drift log(NOPEprop)	IGC	Training	0.0112	0.0102	4.44	0.749	0.69
ETS(A,A,N) log(NOPE)prop	IGC	Test	0.00988	0.00788	4.47	0.577	0.609
ARIMA(0,1,0) w/drift log(NOPEprop)	IGC	Test	0.0111	0.00788	4.39	0.577	0.686

Table 8.14. Report on the accuracy of the point forecasts for 2021 – 2022 generated with ETS(A,A,N) and ARIMA(0,1,0) with drift. The ETS model appears to be marginally more accurate.

The point forecasts of the ETS(A,A,N) and ARIMA(0,1,0) with drift are minimally different. The former predicts the proportion of examples without $\mathbf{P_1}$ to be ca. 16.6% in 2021 and 15.9% in 2022. The latter predicts examples without $\mathbf{P_1}$ to make up ca. 16.4% of the data in 2021 and 15.7% in 2022. Observed values are 17.9% (2021) and ca. 17.5% (2022) and lie fully within the 95% confidence intervals of the ETS(A,A,N) and ARIMA(0,1,0) with drift models.

An evaluation of the forecasts distribution of all the methods discussed above using quantile scores, Winkler scores, Continuously Ranked Probability Scores and Scale-free comparison using skill scores, suggests that the ETS(A,A,N) performs about 16.6% better than the Naïve model when skill scores (based on crps) are considered. The ARIMA(0,1,0) with drift is only estimated to be around 1.9% better, although keeping in mind that the evaluation is based on a very small (2 steps ahead) test set.

IGC, *á bak við*						
Model	source	type	qs probs = 0.1	winkler level 80	crps	skill
Drift	IGC	Test	0.00667	0.0566	0.00931	-0.31
Mean	IGC	Test	0.0122	0.212	0.0488	-5.87
Naïve	IGC	Test	0.00451	0.0502	0.0071	0
ETS(A,A,N) log(NOPEprop)	IGC	Test	0.00361	0.0355	0.00592	0.166
ARIMA(0,1,0) w/drift log(NOPEprop)	IGC	Test	0.00413	0.0489	0.00697	0.0193

Table 8.15. Based on quantile, Winkler, crps and skill scores, it looks like the ETS(A,A,N) model has the best forecast distribution for the test period.

With both the ETS(A,A,N) and ARIMA(0,1,0) with drift performing better than the benchmark models, these models were fitted to the whole time series and used to predict 20 steps into the future. The results suggest that for the next several years, the proportion of examples without $\mathbf{P_1}$ will continue to decrease. The point forecasts of the two models are very similar, with the ETS ETS(A,A,N) predicting a decrease from 15.45% (2022) to 7.1% (2041) and the ARIMA(0,1,0) with drift 15% (2022) to 6.8% (2041)

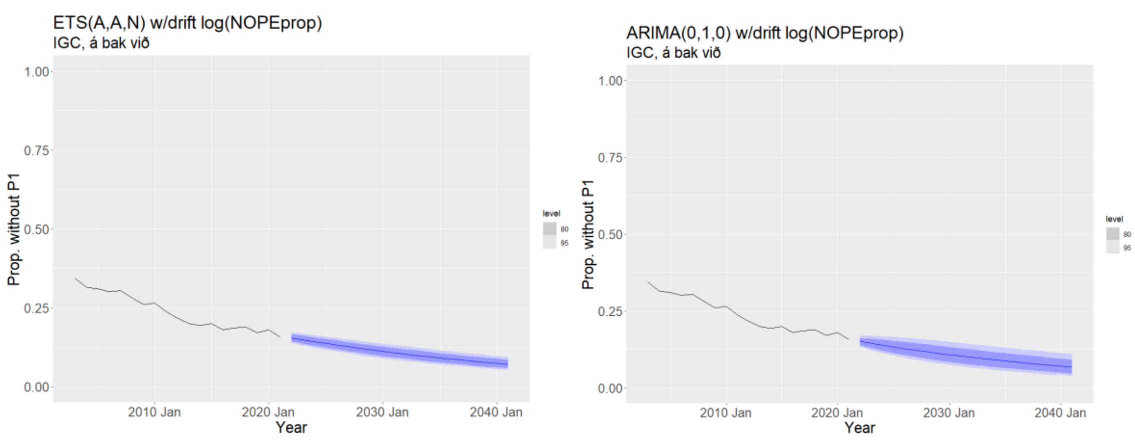


Figure 8.31. Forecasts 20-step ahead using ETS(A,A,N) and ARIMA(0,1,0) with drift. Both models predict a decrease in examples without P1 in the coming years.

As noted above, an STL decomposition of the time series suggested that the remainder component was relatively small. Nevertheless, it is possible to remove the remainder component, hypothesizing it does not hold meaningful information about the proportion of examples without $\mathbf{P_1}$ over time. An ETS(A,A,N) model was fitted to the trend, seeing that this model gave slightly more accurate point forecasts and forecast distribution for the test series than other models, and used to predict 30 steps ahead. The results are similar to using an ETS(A,A,N) for the non-decomposed series in that they indicate a decrease in the proportion of examples without $\mathbf{P_1}$ over time. The forecast intervals are now narrower which has an effect on the point predictions (these are based on the mean of the forecast distribution) which now show a decrease from 15.5% (2022) to 4.7% (2022). The 30-step ahead forecast is shown in Figure 8.32.

Note that in the case of this particular time series, using the trend from an STL decomposition (trend window = 13) to fit a model and predict the future might be slightly problematic. Forecast distribution and intervals are usually based on error terms in the series. Removing the remainder component from an already relatively smooth series might result in the forecast distribution to be narrower than is reasonable, i.e., it may result in removing an informative part of the series. For series with more "noisy" remainder components, using the trend might be more feasible. Essentially, the question is how to balance what constitutes an informative signal in the series about propagation of language change and what is simply noise that is in the way.

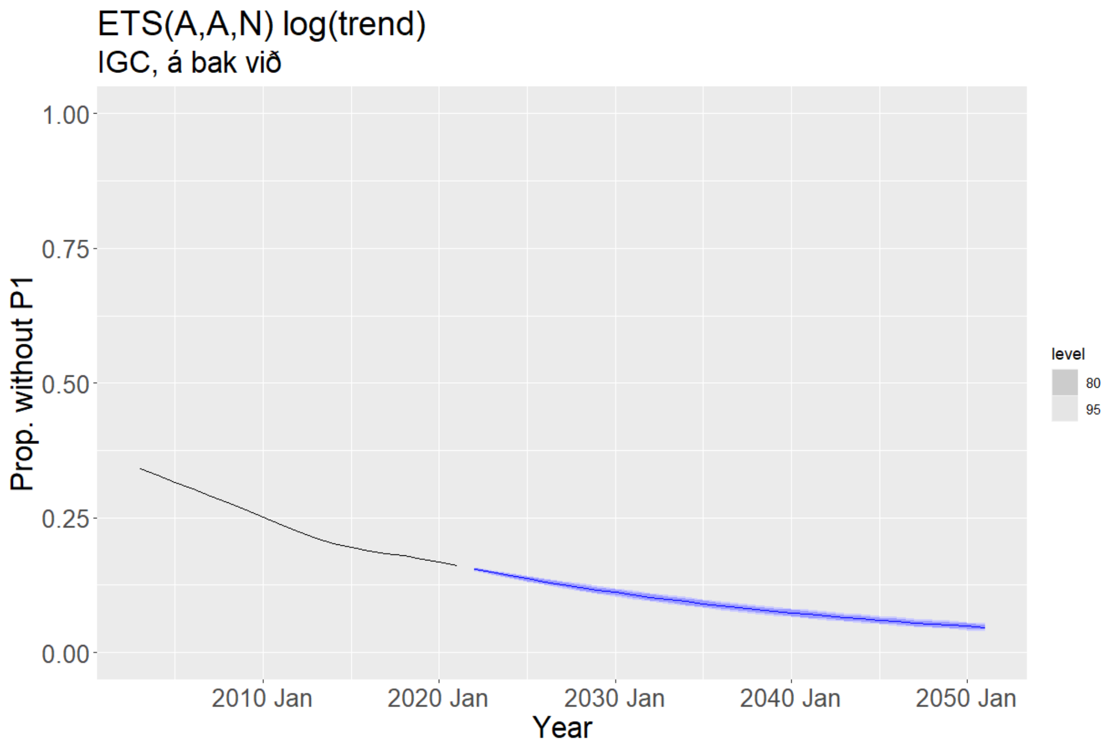


Figure 8.32. An ETS(A,A,N) model fitted to the trend of the whole series and used to predict the proportion of examples without $\mathbf{P_1}$ in the complex preposition *á bak við* from 2022 to 2051.

The general picture that emerges from studying the proportion of examples without $\mathbf{P_1}$ in the complex preposition *á bak við* in data from IGC, is that the proportion of examples without $\mathbf{P_1}$ decreases over time. In 2003 the proportion is ca. 34% but in 2021 ca. 15.7%. In addition to three simple models (Naïve, Drift and Mean model), an ETS(A,A,N) and ARIMA(0,1,0) were fitted to the training data. Both of these resulted in better point forecasts and forecast distributions for the training period than the simple models. Although the ETS(A,A,N) and ARIMA(0,1,0) gave similar results, the former was marginally better at predicting the test data than the latter. Using these models to prejudice a 20-step ahead forecast, the expectations that emerge are as follows: Both models predict a continued decrease in examples lacking $\mathbf{P_1}$ with ETS(A,A,N) suggesting a decrease from 15.45%

(2022) to 7.1% (2041) and the ARIMA(0,1,0) with drift 15% (2022) to 6.8% (2041). Prediction intervals are generally fairly narrow. Expectations generated from the forecasts in this section are discussed further in Section 8.6.

8.5.5 Twitter *á bak við*

The Twitter time series containing information on the proportion of examples of *á bak við* lacking $\mathbf{P_1}$ consists of 44 (quarterly) observations from 2012 to 2022. The series appears to be relatively stable over time, with around 32% of examples in Q1 2012 lacking $\mathbf{P_1}$ and 32.3% in Q4 2022. The mean of the whole series is 35%, although it appears that glancing at the series there might be a slight change in the level over time. Thus, only considering data from up until 2018, the mean is just above 37%, with data from 2019 onwards having a mean of 32.5%. The trend strength, as computed based on the whole series is 0.7542277, linearity -0.1702548, curvature -0.0668433. As witnessed by the ACF, where the four initial coefficients lie outside the expected limits (see Figure 8.33), and by the Ljung-Box statistics (= 50.0, p-value = 0.000000263, lag = 10) the series is not a white noise series. Related to that, the series is non-stationary and would have to be differenced once to make it stationary.

Given that the series consists of quarterly observations, seasonality might be detected. This is in fact the case. An STL decomposition (see Chapter 7, Section 7.2.2) of the series (trend window = 13), where each observation is assumed to consist of a trend component, a seasonal component and a remainder component (cf. Figure 8.33), shows an extremely small seasonal component between -0.01 and slightly above 0.01. Note that the seasonal window has been set to periodic to prevent the seasonal element from changing over time. The remainder component is larger than the seasonal component. The STL decomposition also shows a negative trend from around 2014 and onwards. The window for the trend component was specified as 13 to minimize the amount of noise. The gray boxes in each of the panels in Figure 8.33 show the scale i.e., they are the same size.

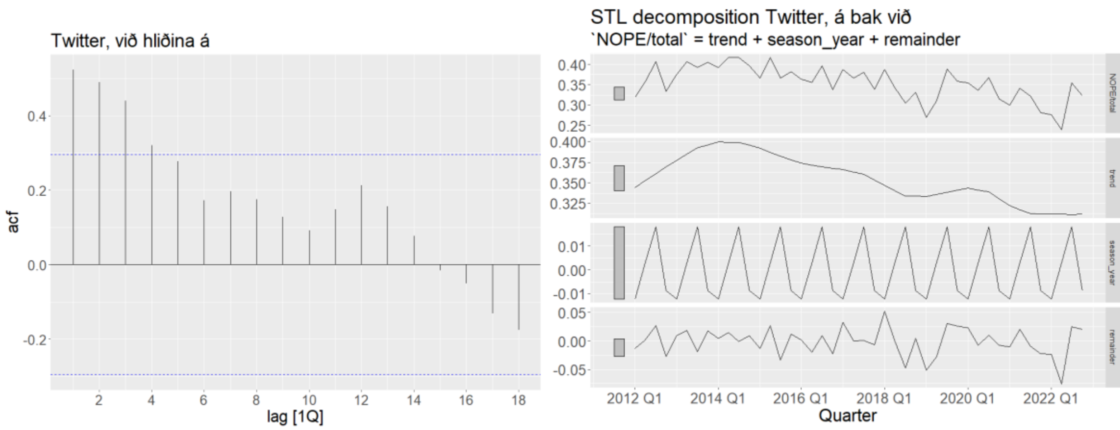


Figure 8.33. Some features of the Twitter time series that contain information on the proportion of examples of *á bak við* that lack $\mathbf{P_1}$. Both the ACF and the STL decomposition suggest the series has a slight trend.

The series was divided into a training and test set. As noted earlier (Section 8.5.1), the training set contained 40 observations from Q1 2012 – Q4 2021, and the test set 4 observations from Q1 – Q4 2022.

Four simple models were fitted to the Twitter first person training data: i) a Naïve model which assumes that all future observations will be the same as the last recorded observation in the series, ii) a Seasonal naïve model which assumes each quarter in the future will be the same as last reworded quarter of the same type, iii) a Mean model which assumes that all future values will be equal to the mean of the whole series, and iv) a Drift model where changes over time are assumed to be average change in the historical data (see further in Chapter 7, Section 7.2.3). The point-predictions of each of the models for the test period are shown in Figure 8.34 along with the whole series. None of the models appear particularly good. Of the four, the Naïve model appears to be the most accurate, predicting 28% of examples lacking **P**$_1$ in Q1 2022 – Q4 2022. Observed values lie between 27.6% and 35.5%. A summary of the fit of each of the models to the training data as well as how accurate the point forecasts are, is provided in Table 8.16.

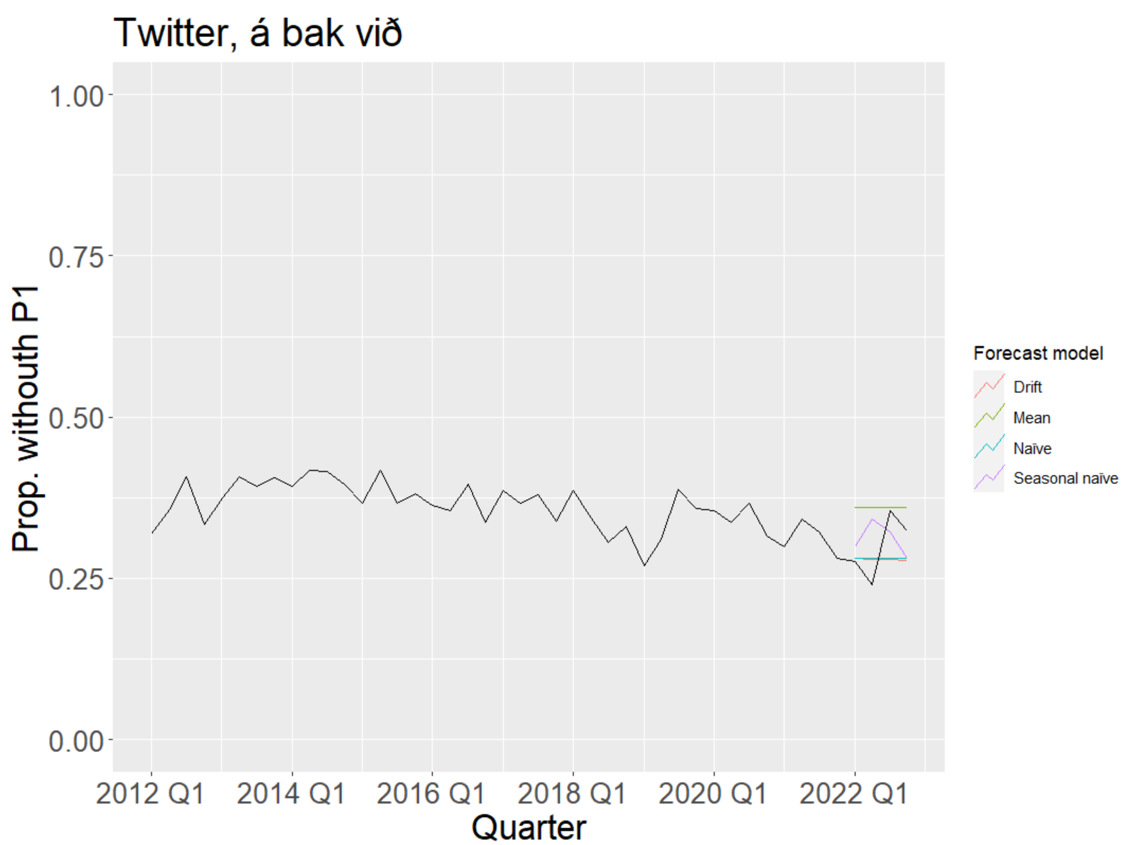


Figure 8.34. Four simple models, a Drift model, a Naïve model, a Seasonal naïve model and a Mean model, were fitted to the training data of the Twitter time series containing information proportion of examples of *á bak við* lacking **P**$_1$. The model was used to forecast Q1 – Q4 2022 with the Naïve model performing the best.

Twitter, *á bak við*							
Model	source	type	RMSE	MAE	MAPE	MASE	RMSSE
Mean	Twitter	Train	0.0377	0.0312	8.99	0.925	0.867
Naïve	Twitter	Train	0.0381	0.0334	9.51	0.992	0.878
Seasonal	Twitter	Train	0.0434	0.0337	9.83	1	1
Drift	Twitter	Train	0.0381	0.0333	9.46	0.988	0.878
Drift	Twitter	Test	0.0487	0.0416	13.5	1.23	1.12
Mean	Twitter	Test	0.0754	0.0612	23.2	1.81	1.74
Naïve	Twitter	Test	0.0472	0.0406	13.3	1.2	1.09
Seasonal naïve	Twitter	Test	0.0587	0.05	18.3	1.48	1.35

Table 8.16. Report on the fit of the simple models to the training data and the accuracy of the models given how they did in predicting Q1 – Q4 2022.

In addition to the four simple forecasting models mentioned above, an ETS and an ARIMA model were also fitted to the training data. The type of ETS and ARIMA model along with initial states was determined using the ETS() and ARIMA() functions from the fable package. While the ETS() function selects a model by minimizing the AICc (Hyndman & Athanasopoulos 2021:254), the ARIMA() function relies on the Hyndman-Khandakar algorithm for automatic ARIMA modeling (Hyndman & Athanasopoulos 2021:286).

An automatic selection of an ETS model suggested simple exponential smoothing with additive eros, or ETS(A,N,N). The smoothing parameters were alpha = 0.3374196 the initial state l = 0.3559784. The residuals are within acceptable limits (mean = -0.003235794), are normally distributed and can plausibly be regarded as white noise (Ljung-Box statistics 3.32, p-value 0.973).

An automatic selection of ARIMA resulted in ARIMA(0,1,1) with ma1 = -0.6397 (s.e. 0.1426). The residuals have desirable characteristics (mean = -0.001545076), are normally distributed and can be considered white noise (Ljung-Box statistics 2.82, p-value = 0.985).

Point forecasts for the ETS(A,N,N) and ARIMA(0,1,1) with prediction intervals are shown in Figure 8.35. Both models produce a very similar "flat" forecast with the ARIMA(0,1,1) predicting proportion of examples without **P1** being ca. 31.1% for the whole test period (Q1 2022 – Q4 2022) and the ETS(A,N,N) 31.2%. Observed values lie between ca. 23.9% and 35.5%, with the lowest value falling outside the 95% prediction interval of both the models. The accuracy of the fit of the two models and the accuracy of the point forecasts are summarized in Table 8.17. Note that the accuracy of the point predictions is almost identical for the two models, with the ARIMA(0,1,1) returning marginally better results when taking into account the RMSE (Root Mean Square Error).

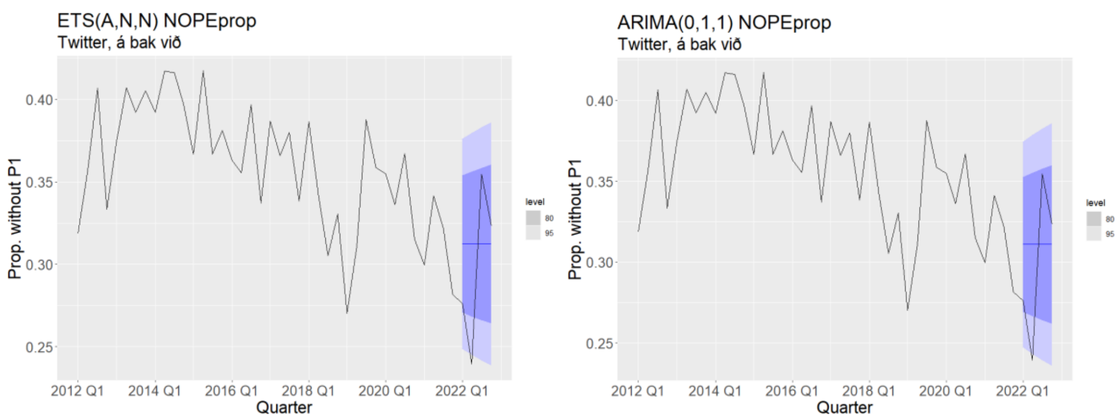


Figure 8.35. ETS(A,N,N) and ARIMA(0,1,1) were fitted to the training data and used to predict the test data (Q1 – Q4 2022).

Twitter, *á bak við*								
Model	source	type	RMSE	MAE	MAPE	MASE	RMSSE
ETS(A,N,N) NOPEprop	Twitter	Training	0.0317	0.0267	7.67	0.791	0.73
ARIMA(0,1,1) NOPEprop	Twitter	Training	0.0317	0.0262	7.49	0.777	0.73
ETS(A,N,N) NOPEprop	Twitter	Test	0.0462	0.0406	14.7	1.2	1.06
ARIMA(0,1,1) NOPEprop	Twitter	Test	0.0458	0.0406	14.7	1.2	1.06

Table 8.17. Accuracy of the fit of the ETS(A,N,N) and ARIMA(0,1,1) models and the accuracy of the point forecast for the test data.

Neither the ETS(A,N,N) nor the ARIMA(0,1,1) model appear particularly good at predicting the test data. Nevertheless, an evaluation of the forecasts distribution of all the methods discussed above using quantile scores, Winkler scores, Continuously Ranked Probability Scores and Scale-free comparison using skill scores (cf. Figure 8.36), suggests the ARIMA(0,1,1) performs 19% better than the Seasonal naïve model and the ETS(A,N,N) ca. 18% better when skill scores (based on crps) are considered. The best forecast for the test period, taking into account both point forecast and forecast distribution, seems to come from the Naïve model which assumes all future values are equal to the last observed value of the series. Keep in mind, however, that evaluations are based on a relatively small test data (4 observations). For a comparison of the accuracy of the forecast distribution of various models see Table 8.18.

Twitter, *á bak við*						
Model	source	type	qs probs = 0.1	winkler level 80	crps	skill
Drift	Twitter	Test	0.0194	0.155	0.0272	0.21
Mean	Twitter	Test	0.05	0.361	0.0478	-0.391
Naïve	Twitter	Test	0.0184	0.15	0.0265	0.23
Seasonal naïve	Twitter	Test	0.0318	0.227	0.0344	0
ETS(A,N,N) NOPEprop	Twitter	Test	0.0207	0.163	0.0281	0.182
ARIMA(0,1,1) NOPEprop	Twitter	Test	0.0202	0.159	0.0279	0.19

Table 8.18. A comparison of the forecast distribution of the simple models and that of ETS(A,N,N) and ARIMA(0,1,1). Note that the best model in this case is the Naïve model which is considered 23% better than the Seasonal naïve model.

Seeing that the Naïve model received the best evaluations for point forecasts and forecast distribution, this model was fitted to the whole dataset and used to generate a 15-step ahead forecast, i.e., from Q1 2022 to Q3 2025 (Figure 8.36).[87] Out of a desire to also try a more complex model, an ARIMA(0,1,1) was furthermore fitted to the whole dataset and used to generate a 15-step ahead forecast. The point forecasts of both models assume variation around a constant mean. The Naïve suggests a mean of 32.3% for all future periods and the ARIMA(0,1,1) 31.1%. The forecast intervals for the ARIMA(0,1,1) model are much narrower than of the Naïve model, cf. Figure 8.36.

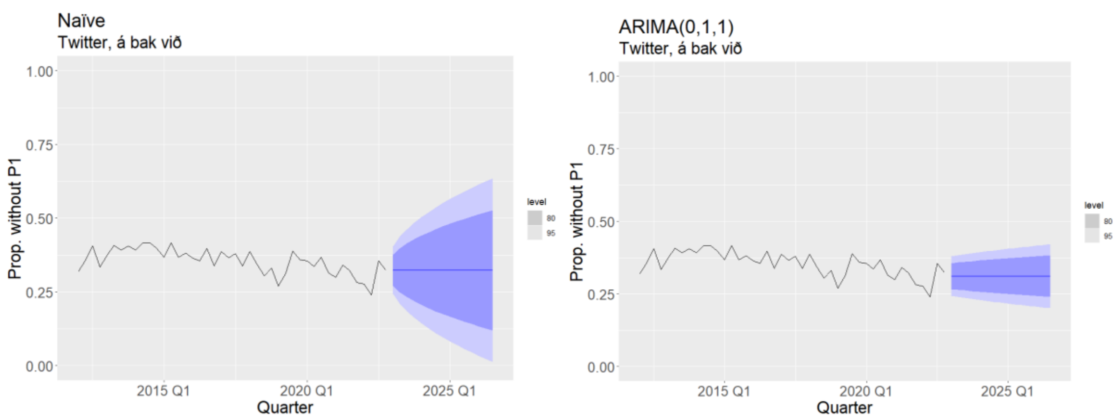


Figure 8.36. Two models, a Naïve model and ARIMA(0,1,1) fitted to the whole series and used to produce a 15-step ahead forecast for the period Q1 2023 – Q3 2026.

As noted above, an STL decomposition of the time series suggested that the remainder component was relatively large. Assuming the remainder component reflects non-meaningful “noise”, it is possible to generate a forecast relying on the trend component only. This has been done in Figure 8.37 where an ARIMA(0,1,1) model was fitted to the trend of the whole time series (after the seasonal and remainder components had been removed) and used to produce a forecast 20 steps into the future or until Q4 2027. The point forecast suggests that the proportion of examples of *á bak við* lacking $\mathbf{P_1}$ will be

[87] Forecasts for other series discussed in Chapter 8 and Chapter 9 typically included 20-step ahead predictions.

around 31.2% in the coming years and will not go below 25% or above 35%. These values are almost identical to a forecast relying on the raw time series; the forecast intervals are, however, slightly narrower.

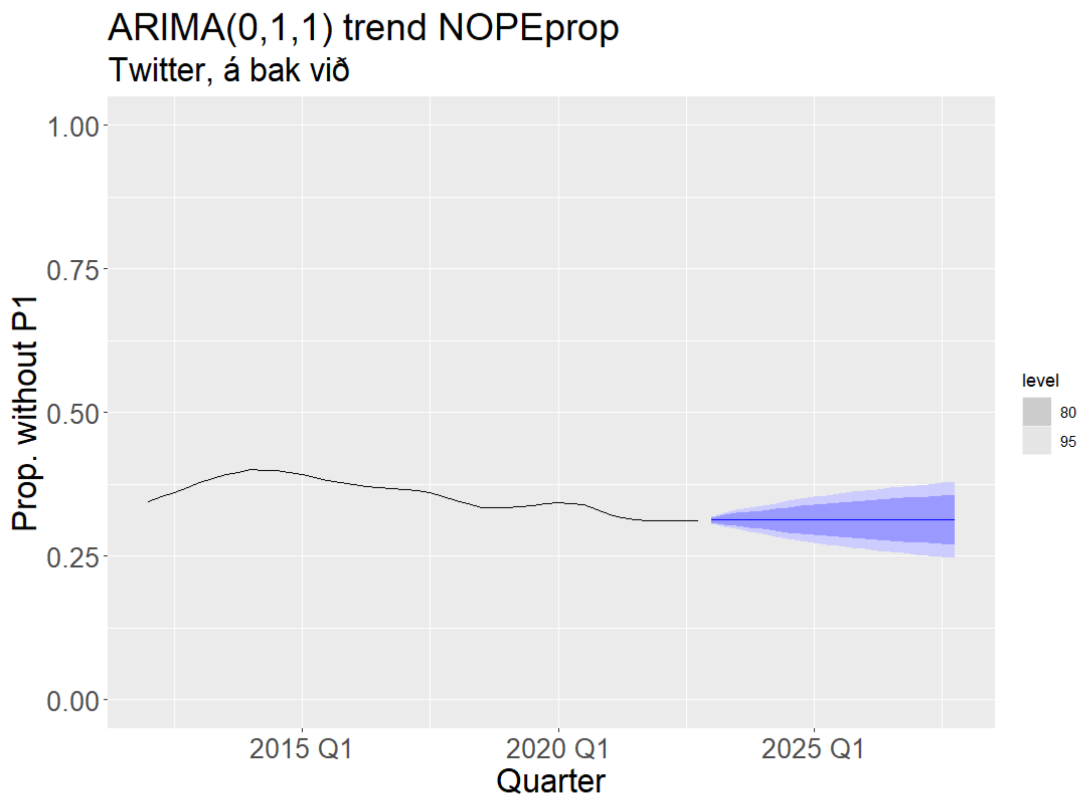


Figure 8.37. An ARIMA(0,1,1) model fitted to the trend of the whole series and used to predict the proportion of oblique first-person subjects from Q1 2023 to Q4 2027. The forecast suggests that the proportion of examples of *á bak við* lacking $\mathbf{P_1}$ will be around 31.2% for the foreseeable future.

The general picture that emerges from studying the proportion of examples without $\mathbf{P_1}$ in the complex preposition *á bak við* in data from Twitter, is that the proportion of examples without $\mathbf{P_1}$ stays relatively stable over time. In Q1 2012 the proportion is ca. 32%, and in Q4 2022 it is 32.3%. In addition to four simple models (Naïve, Seasonal naïve, Drift and Mean model), an ETS(A,N,N) and ARIMA(0,1,1) were fitted to the training data. Both of these resulted in worse point forecasts and forecast distributions for the training period than the simple models. Note however, that the values for ETS(A,N,N), ARIMA(0,1,1) and the Naïve model gave similar results.

The Naïve model and ARIMA(0,1,1) were fitted to the whole time series and used to produce 15-step ahead forecasts. These give rise to the following expectations: Both models give a very similar "flat" forecast, with predictions slightly above 31%, although the forecast distribution is slightly different. The ARIMA(0,1,1) has narrower forecast intervals. When applied to the trend of the series, as opposed to the raw series, the ARIMA(0,1,1) model suggests that the proportion of examples of á bak við lacking $\mathbf{P_1}$ will fall somewhere within the range of 25% to 35%. Expectations generated from the forecasts in this section are discussed further in Section 8.6.

8.5.6 Interim summary

Summarizing the time series and the forecasts from Sections 8.5.1 – 8.5.5, it can be noted that the two complex prepositions *við hliðina á* and *á bak við* behave differently when it comes to dropping the initial $\mathbf{P_1}$ (*við* or *á*), with $\mathbf{P_1}$ generally dropped less in *við hliðina á* than *á bak við*. Data from IGC and Twitter furthermore show slightly different proportions of dropping $\mathbf{P_1}$.

In the case of *við hliðina á*, data from IGC show that the initial $\mathbf{P_1}$ *við* is dropped less than 12% of the time, with the proportion going down in 2020 (2.6%) and 2021 (0%). When fitting a Drift model to the series, the model picks up on the lowered proportion in recent years and suggests that future observations will be close to zero. Data from Twitter shows a slightly higher proportion of examples without $\mathbf{P_1}$ or between 14% and 26.4%.

The Twitter series appears to be relatively stable over time. Using a Mean model, the expectation is that examples of *við hliðina á* in the future will lack $\mathbf{P_1}$ around 12.4% of the time. Using an ARIMA model on the trend of the Twitter series, the expectation is that the proportion of examples without $\mathbf{P_1}$ might increase slightly in the future, going from 14.7% in Q1 2023 to ca. 17.7% in Q4 2027.

In the case of *á bak við*, data from IGC shows that the initial $\mathbf{P_1}$ *á* is dropped around 34% of the time in 2003, and around 15.7% in 2021. Over the whole period there is a consistent negative trend. An ETS(A,A,N) model fitted to the series picks up on the downward trend and predicts a continued decrease in dropping of $\mathbf{P_1}$, from 15.45% (2022) to 7.1% (2041). Similarly, an ARIMA(0,1,0) with drift also suggests a continued decrease in dropping of $\mathbf{P_1}$ with the proportion being around 15% in 2022 and 6.8% in 2041. Data from Twitter shows both a higher proportion of examples without $\mathbf{P_1}$ and more stability over time. A Naïve model turned out to have the best prediction for the test period and fitting this type of model to the whole series, the proportion of examples lacking $\mathbf{P_1}$ is expected to be around 32% for all future periods. An ARIMA(0,1,1) model provides narrower forecast intervals and also converges on a fairly flat forecast with future observations being hypothesized to lack $\mathbf{P_1}$ about 31.2% of the time.

8.6 Summary and discussion

The complex prepositions *á bak við* ‘behind’ and *við hliðina á* ‘next to’ first appeared in the 18th (*á bak við*) and 19th (*við hliðina á*) centuries in Icelandic. They replaced earlier ways of expressing location behind and next to something (see Section 8.3.4). In Modern Icelandic the complex prepositions are sometimes encountered in a compact (or simplified) form, namely *bakvið* 'behind' and *hliðiná* 'next to', written in one word and lacking the initial **P_1**. The simplified forms have been claimed to be innovative (Friðjónsson 2004, 2007; Rögnvaldsson 2021; see also discussion in sections above), showing the direction of change stated in (8.83).

(8.83)	a.	*á bak við* $(+N)_{ACC}$	>	*bakvið* $(+N)_{ACC}$
	b.	*við hliðina á* $(+N)_{DAT}$	>	*hliðiná* $(+N)_{DAT}$

Considering the full forms *á bak við* and *við hliðina á*, these appear to consist of a sequence of a preposition, a noun and a second preposition (**$P_1N_1P_2$**) where the second preposition assigns either accusative or dative case to its complement, depending on the preposition. It was argued here that these have already been grammaticalized in the modern language and are understood as a single chunk, even though writing may still indicate three separate words. Evidence of grammaticalization was presented in the form of loss of connection to relevant lexical nouns, innovative written variants and generally usage of the prepositions (Section 8.3.2), as well as in the form of syntactic inflexibility of the sequences (Section 8.3.3). It was suggested that the absence of **P_1** should be understood in terms of phonological erosion, and this is the only part of the change that is still ongoing. Speakers can be divided into two groups w.r.t. to the ongoing change, namely i) those who always include **P_1** and ii) those who use forms with and without **P_1** in a seemingly interchangeable way. Needless to say, these changes deserve more thorough documentation than what has been provided here, for instance w.r.t. the phonological signal itself as well as the age of individuals dropping $\mathbf{P_1}$ versus those who are unable to drop it. That is a task for future research.

A change from a multi-word string into a single preposition is quite common in the world's languages and may even take place more than once at different times within the same language. Given the directionality of the grammaticalization, which consists of semantic change, downgrading analysis (or reanalysis) and phonological erosion, the expectation towards *á bak við* and v*ið hliðina á* is that grammaticalized $\mathbf{N_1P_2}$ variants without $\mathbf{P_1}$ should eventually win out.

To document and forecast changes in *á bak við* and *við hliðina á*, four time series were constructed, two for each complex preposition. These took into account the proportion of examples lacking $\mathbf{P_1}$ and were based on material from the Icelandic Gigaword Corpus (informal and semi-formal material), covering the period 2003 – 2021, and on material from Twitter, covering the period 2012 – 2022. The IGC data was projected into a yearly time series with 19 observations and the Twitter data into a quarterly series with 44 observations. A number of decisions were made when constructing the time series, such as what type of data to use, which examples to include in the data set, how to annotate the data, what the frequency of the series should be, how to split each series into a training and test set etc.

Two general comments can be made about the four time series. First, that $\mathbf{P_1}$ is more frequently dropped in the complex preposition *á bak við* than *við hliðina á*. The mean of the IGC series for *við hliðina á* over the 19-year period is around 8% while the mean for *á bak við* covering the same period is 23.6%. The mean of the Twitter series for *við hliðina á* is 12.4% while the mean for *á bak við* is 35%. This suggests that á *bak við* might generally be further along in the grammaticalization (including phonological erosion) and fits with this complex preposition being older than *við hliðina á* and showing signs of changes earlier. Second, data from Twitter generally shows a higher proportion of dropping $\mathbf{P_1}$ than informal and semi-formal data from IGC. This is already visible when comparing the mean of the four series (for *við hliðina á*: Twitter 12.4% and IGC 8%, for *á bak við*: Twitter 35% and IGC 23.6%). The difference between IGC and Twitter might be traced to the fact that part of the IGC material contained semi-formal data which includes written texts such as online newspaper articles that are intended to be read by other individuals but may not necessarily go through rigorous copy editing or proofreading. As noted in Section 8.4.3, the semi-formal material appeared to have a similar proportion of forms lacking $\mathbf{P_1}$ as formal material while informal material had a higher proportion of examples without $\mathbf{P_1}$. The reasons for including semi-formal material in the IGC time series was that it provided more material to work with (more examples for each time period) and the series can be compared to those in Chapter 9 which focus on oblique subjects with the predicate *hlakka til* 'look forward to' in Modern Icelandic.

Interestingly, the IGC time series do not all fully conform to expectations towards the trajectory of the change in *á bak við* and við *hliðina á.* While the series for *við hliðina á* is relatively stable over time, it shows a sudden decrease in the omission of $\mathbf{P_1}$ in 2020 (2.6%) and 2021 (0%). The last observation might be due chance since it is only based on 27 examples with no instance of a dropped $\mathbf{P_1}$. However, the observation for 2018 is only based on 44 examples but still in line with the rest of the series, showing the proportion of no $\mathbf{P_1}$ to be ca. 6.4%. Additionally, the observation for 2020 is based on a total of 3,319 examples and still shows a drop to 2.6%. So, perhaps the figures really suggest a decrease in dropping of $\mathbf{P_1}$ and an increase in the use of the full $\mathbf{P_1N_1P_2}$ sequence. This may be compared to the IGC series for *á bak við*, which shows a consistent decrease in dropping $\mathbf{P_1}$ over time. In 2003 the proportion of examples lacking the initial $\mathbf{P_1}$ is around 34%, but in 2021 it is down to ca 15.7%. Fitting an ETS(A,A,N) and ARIMA(0,1,0) to the series for *á bak við* (these were the models that performed the best on predicting the test data) suggests a continued decrease in the use of $\mathbf{P_1}$ in the near future. The proportion might even go down to ca. 6–7% in 2041. As for the IGC series for *við hliðina á*, no model (neither a simple one nor a more complex one) did particularly well at predicting the test period. The model that performed the "best" was the Drift model which is based on averages of changes in the whole series. Poor predictions for the test period are very likely due to the 2020 and 2021 observations that show a sudden decrease in dropping of $\mathbf{P_1}$. One may choose to ignore these and predict that all future values will lie somewhere around the mean of the whole series, i.e., that the proportion of examples without $\mathbf{P_1}$ will be ca. 8%. Another option is to go with the Drift model, in which case future values will fall somewhere close to 0%.

Turning to data from Twitter, the series for *við hliðna á* appears relatively stable over time, with a mean around 12.4%. The model that best predicted unobserved values in the test was the Mean model. Using this model on the whole series, future values are expected to lie close to the 12.4% mean. If an STL decomposition (trend window = 13) is used on the series to look at the trend component, a mild cyclic pattern appears, with the proportion of examples without $\mathbf{P_1}$ decreasing and increasing over time. Using an ARIMA model on the trend, suggests a slight increase in the proportion of examples lacking $\mathbf{P_1}$, going from 14.7% in Q1 2023 to ca. 17.7% in Q4 2027. Whether one chooses to follow predictions based on the mean of the raw series, or predictions based on the trend component of the series, the forecast is reasonable given expectations born out of ideas of grammaticalization. Over an 11-year period very little change is observed. The models see that and predict either no change or a slight increase in dropping the initial $\mathbf{P_1}$ element. The results for the Twitter data for *á bak við* show a similar pattern. This series is also relatively stable although a minor change in level may be observed (from 2012 to 2018 the mean of the series is just above 37%, but from 2019 to 2022 it is 32.5%). Of the four simple models fitted to the data (Naïve, Seasonal naïve, Drift and Mean), the Naïve model produced the most accurate point forecast for the test period. An ARIMA(0,1,1) produced very similar results. If these are used to predict future periods, observations from Q1 2023 to Q3 2026 are expected to be around 31.1%–32.3%. A similar result is obtained when the trend component from the series is extracted (STL decomposition, trend window = 13), resulting in future value being hypothesized to be around 31.2% and not go below 25% or higher than 35%.

An important thing to note about the four time series created and used for the present study, is that they cover a relatively short time period in the context of language change, i.e., between 11 (Twitter) and 19 (IGC) years. It is unclear how much change should be expected within such a short period. Stability would not be unsurprising, but changes might also occur. In the case of the complex prepositions *á bak við* and *við hliðina á*, known directionality in grammaticalization gives rise to the expectation that dropping of $\mathbf{P_1}$ elements should increase, i.e., if loss of $\mathbf{P_1}$ is attributed to phonological erosion. Predictions that suggest less dropping of $\mathbf{P_1}$ are not in line with such expectations and need to be explained. Both of the IGC time series show such unexpected directions, especially the series focusing on the complex preposition *á bak við* which indicates a consistent decrease in dropping **P_1**. One reason might be that the change is heading in the opposite direction of what is expected, namely restoring **P_1** instead of omitting it. This could be due to the initial hypothesis being wrong, namely that loss of **P_1** is not attributable to phonological erosion or grammaticalization in a wider sense. Another explanation would be that it relates to some external factor. As noted earlier (Section 8.2.1), Modern Icelandic has a peculiar tendency to use multi word prepositions in certain contexts where other Germanic languages use single word prepositions (Berthele et al. 2015:88; see also Friðjónsson 2005:26). Perhaps this tendency is causing the complex prepositions *á bak við* and *við hliðina á* to retain and restore the initial **P_1**. If this is the case, then we might be witnessing changes that relate to the typological aspect of Icelandic. Before being carried away with this explanation, it is important to remember that neither of the Twitter time series exhibited a similar decrease in forms lacking the initial **P_1** is. In fact, the Twitter data pointed mostly towards a stable future. This might indicate that the changes observed in the IGC series for *á bak við* are not due to actual decrease in the dropping of **P_1**. Rather, they might reflect some aspect of the nature of the IGC data, for instance that it was partially based on semi-formal language, so normative pressure and writing standards may have had an effect. If that is the case, the questio remains as to what the gap is between written material and the actual situation in the language community. Hopefully these are all factors that might be quantifiable in the future, allowing for the incorporation of them into forecasting models. For now, we are just left wondering what the gap is between convenient written E-language data and the actual situation in the language community.

To conclude, the complex prepositions discussed here, *á bak við* and *við hliðina á*, certainly deserve more study and more thorough documentation than what has been presented here. The creation of time series based on data from recent years allowed for studying some aspects of these changes. When applying forecasting models to the time series, the models pick up on various patterns and project them into the future. When stability is observed, stability is predicted. When trends are observed, a continuation of those trends is predicted. These projections give rise to various expectations regarding the trajectory of the changes under discussion. Whether these expectations will be met or not remains to be seen.

9. Changes in subject case marking in Icelandic

9.1 Introduction

In addition to regular nominative subjects, Icelandic has been shown to have subjects in oblique cases with certain predicates (e.g., Andrews 1976, 1982; for an overview see Thráinsson 2007:146–150). While some of these are old in the language, and may be of Scandinavian, Germanic or Proto-Indo European origin (on oblique subjects representing an archaic layer, see especially Eythórsson and Barðdal 2005; Barðdal & Eythórsson 2003, 2009, 2012; on oblique-subjects within other Scandinavian languages see Jónsson & Eythórsson 2011:234; Falk 1997:54), others have appeared within the recent history of Icelandic (e.g., Sigurðardóttir & Eythórsson 2022). The predicate *hlakka til* ‘look forward to’ falls in the latter category, originally appearing with a nominative subject as in (9.1). In Modern Icelandic, however, it is often encountered with either an accusative or dative subject, see (9.2), in addition to the original nominative.

(9.1) ***Ég*** *hlakka* *til* *sumarsins*
I-NOM look.forward to the.summer
‘I look forward to the summer.’

(9.2) a. ***Mig*** *hlakkar* *til* *sumarsins*
me-ACC look.forward to the.summer
‘I look forward to the summer.’

b. ***Mér*** *hlakkar* *til* *sumarsins*

me-DAT look-forward to the.summer

'I look forward to the summer.'

The change from (9.1) to (9.2) has been well documented, both with respect to its origin (Sigurðardóttir & Eythórsson 2022) and its propagation through the language community (e.g., Svavarsdóttir 1982; Jónsson & Eythórsson 2003; Nowenstein 2023). The change is properly termed oblique case substitution (Sigurðardóttir & Eythórsson 2022), although sometimes the label 'Dative Sickness' has been used (e.g., Jónsson 1997–1998:29; 2003:155; Thráinsson, Eythórsson, Svavarsdóttir & Blöndal 2015:33; Óladóttir 2017:236; Nowenstein 2023:49).

The change from (9.1) to (9.2) goes against the main trend in changes in case marking in Modern Icelandic, namely that from oblique to the nominative. However, it is understandable in light of the fact that many experiencer predicates occur with an oblique subject, for instance *vanta* 'need', *dreyma*, 'dreem', *blöskra* 'be shocked, be horrified' and *leiðast* 'be bored' to name only a few.[88] Documentations of the change shows that the proportion of speakers that use oblique subjects has been increasing over time, even though normative pressure and intense prescriptivism in schools demands nominative (see discussion in Óladóttir 2017:120–125, 236) . Studies also show that subjects that are first person pronouns tend to be treated differently from other types of subjects. This is likely due to prescriptive rules and school teaching targeting the first person (cf. Óladóttir 2017), although it has also been suggested that children may be acquiring a split case marking

[88] Both *vanta* and *dreyma* traditionally take an accusative subject. However, some speakers now use dative with *vanta* and some speakers use nominative with *dreyma*.

with certain predicates, e.g., nominative for the first person subjects and oblique for other types of subjects (Nowenstein 2014a,b, 2023).

The current study presents a novel type of data for studying the change from nominative to oblique with *hlakka til*. The data comes in the form of a regular time series where each observation is based on examples attested in written sources and contains information on the proportion of oblique subjects at the relevant point in time. Examples were obtained from two sources, namely the Icelandic Gigaword Corpus (IGC, rmh=2019, rmh=2022 cf. Steingrímsson et al. 2018) and Twitter (https://twitter.com/). Time series constructed based on the former source covered the period 2000–2021 although only years 2003–2021 were used for time series analysis and forecasting. Time series based on material from Twitter contained examples from Q1 2009 to Q4 2022 although only Q1 2012 to Q4 2022 were used when making forecasts. The reason for not using the whole series was that early observations are based on few examples and introduce considerable noise into the series (see Section 9.5.1).

The novelty of the current study does not only lie in documentation through regular time series, but also through containing predictions about the future regarding oblique case marking of subjects of *hlakka til*. In some instances, the predictions reach the year 2040. Each time series that was taken into consideration was split into a training and test set. Initially, three to four simple models were fitted to the training set (a Naïve model, a Seasonal naïve model, a Mean model and a Drift model) and used to predict observations in the test set. Next, slightly more complex models were fitted to the training set and used to predict attested observations. The more complex models involved methods such as exponential smoothing (ETS models) and autoregression (ARIMA models). In some instances, the simple models turn out to generate more accurate point predictions and forecast distribution for the test set than the more complex models. Forecasts for future periods were generated using an ETS, an ARIMA, or one of the more simple models. Generally, two or more models were used to predict future periods, in which case differences and similarities between predictions are discussed.

The main results from the present study are as follows. First, by looking at the regular time series it becomes apparent that there is considerable gap between the proportion of subject case marking with *hlakka til* in written material and what has previously been documented through surveys (e.g., Svavarsdóttir 1982, Eythórsson & Jónsson 2003). While surveys have shown up to 80% of children tested between 1980 and 2019 using an oblique subject when the subject is not a first person pronoun, the proportion of oblique in written material is generally around or below the 50% level (IGC) or anywhere between 20% and 80% (Twitter). First person subjects show a different pattern, with the proportion of oblique decreasing over time. In IGC material it goes from 32% in 2003 to little less than 6% in 2021. On Twitter, the proportion is around 18% in Q1 2012 and goes down to ca. 8.5% in Q4 2022. A survey documenting the use of oblique with first person subjects suggests almost 50% oblique (Svavarsdóttir 1982).

Second, extrapolating documented patterns into the future suggests that the proportion of oblique will continue to go down when the subject is a first person pronoun. A bold prediction shows a future (IGC data) with almost no oblique first person subjects in 2040. The future for other types of subjects with *hlakka til* looks slightly different. Forecast intervals in this case are quite wide, but point predictions suggest a mean of 44% of oblique subjects for future periods into 2039 (IGC data) or that the observations will fall somewhere between 40%–70% (Twitter data). It should be noted that observations for proportion of oblique non-first person subjects are based on relatively few examples, especially in case of data from Twitter. Time series documentations of oblique subjects with *hlakka til* as well as forecasts are discussed more thoroughly in relevant sections of the chapter.

The structure of the chapter is as follows. Section 9.2 provides a general overview of subject case marking in Icelandic, the properties of oblique subjects (9.2.1), and direction of changes in case marking (Section 9.2.2). Section 9.3 takes a closer look at subject case marking with *hlakka til* 'look forward to' in Icelandic, discussing factors such as variation and ambiguity in case marking (Section 9.3.1) , previous documentation of changes in case marking with *hlakka til* (9.3.2), normative pressure which seems to mostly affect uses of first person subjects with the predicate (9.3.3), and expectations towards the continued propagation of the change and what the future might look like (9.3.4). Section 9.4 is dedicated to describing the data used for the forecasting, how it was obtained, how annotations were made, any interesting observations and projections into time series. Four time series were taken into consideration, two for each data source. One series contains information on the proportion of oblique when the subject is a first person pronoun and the other information on the proportion of oblique when the subject was of a different type. Descriptions of the four time series are found in Section 9.5.1 and fitting of models and forecasts in 9.5.2–9.5.4. Results from the study are discussed in Section 9.6.

9.2 Subject case marking in Icelandic and directions of change

9.2.1 The case of subjects

Subject arguments in Icelandic can appear in any of the four cases: nominative, accusative, dative or genitive.[89] Of these, the nominative is by far the most common and most productive case. Newly coined predicates, borrowings, calques and neologisms typically take a nominative subject (Barðdal 2001 and later). Predicates appearing with other subject cases are fewer. An estimated distribution of subject case in Icelandic, based on a statistical analysis of selected tests, suggests that roughly 94% of subjects in Modern Icelandic occur in nominative case, about 4% in dative, around 1% in the accusative and less than 1% in genitive (Barðdal 2001:180; cf. also Thráinsson 2007:156).

Oblique subjects exhibit the same syntactic properties as nominative subjects except for agreement. Numerous subject tests have been proposed on the basis of the behavior of nominative subjects and these also work for oblique subjects. Tests for subjecthood include inversion with the finite verb, control infinitives (PRO-infinitives), conjunction reduction, raising to object (ECM) and raising to subject (see e.g., Zaenen et al. 1985; Sigurðsson 1989, 1997, 2002a; Rögnvaldsson 1997; Barðdal 2002; Barðdal & Thórhallur Eythórsson 2003a; for an overview see Thráinsson 2007:146–167). Examples of subject inversion with the finite verb are provided in (9.3)–(9.5), with (9.3) showing nominative subject (9.4) showing accusative and (9.5) dative.

[89] The content of Section 9.2 is largely based on joint work with Thórhallur Eythórsson on the emergence of oblique subjects in Icelandic (see Sigurðardóttir & Eythórsson 2022).

(9.3) a. ***Arnaldur*** *hafði* *lesið* *bókina* *í fyrra*.

Arnaldur-NOM had read the.book last.year

‘Arnaldur$_{\text{NOM}}$ had read the book last year.’

b. Í fyrra hafði **Arnaldur** lesið bókina.

last.year had Arnaldur-NOM read the.book

‘Last year, Arnaldur$_{\text{NOM}}$ had read the book.’

(9.4) a. ***Arnald*** *hafði* *vantaði* *upplýsingar* *í fyrra*.

Arnald-ACC had needed information last.year

‘Arnaldur$_{\text{ACC}}$ had needed information last year.’

b. *Í fyrra* *hafði* ***Arnald*** *vantað* *upplýsingar*.

last.year had Arnaldur-ACC needed information

‘Last year, Arnaldur$_{\text{ACC}}$ had needed informaiton.’

(based on Thráinsson 2007:162, ex. (4.36))

(9.5) a. ***Arnaldi*** *hafði* *hlakkað* *til* *ferðalagsins* *í fyrra*.

Arnaldur-DAT had look.forward to the.trip last.year

‘Arnaldur$_{\text{DAT}}$ had looked forward to the travel last year.’

b. *Í fyrra* *hafði* ***Arnaldi*** *hlakkað* *til* *ferðalagsins*.

last.year had Arnaldur-DAT look.forward to the.trip

'Last year, Arnaldur$_{DAT}$ had looked forward to the travel.'

A further example, showing the occurrence of nominative, accusative and dative subjects in subject-to-object raising are provided in (9.6) and (9.7). Note that the dative in (9.7b) keeps its subject case marking from *hlakka til*, while both the nominative in (9.6) and the accusative in (9.7a) appear in the accusative.

(5.6) *Ég* *tel* ***Arnald*** *hafa* *lesið* *bókina*.

I believe Arnaldur-ACC have read the.book

'I believe Arnaldur$_{ACC}$ has read the book.'

(9.7) a. *Ég* *tel* ***Arnald*** *hafa* *vantað upplýsingar*.

I believe Arnaldur-ACC have needed information

'I believe Arnaldur$_{ACC}$ has needed information.'

b. *Ég* *tel* ***Arnaldi*** *hafa* *hlakkað* *til* *ferðalagsins*

I believe Arnaldur-DAT have look.forward to the.journey

'I believe that Arnaldur$_{DAT}$ has looked forward to the journey.'

Despite showing the same syntactic properties as nominative subjects, oblique subjects differ from nominative ones in that the finite verb does not agree in person or number with them. Rather, the finite predicate is always in third person singular when the subject is in an oblique case (9.9).

(9.8)	*Þið*	*Arnaldur*	***hafið***	*lesið*	*bókina*.
	you-PL.NOM	Arnaldur-NOM	have-2P.PL	read	the.book

‘You and Arnaldur have read the book.’

(9.9)	*Ykkur*	*Arnaldi*	***hefur***	*hlakkað*	*til*	*ferðalagsins*.
	you-PL.DAT	Arnaldur-DAT	have-3P.SG	look.forward	to	the.trip

‘You and Arnaldur have looked forward to the trip.’

A further difference can be argued to be tied to the lexical semantics of the subjects. Nominatives are unspecified for lexical semantics and can have any thematic role. They can be agents, experiencers, themes etc. Oblique subjects are never agents. Dative subjects typically denote experiencers and goals (including beneficiaries and recipients) and accusative subjects are usually experiences and themes (and patients).[90] Genitive subjects, which occur with about ten predicates, have no clear link to specific semantic roles aside from not being agents (see Jónsson 1997–98, 2003; Barðdal & Eythórsson 2009 among others).

[90] Experiencers are here taken to be a broad category consisting of subcategories such as verbs of emotion, e.g., *fýsa* ‘want’, *langa* ‘want’, and verbs of bodily functions such as *verkja* ‘feel pain’ and *hrylla við* ‘be disgusted by’ (see e.g., Jónsson 1997–98, Barðdal 2001).

9.2.2 Directions of change in subject case marking

Changes in subject case marking within Icelandic are of several types. The most common is Nominative Substitution (Ice. *nefnifallshneigð*) which involves the replacement of oblique case by nominative. The change typically affects theme subjects and is unsurprising given that nominative is the most common subject case in Icelandic (see 9.2.1). An example of Nominative Substitution is provided in (9.10) where (9.10a) represents an older variant and (9.10b) an innovation.

(9.10)	a.	***Bátnum***	*hvolfdi*.
		the.boat-DAT	capsized
	b.	***Báturinn***	*hvolfdi*.
		the.boat-NOM	capsized
		‘The boat capsized.’	

A second type of change is Dative Substitution (Ice. *þágufallshneigð*), sometimes referred to as “Dative Sickness” (Ice. *þágufallssýki*). Here, a dative (9.11b) replaces an earlier accusative (9.11a). Importantly, Dative Substitution only affects subjects denoting experiencers. The term has sometimes been used to refer to change from nominative to dative (e.g., Jónsson 1997–1998:29, 2003:155; Thráinsson, Eythórsson, Svavarsdóttir & Blöndal 2015:33; Óladóttir 2017:236; Nowenstein 2023:49), but in the present work it is reserved for changes from accusative to dative.

(9.11) a.

Mig	*langar*	*í*	*nammi*.
me-ACC	wants	PREP	candy

b.

Mér	*langar*	*í*	*nammi*.
me-DAT	wants	PREP	candy

'I want candy.'

Genitive Avoidance (Ice. *eignarfallsflótti*) is a third type of change which involves the replacement of genitive with another case, typically dative as in (9.12b). Although Genitive Avoidance primarily affects objects, a few examples involving subjects have been reported (Jónsson 2017).

(9.12) a.

Þeirra	*bíður*	*erfitt*	*verkefni*…
them-GEN	awaits	difficult-N	task-NOM

b.

Þeim	*bíður*	*erfitt*	*verkefni*…
them-DAT	awaits	difficult-N	task-NOM

'A difficult task awaits them…'

(Online news article 2017, November 11th)

A fourth type of change, termed Oblique-Case Substitution (Ice. *aukafallshneigð*), involves a change in subject case marking whereby predicates originally taking a nominative subject start appearing with either an accusative or dative subject. Thus, at one point in the history of Icelandic the subject occurs in the nominative with the relevant predicate, and at a later point it occurs with the accusative or dative. Sometimes Oblique-Case Substitution has been treated together with Dative Substitution, described above. However, since Dative Substitution involves a change from one oblique to another oblique and the Oblique-Case Substitution a change from nominative to oblique, there are reasons to keep these separate.

Oblique-Case Substitution seemingly only affects a handful of experience predicates that do not have a transitive (causative) variant.[91] This includes at least the predicates *hlakka til* 'look forward to', *kvíða* (*fyrir*) 'be anxious about', *kenna í brjósti um* 'feel sorry for', *kenna til* 'feel pain', *finna til* 'feel pain', and *skjöplast* 'be mistaken' (e.g., Friðjónsson 1989:13; Halldórsson 1982; Svavarsdóttir 1982). Of these, the predicates *hlakka til*, *kvíða fyrir* and *finna til* have probably gained the most attention, both in scholarly literature and in the eye of the public.

The four changes discussed above, Nominative Substitution, Dative Substitution, Genitive Avoidance and Oblique-Case Substitution, can be summarized and visualized as in Figure 9.1 and Figure 9.2. The Former shows changes to and from nominative and the latter shows changes from one oblique case to another.

[91] The lack of transitive variant is relevant for Oblique-Case Substitution as it guarantees that the change is directly from nominative to oblique. There are instances of predicates that exist in both transitive and intransitive use where the intransitive (anticausative) structure originally had a nominative subject, but later showed up with an oblique subject. Interestingly, the case of the innovative oblique subject matches that of the object in a relevant transitive (causative) structure. For these reasons, the change might be argued to stem from the anticausativization strategy used for creating the intransitive (anticausative), rather than representing a straightforward change from nominative to oblique (see further in Sigurðardóttir & Eythórsson 2022).

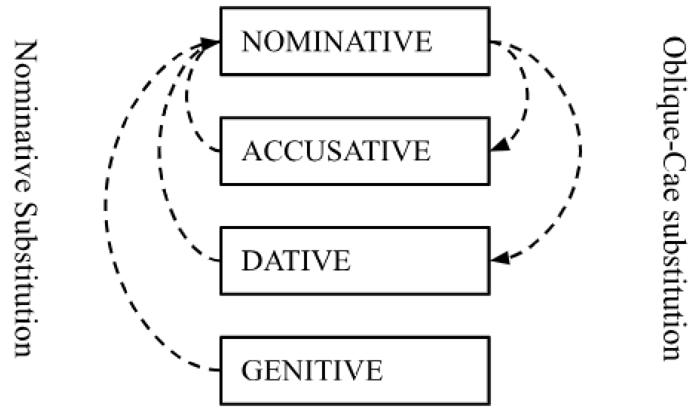


Figure 9.1. Changes in subject case marking in Icelandic to and from nominative. The predicate *hlakka til* ‘look forward to’ is subject to Oblique-Case Substitution.

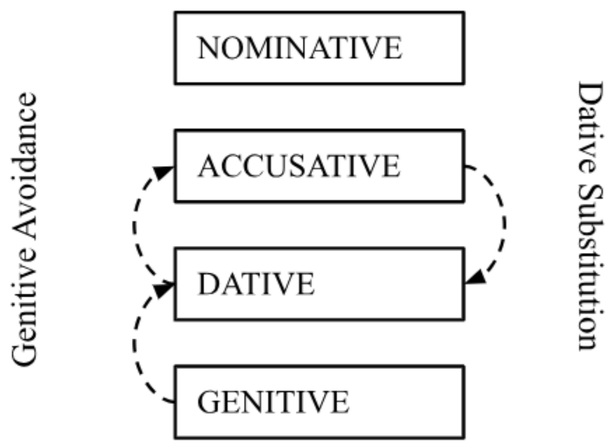


Figure 9.2. Changes in subject case marking in Icelandic where an oblique case is replaced by another oblique case. Note that a change from genitive to nominative is usually subsumed under Genitive Avoidance, but is here treated with Nominative Substitution as in Figure 9.1.

The changes in subject case marking discussed above and depicted in Figure 9.1 and Figure 9.2 do not appear to be of equal status. Some are more frequently attested and can be claimed to be more “understandable” than others. For instance, given that nominative is the subject case par excellence in Icelandic, it is not surprising that predicates previously taking oblique subjects might start appearing with a nominative one.

On the basis of evidence in Icelandic and Faroese, Eythórsson (2002, 2015a, 2015b; see also Eythórsson & Thráinsson 2017) has proposed a Case Directionality Hypothesis (CDH) which accounts for the majority of changes in subject case marking in Icelandic. The CDH, stated in (9.13), draws on the distinction between *structural case* and *lexical case*, of which the latter can be either *regular* (*thematic*) *lexical* or *idiosyncratic lexical* (see Yip, Maling & Jackendoff 1987, Jónsson 1997–98, 2003 and Eythórsson & Thráinsson 2017). The hypothesis predicts that lexical case can be replaced by structural case and idiosyncratic lexical case can be replaced by regular lexical case.

(9.13) Case Directionality Hypothesis (CDH)

a. Lexical case → structural case

b. Idiosyncratic lexical case → thematic lexical case

(from Eythórsson and Thráinsson 2017:60)

The division into structural and lexical case derives from the work of Yip, Maling & Jackendoff (1987), who propose that (surface) case assignment is either due to the grammatical function of the argument, where case is assigned in a certain position within the structure, or it that is lexically conditioned. If a case is lexically conditioned, it may be either predictable based on the thematic role of the argument (regular lexical case) or it may be idiosyncratic and unpredictable (Eythórsson & Thráinsson 2017:57–58).

Under the CDH, Nominative Substitution can be regarded as the result of structural pressure within the syntax. Case assignment that used to be lexically conditioned, i.e., dependent on the relevant lexical predicate, is now structurally conditioned. This makes Nominative Substitution different from changes where one oblique case is replaced by another oblique case. In these instances, the change must be motivated by the thematic role of the relevant argument and the expectation is typically that idiosyncratic lexical case will yield to regular (thematic) lexical case (Eythórsson 2002, Eythórsson & Thráinsson 2017:57–58).

In light of the CDH, changes in case marking from nominative to oblique are unexpected. The CDH essentially predicts that system-internal pressure should prevent the nominative case on subjects from changing. However, given that oblique subjects are often experiencers, one might view Oblique-Case Substitution as an attempt to regularize case marking with experiencer predicates in Icelandic. In fact, it has sometimes been referred to as analogical change (termed "morphosyntactic leveling") under the influence of predicates such as *langa* 'want' and *vanta* 'need' that were always attested with an oblique subject (Eythórsson 2002, 2015a, 2015b, Jónsson & Eythórsson 2005). If this is true, the question may be raised why other nominative experiencers are not typically replaced by oblique ones. Observe for instance (9.14) where an oblique subject is not attested.

(9.14) a. ***Ég*** *þjáist.*
I-NOM suffer
'I am suffering'

b. ***Ég*** *elska* *þig*.

I-NOM love you-ACC

'I love you.'

Partially answering the question, it may be noted the absence of oblique (accusative) subjects with intransitive *st*-predicates like *þjást* 'suffer' are not surprising given that the *st*-morpheme originally derives from a reflexive *sik, sig* 'self' (Ottósson 1992; Ottosson 2008). The only possible binder for a simple reflexive argument in a simple clause in Icelandic is nominative. However, this fact does not account for why the case of nominative subjects such as in (9.14a) is not prone to change to oblique once binding conditions no longer need to be met (Sigurðardóttir & Eythórsson to appear). As for predicates like *elska* 'love', these may involve different kinds of experiencers than predicates like *vanta* 'need', *langa* 'want' and *hlakka til* 'look forward to', causing them to retain nominative case marking. Jónsson (2003:137) has suggested that "Psych-verbs denoting strong positive feelings cannot have an oblique subject".

Leaving aside nominative experiencers and viewing Oblique-Case Substitution as an attempt to regularize case-marking with experiencer predicates (on this see also Nowenstein 2023), the expectation is that oblique case marking should eventually win out with experiencer predicates. This includes oblique case marking on subjects with *hlakka til* 'look forward to'.

9.3 Variation in subject case marking with *hlakka til* 'look forward to'

9.3.1 Variation and ambiguity in case marking in Modern Icelandic

The predicate *hlakka til* 'look forward to' can appear with either a nominative, accusative or dative subject in Modern Icelandic.[92] Of these, the nominative represents the original state of affairs (9.15) with accusative and dative subjects being innovative (9.16). As it appears, the uses of oblique are becoming more prominent (Section 9.3.2).

(9.15)		***Ég***	*hlakka*	*til*	*sumarsins.*
		I.NOM	look.forward.1P.SG	to	the.summer
		'I look forward to the summer.'			

(9.16)	a.	***Mig***	*hlakkar*	*til*	*sumarsins.*
		I-ACC	look.forward-3P.SG	to	the.summer
	b.	***Mér***	*hlakkar*	*til*	*sumarsins.*
		I-DAT	look.forward-3P.SG	to	the.summer
		'I look forward to the summer.'			

Although it is usually straightforward to determine whether the subject with *hlakka til* occurs in the nominative or not, there are instances where matters are not so simple. Due to case syncretism in the nominal system, the form of the subject is sometimes morphologically ambiguous between one or more cases. Usually this does not cause problems in determining the case marking. Note that the finite verb generally agrees in

---

[92] The predicate is formed from the verb *hlakka* 'cry (of a bird of pray)' and the preposition *til* 'towards'.

person and number with the subject when it is nominative (see section above) but always has the 3rd person singular form when the subject is an oblique.[93] Observe for instance (9.15) where the nominative *ég* 'I' and the predicate *hlakka* match in person and number*;* in (9.16) this is not the case. The present and past tense indicative of the predicate *hlakka* (*til*) in all persons (1st, 2nd and 3rd) and numbers (singular and plural.) is shown in table Table 9.1.

	Present tense			**Present tense**	
	SG	**PL**		**SG**	**PL**
1p	*hlakka*	*hlökkum*	**1p**	*hlakkaði*	*hlökkuðum*
2p	*hlakkar*	*hlakkið*	**2p**	*hlakkaðir*	*hlökkuðuð*
3p	*hlakkar*	*hlakka*	**3p**	*hlakkaði*	*hlökkuðu*

Table 9.1. The present and past tense indicative of the predicate *hlakka* in all persons and numbers.

When encountering a morphologically ambiguous subject form, it is often possible to determine the case of the subject simply by looking at the form of the predicate. Take the plural definite form *börnin* 'the children' of the neuter noun *barn* 'child' as an example. Although the form is morphologically ambiguous between nominative and accusative plural, it is possible to claim that it is nominative in (9.17a), and accusative in (9.17b) due to the form of the predicate. In (9.17a) the predicate agrees with the subject but in (9.17b) it does not. For a comparison, (9.17c) shows the dative plural (*börnunum* 'the children') where the predicate is also in 3rd person singular, just like in (9.17b).

[93] A notable exception from this involves DAT-NOM structures where the nominative is plural. In these cases, the finite predicate may (but does not have to) agree with the nominative. On this type of agreement, see Ussery (2013).

(9.17) a. Nominative subject

Börnin	*hlakka*	*til*	*jólanna.*
the.children-NOM	look.forward	to	the.christmas

b. Accusative subject

Börnin	*hlakkar*	*til*	*jólanna*.
the.children-ACC	look.forward	to	the.christmas

c. Dative subject

Börnunum	*hlakkar*	*til*	*jólanna*.
the.children-DAT	look.forward	to	the.christmas

Occasionally, it is not possible to accurately determine the case of the subject. This usually happens when the predicate and the subject are in the 3rd person singular and the subject is morphologically ambiguous, either due to syncretism or due to two lexical items having identical forms that make identifying the case difficult.[94] Three examples are provided in (9.18)–(9.20). In (9.18), the third person masculine form *hann* 'he' is ambiguous between nominative and accusative singular, in (9.19) the oblique form *Láru* (personal name, nominative *Lára*) is ambiguous between accusative and dative, and in (9.20) the foreign name *Biden* is ambiguous between three cases: nominative, accusative and dative. In none

[94] The nominative and accusative singular of the third person masculine pronoun *hann* 'he' are syncretized. An example of a non-syncretized ambiguous form is *Gunna* (personal name), which could either be nominative of the female name *Gunna* or oblique of the masculine name *Gunni*.

of the examples is the form of the finite verb informative for determining the case of the subject.[95]

(9.18) Nominative or accusative subject

Hann	*hlakkar*	*til*	*jólanna*
he-NOM/ACC	look.forward-3P.SG	to	the.Christmas

'He looks forward to Christmas.'

(9.19) Accusative or dative subject

Láru	*hlakkar*	*til*	*jólanna*
Lára-ACC/DAT	look.forward-3P.SG	to	the.Christmas

'Lára looks forward to Christmas.'

(9.20) Nominative, accusative or dative subject

Biden	*hlakkar*	*til*	*jólanna*.
Biden-NOM/ACC/DAT	looks.forward-3P.SG	to	the.Christmas

When it comes to selecting a case for the subject with *hlakka til*, speakers are not always consistent and sometimes make use of more than one case. Intra-speaker variation has been documented and discussed in the literature (Nowenstein 2014,a,b, 2017, 2023). The variation in case marking can appear in two ways: i) with speakers sometimes selecting

[95] The form *Láru* in (9.19) is actually ambiguous between accusative, dative and genitive. Similarly, the form Biden (9.20) might be regarded as morphologically either nominative, accusative, dative or genitive. However, since genitive is not really an option for subject case marking with *hlakka til*, this option is excluded.

one subject case and sometimes another, and ii) with speakers mixing cases in the same utterance. Examples of the latter are shown in (9.21) where a speaker uses dative (2nd person pronoun) and nominative (1st person pronoun) in the same utterance.

(9.21)	*Eins og*	***þér***	*hlakka*	***ég***	*til…*
	Like	you-DAT	look-forward	I-NOM	to
	'Like you$_{\text{DAT}}$, I$_{\text{NOM}}$ look forward to…'				

Although variation between nominative and oblique may originally be due to prescriptivism (see 9.3.3), it has been argued that speakers may adopt variation in case marking as a part of their grammar during language acquisition (Nowenstein 2014a,b, 2023). Whether this holds true for *hlakka til* or not, a considerable intra-speaker variation is observed with subject case marking with *hlakka til*. Interestingly, the nominative surfaces most frequently in 1st person uses of the predicate which happens to be the most common form used by speakers – apparently, people like to talk about themselves when looking forward to something (see further in Section 9.4).

In the examples above, *hlakka til* always occurs as a finite predicate in a simple clause. Of course, this is not representative of all attested uses of *hlakka til*. The predicate also appears in combination with other verbs, either auxiliaries or in raising structure. In these cases *hlakka* is present in an infinitival form (with or without infinitival marker *að*) or in the supine (*hlakkað*). There is a slight chance that these might interfere with case marking of the subject, even though *hlakka til* remains the main predicate that is assumed to assign a thematic role and case to the subject.

Examples with a single auxiliary, *munu* 'will' denoting a future event and *hafa* 'have' referring to a past event, are provided in (9.22)–(9.23). In (9.22), *hlakka* appears in the infinitive; in (9.23) it is in the supine.

(9.22)	a.	***Börnin***	*munu*	*hlakka*	*til*	*jólanna*.
		the.children	will	look.forward	to	the.Christmas
	b.	***Börnin/Börnunum***	*mun*	*hlakka*	*til*	*jólanna*.
		the.children-ACC/DAT	will	look.forward	to	the.Christmas
		'The children will look forward to Christmas.'				

(9.23)	a.	***Börnin***	*hafa*	*hlakkað*	*til*	*jólanna*.
		the.children-NOM	have	look.forward	to	the.Christmas
	b.	***Börnin***/***Börnunum***	*hefur*	*hlakkað*	*til*	*jólanna*.
		the.children-ACC/DAT	have	look.forward	to	the.Christmas
		'The children have looked forward to Christmas.'				

Further examples are provided in (9.24) and (9.25) where *hlakka til* occurs with an infinitival marker *að* and the verbs *byrja* 'start, begin' and *fara* 'start, begin'. Although these verbs are "*transparent* to the properties of the main verb they take and do not assign case or select a theta role" (Sigurðsson 1989[1992]:67), they have been reported to show a dual behavior, sometimes behaving like regular raising predicates and sometimes like control verbs (Sigurðsson 1989[1992]:56–71, see also Sigurðsson 2017:244–247).[96] In the examples in (9.24) and (9.25), *hlakka* is assumed to remain the main (lexical) predicate that assigns case to the subject. Provided a speaker accepts an oblique subject, the b-examples are grammatical.

(9.24)	a.	***Börnin***	*byrjuðu*	*að*	*hlakka*	*til*
		the.children-NOM	started	to	look.forward	to
		jólanna.				
		the.Christmas				

	b.	***Börnin/Börnunum***	*byrjaði að*	*hlakka*	*til*	
		the.children-ACC/DAT	started to	look.forward	to	
		jólanna.				
		the.Christmas				
		'The children have started looking forward to Christmas.'				

(9.25)	a.	***Börnin***	*fara*	*að*	*hlakka*	*til*	*jólanna*.
		the.children-NOM	start	to	look.forward	to	the.Christmas

	b.	***Börnin/Börnunum***	*fer*	*að*	*hlakka*	*til …*
		the.children-ACC/DAT	start	to	look-forward	to
		'The children start looking forward to…'				

[96] When data was gathered and prepared for forecasting (see Section 9.4), these were treated as regular raising predicates that contributed aspects to the meaning and did not risk affecting the case marking.

The auxiliaries and modals in (9.22)–(9.25) may be claimed to be fairly neutral in meaning, simply contributing aspects, i.e., that something occurred in the past (*hafa* ‘have’), will occur in the future (*munu* ‘will’), or has started occurring (*byrja, fara* ‘start’). Other auxiliaries also exist, perhaps with a slightly less neutral meaning than the ones in (9.22)–(9.25). For instance, the modals *mega* ‘may, can, be able to, *geta* ‘can, be able to’ and *kunna* ‘may/might’ indicate a possibility (or ability) to to something and *eiga* ‘ought’, *þurfa* ‘need’, *verða* ‘have to’ and *skulu* ‘should’ add an obligation towards the main predicate; *vilja* evokes the notion of an independent will ‘want to’. Aside from *vilja*, these normally count as raising predicates when used as auxiliaries. However, they (aside from *skulu* ‘should’) also exist as independent lexical verbs, assigning nominative case to their subjects.[97]

Since *hlakka til* exhibits variation in subject case marking, it is possible that combining the predicate with modals such as *eiga*, *þurfa*, *verða*, *skulu*, *mega*, *geta*, *kunna* and *vilja* might raise the likelihood of nominative being used. Although this has not been systematically investigated, it has been pointed out (Svavarsóttir 2023) that non-nominative subject case is sometimes lost in certain raising structures (on some modals

[97] The meaning of the lexical predicates *mega*, *geta*, *kunna*, *eiga*, *þurfa* and *vilja* are not always identical to the modal use. Thus, *mega* means ‘be allowed to’ instead of modal ‘be able to’ and *kunna* means ‘know’ instead of ‘may/might’.

(i)
a. *Ég má þetta*
I can this
‘I am allowed to do this’

b. *Ég get þetta*
I can this
‘I am able to do this’

c. *Ég kann þetta*
I know this
‘I know this.’

d. *Ég á þetta*
I own this
‘I own this.’

e. *Ég þarf þetta*
I need this
‘I need this’

f. *Ég vil þetta*
I want this
‘I want this’

sometimes behaving like raising predicates and sometimes like control predicates see Sigurðsson 1989[1992]:56–71; Thráinsson & Vikner 1995; see also Sigurðsson 2017:244–247).[98]

In addition to complex clauses with a single auxiliary, complex clauses with multiple auxiliaries may occur. The use of multiple modals and auxiliaries increases the surface distance between the subject and the lexical verb, which might interfere with the subject case marking.[99] Examples (9.26)–(9.29) show *hlakka til* occurring with two to four auxiliaries and/or modals. Even though all these show case preservation in subject raising, it remains to be documented whether and up to what extent they might affect changes in case marking of predicates like *hlakka til,* especially when many of them occur together such as in (9.26)–(9.29). This question is left for future research.

(9.26) 2x: *þurfa*, *geta*

a.	*Ég*	***þarf***	*að*	***geta***	*hlakkað*	*til*.
	I-NOM	need	to	be.able	look.forward	to

b.	*Mig/Mér*	***þarf***	*að*	***geta***	*hlakkað*	*til*.
	me-ACC/DAT	need	to	be.able	look.forward	to

‘I need to be able to look forward…’

[98] Of course, this raises the question whether relevant structures are indeed always raising structures or whether some speakers treat them as control. A different question might be whether the non-preservation of case might reflect changes in subject case marking with a lower predicate.
[99] As discussed by e.g., Corbett (2006:170) and Gold et al. (2018), linear distance and linearity in general may affect agreement, with elements linearly close together agreeing.

(9.27) 3x: *mun*, *geta*, *fara*

a.

Þú	***munt***	***geta***	***farið***	*að*	*hlakka*	*til*.
you-NOM	will	be.able	start	to	look.forward	to

b.

Þig/Þér	***mun***	***geta***	***farið***	*að*	*hlakka*	*til*.
you-ACC/DAT	will	be.able	start	to	look.forward	to

(9.28) 3x: *hafa*, *vera*, *farinn*

a.

Hún	***hefur***	***verið***	***farin***	*að*	*hlakka*	*til.*
she-NOM	has	been	started	to	look.forward	to

b.

Hana/Henni	***hefur***	***verið***	***farið***	*að*	*hlakka*	*til.*
she-ACC/DAT	has	been	started	to	look.forward	to

'She must have been looking forward to Christmas.'

(9.29) 4x: *hljóta*, *hafa*, *vera*, *búinn*

a.

Hún	***hlýtur***	*að*	***hafa***	***verið***	***búin***	*að*	*hlakka*
she-NOM	must	to	have	been	done	to	look.forward
til	*lengi*						
to	long						

'She must have been looking forward (to something) for a long time.'

b. *Henni* **hlýtur** *að* **hafa** **verið** **búið** *að* *hlakka*
she-DAT must to have been done to look.forward
til *lengi*
look.forward to long
'She must have been looking forward (to something) for a long time.'

There are certain predicates that may be on the borderline of whether they allow raising to subject position or whether they themselves assign a case. The verb *virðast* 'seem' is typically considered a raising predicate (Thráinsson 2007:441, Sigurðsson 2017:239), but occasionally it does not preserve the case of an embedded oblique-subject predicate. Svavarsdóttir (2023:30) notes that non–preservation of cases occurs in about 15% of examples with *virðast* 'seem' in data from the Icelandic Gigaword Corpus.

(9.30) a. *Hún* *virðist* *hlakka* *til*
she seems look.forward to

b. *Henni* *virðist* *hlakka* *til*
she seems look.forward to

The predicate *segjast* 'say of oneself, claim, be said' also shows a somewhat odd behavior. Apparently, speakers generally "differ as to which case they find better for the matrix subject, dative or nominative" (Wood 2015:294–298). Since *hlakka til* originally appeared with a nominative subject, some interesting interactions with *segjast* might be expected. In any case both nominative and oblique may be found, (9.31).[100]

(9.31)	a.	***Hún***	*segist*	*hlakka*	*til*
		she-NOM	says	look.forward	to
	b.	***Henni***	*segist*	*hlakka*	*til*
		her-DAT	says	look.forward	to
		'She says she looks forward to…'			

Finally, *hlakka til* may be combined with predicates that control the case of the subject such as *vera sagður* 'be said' (9.32) and *ætla* 'intend' (9.33), of which the latter may attribute intentionality to the subject.

(9.32)	a.	***Hún***	*var*	*sögð*	*hlakka*	*til*	*jólanna*.
		she-NOM	was	said	look.forward	to	the.Christmas
		'She was said to be looking forward to Christmas.'					

[100] As someone who uses dative with *hlakka til*, examples of nominative come surprisingly naturally to me when the predicate is embedded under *segajst*. Examples are further improved when including aspectual verbs such as *hafa* 'have' or *munu* 'will'.

(i)	a.	*Hestamaðurinn*	*sagðist*	*hlakka*	*til.*	
		the.horseman-NOM	said	look.forward	to	
	b.	*Hestamaðurinn*	*hefur*	*sagst*	*hlakka*	*til.*
		the.horseman-NOM	has	said	look.forward	to

	b.	****Hana/Henni***	*var*	*sagt*	*hlakka*	*til*	*jólanna*
		she-ACC/DAT	was	said	look.forward	to	the.Christmas

(9.33)	a.	***Hún***	*ætlar*	*að*	*hlakka*	*til*	*jólanna.*
		she-NOM	going	to	look.forward	to	the.Christmas
		'She intends to look forward to Christmas.'					

	b.	?***Hana/Henni***	*ætlar*	*að*	*hlakka*	*til*	*jólanna.*
		she-ACC/DAT	going	to	look.forward	to	the.Christmas

Although both nominative and oblique (accusative or dative) are possible with *hlakka til* when combined with *ætla* 'going to', the use of different cases invokes slightly different senses of the auxiliary. In (9.33a), the use of *ætla* makes the nominative subject *hún* 'she' appear to have the intention of looking forward to Christmas. In (9.33b), intentional reading is less likely although it may still be marginally possible. A more likely interpretation is that *ætla* conveys the meaning of 'seem, appear'. Compare (9.33b) to (9.34) where the meaning is semantically bleached, and no intentionality whatsoever can be attributed to the subject.[101]

(9.34)	a.	*Það* ***ætlar***	*ekki*	*að*	*stytta*	*upp.*
		EXPL going.to	not	to	shorten	up
		'It looks like the rain is not going to stop'				

[101] In the case that *ætla* appears with a reflexive *sér* 'self', it always takes a nominative subject and behaves like a control verb rather than a raising predicate (Wood 2022).

b.	*Mér*	***ætlaði***	*aldrei*	*að*	*hlýna.*
	me-DAT	going.to	never	to	get.warm

‘It seemed like I was never going to get warm.’

Intentionality is strongly linked to agency which, in turn, is associated with nominative case marking on subjects (Jónsson 1997–98, 2003). As mentioned above (Section 9.3.2), oblique subjects are never agents. Nominatives, on the other hand, can be agents although they need not be. Since *ætla* is often (although not always) associated with intentionality, there is a chance that some speakers may be biased towards accepting and using nominative more frequently than otherwise expected.

Subject case marking with *hlakka til* in complex clauses with multiple auxiliaries/modals has not been systematically investigated. Usually, subject case marking with *hlakka til* is surveyed using simple clauses where the subject is immediately adjacent to the inflected predicate *hlakka til*, so there is no doubt as to where the case comes from. As discussed in Section 9.4.2, more complex structures are found in attested data on the internet. These include structures where the predicate occurs without an overt subject, in which case it can be difficult to properly establish the subject case marking, and instances of constructions where two or more predicates might compete for assigning case to the relevant noun phrase. For further discussion on case marking and mixed case marking with *hlakka til*, see the section on prescriptivism (Section 9.4.2) and the section where the data used for forecasting is described (Section 9.3.3).

9.3.2 Diachrony and previous documentation

As already mentioned (Sections 9.1 and 9.3.1), *hlakka til* 'look forward to' was originally used with a nominative subject. The earliest attested example with an oblique case is from the late 19th century where the predicate appears with an accusative subject, (9.35a). A dative subject is attested somewhat later, or around the mid-20th (9.35b) (for an overview and documentation of early examples of *hlakka til* with oblique subjects see e.g., Sigurðardóttir & Eythórsson 2022).

(9.35) a.

Mig	*hlakkar*	*til,*	*að*	*fá*	*að*	*verða*	*félagi*	*þinn*
me-A	look.forward	to	to	get	to	be	partner	your
og	*sessunautur*.							
and	companion							

'I look forward to be allowed to be your partner and companion.'

(*Þjóðólfur* 1892(1):13)

b.

Einnig		*hefi*	*ég*	*heyrt*	*suma*	*segja: „****Mér***	*hlakkar*
furthermore		have	I	heard	some	say	I-DAT look.forward
svo	*mikið*	*til*	*að*	*komast*	*í*	*berjatúrinn*."	
so	much	to	INF	come	to	the.berry.picking.tour	

'Furthermore, I have heard some people say: "I look so much forward to being able to go on the berry-picking tour"' (*Unga Ísland* 1941(1):3)

Examples of accusative and dative subjects with *hlakka til* appear somewhat sporadically until after the mid 20th century. Many of the early examples occur in articles that openly point out the innovative variants, often declaring them to be incorrect. This is, for instance, the case in (9.35b), where a young girl expresses her concern about children using language incorrectly. The sole function of her writing is to encourage children to think more about how they speak.

The strong prescriptive attitude against the use of oblique cases with *hlakka til* (see also Section 9.3.3) may account for why oblique has a slow uptake in published material. Using the online portal tímarit.is, an overview of uses of nominative and oblique cases with *hlakka til* in published periodicals may be obtained. Targeting uses in first person singular, a search was made for three different strings: *ég hlakka til*, *mig hlakkar til* and *mér hlakkar til*.[102] The results from the search are provided in 9.2.

While the earliest use of an accusative *mig hlakkar til* (see example (9.35a) above) is from a short story translated into Icelandic, other early examples mainly appear in articles discussing correct and incorrect language use. Thus, examples from between 1930 and 1969 containing accusative or dative subjects are all from articles where good and bad language use is mentioned. It stands to reason that published, proofread and copy-edited material may not give an accurate picture of the situation within the language community. A survey targeting what individuals do in their daily life may be more informative of the status of the subject case marking with *hlakka til*.

[102] Note that this search does not capture uses where the subject and the preposition til are non-adjacent to the predicate *hlakka*. It just picks out the three strings mentioned (*ég hlakka til*, *mig hlakkar til*, *mér hlakkar til*).

Period	NOM	ACC	DAT	Total OBL	Total
1890–1899	3	1	0	1 (25%)	4
1900–1909	10	0	0	0 (0%)	10
1910–1919	11	0	0	0 (0%)	11
1920–1929	12	0	0	0 (0%)	12
1930–1939	52	1	0	1 (ca. 1.89%)	53
1940–1949	134	1	2	3 (ca. 2.19%)	137
1950–1959	311	3	4	7 (ca. 2.20%)	318
1960–1969	470	3	2	5 (ca. 1.30%)	475
1970–1979	457	5	1	6 (ca. 1.30%)	463
1980–1989	794	14	11	25 (ca. 3.04%)	819
1990–1999	1150	17	11	28 (ca. 2.38%)	1178
2000–2009	2026	18	18	36 (1.75%)	2062
2010–2019	2183	8	10	18 (0.82%)	2201
2020–2029	297	3	1	4 (1.33)	301

Table 9.2. An overview of distribution of the strings *ég hlakka til*, *mig hlakkar til* and *mér hlakkar til* as they appear on tímarit.is. Leaving aside the earliest attested example with an accusative subject, the proportion of oblique cases in the first person singular in published material stays between ca. 0% to 3%.

The first large-scale study of subject case marking with *hlakka til* was conducted in the early 80s (Svavarsdóttir 1982). The study focused on documenting changes in subject case marking among 11-year-old students in elementary schools all over Iceland. Several verbs subject to Dative Substitution (a change from accusative to dative subject, cf. Section 9.2.2 above) were studied. In addition, two predicates that historically had a nominative but have a tendency to appear with either accusative or dative subject were investigated. These were *hlakka til* and *kvíða fyrir*. Around 200 eleven-year-old children from 11 randomly selected schools were tested, i.e., about 5% of all children in 5th grade in elementary schools in the country between the years 1980–1981. The survey was conducted in two parts. Part A looked at uses of the 3rd person singular feminine pronoun *hún* 'she' with all the predicates under investigation. Part B used the same text, but focused on the 1st person singular. Only one instance of the predicate *hlakka til* appeared in each part (for more thorough discussion of methods etc. see Svavarsdóttir 1982:26–28). The overall results for the whole country are presented in Table 9.3.

Part A: 3rd person singular feminine Total completed: 202			**Part B**: 1st person singular Total completed: 198		
NOM	**ACC**	**DAT**	**NOM**	**ACC**	**DAT**
19.30%	60.90%	19.80%	45.60%	40.10%	11.90%

Table 9.3. Proportion of nominative, accusative and dative subjects with *hlakka til* in part A and B of Svavarsdóttir's (1982:31) survey. The children had only one opportunity in each part to provide a subject for the predicate. Note that the proportions provided for the B part assume that only 98.0 of the total (202) participants completed the task.

It is interesting to note that the use of oblique case is more prominent in the 3rd person singular feminine (combined total 80.7%) than the 1st person singular (combined total 52%). The difference is likely to be traced to prescriptive grammar teaching in school, something that is discussed further below (Section 9.3.3). Another point worth mentioning is that according to Svavarsdóttir's study (1982:20, 46), innovative subject case marking can be linked to children's performance at school and the parent's socio-economic status. Children who have educated parents and do well at school are less likely to use innovative case marking.

A large-scale follow-up study to Svavarsdóttir's research in the early 80s (Svavarsdóttir 1982) was conducted in the fall of 2001 by Jónsson and Eythórsson (2003). Jónsson and Eythórsson followed similar procedures as Svavarsdóttir and tested variation in subject case marking of 26 predicates (including *hlakka til*), focusing on the case of the 3rd person singular feminine pronoun *hún* 'she'. In total, 900 eleven-year-old children (born in 1990) in 6th grade of elementary school were tested. Of these, 845 answer sheets were used for analysis. A summary of variation in subject case marking with *hlakka til* is shown in Table 9.4. The category “other” may refer to unanswered questions or forms that could not be identified as either nominative, accusative or dative.

The use of 3rd person singular feminine Total included in the analysis: 845			
NOM	**ACC**	**DAT**	**OTHER**
14.90%	41.40%	43.20%	0.50%

Table 9.4. Proportion of nominative, accusative and dative subjects with *hlakka til* in Jónsson’s and Eythórsson’s (2003) survey conducted in the fall of 2001. The survey was modeled after that of Svavarsdóttir (1982) from the early 80s.

In the 20-year period from 1980 to 2001, some changes can be detected. Notably, the proportion of nominative subject marking (when the subject is the 3rd person feminine pronoun *hún* ‘she’) has dropped from roughly 19.3% to 14.9%. The use of oblique cases has increased from a total of 80.7% to 84.6%, with the dative appearing more frequently than the accusative in the survey from 2001.

A third large-scale study on variation in subject case marking with selected predicates in Modern Icelandic was carried out within the project *Variation in Icelandic Syntax* (Thráinsson 2013; Thráinsson, Svavarsdóttir, Rögnvaldsson, Jónsson, Sigurjónsdóttir & Blöndal 2013). The project was active in 2004, when pilot studies were conducted, and between 2005–2007, when the majority of the material was gathered in three written surveys in 30 places around the country. Instead of only focusing on young children, four different age groups were tested. These included teenagers in 9th grade of elementary school (ca. 14–15 years old), individuals between 20–25, adults between 40–

45 and older individuals between 65–70. Two methods were used to elicit responses. The first method involved participants providing grammaticality judgments using three available choices: i) *YES = Normal sentence. I can say this*, ii) *NO = Ungrammatical sentence. I cannot say this*, and iii) *? = Dubious sentence. I can hardly say this* (Thráinsson, Angantýsson, Sigurðsson, Steingrímsdóttir & Eythórsson 2013:33). For testing subject case marking with *hlakka til*, each participant had to evaluate three sentences, one with a nominate, one with accusative and one with dative. The second method for inquiring about subject case with *hlakka til* was in accordance with the procedure of both Svavarsdóttir (1982) and Jónsson and Eythórsson (2003). For this task, participants had to fill in a blank space the appropriate form of a 3rd person singular feminine pronoun *hún* 'she'. As in previous studies, participants only had to select the appropriate subject form for *hlakka til* once (for more detailed discussion of methods and test material see Thráinsson et al. 2013:19–46; Svavarsdóttir 2013:90–92). A summary of results for *hlakka til* for both taks types and all age groups are shown in Table 9.5.

	Judgment task			Fill-in-the-blank: 3rd person singular feminine pronoun hún 'she'		
GROUP	**NOM**	**ACC**	**DAT**	**NOM**	**ACC**	**DAT**
9th grade	32.50%	66.20%	63.80%	15.20%	38.60%	46.20%
20–25	45.80%	66.80%	58.40%	27.50%	47.00%	25.50%
40–45	54.20%	59.80%	31.80%	31.60%	44.60%	23.80%
65–70	66.00%	42.30%	16.60%	45.20%	33.80%	21.00%

Table 9.5. A summary of results for case marking with *hlakka til* from the studies in 2005–7, based on tables from Thráinsson, Eythórsson, Svavarsdóttir & Blöndal (2015:42, 44). Around 740–748 answer sheets were used for the analysis.

Thráinsson et al. (2015:44) point out that even though the judgment task and fill-in-the-blank production task show slightly different results for each age group, the results are nevertheless consistent. The youngest group has the lowest nominative acceptance and usage in both instances (32.5% for judgment tasks, 15.2% for production) and the oldest group has the highest (66.0% for judgment tasks and 45.2% for production). Viewing the results in light of the apparent time approach, there are indeed indications that the use of oblique subjects with *hlakka til* is indeed increasing. The younger the speaker is, the more likely they are to use an oblique case. Viewing the results in light of real time studies, the age group involving 9th graders is the most comparable to participants of the previous two studies which involved 11-year-old children (Svavarsdóttir 1982; Jónsson & Eythórsson 2003). Comparing the results with the more recent of these (Jónsson & Eythórsson 2003), no drastic differences are observed. The nominative in the 2005–7 survey is 15.2% as opposed to 14.9% earlier, the accusative is 38.6% which is slightly lower than the earlier 41.4%, and the dative has seemingly increased to 46.2% from 43.2% earlier. Given that there were only a few years between the two surveys (one taking place in 2001 and the other in 2005–7) and that the participants were slightly older in the more recent study (around 14–15 instead of 11 years old) the results are understandable. A higher proportion of nominative could be due to participants having reached an age where they are better at incorporating prescriptive rules. As to the oblique cases, dative seems to be gaining more ground.

A fourth study on case marking with predicates affected by Dative Substitution and Oblique-Case Substitution was conducted in 2016–2019 within the MoLiCoDiLaCo-project, led by Sigríður Sigurjónsdóttir and Eiríkur Rögnvaldsson (2018, cf.

https://molicodilaco.hi.is/). The goal of the project was to investigate the influence of English on Icelandic language acquisition and grammar (Sigurjónsdóttir & Nowenstein 2021). Results on variation in case marking are extensively discussed in Nowenstein (2023), where the focus is on syntactic and semantic bootstrapping during the acquisition of case marking. Due to the study targeting younger individuals, aged 3–13, the number of participants is not as high as in older studies on variation in case marking. Nowenstein (2022:188) notes that 101 children participated in the forced-choice task, and these were then split up into four age categories, 3–5;11 years old, 6–8;11 years old, 9–11;11 years old and 12–13;11 years old (see Nowenstein 2022:89). The forced-choice task tested the use of case with predicates that are already attested in the language in addition to novel made-up predicates. The verb *hlakka til* falls into the former category. The results for participants in the oldest group are provided in Table 9.6. For a complete overview and discussion of other age groups and other predicates, see Nowenstein (2023, in particular pages 91, 93).

Forced-choice task 21 participants aged 12–13;11 years old			
NOM	**ACC**	**DAT**	**OTHER**
19%	33%	43%	5%

Table 9.6. A summary of results for variation in subject case marking with *hlakka til* among 12–13;11 year old children. Based on data obtained in 2016–2019 in the MoLiCoDiLaCo-project (Sigurjónsdóttir & Rögnvaldsson 2018; see also Nowenstein 2023).

It is worth noting that the "measurements" of uses of nominative and oblique with *hlakka til* from the MoLiCoDiLaCo-project are not necessarily fully comparable to those in previous studies. As already noted, only 21 participants make up the group of individuals between 12 and 13;11 years old. This is considerably less than in Svavarsódttir's (1982) study which featured around 200 children, Jónsson and Eythórsson's study which had roughly 845 participants and the study conducted within the project Variation in Icelandic Syntax where over 700 individuals participated. Furthermore, while the three older studies focused on having children produce forms of the 3rd person singular feminine pronoun *hún* 'she', the forced-choice task in MoLiCoDiLaCo study relied on participants selecting forms of the feminine noun *stelpan* 'the girl'. Nevertheless, it is worth comparing the four studies as in Table 9.7.

Study	**NOM**	**ACC**	**DAT**	**OTHER**
Svavarsdóttir, 1980–1981	19.30%	60.00%	10.80%	NA
Jónsson & Eythórsson, fall 2001	14.90%	41.40%	43.30%	0.50%
Variation in Icelandic Syntax, 2005–2007	15.60%	38.60%	46.20%	3%
MoLiCoDiLaCo, 2016–2019	19%	33%	43%	5%

Table 9.7. Results from four different studies on subject case marking with *hlakka til* compared. The first three studies focused on forms of the 3rd person singular feminine pronoun *hún* 'she', while the last study involved forms of the feminine noun *stelpa* 'girl'.

The direction of change from Svavarsdóttir's (1982) research to that of Jónsson and Eythórsosn (2003) and *Variation in Icelandic Syntax* (Thráinsson 2013, Thráinsson et al. 2013) seems to be towards abandoning nominative in favor of oblique case marking, in particular dative. However, the direction away from nominative is not as clear if one compares the earliest study to the most recent one. In that case, the situation may look somewhat stable with respect to uses of the nominative. On the other hand, a diminished use of accusative is still observed.

Another point worth mentioning is that the situation described in the four studies above do not seem to match the proportions of nominative and oblique cases as found in published material through the library portal tímarit.is. As already noted (see Table 9.2), examples with an oblique case make up between 0%–3% of attested examples from each decade when the use of the 1st person singular is investigated.[103] In the part of Svavarsdóttir's study (1982) that looked at the 1st person singular, the dative appeared in 11.9% of the cases and the accusative in around 40%. Combined, this makes up about 52% of the examples. Children that were 11 years of age in 1980 should have been around 40 to 50 years in 2010–2019. It seems reasonable to think that some of them contributed to published material at that time. If this is the case, one should expect higher proportions of oblique cases in published material than what is found. A gap from 52% of examples having an oblique case (children in the early 1980s) to 0.82% of the attested examples on tímarit.is between 2010–2019 having an oblique case, seems quite large and demands an explanation. Picking up on Svavarsdóttir's (1982:20, 42–46) observation that the results suggest a correlation between how well children did at school and which class in society their parents belonged to and their use of case with predicates like *hlakka til*, one might think that maybe only the well-educated from higher classes in society went on to write material that would be published in journals and papers. Another explanation is that we are observing the result of copy editing and, more generally, prescriptivism. These are discussed briefly in the following section (9.3.3).

9.3.3 Effects of prescriptivism

As the earliest attested examples indicate (see section 9.3.2), using an oblique subject with *hlakka til* has traditionally been viewed negatively (Svavarsdóttir 1982:23–25; Pálsson 1979). In fact, variation in case marking with *hlakka til* appears to be one of the most widely

[103] Note again that the search only targeted three strings where the subject is immediately adjacent to the finite predicate.

discussed changes within Icelandic (Svavarsdóttir 1982:23; Óladóttir 2017:27). Time and time again, the nominative has been declared the "correct" case to use with this predicate, and thus, normative pressures in the society strongly disfavor the use of an oblique.

Within the Icelandic school system, in particular in elementary and secondary schools, there have been pervasive efforts to root out the use of oblique subjects with *hlakka til*. Children and teenagers are explicitly taught not to use dative or accusative with the predicate. These endeavors have not been very successful. Svavarsdóttir (1982) noted already in the early 80s that fights against various widespread changes were by some (e.g., Pálsson 1979) deemed hopeless and harmful, and that they should be discontinued. Nevertheless, the "battle" is still in full swing. Commonly used teaching material in elementary school features discussions on the incorrectness of oblique subjects with *hlakka til* (see discussion in Óladóttir 2017:120–125, 236). Additionally, teachers typically advise students to learn to use the nominative case with the predicate. Questions about "correct" case marking with *hlakka til* appear extremely often on exams (Óladóttir 2017:2). Thus, despite over 50 years of nominative being declared the only "correct" subject case with *hlakka til*, the use of oblique cases still prevails. Additionally, it is safe to say that no one goes through the Icelandic school system without having heard that using an oblique subject with *hlakka til* is "incorrect". As Óladóttir (2017:28) points out, it is likely that the prescriptive battle against uses of oblique case has "greatly affected the situation found among many language users today". So how exactly are the prescriptive rules incorporated in teaching? And what kind of effects do they have on the use of *hlakka til*?

Both Svavarsdóttir (Svavarsdóttir 1982:37, 2013:107-108) and Óladóttir (2017:251) note that when subject case marking with *hlakka til* is discussed in the school system, the focus is first and foremost on uses of the 1st person singular pronoun *ég* ‘I’ . Phases such as “One should not say *mig hlakkar* til … but rather *ég hlakka til*” are frequently encountered (see Svavarsdóttir 1982:37). This seems to have had several consequences.

First, Svavarsdóttir’s results (1982) showed a considerable difference in the use of subject case with *hlakka til*, depending on whether the subject was the 1st person pronoun *ég* ‘I’ or the 3rd person pronoun *hún* ‘she’. While about 19.3% of examples appeared in nominative when the subject was in the 3rd person, this number rose to 45.6% when the subject was in the 1st person. Similarly, Óladóttir (2017:249–250) notes that students find the use of nominative with *hlakka til* more acceptable when the subject is the 1st person pronoun *ég*, rather than any other pronoun or a noun.[104]

Second, the use of nominative has, in the mind of students, come to be strongly linked to exams where students are frequently asked to pick out the “correct” use of *hlakka til*. In some cases, the examiners attempt to trick students by using ambiguous morphological forms in test sentences. Observe (9.36) which appeared on an official test question from the year 2000, relying on the personal names *Siggi* and *Sigga*. Only the nominative would be considered correct (Óladóttir 2017:107–108).[105]

---

[104] Investigating grammar teaching in elementary schools, Óladóttir conducted interviews with both students and teachers in 2010.

[105] The question, which is a multiple-choice question, includes four options. Only one option is considered correct, namely the option where the subject is nominative, (9.36b).

(9.36) a. ***Sigga*** ***litla*** *hlakkar* *til* *jólanna.*

Siggi-M.ACC small-M.ACC look.forward to the.Christmas.

Hann *er* *þriggja ára*.

He-M is three years.

'Little Siggi looks forward to Christmas. He is three years old.'

b. ***Sigga*** ***litla*** *hlakkar* *til* *jólanna.*

Sigga-F.NOM small.F.NOM looks.forward to the.Christmas.

Hún *er* *þriggja ára*.

She-F is three years

'Little Sigga looks forward to Christmas. She is three years old.'

Needless to say, attempts to eradicate oblique subjects with *hlakka til* has caused considerable frustration among students in the school system. It may even have resulted in widening the gap between written and spoken language as the nominative becomes associated with written material, e.g., exams, teaching material and so on (Óladóttir 240–242).

Third, the overemphasis on using a nominative (and not accusative or dative) with *hlakka til*, has resulted in language users becoming extremely aware of how they use the predicate (Óladóttir 2017:244). It has also upset the intuition that speakers have about case marking with this particular predicate, causing a lot of insecurity. Speakers are caught between trying to follow their intuitions and wanting to protect the language. It does not help that when an oblique case is used, people are generally perceived as speaking "incorrectly" (Óladóttir 2017:243–244, 254). This may cause speakers to consciously attempt to correct their use of the predicate for *hlakka til*. As some examples found on social media suggest, these corrections are not always completely followed through, resulting in half-corrected utterances with mixed case marking or examples where only half of the utterance has been corrected (see discussion on data in Section 9.4.2).

Fourth, it appears to matter how frequently speakers encounter nominative subjects marked with *hlakka til*. The more frequently they are encountered, the more readily speakers accept and use the nominative (Óladóttir 2917:145). It has even been suggested that due to children being exposed to nominative in the 1st person singular more frequently than in other instances, they may acquire a mixed case marking (see Nowenstein 2014a,b on children acquiring mixed case marking with predicates originally occurring with accusative subject).[106]

9.3.4 Expectations

While the Case Directionality Hypothesis predicts changes from oblique case marking to nominative with subjects, it has been pointed out that experiencer predicates are often characterized by non-nominative case marking. Given this, the experiencer nature of *hlakka til* may be the primary drive for changes in subject case marking towards oblique. This has, however, not prevented long-standing attempts to eradicate the use of oblique subjects with *hlakka til* (Section 9.3.3). Judging from the discussion above, oblique case

[106] Nowenstein (2014a,b) investigates mixed case marking with predicates originally occurring with an accusative subject. She argues that intra-speaker variation is affected by factors such as whether the subject is in the 1st person or not and whether syncretism is found in the paradigm. She furthermore claims that children may acquire mixed case marking in accordance with the proportion of innovative and traditional case marking found in their Primary Linguistic Data (PLD).

marking with *hlakka til* prevails and seems to be on the rise, with dative being the most favored choice (Section 9.3.2). Based on these facts, the expectations for the propagation of the change are as follows. Oblique case marking on the subject of *hlakka til* is expected to continue to increase at the expense of nominative. Due to overemphasis on the 1st person singular with *hlakka til* in the school system, nominative subjects are likely to survive the longest there. Furthermore, since changes in case marking with *hlakka til* are widely discussed, copy-edited and proofread material is expected to have a higher proportion of nominative than other text types and other areas of language use. Casual language, possibly on social media, is more likely to contain examples of accusative and dative with *hlakka til*. However, individuals may still be aware of how they use the predicate and, possibly, be prone to consciously (or semi-consciously) "correct" their usage. Finally, there is a possibility that children may acquire a nominative case for the first person singular of the predicate and another case for other instances.

9.4 Data annotation and description

9.4.1 Data Source

Quantitative data for documenting variation in Modern Icelandic and generating predictions for subject case marking with *hlakka til* 'look forward to' come from two different sources: Twitter (Twitter API v2) via the R-package *academictwitteR* (Barrie & Ho 2021) and The Icelandic Gigaword corpus (rmh=2019, rmh=2022 cf. Steingrímsson et al. 2018). These sources contain material from different registers and cover slightly different periods of time.

Data from Twitter is taken to represent semi-informal, non-proofread language. Tweets can be posted by anyone with a Twitter account and they do not go through a formal review or approval before being posted online.[107] However, due to language purism and prescriptivism being widespread in Iceland, individuals who tweet may try to conform somewhat to preconceived ideas of the standardized language and general rules of punctuation. It is mostly for these reasons that Twitter data is not taken to reflect fully informal language.

When obtaining data from Twitter, the target language was set to Icelandic and retweets were excluded from the query. A search was made for various forms of the predicate *hlakka*, shown in Table 9.8, without including the preposition *til*.

Form	**Possible case marking on the subject**
hlakka, hlakkar, hlakkaði, hlakki, hlakkað	NOM or OBL
hlökkum, hlökkuðum, hlakkið, hlakkið, hlakkaðir, hlökkuðuð, hlökkuðu	NOM or OBL

Table 9.8. A list of forms searched for on Twitter (API v2). Only 3rd person singular forms, the infinitive and the supine can occur with an oblique subject.

Although Twitter contains data from 2006 onwards, Icelandic tweets are relatively few until around 2009. Due to this, data prior to 2009 is not well suited for a time-series analysis. As a result, the Twitter data used here only covers the period from 2009 to 2022, both years included. An example of a query made on the 27th of January 2023 is provided

[107] Although anyone and everyone can sign up for a Twitter account, not everyone does. This results in data from Twitter being subject to self-selection bias such that only tweets from individuals who have chosen to be active users of Twitter figure into the dataset. The assumption here is nevertheless that this type of data remains relatively consistent over time and that it is informative of language use of some part of the population.

in (9.37). As can be seen, the target string in this case was “hlakkar” within Icelandic (lang:is). The search covered the period from January 1st 2022 to January 1st 2023 with retweets being excluded.[108]

(9.37) hlakkar -is:retweet
2022-01-01T00:00:01Z
2023-01-01T00:00:00Z

The second source of data was The Icelandic Gigaword Corpus (IGC), of which two versions were used (rmh=2019, rmh=2022, cf. Steingrímsson et al. 2018). These contain about 1,532 (IGC-2020) and 2,429 (IGC-2022) million running words and span several different text types, everything from online forum threads and blog posts (informal language) to printed news media (formal language). The majority of the material in IGC is from around and after the year 2000. In order to keep a consistency in the text types over time, only data from 2000 to 2019 (rmh=2019) and from 2020 to 2022 (rmh=2022) was used in the current study. The reason for two versions of the IGC being used is that data gathering and annotation had already begun when the IGC-2022 was released. Since the older version only contained material up until 2019, a more recent version of the corpus was used to fill in the period 2020–2021 for selected news media and for the period of 2000–2021 for online discussion threads. IGC-2022 was also used to obtain data from the period 2000–2021 from the source Bland.is (an online discussion thread), which had not been included in the original search.

[108] Data for the period January 1st, 2009, to December 31st, 2021, had been abstracted previously.

The searches in IGC targeted the lemma of *hlakka*. The intention was to only include material with a timestamp containing information on the year and month of writing.[109] This resulted in some sources being excluded, for instance published books and academic journals which may not have been created the year they were published. A list of the type of material obtained from IGC is presented in (9.38) and (9.39). Note that for the more recent version of IGC (rmh=2022), only four sources were used to fill in the period from 2020–2021.

(9.38) **IGC-2019 SOURCES** (55/77)

News media: printed newspapers, online newspapers

Social media: blog posts (heimur.is, Silfur Egils)

Special material: various local news media, specialized news material (agriculture and fish news), sports news, tabloids, cultural material

Administrations: Speeches from Althingi

(9.39) **IGC-2022 SOURCES** (4/87)

News media: printed newspapers (Morgunblaðið), online newspapers (Mbl.is)

Social media: Bland.is

Special material: sports news (fótbolti.is)

[109] Originally, the intention was to only obtain data that included information on year and month of writing. However, some of the timestamps turned out to have a placeholder “00” for both month and day such that only the year (and not month) of creation could be established.

9.4.2 Data annotation

Data from Twitter and the Icelandic Gigaword corpus was cleaned and carefully annotated using Microsoft Excel (version 2308, Microsoft Corporation 2022). Various methods were used to clean the data and exclude examples that did not contain the predicate *hlakka til*. When the prepositions *í* 'in' and *yfir'* over' occurred directly after any form of *hlakka*, the example was taken to be an instance of *hlakka í* 'chuckle, laugh quietly or inwardly' and *hlakka yfir* 'exhult or gloat over'. Duplicates of *hlakka til* were removed based on examples being identical. Examples were deemed identical in IGC if Excel judged the match (the form of the predicate), left and right context and original source to be the same. Twitter-examples were deemed identical if the text of the tweet was identical.[110] The essential criteria for examples to be included in analysis of subject cae marking was that the example must be an instance of *hlakka til* and contain an overt subject whose case was assigned by *hlakka til*.

(9.40) **Essential Criteria for examples to be analyzed**

(i) They contain an instance of the predicate *hlakka til*

(ii) They contain an overt subject whose case was assigned by *hlakka til*

Criteria (ii) excludes examples such as in (9.41) where the subject of *hlakka til* is assigned a case by a higher predicate, e.g., *láta* 'make' in (9.41a) or *segja* 'say' in (9.41b). It furthermore excludes examples such as in (9.43) where the predicate occurs in an infinitival clause, and examples as in (9.42) where *hlakka til* appears without an overt subject.

[110] Twitter search in R already excluded retweets. However, they did not exclude tweets tweeted more than once by the same individual.

Although it is sometimes possible to determine the case of covert subjects, e.g., by referencing the form of the predicate (9.42a), this is not always so. The case of null subjects that are 3rd person singular can, for instance, not be determined (9.42b). Instead of risking introducing a considerable amount of ambiguity in the annotation, with nominative more easily identifiable than oblique, all examples with covert subjects were marked NULL and excluded from further analysis.

(9.41)	a.	*Við*	*skulum*	*láta*	***okkur***	*hlakka*	*til*
		we	shall	make	us	look.forward	to

‘We’ll make us look forward to…’

b.	*Ég*	*sagði*	***manninn***	*hlakka*	*til*
	I	said	the.man-ACC	look.forward	to

‘I said that the man looked forward to…’

(9.42)	a.	*Ég*	*kem*	*heim*	*um*	*jólin*	*og*	___
		I	come	home	around	the.Christmas	and	NULL.NOM

hlakka	*til*	*að*	*sjá*	*þig.*
look.forward	to	INF	see	you

‘I’m coming home for Christmas and (I) look forward to seeing you.’

b. *Jón* *kemur* *heim* *um* *jólin* *og* ___
John comes home around the.Christmas and NULL.AMB
hlakkar *til* *að* *sjá* *þig*.
look.forward to INF see you
'John comes home for Christmas and (he) looks forward to seeing you.'

(9.43) …*mér* *finnst* *líka* *bara* *svo* *gaman* *að* *hlakka* *til*
I find also just so fun to look.forward to
einhvers *svona*.
something of-that-type
'… I just find it so much fun to look forward to something like this.'

(IGC, Bland.is 20110101)

Examples of *hlakka til* containing an overt subject and confirming to criteria (9.40-ii) were divided into two categories, marked REG and REG2. The first category (REG) contains instances of *hlakka* as a finite predicate in simple clauses, as well as *hlakka til* in the infinitival or supine form combined with the auxiliaries/modals *byrja* 'begin', *fara* 'begin', *hafa* 'have' and *munu* 'will', which were deemed highly unlikely to affect the case marking of the subject (see also Section 9.3.1 above). To make sure these could be later identified if needed, a comment noting which auxiliary was used was included in the annotation in a separate column. The second category (REG2) contained *hlakka til* in an infinitival or supine form combined with auxiliaries that were deemed to have the potential to interfere with regular case marking of the subject. These included predicates such as *virðast* 'seem', *hljóta* 'must', and *ætla* 'intend to'. An overview of the type of examples of *hlakka til* with an overt subject is provided in Table 9.9.

Subject	Type	Comments
YES	REG	*hlakka til* occurs as a finite predicate in a simple clause or combines with *byrja*, *fara*, *hafa* or *munu*
	REG2	*hlakka til* combines with one or more of the following: *vera búinn*, *vera farinn*, *eiga*, *geta*, *hætta*, *hljóta*, *kveðast*, *mega*, *segjast*, *skulu*, *þurfa*, *vera*, *verða*, *virðast* and/or *ætla*

Table 9.9. An overview of the annotation of examples of *hlakka til* containing an overt subject.

The Icelandic Gigaword Corpus (*rmh=2019*, rmh=2022 cf. Steingrímsson et al. 2018) already contains an automatic morphological annotation. However, this annotation was ***not*** used for identifying subject case marking as it treats ambiguous forms in a problematic way, e.g., by always favoring nominative over oblique. Furthermore, the annotation available in IGC does not identify subjects. Therefore, it was considered appropriate to annotate subjects and case marking from scratch. A second argument in favor of that is that instead of two "layers" of possible misparses of the case of subjects, i.e., from IGC and the author of the present work, errors could only be introduced by the author of the present work.

Identification of subjects and annotation of subject case marking was done semi-automatically and by hand. The semi-automatic method involved searching for particular strings using the IF function in Excel and annotating relevant examples in batches. Nouns and pronouns occurring immediately before or after *hlakka* were generally hypothesized to be subjects. For example, the morphological form *ég* 'I (nom)' occurring either before or after *hlakka* was annotated as a nominative subject. In most cases, the semi-automatic annotation was manually verified.

Annotation of subject case marking took into consideration morphological ambiguity. If the case of a subject could not be accurately determined, it was marked in such a way that the nature of the ambiguity was clear. Thus, ambiguity between nominative and accusative was marked differently from ambiguity between accusative and dative. In total, six labels were used: NOM, NOMACC, ACC, ACCDAT, DAT and AMB, (9.44).

(9.44) Basic annotation of subject case with *hlakka til*

NOM	The subject is in the nominative case.
NOMACC	The subject is either in the nominative or accusative case.[111]
ACC	The subject is in the accusative case.
ACCDAT	The subject is either in the accusative or the dative case.
DAT	The subject is in the dative case.
AMB	The subject is three-way ambiguous in case marking. It could be nominative, accusative or dative.

Examples of morphologically ambiguous subjects are provided in (9.45)–(9.47).

[111] In addition to forms that are … forms of masculine nouns such as *Valtýr, Víðir* which are traditionally only nominative are marked as ambiguous between nominative and accusative. This was done because many speakers (including the other of this work) take the -*r* to be part of the stem, resulting in using the same form for nominative and accusative.

(9.45)

Starfsfólkið	*hjá*	*Prentmet*	*hlakkar*	*til*
staff	at	Prentment	looks.forward	to

‘The staff at Prentment looks forward to…’ (Twitter, 2011)

(9.46)

Gleðiegt [sic]	*nýtt*	*ár,*	***okkur***	*hlakkar*	*til*
happy	new	year,	we-ACC/DAT	look-forward	to

að	*byrja*	*nýja*	*árið*
to	start	the.new	year

‘Happy new year, we look forward to starting the new year..’ (Twitter, 2010)

(9.47)

Gerrard	*hlakkar*	*til*	*að*	*vinna*	*með*	*Hodgson*
Gerrard-AMB	looks.forward	to	INF	work	with	Hodgson

‘Gerrard looks forward to work (Twitter, 2012)

Whenever possible, ambiguity was resolved. This was done with reference to the morphological form of the predicate or to a morphologically ambiguous phrase that should agree with the subject in case and number. For instance, example (9.48) was annotated as having an accusative subject even though the forms *stjórn* ‘management’ and *starfsfólk* ‘employees’ are morphologically ambiguous between nominative and accusative. The form of the predicate, namely *hlakkar*, suggests that the subject should be in an oblique case. A nominative subject in the plural would be expected to occur with the form *hlakka*, agreeing with it in number (see also discussion in Section 9.3.1).

(9.48) ***Stjórn*** ***og*** ***starfsfólk*** *sjóðsins*

management-ACC and employees-ACC fund-GEN

hlakkar *mikið* *til…*

look.forward-3P.SG much to

'The management and employees of the fund look greatly forward to…'

(IGC, mbl.is, 20210514)

The examples in (9.49)–(9.50) show how ambiguity can be resolved with reference to an unambiguous phrase that is expected to agree with the subject in case and number. In examples like (9.49) the form *okkur* (oblique of the pronoun *við* 'we') is morphologically ambiguous between an accusative or dative plural. Relying on the forms *öllum* (a dative plural of *allir* 'everyone') and *alla* (an accusative plural of *allir* 'everyone') the case can be argued to be dative in (9.49a) and accusative in (9.49b). In (9.50) the name *Jason Koumas* could be taken as either nominative, accusative or dative. The attributive *nýjum leikmanni*, a dative of *nýr leikmaður* 'new player', provides evidence for analyzing it as dative.

(9.49) a. ***Okkur*** *hlakkar* *öllum* *til…*

us-DAT look.forward all-DAT to

b. ***Okkur*** *hlakkar* *alla* *til…*

us-ACC look.forward all-ACC to

'We all look forward'

(9.50)	*Jason*	*Koumas*		***nýjum***		***leikmanni***		*Wigan*
	Jason	Koumas		new-DAT		player-DAt		Wigan
	hlakkar	*til*	*að*	*hefja*	*að*	*leika*	*með*	*félaginu*
	look.forward	to	INF	begin	INF	play	with	the.team

‘Jason Koumas, new player at Wigan, looks forward to playing with the team.’

(IGC, Fótbolti.net, 2007)

In occasional examples, there was a mismatch between the morphological form of the subject and the form of the predicate. Thus, the nominatives *vinkonurnar Hildur og Kolbrún* ‘the friends Hildur and Kolbrún’ in (9.51a) and *Stefanía og Steinunn* ‘Stefanía and Steinunn’ in (9.51b) are expected to appear with a plural form of the predicate, namely *hlökkuðu* for (9.51a) and *hlakka* for (9.51b). Instead, the predicate is in the 3rd person singular in both examples. As already noted (see Section 9.3.1) this would be appropriate for an oblique subject.

(9.51) The predicate points to oblique, but the subject is morphologically nominative

a.	***Vinkonurnar***	***Hildur***	*og*	***Kolbrún***	<u>*hlakkaði*</u>
	friends-NOM	Hildur-NOM	and	Kolbrún-NOM	look-forward.3P.SG
	mikið til				
	much to				

‘The friends, Hildur and Kolbrún, looked much forward to (something)’ b.

Stefanía	*og*	***Steinunn***	<u>*hlakkar*</u>	*til…*
Stefanía-NOM	and	Steinunn-NOM	look.forward-3P.SG	to

'Stefanía and Steinunn look forward to…'

The situation in (9.52) is the opposite of that in (9.51). Here, the subject is clearly morphologically dative, but the form of the predicate matches what would be expected with a morphologically nominative subject.

(9.52) The predicate points to nominative, but the subject is morphologically dative

a.

Djöfull	***hlakka***	***mér***	*til*
devil	look.forward	me-DAT	to

'Fuck how much I look forward to…' (Twitter, 2011)

b.

mer	***er***	***sjálf***	***farin***	***að***	***hlakka***
me.DAT	is	self.NOM	started	INF	look.forward

til	*þegar*	*hún*	*verður*	*18…*
to	when	she	becomes	18

'I myself have started looking forward to when she turns 18…'

(IGC, Bland.is, 2007)

In cases where there was a mismatch between the morphological form of subject and the predicate as in (9.51) and (9.52), the annotation was biased towards oblique. The argument for doing so was that prescriptivism is very prominent with *hlakka til* and individuals may consciously attempt to choose a form that conforms with what is considered "good language". It is not unlikely that such corrections were attempted for examples like those in (9.51) and (9.52), with only a partial success. A conscious correction of subject case marking with *hlakka til* may also be responsible for examples where mixed case marking is observed. These can be as in (9.53) where there the language user switches from a dative (*þér* 'you$_{\mathrm{DAT}}$' to a nominative (*ég* 'I$_{\mathrm{NOM}}$'), or as in (9.54) where individual lexical items that are part of subject do not agree in case.

(9.53)	*Eins og*	***þér***	*hlakka*	***ég***	*til…*
	Like	you-DAT	look.forward	I-NOM	to

'Like you$_{\mathrm{DAT}}$, I$_{\mathrm{NOM}}$ look forward to…'

(9.54)	*a.*	***Okkur***	*hlakkar*	***öll***	*svo*	*til að*	*fá*	*hann…*
		us-ACC/DAT	look.forward	all-NOM	very	to to	receive	him

'We all look very much forward to receiving him.' (IGC, Bland 2007)

b.	*…að*	***við***	***systrunum***	*hlökkuðum*	*liggur við*	*til*
	that	we-NOM	sisters-DAT	look-forward	almost	to
	að	*fara*	*til*	*hans.*		
	to	go	to	him		

'... that we, the sisters, almost looked forward to going to him.'

(IGC, Bland, 2008)

The number of examples where case marking was mixed or where the case marking of the subject did not match the expected form of the predicate were at least 20.

Few examples (less than 10) contained the subject *minns* 'I' which is a type of diminutive sometimes used to refer to oneself or a character one is playing in a game. It appears to be formed from a possessive pronoun *minn* 'mine' with an extra *-s* at the end (for a discussion of this form and further examples of its use, see Wood & Sigurðsson 2011). Since the form *minns* can technically be used for any case, it was annotated as AMB.[112] Note that the form of the finite verb (*farið*) in (9.55) suggests that *minns* is in an oblique case.

(9.55)	***Minns***	*er*	*farið*	*að*	*hlakka*	*til*	*sko.*
	mine-M.AMB	is	started	INF	look.forward	to	PARTICLE
	'I have started looking forward'				(IGC, Bland 2006)		

Finally, some examples obtained from both Twitter and IGC occurred in discussions of the predicate *hlakka til* and which case its subject should (prescriptively) be in. While some speakers do not hesitate to inform others that the subject of the predicate should be nominative (9.56a), others are confident that it should be accusative (9.56b).

[112] The form *minns* can can be used as a subject, e.g., in (i) where it occurs with a predicate assigning nominative case, or it can be used as an object, cf. (ii) where it occurs with a predicate (*lemja* 'strike') that generally assigns accusative case to its object.

(i)	*Minns*	*fór*	*í*	*kolaportið*	*í dag…*	
	mine	went	to	Kolaportið	today	
	'I went to Kolaportið today.' (http://alnaemi.blogspot.com/)					
(ii)	*…það*	*lamdi*	*minns*	*bara*	*í*	*framan.*
	it	struck	me	just	in	face.
	'It [the hail] hit me in the face.' (http://skrifar-inn.blogspot.com/2005/11/vottaplani.html)					

(9.56) a. *Maður segir* ***ÉG*** *HLAKKA TIL!!!*

man says I-NOM look.forward to

‘One says I$_{\text{NOM}}$ LOOK FORWARD TO.’ (IGC, Bland.is 20051001)

b. *Nei, það er* ***''hestinn*** *hlakkar til'' alveg*

no it is horse-ACC looks.forward to just

eins og ***''manninn*** *hlakkar til'', ekki*

like and the.man-ACC looks.forward to not

''maðurinn hlakkar til''.

the.man-NOM looks-forward to

“No, it is “the horse$_{\text{ACC}}$ looks forward to”, just like “the man$_{\text{ACC}}$ looks forward to”, not “the man$_{\text{NOM}}$ looks forward to.” (IGC, Bland.is, 2006)

Examples occurring in discussion on appropriate use of case were not systematically annotated as such, nor were they counted. Provided the predicate had a subject, the example was treated like any other example and annotated accordingly.

9.4.3 General overview of data

A total of 41,987 examples were retrieved from Twitter. Of these, 588 were not an instance of the predicate *hlakka til* and in 918 examples *hlakka til* was not responsible for the case assignment of the subject. Of the remaining 40,481 examples, 29,201 did not have an overt subject. This leaves 11,280 examples with an overt subject, falling into the categories REG (10,479) and REG2 (801) as described above.

Raw and unannotated examples retrieved from IGC, not counting data from the subsource Bland, were in total 45,546, with 42,523 being from rmh=2019 and covering the period up until end of 2019, and 3,023 being from rmh=2022 and covering the period 2020–2021. Duplicates (total of 2,610) and examples that did not meet the criteria set out in (9.40) were removed, leaving 31,145 examples with an overt subject whose case was assigned by *hlakka til*. Some additional examples were removed including those occurring in material before 1999 (a total of 257 examples) and those from Alþingisræður 'Parliament speech' (additional 1,471 examples), Hæstiréttur 'high court' (26 examples) and Dómstólar 'courts' (23 examples). The reason for removing examples prior to 1999 is that they are too few and too irregularly spaced in time to be used for regular time series analysis. Examples from Alþingisræður, Hæstiréttur and Dómstólar were removed because these are of a different nature than written material from news media and they were not supposed to have been included in the original search. Removing these left a total of 29,368 examples.

IGC examples retrieved from the subsource Bland.is, covering the period 2000–2021, were in total 30,037. Once duplicates (1,335) and examples that did not meet the criteria in (9.40) had been removed (total of 15,878) 12,824 examples remained.

Once data from IGC had been properly cleaned and combined into a single file, the total number of examples containing an overt subject were 42,192, falling into the categories REG (36,962) and REG2 (5,230) as described above (see Table 9.10). The number of examples of each type of case marking in IGC and Twitter, not taking into account the type of example, were as in Table 9.10.

Source	subject case with *hlakka til*						
	NOM	**NOMACC**	**ACC**	**ACCDAT**	**DAT**	**AMB**	**TOTAL**
Twitter (2009–2022)	9,452 (c. 83.8%)	110 (c. 1%)	796 (c. 7%)	251 (c. 2.2%)	450 (c. 4%)	221 (c. 2%)	11,280 (100%)
IGC (2000–2021)	32,688 (c. 77.5%)	3,018 (c. 7.2%)	3,085 (c. 7.3%)	868 (c. 2%)	1,761 (c. 4.2%)	772 (c. 1.8%)	42,192 (100%)

Table 9.10. Overview of subject case marking with *hlakka til* as found in IGC and on Twitter. Unambiguous nominative is the most common case marking, making up between 77% and 84% of the total examples.

Type of example	subject case with *hlakka til*						
	NOM	**NOMACC**	**ACC**	**ACCDAT**	**DAT**	**AMB**	**TOTAL**
IGC REG	29,117 (c. 78.9%)	2,268 (c. 6.1%)	2,824 (c. 7.6%)	820 (c. 2.2%)	1,517 (c. 4.1%)	416 (c. 1.1%)	36,962 (100%)
IGC REG2	3,571 (c. 69.6%)	750 (c. 14.6%)	261 (c. 5.1%)	48 (c. 0.9%)	244 (c. 4.8%)	256 (c. 5%)	5,230 (100%)

Table 9.11. Overview of subject case marking in categories REG and REG2 in IGC.

As shown in Table 9.10 and Table 9.11, the most common subject case with *hlakka til* is nominative which accounts for between 70% and 84% of attested examples with an overt subject. This is a very different picture from what studies on subject case marking with *hlakka til* show when the use of case was investigated among 11-year-old school children (Svavarsdóttir 1982, Jónsson & Eythórsson 2003). In those studies, nominative occurred in less than 20% when the subject was a 3rd person pronoun (three studies) and around 45% when the subject was the 1st person singular pronoun *ég* ‘I’ (one study, Svavarsdóttir 1982). Seeing that nominative is more common with 1st person singular subjects than other types, it is tempting to partially attribute the high proportion of nominative in written material to frequent usages of the predicate with a 1st person subject. Other factors that might contribute towards nominative being so common include conscious choice of case marking due to prescriptivism, copy editing and proofreading.

The majority of overt subjects with *hlakka til* are pronominal 1st person subjects, suggesting that individuals mostly use the predicate to talk about themselves. On Twitter, roughly 89% of instances with an overt subject are pronominal 1st person subjects. The proportion in IGC is about 83%. An overview of types of subjects with *hlakka til* is shown in Figure 9.3 and Figure 9.4. Note that the label UNID in Figure 9.4 refers to unidentified subjects, i.e., the subject case marking has been recorded but the type of subject (whether it is pronominal, common noun, a name etc.) has not.

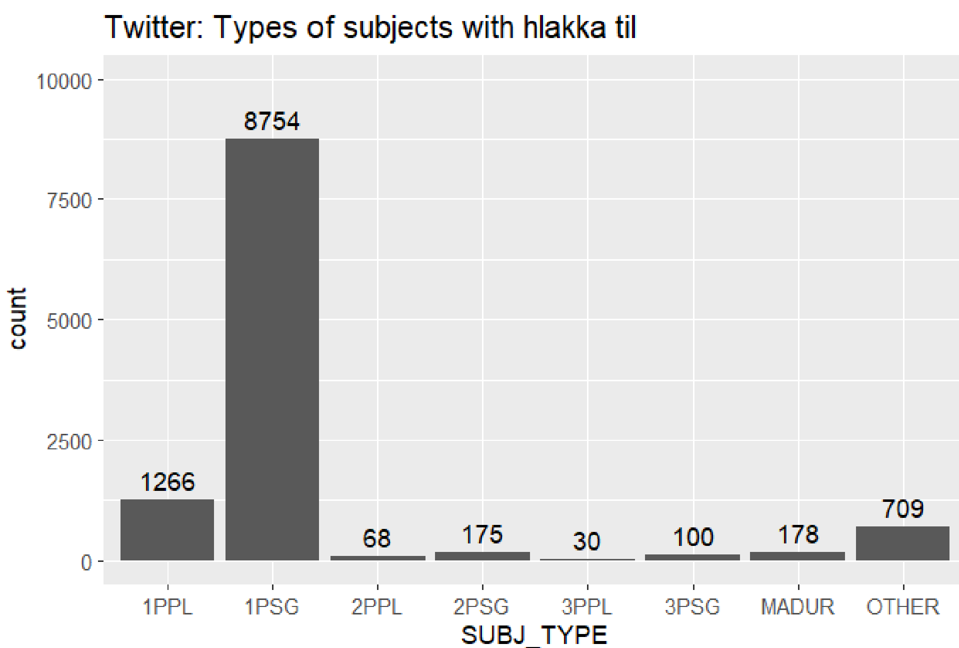


Figure 9.3. Overview of types of subjects with *hlakka til* on Twitter. Pronominal subjects in the first person are 10,020 (8,754 in the singular, 1,266 in the plural) which is about 89% of total examples with an overt subject.

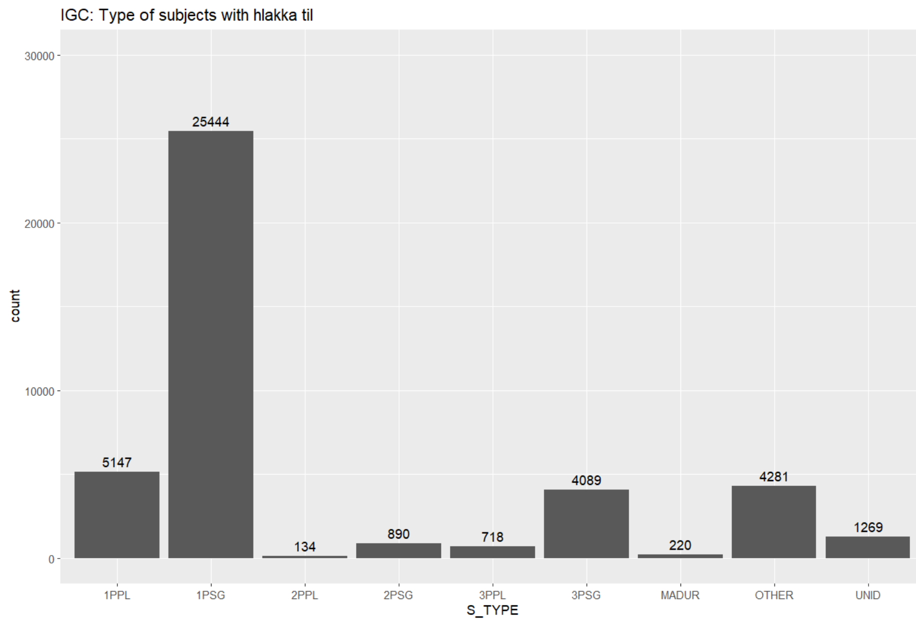


Figure 9.4. Overview of types of subjects with *hlakka til* in IGC. Pronominal subjects in 1st person are 30,591 (25,444 in the singular, 5,147 in the plural) which is about 83% of total examples with an overt subject.

As already mentioned, data from IGC consists of various types of material, everything from copy-edited newspapers to online discussion threads. Subject case marking might naturally vary depending on the nature of individual sources in the IGC. Figure 9.5 provides an overview of case marking, taking into account the different sub sources within IGC. Note how accusative and dative case marking appears to be more common in certain resources in the corpus.

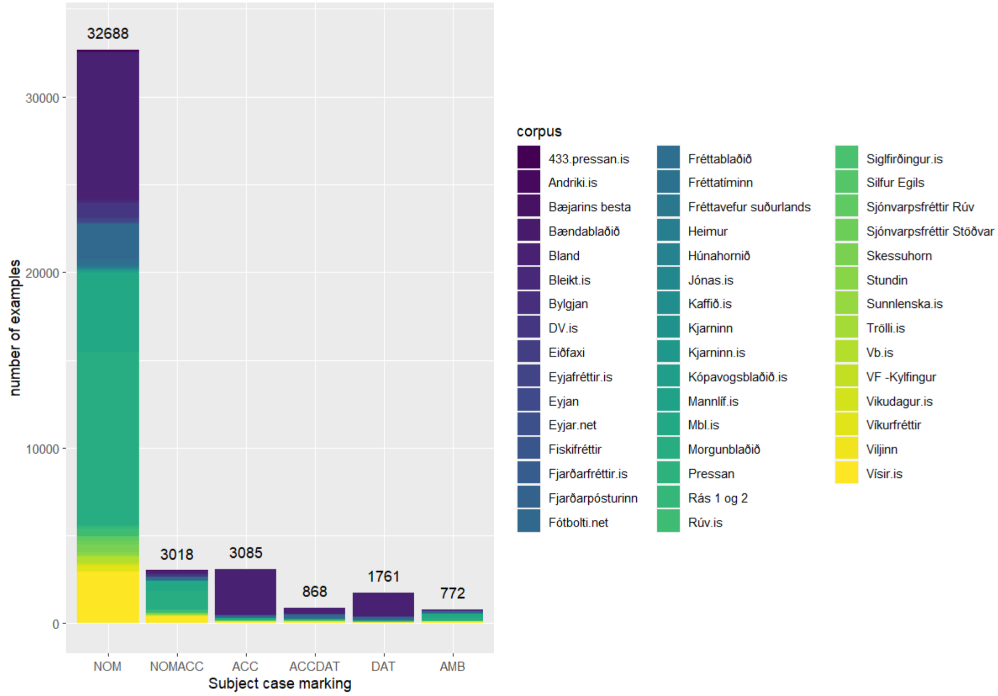


Figure 9.5. Subject case marking within various sources in IGC. Note that accusative and dative subjects appear to be more common in certain sources within the corpus.

In an attempt to capture the observations that the use of oblique subjects might be tied to language register, individual sources in IGC were labeled either as representing formal, semi-formal or informal registers. As with the case study in Chapter 8, formality-labeling was based on i) whether the material was likely to be proofread or not, ii) how the material was mediated, i.e., via printed publication, radio, or online, and iii) which group of people the material was targeted at, i.e., the general population or just a few individuals. The labeling of individual sources within the IGC is shown in (9.57).[113]

[113] *Fréttatíminn* appears to be listed twice, once in formal material and once in semi-formal material. The one listed under formal material is a printed newspaper (published and distributed between 2010 and 2017) while the other (under semi-formal material) is an online new media page.

(9.57) The annotation of three types of registers

Form: Formal language, printed or broadcasted material. The use of "standard" and "good" language is considered very important. The material is likely proofread.

Sources: *Bændablaðið*, *Eiðfaxi*, *Fjarðarpósturinn*, *Fréttablaðið*, *Fréttatíminn*, *Kjarninn (blað)*, *Kópavogsblaðið*, *Morgunblaðið*, *Rúv.is*, *Siglfirðingur*, *Sjónvarpsfréttir RÚV*, *Sjónvarpsfréttir Stöðvar 2*, *Stundin*

SForm: Semi-formal language, mostly online material. The material is typically not proofread although the use of "good" language is considered appropriate.

Sources: *Bæjarins besta*, *Bb.is*, *Bleikt.is*, *Bóndi.is*, *Bylgjan*, *Dv.is*, *Eiðfaxi*, *Eyjafréttir.is*, *Eyjan*, *Eyjar.net*, *Fiskifréttir*, *Fjarðarfréttir.is*, *Fjarðarpósturinn*, *Fótbolti.net*, *Fréttablaðið.is*, *Fréttatíminn*, *Fréttatíminn.is (gamli vefur)*, *Fréttavefur Suðurlands*, *Heimur.is*, *Húnahornið*, *Húni*, *Jonas.is*, *Kaffið.is*, *Kjarninn.is*, *Kópavogsblaðið.is*, *Kylfingur.vf.is*, *Mannlíf.is*, *Mbl.is*, *Pressan*, *Rás 1 og 2*, *Siglfirðingur*, *Silfur Egils*, *Skessuhorn*, *Stundin.is*, *Sunnlenska.is*, *Trölli*, *Vb.is*, *Vf.is*, *Vikudagur.is*, *Víkurfréttir*, *Viljinn.is*, *Vísir.is*, *433.pressan.is*

IForm: Informal language, online material. The material is not proofread and conforming to a "standard" or "proper" language is typically not considered crucial as the material is not necessarily intended to be read by everyone.

Sources: *Spjallvefur Bland*

Figure 9.6 shows subject case marking with *hlakka til* in the three different registers mentioned in (9.57). Observe how informal material, which only includes the source Bland.is, has the most attestations of accusative and dative case marking. Material listed as formal (Form) has hardly any unambiguous oblique subjects; there are quite a few ambiguous nominative or accusative subjects and some three-way ambiguous examples (AMB). Semi-formal language appears to be a mixed bag, including some oblique subjects but not nearly as many as informal language. Figure 9.7 breaks down the distribution of examples within the IGC, showing that the majority of the data comes from six sources, namely Bland.is (informal), Morgunblaðið (formal), Mbl.is (semi-formal), Vísir.is (semi-formal), Fótbolti.net (semi-formal) and DV.is (semi-formal).

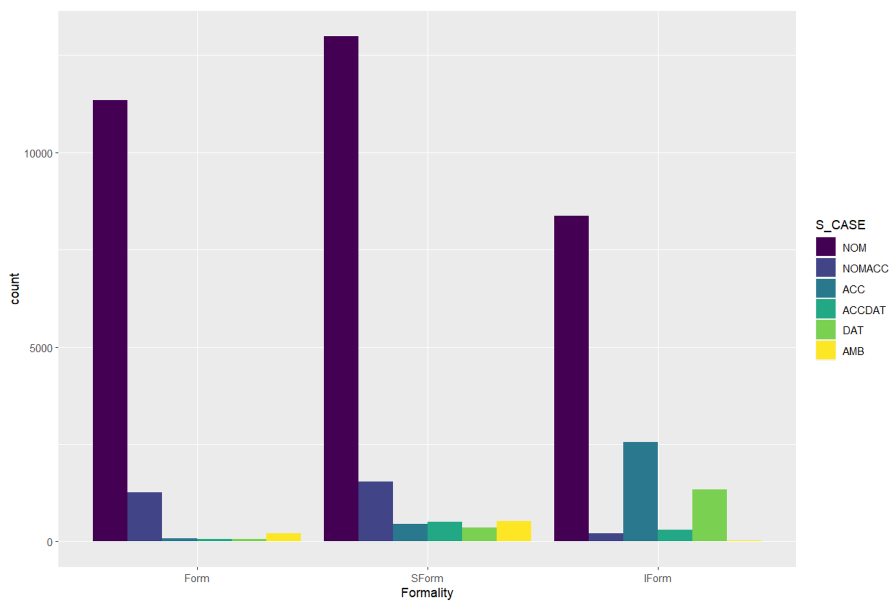


Figure 9.6. Subject case marking with *hlakka til* in formal (Form) semi-formal (SForm) and informal (IForm) sources in IGC. Oblique subjects appear to be more prominent in informal sources than in formal and semi-formal sources.

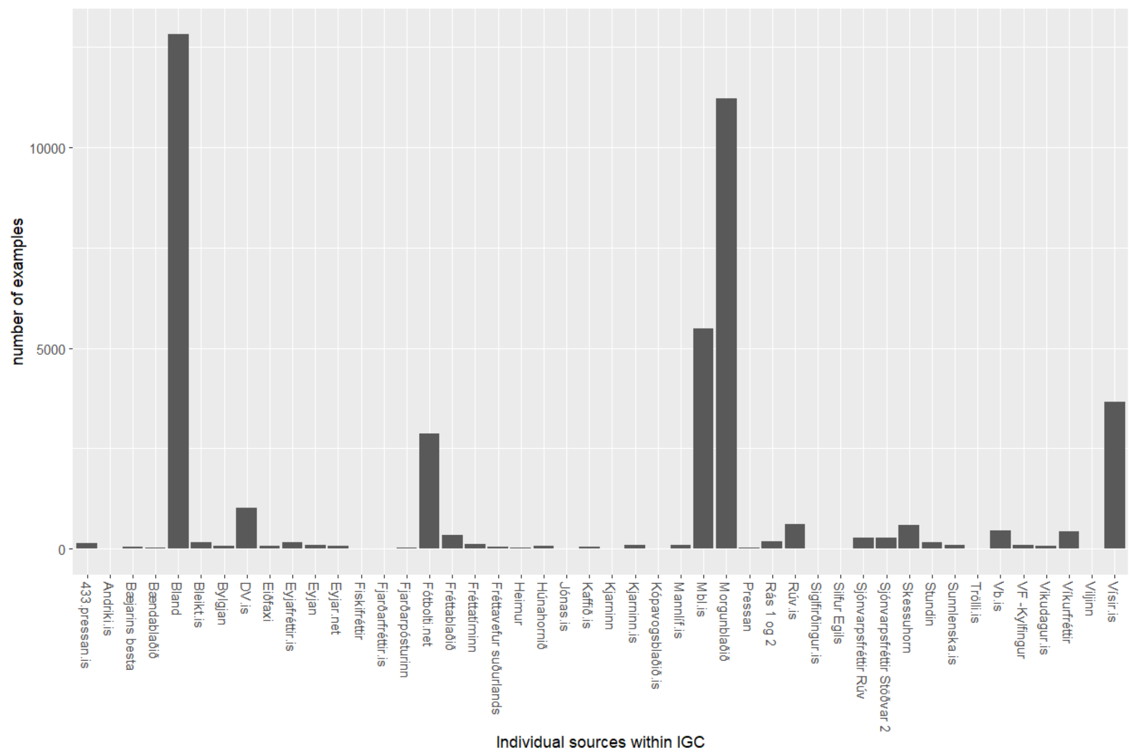


Figure 9.7. Overview of the amount of data from the various sources within IGC. Six sources contribute most of the examples: Bland.is (12,824 examples, informal language), Morgunblaðið (11,216 examples, formal language), Mbl.is (5,486 examples, semi-formal language), Vísir.is (3,668 examples, semi-formal language), Fótbolti.net (2,862 examples, semi-formal language) and DV.is (1,028 examples, semi-formal language).

Since the present study is not so much concerned with which exact case of subjects, but rather whether they were nominative or oblique, the six-way basic annotation in (9.44) was used to create two types of additional annotations. These were i) biased towards oblique and ii) towards nominative. For oblique-biased annotation, it was assumed that every example that could potentially involve an oblique did have an oblique. For the nominative-biased annotation, it was assumed that every example that could potentially involve a nominative, did indeed have a nominative. A summary of the basic annotation along with oblique-biased and nominative-biased ones is shown in Table 9.12.

Annotation	Obl-biased	Nom-biased
NOM	NOM	NOM
NOMACC	OBL	NOM
ACC	OBL	OBL
ACCDAT	OBL	OBL
DAT	OBL	OBL
AMB	OBL	NOM

Table 9.12. A summary of how the original annotation was used to create two additional annotations, one biased towards oblique case marking and the other towards nominative case marking.

Oblique-biased and nominative-biased annotations give an idea of an upper and a lower limit of oblique subjects, i.e., how often the subjects are without a doubt oblique (nominative-biased annotation) and how often they might be interpreted as an oblique (oblique-biased annotation). These are summarized in Figure 9.8–Figure 9.9. Interestingly, ambiguous case marking is more common in formal and semi-formal sources than in informal sources. This leads to the difference between nominative- and oblique-biased annotation being greater for those sources. The reason for the difference can be attributed to nominative/oblique ambiguous 3rd person subjects (especially the pronoun *hann* 'he' which can morphologically be either nominative or accusative) and ambiguous foreign names being more commonly found in news sources such as Morginblaðið, Mbl.is and Fótbolti.is than in online discussion threads such as Bland.is. Tweets from Twitter show a similar pattern as material from Bland.is. Ambiguous nominative/oblique case marking is not common on Twitter.

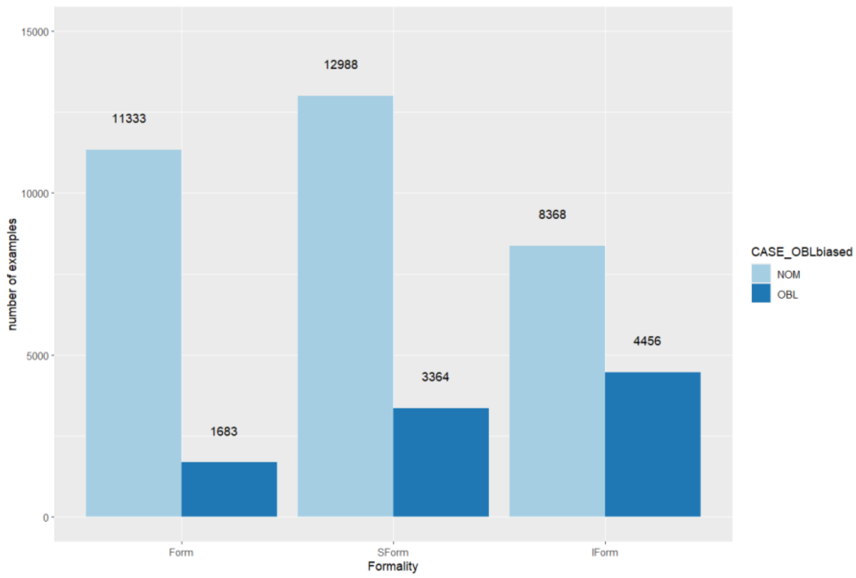


Figure 9.8. Biasing ambiguous case marking towards oblique gives an idea of how often the subject of *hlakka til* might be inferred as being in an oblique case. Oblique subjects in formal material are almost entirely drawn from examples with ambiguous nominative-oblique case marking.

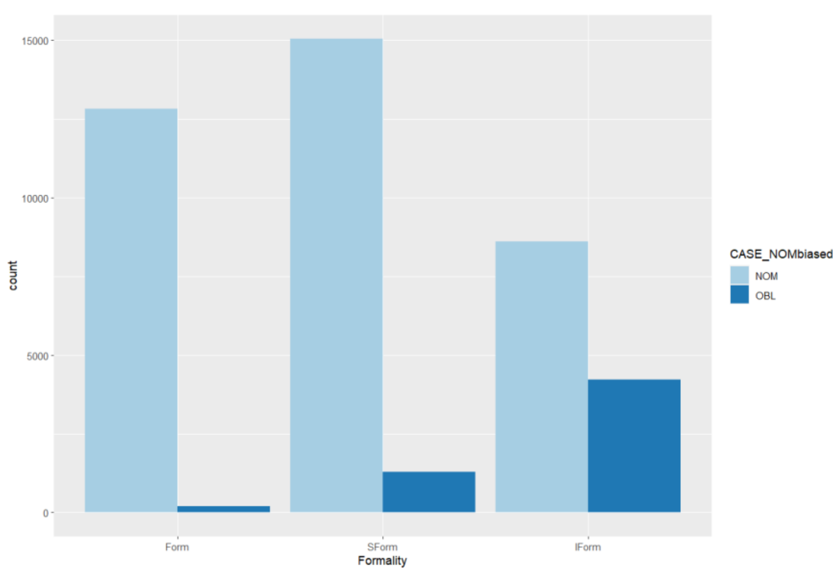


Figure 9.9. Biasing ambiguous case marking towards nominative gives an idea of how frequently oblique subjects are without a doubt encountered. It also shows more clearly the difference between formal, semi-formal and informal language in IGC. The informal language has by far the most examples of unambiguous oblique subjects.

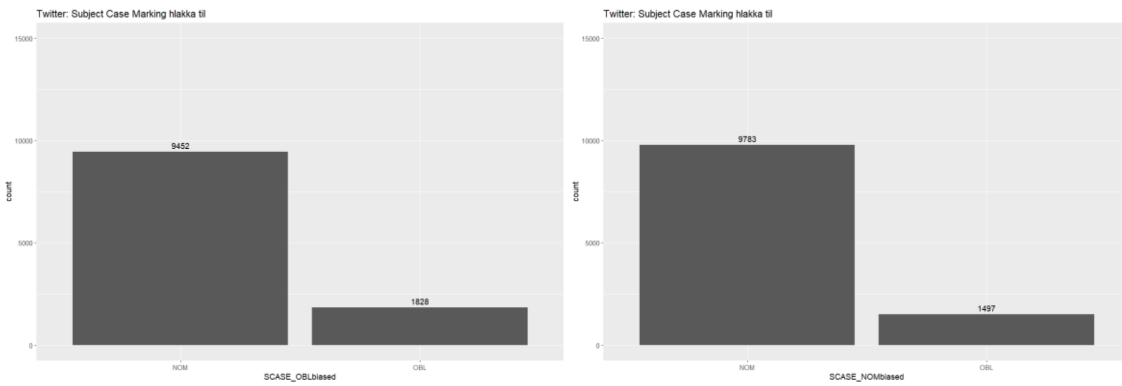


Figure 9.10. Subjects whose case is ambiguous between nominative and oblique are not very common on Twitter.

9.4.4 First person subjects vs. other subjects

Given that prescriptivism generally targets towards the use of first-person subjects of *hlakka til*, and given that previous research (e.g. Svavarsdóttir 1982) has shown a difference in proportion of nominative and oblique with first person subjects as opposed to other types, it is reasonable to take a closer look at the distribution of subject types with the predicate in the relevant sources.

When looking at subject types in IGC and Twitter, it quickly becomes apparent that first person subjects are significantly more common than other subject types, possibly reflecting the fact that individuals like to talk about themselves when employing this predicate online. Furthermore, when examining the usage of nominative and oblique cases with various subject types, it becomes evident that nominative is much more prevalent in first-person subjects compared to other subject types. Although a higher proportion of nominative than oblique with first person subjects is in line with prescriptivism and previous literature, the paucity of oblique subjects is nevertheless interesting.

Figure 9.11 summarizes the number of examples of each type of subject found in IGC. The label MADUR refers to the indefinite pronoun *maður* which is homonymous to the noun *maður* ‘man’. The label UNID covers non-first person subjects whose precise category has not been determined. In Figure 9.12 all subjects that are not first person pronouns have been grouped together, showing even more clearly the difference in choice of case with first person subjects versus other types of subjects.

Since the type of material found in IGC, i.e., whether it was considered formal, semi-formal or informal language, was found to matter in how often oblique subjects were found (see Section 9.4.2), this should also be taken into consideration when projecting the data into time series. As Figure 9.13 shows, first person subjects are almost never found in formal material. The reason is likely normative pressure in combination with copy-editing and proofreading. The situation is slightly different with non-first person subjects as shown in Figure 9.14. Here oblique appears to make up around one third of the examples in formal material. For semi-formal and informal material, they are more frequent than nominative subjects. Note, however, that both Figure 9.13 and Figure 9.14 are based on oblique-biased annotation, where morphologically ambiguous forms are considered to be oblique. Thus, the high number of non-first person oblique examples in informal material is likely due to syncretized forms, for instance personal (often foreign) names that do not have obvious case-marking and the third person pronoun *hann* ‘he’ that looks identical in nominative and accusative.

Just like in IGC, examples from Twitter show a clear difference in case marking between first person subjects and other subjects. Figure Figure 9.15 summarizes the number of examples of each subject type on Twitter. As can be seen, the amount of non-first person subjects is minimal in comparison with first person subjects. Nominative case marking is also more common with first person subjects than other types. This is more obvious in Figure 9.16 where all non-first person subjects have been grouped together. Although the proportion of nominative and oblique is almost the same for these types of subjects, oblique is slightly more frequent. For first person subjects, nominative is way more common than oblique.

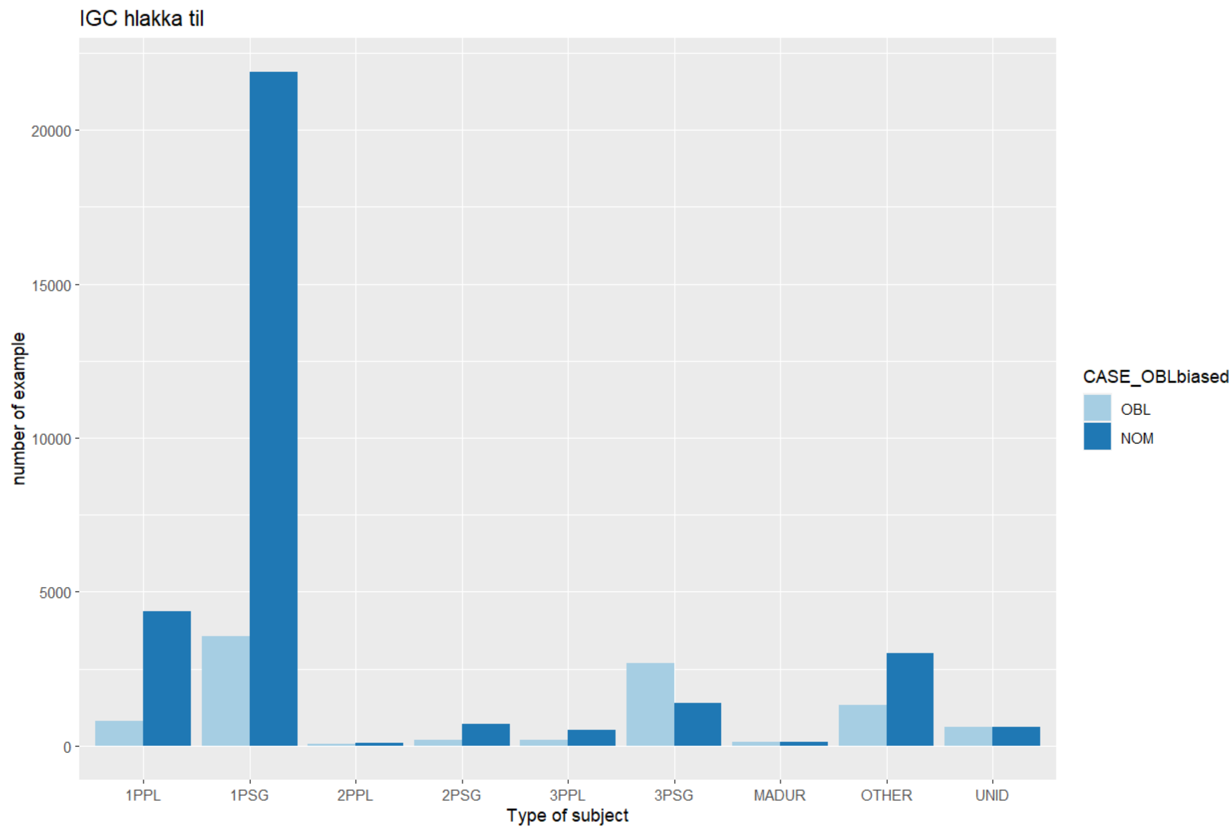


Figure 9.11. An overview of subject-types found in IGC. Only the category OTHER includes nominal subjects. The label UNID refers to non-first person subjects whose exact type has not been recorded.

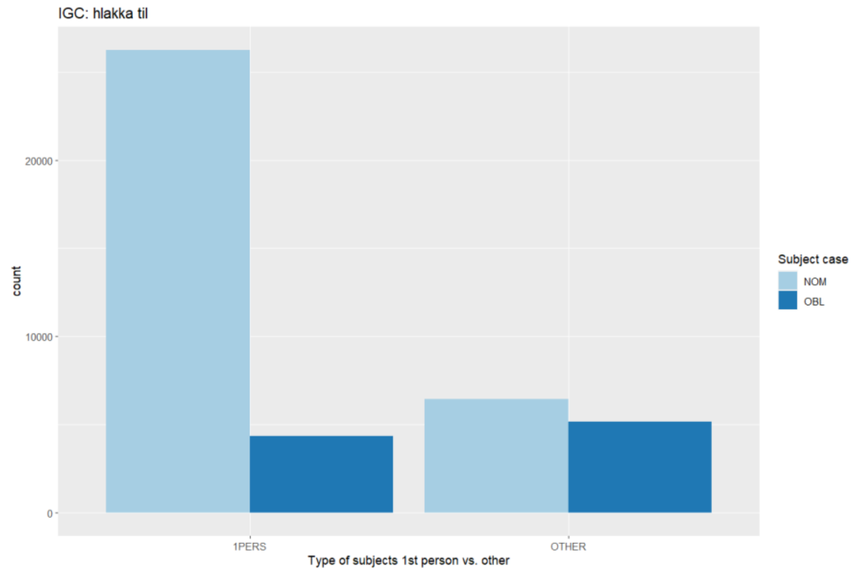


Figure 9.12. First person subjects are less likely to be in an oblique case than other subjects. This might be the result of prescriptivism.

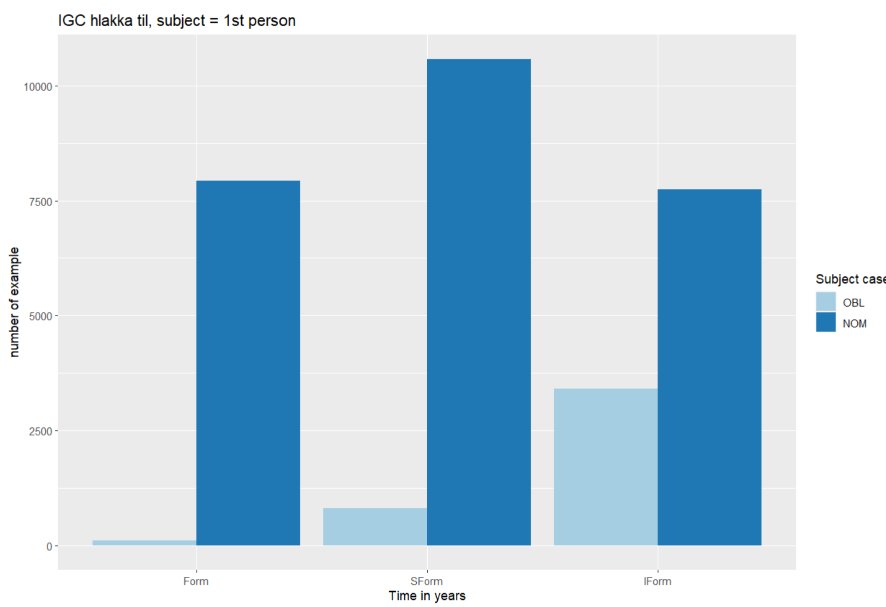


Figure 9.13. Formal material in IGC hardly has any 1st person subjects in an oblique case. Note that the figure is based on oblique-biased annotation.

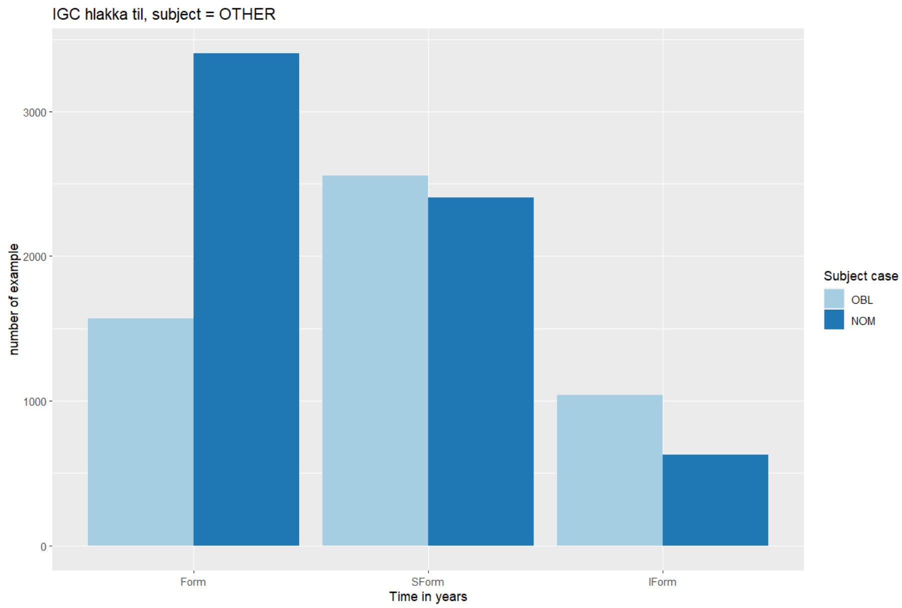


Figure 9.14. Subjects that are not in the 1st person pronouns. Note how oblique is less common than nominative in formal material (newspaper articles) than in semi-formal and informal material. Note that the figure is based on oblique-biased annotation.

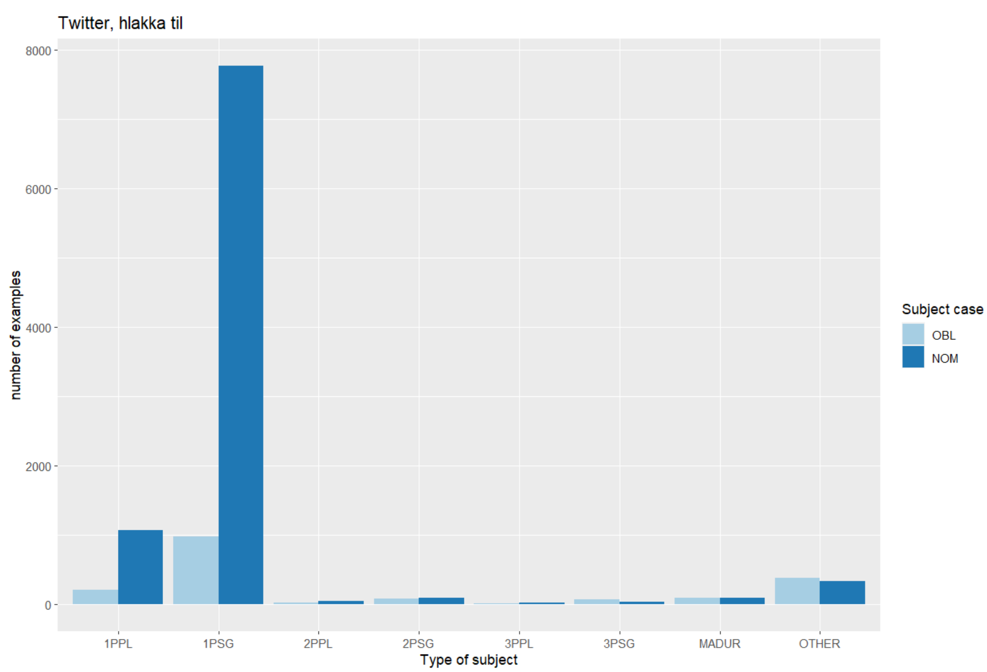


Figure 9.15 An overview of subject-types found on Twitter. Only the category OTHER includes nominal subjects. The label UNID refers to non-first person subjects whose exact type has not been recorded.

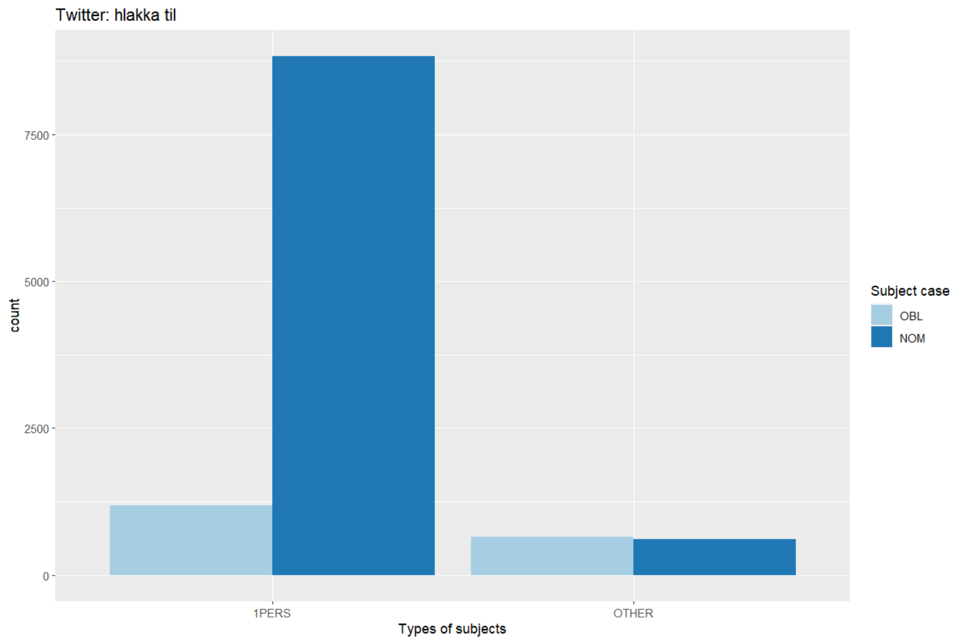


Figure 9.16. Majority of 1st person subjects occur in the nominative case. For other types of subjects, the proportion of nominative and oblique is very similar.

The clear difference between first person subjects and other subjects when it comes to the choice of case suggests that these should be treated separately, i.e., they should be projected into two time series rather than one. This decision is not only based on the appearance of data in IGC and Twitter, but also on previous documentation that has shown that nominative is likelier to surface on first person subjects than other subjects (Svavarsdóttir 1982). It is also in line with normative pressures targeting the first person and speakers therefore being more aware of case marking when using *hlakka til* with first person subjects than other subjects. In addition to taking subject type into consideration when attempting to forecast the propagation of the change, it is also important to be aware of the type of material found in the IGC. Since formal language shows minimal use of oblique with first person pronouns (likely due to prescriptivism, copy-editing and proofreading), this data might skew the predictions. Due to this, only semi-formal and informal sources from IGC will be taken into consideration. However, before taking a closer look at the individual time series used for forecasting, it is worth explaining briefly how the data is projected into time series (Section 9.3.5).

9.4.5 Projecting data into regular time series

So far, data from IGC and Twitter has been treated as coming from a single uniform period (Twitter 2009–2022, IGC 1999–2009). This data can be converted into regular time series. For doing this, oblique-biased annotation was used. Data from IGC was converted into yearly time series and data from Twitter into quarterly time series, cf. Figure 9.17. The reason for choosing quarterly series for Twitter was that the data only covers 14 years which means that yearly time series would barely meet the minimum requirements of number of observations for a time series analysis. Quarterly series on the other hand results in 56 observations.

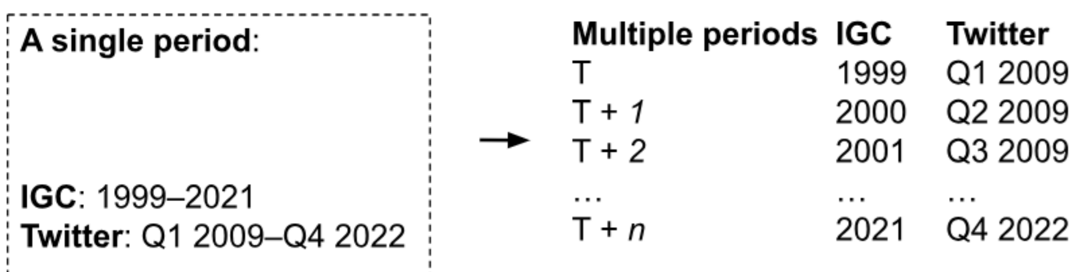


Figure 9.17. Taking data from a period of 14–22 years and creating yearly (IGC) and quarterly (Twitter) time series.

Note that the term *observation* refers to the proportion of oblique subjects at a given point in time, whether it be a quarter or a year. Behind each observation are multiple examples.

Figures 9.18 to 9.21 summarize nominative and oblique case marking with *hlakka til* over time. The figures include all types of subjects, i.e., 1st person subjects and other subject types. In the case of data from IGC, the proportion of nominative and oblique is shown for each year. For data from Twitter data, proportion of nominative and oblique is shown for each quarter of the year. The Visualization is done with bar plots. For visualization of the time series used for forecasting a line plot is used; these are shown in Section 9.5.

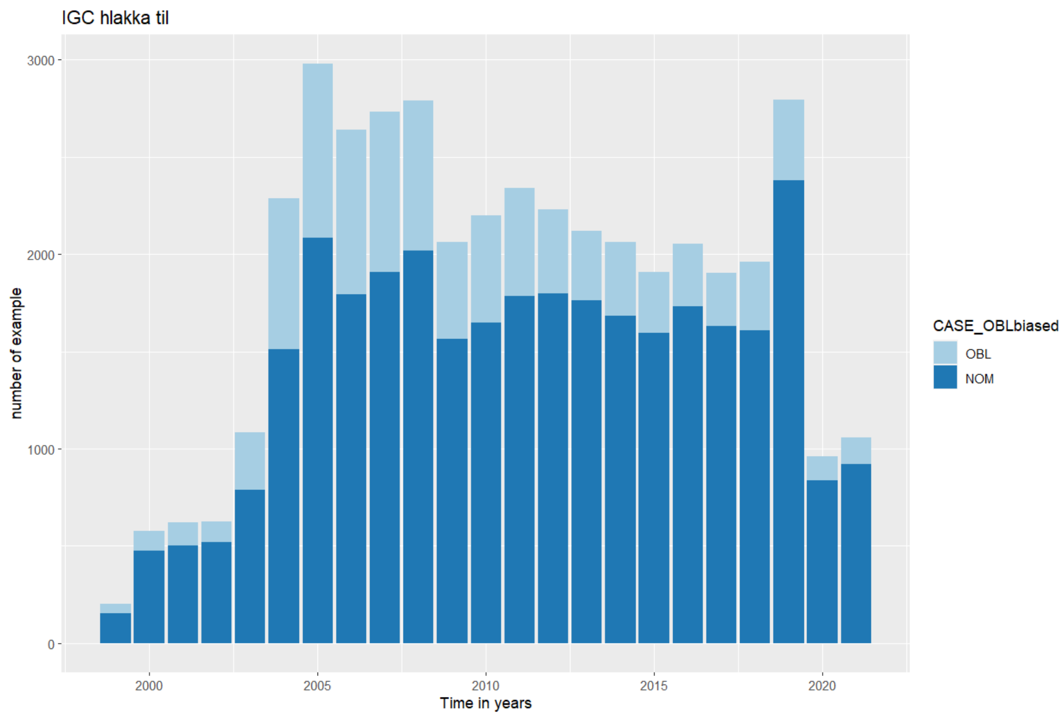


Figure 9.18. Overview of nominative and oblique subjects with *hlakka til* in IGC over time. The figure includes all types of texts (formal, semi-formal and informal), all types of subjects, i.e., 1st person subjects and other subjects, and is based on oblique-biased annotation.

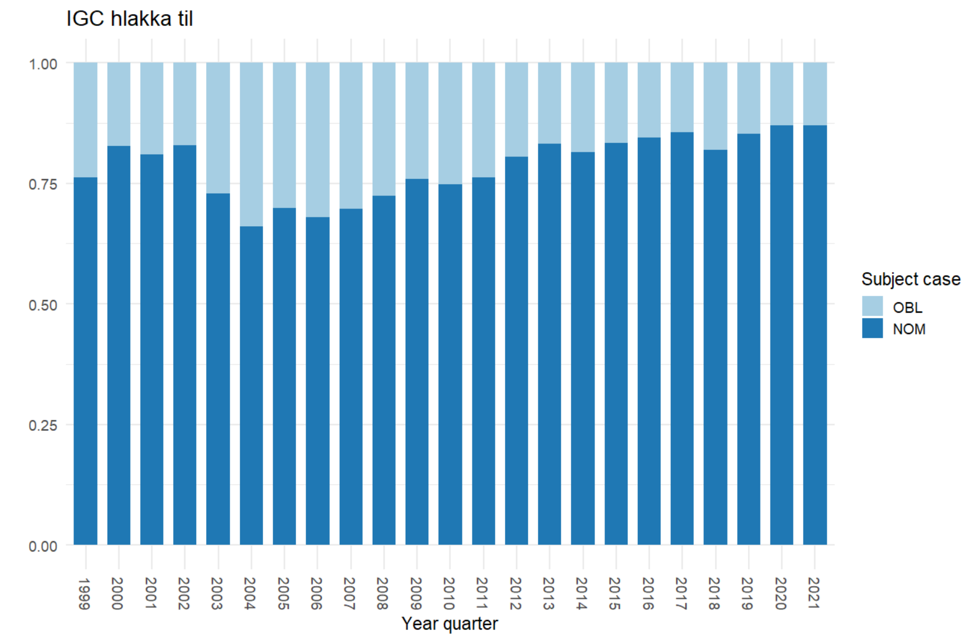


Figure 9.19. Overview of the proportion of nominative and oblique subjects with *hlakka til* in IGC over time. The figure includes all types of texts (formal, semi-formal and informal), both types of subjects (1st person and other subjects) and is based on oblique-biased annotation.

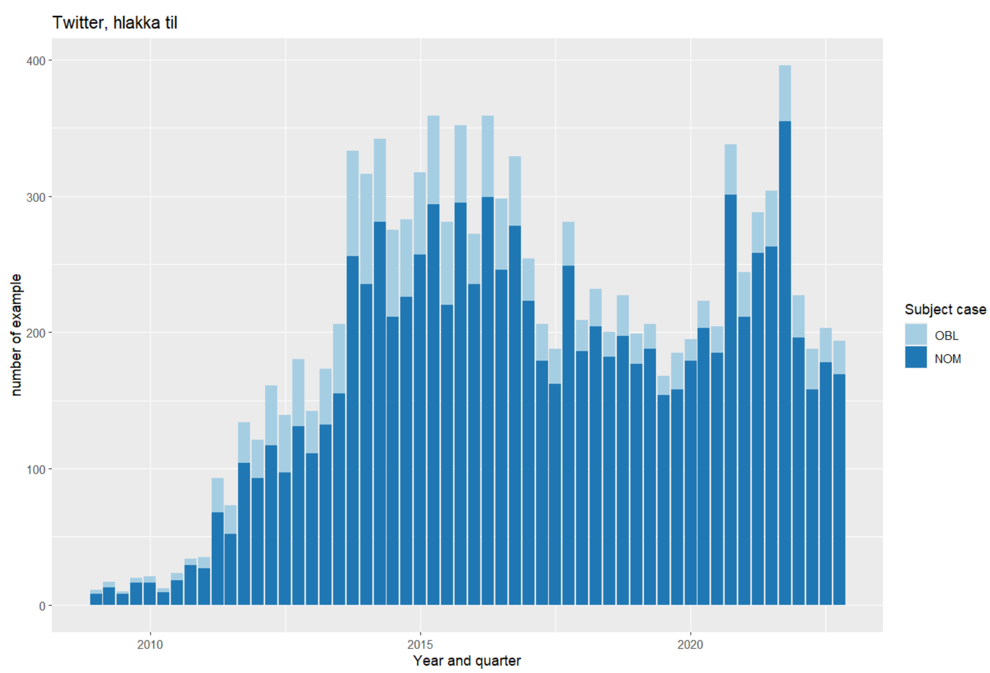


Figure 9.20. Overview of nominative and oblique subjects with *hlakka til* on Twitter over time. The figure includes both types of subjects (1st person and other subjects) and is based on oblique-biased annotation.

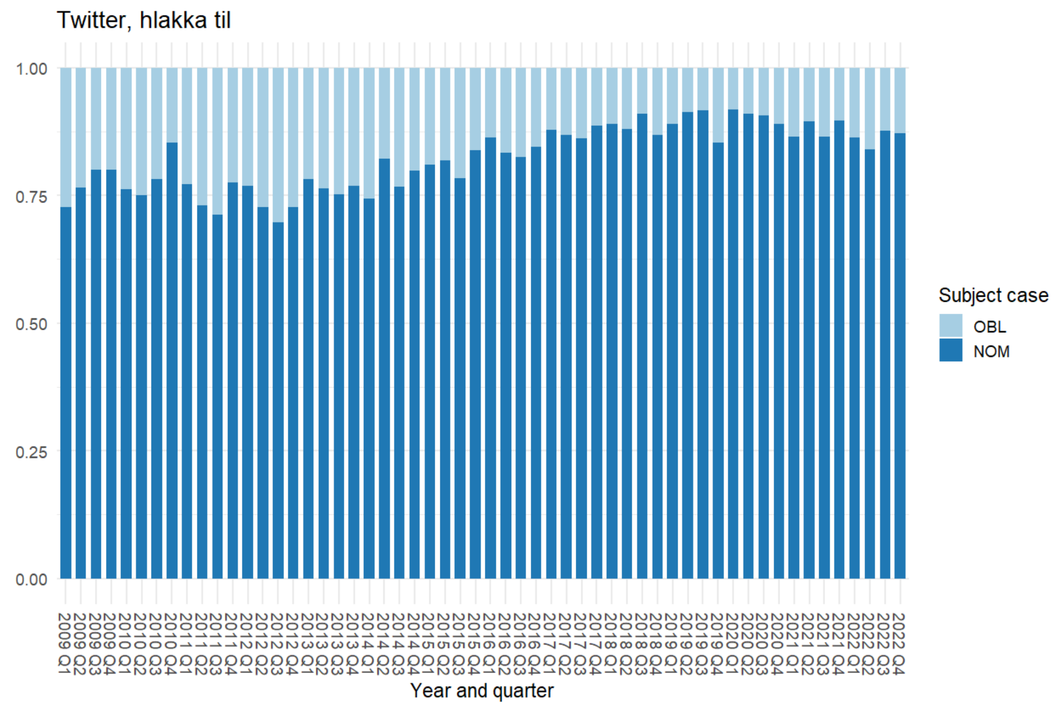


Figure 9.21. Overview of the proportion of nominative and oblique subjects with *hlakka til* on Twitter over time. The figure includes all types of subjects, i.e., 1st person subjects and other subjects, and is based on oblique-biased annotation.

By glancing at the figures in Figure 9.18–Figure 9.21, two things may be noted. First, there is overall less data in the earlier periods, i.e., before 2004 in IGC and before 2012 on Twitter. Second, it seems as if nominative case marking is generally increasing with in uses of *hlakka til*. This is somewhat unexpected in light of oblique case marking having been noted to be on the rise (see discussion in Section 9.3.2), but might be explained in terms of prescriptivism (see also Section 9.3.3). For the data presented in Figure 9.18–Figure 9.21 above, no distinction is made between different types of subjects, i.e., 1st person subjects are treated together with other subjects. This might be problematic as prescriptive discussions on *hlakka til* tend to focus on uses of the *hlakka til* with a 1st person subject so these might be responsible for diminishing oblique over time. Up to the point of speakers being aware of the use of case with the predicate, they may be more alerted to uses including 1st person pronouns. Another explanation might be that children have started to acquire nominative case marking with *hlakka til* when the subject is a 1st person pronoun, but revert to different case marking for other instances. Whatever the reason is, it remains to be seen whether nominative is indeed increasing with 1st person subjects.

Given the difference in proportion of nominative versus oblique case marking in first person subjects and other kinds of subjects (see also Section 9.4.4), it is reasonable to treat these separately. A notable downside of doing so is that the majority of examples from both IGC and Twitter have a first person subject, resulting in time series for subjects that are not in the first person being based on relatively few observations. For the total 42,192 of IGC examples, 30,591 (around 72.5%) have a first person subject while 11,601 have a subject of any other type (around 27.5%). As for Twitter, 10,102 examples (around 88.8%) have a 1st person pronoun, while only 1,260 examples (around 11.2%) have a different type of subject. It is self-evident that spreading 1,260 examples over a 12-year period, with each year having four quarters, results in each observation in the time series being based on relatively few examples. However, treating first person subjects separately from other types of subjects is reasonable and is done in the section on forecasting. The time series and their patterns in time series used for forecasting are discussed more thoroughly in Section 9.5.

9.5 Time series analysis and forecasting

9.5.1 The time series

Having described the data obtained from both IGC and Twitter in Section 9.4, it is appropriate to take a closer look at the time series used for forecasting. Four time series are taken into consideration, two for each data source. These included information on the proportion of oblique subjects when the subject is a 1st person pronoun, and information on the proportion of oblique subjects where the subject is not a 1st person pronoun.

Data from semi-formal and informal sources in IGC (based on a total of 29,176 examples) was projected into two yearly time series, one focusing on 1st person subjects (based on 22,550 examples) and the other on other types of subjects (6,626 examples). These covered the period from 1999 to 2021, both years included, giving a total of 23 observations sequenced in time for each series. Data from Twitter was also projected into two quarterly time series, one focusing on 1st person subjects and the other on other subjects. These covered the period from Q1 2009 to Q4 2022, giving a total of 56 observations sequenced in time. Table 9.13 summarizes relevant factors for each of the fore time series, i.e., which subject-case annotation it is based on, the number of observations for each of the series, and number of examples the series are based on. The four time series are shown in Figures 9.22; The y-axis shows proportion from 0 (0%) to 1 (100%).

Overview of data for *hlakka til*						
Source	**Period covered**	**Observations**	**Annotation**	**N examples for subject type**		
				1st person	*other*	*total*
IGC	1999–2021	23 yearly	OBLbiased	22,550	6,626	29,176
Twitter	Q1 2009–Q4 2022	56 quarterly	OBLbiased	10,020	1,260	11,280

Table 9.13. Summary of the time series, the annotation they are based on, the number of observations and number of examples behind each series. As an example, the IGC time series on subjects of *hlakka til* that are 1st person pronouns contains 22,550 examples.

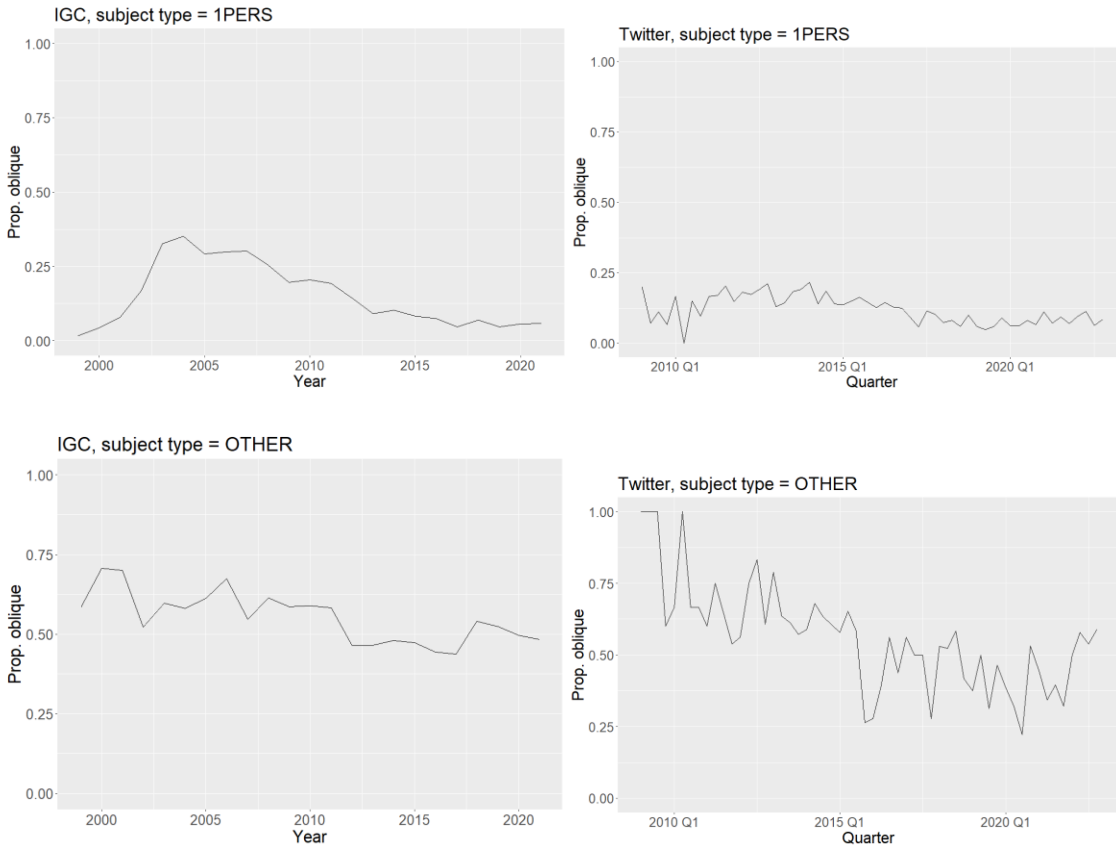


Figure 9.22. Four time series, showing the proportion of oblique subjects with *hlakka til* over time. Data from Twitter contains 56 observations for each series, data from IGC 23. The four time series are shown on the same scale, i.e., the y-axis shows the proportion between 0 (0%) and 1 (100%).

When glancing at the series in Figure 9.22, the initial few observations in each series may appear slightly odd, i.e., they seem to consist of more "noise" than the rest of the series. Recalling that the number of examples behind the first few observations are not as many as for the rest of the series, this is perhaps not surprising. In order to minimize the effect of early observations, these were not taken into further consideration for analysis, model fitting or forecasting. In other words, these were omitted. This was done in the following way. Examples from Twitter are relatively few until around 2012 so Q1 2009 – Q4 2011 were removed from the time series, leaving 44 observations stretching from Q1 2012 to

Q4 2022. For IGC, examples are relatively few until around 2003, so years 1999–2002 were removed, leaving 19 observations stretching from 2003 to 2021. The shortened time series, henceforth referred to as complete time series, are depicted in Figure 9.23. A summary of these series is provided in Table 9.14.

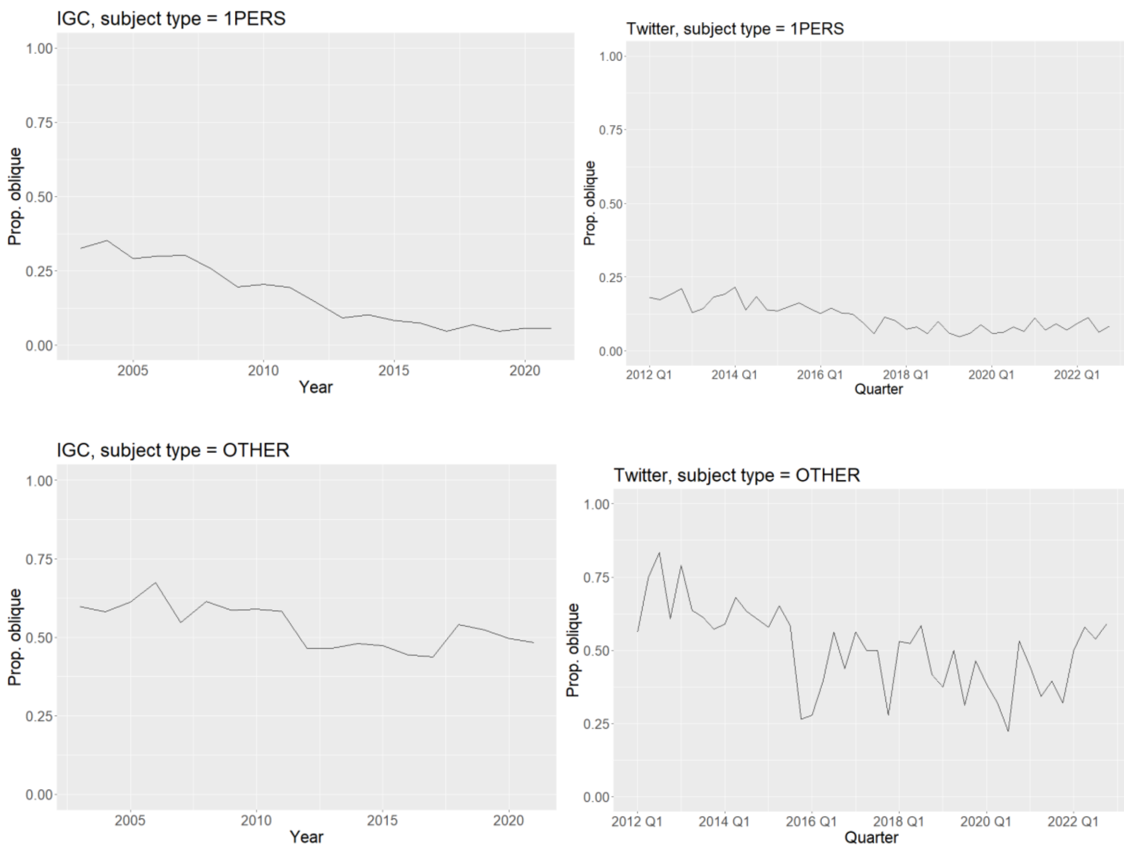


Figure 9.23. The four time series once initial observations have been removed. Data from Twitter now consists of 44 observations for each series, and data from IGC 19.

Overview of data for *hlakka til*						
Source	**Period covered**	**Observations**	**Annotation**	**N examples for subject type**		
				1st person	*other*	*total*
IGC	2003–2021	19 yearly	OBLbiased	22,198	6,376	28,574
Twitter	Q1 2012–Q4 2022	44 quarterly	OBLbiased	9,629	1,177	10,806

Table 9.14. Overview of the time series (henceforth the complete time series), the annotation they are based on, the number of observations and number of examples behind each series. For instance, the IGC time series on subjects of *hlakka til* that are 1st person pronouns contains 22,198 examples.

In general, the overall proportion of oblique subjects varies depending on whether the subject is a 1st person pronoun or not, and whether the data is from Twitter or IGC. When the subject is a first person pronoun, the proportion of oblique is generally lower than when the subject is of some other type. In IGG, first person oblique subjects with *hlakka til* are around 35% in 2003 and drop to little below 6% in 2021. For Twitter, first person oblique subjects with *hlakka til* are around 20% in 2012 and dip below 10% in 2022. As for other types of subjects, IGC shows a proportion between ca. 44% and 67% with a slight decrease over time, and Twitter shows a proportion anywhere from ca. 20% to just above 80% which also appears to decrease over time. It should be noted that the number of examples behind each observation is way lower for series with non-first person subjects than with first person subjects. It appears that individuals use *hlakka til* more frequently in the 1st person than other persons. Note also that the decrease of oblique over time is not a fully expected behavior as oblique subjects are generally expected to increase over time. However, the decrease of oblique in favor of nominative might be understandable in light of prescriptivism which tends to focus heavily on 1st person subjects. The general direction of change as witnessed by the time series and expectations rising from forecasting are further discussed in Section 9.6.

The four time series, i.e., two for each source (IGC and Twitter) and one for 1st person subjects with *hlakka til* and the other for other types of subjects, were then split into training and test sets. Under normal circumstances, 80% of the observations in a time series is used to train a model and 20% to test how well the model perfors (Hyndman & Athanasopoulos 2021:135). However, since the time series dealt with here are already on the shorter side, setting this number of observations aside for testing models will result in models being fitted to fewer observations further back in time. Some of the series exhibit a negative trend early on, until around 2019 and 2020, but then show a slight increase in the proportion of oblique subjects in more recent times. If the amount of data used to train the model mostly includes the negative trend, the model will naturally highlight that feature and assume continuation of it into the future. Finally, since the goal is to produce a forecast for future time periods, it may be beneficial to include as much information as possible in the training set. For data from Twitter, the training set contained 40 observations (= 10 years) or c. 91% of the data, and the test set contained 4 observations (= 1 year) or c. 0.9% of the time series. For IGC, the training set contained 17 observations (= 17 years) or c. 89% of the time series, and the test set contained 2 observations (= 2 years) or c. 11% of the data. A summary of the training period and test period for each time series is shown in Table 9.15.

Source	**Series**	**Training period**		**Test period**		**Whole series**	
	subject type	*Time*	*N observations*	*Time*	*N observations*	*Time*	*N observations*
IGC	1PERS	2003–2019	17 (c. 89%)	2020–2021	2 (c. 11%)	2003–2021	19 (100%)
IGC	OTHER	2003–2019	17 (c. 89%)	2020–2021	2 (c. 11%)	2003–2021	19(100%)
Twitter	1PERS	Q1 2012– Q4 2021	40 (c. 91%)	Q1 2022–Q4 2022	4 (c. 9%)	Q1 2012– Q4 2022	44 (100%)
Twitter	OTHER	Q1 2012– Q4 2021	40 (c. 91%)	Q1 2022–Q4 2022	4 (c. 9%)	Q1 2012– Q4 2022	44 (100%)

Table 9.15. A summary of the training period and test period for each time series. Taking into account specific properties of the series, such as how short they are and the patterns they show, the training set only made up about 10% of the overall series.

The procedure that was followed for each of the time series used to forecast was as discussed in Chapter 7, Section 7.2.1, namely to specify which model was to be used (define a model), fit the model to the training data (train the model, estimate parameters), and check the performance of the model (evaluate) by looking at the residuals and (when applicable) forecast errors. Finally, using a model where the residuals are within acceptable limits a forecast was produced.

Forecasts for future periods, i.e., periods after 2021 for IGC data and periods after Q4 2022 for the Twitter data were done in two ways, i.e., i) by using a model fitted to the training data and projecting into the future, and ii) by fitting the same type of model used for the training data to the whole dataset and then projecting into the future. Needless to say, even though the same type of model is fitted to the training set and the whole series, the exact parameters of the models may be different. Finally, one forecast is made with a decomposed series, applying a model that was considered appropriate for the trend. Having provided a general overview of the four time series, each of them will now be discussed in turn including information on their properties, the models fitted and the results.

9.5.2 IGC 1st person subjects

The IGC time series containing information on case marking with 1st person pronoun subjects consists of 19 (yearly) observations from 2003 to 2021. The series shows a negative trend, with oblique subjects making up more than 30% of the examples in 2003–2004, and around 5% in 2020–2021. The mean of the series is 0.1679382. There are no obvious outliers in the series and since it consists of yearly observations there is no seasonality. As suggested by the autocorrelation function (ACF), cf. Figure 9.24, as well as the Ljung-Box (statistics 56.1, p-value 0.0000000201, lag = 10) statistics, the series is not a white noise series. The series is non-stationary and would have to be differenced once to make it stationary (KPSS statistics 0.699, p-value = 0.0137). An STL decomposition (see Chapter 7, Section 7.2.2) of the series (trend window = 13), where each observation is assumed to consist of a trend component and a remainder component (cf. Figure 9.24), shows the negative trend clearly, with the strength of the trend being 0.976, the linearity -0.426 and the curvature 0.0726. The decomposition also shows the small random noise, which lies between 0.03 and -0.03. Note that the gray boxes in each of the panels in Figure 9.24 show the same scale, i.e., they are of the same "size".

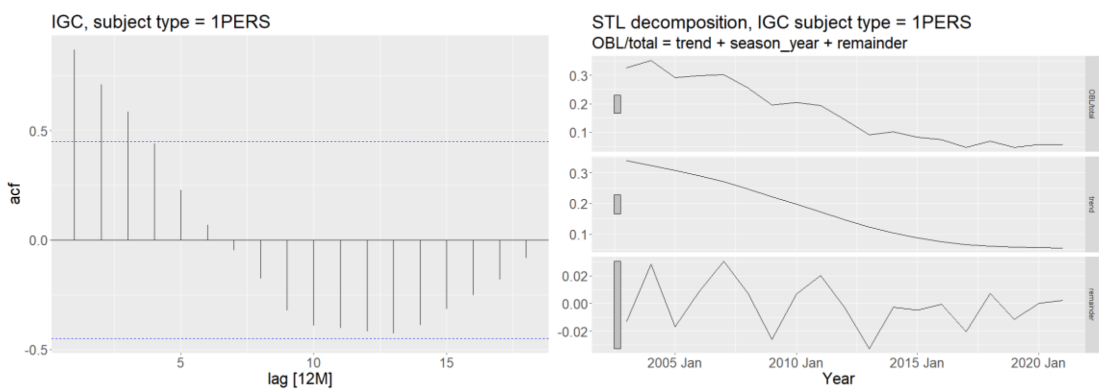


Figure 9.24. ACF and STL of the IGC series for 1st person subjects. The series has a clear negative trend and very little random noise (between 0.03 and -0.03).

Before choosing and fighting forecasting models, the series was divided into a training and test set. As noted earlier (Section 9.5.1), the training set contained 17 observations from 2003–2019, and the test set 2 observations from 2020 – 2021.

Three simple models were fitted to the training data, i.e., i) a Naïve model which assumes that all future observations will be the same as the last recorded observation in the series, ii) a Mean model which assumes that all future values will be equal to the mean of the whole series, and iii) a Drift model where changes over time are assumed to be average change in the historical data (see further in Chapter 7, Section 7.2.3). The point-predictions of each of the models for the test period are shown in Figure 9.25 along with the whole series. The model relying on the mean of the training data appears to give the least accurate prediction, suggesting dative subjects should make up about 18% of examples in 2020–2021 when the observed proportion is closer to 6%. The Drift model and the Naïve model do better, predicting that somewhere between 1–5% of 1st person subjects in the test period will be in an oblique case. A summary of the fit of each of the models to the training data as well as how accurate the point forecasts are, is provided in Table 9.16. Of the three models, the Naïve one appears to give the most accurate point predictions.

IGC, subject type = 1PERS							
Model	source	type	RMSE	MAE	MAPE	MASE	RMSSE
Mean	IGC	Training	0.104	0.0931	86.6	3.42	3.11
Naïve	IGC	Training	0.0334	0.0272	23.1	1	1
Drift	IGC	Training	0.0284	0.0247	17.8	0.907	0.852
Drift	IGC	Test	0.0375	0.0364	64	1.34	1.12
Mean	IGC	Test	0.124	0.124	219	4.56	3.72
Naïve	IGC	Test	0.0102	0.0102	18	0.375	0.306

Table 9.16. A comparison of the accuracy of the fit and the point forecast of each of the three simple models.

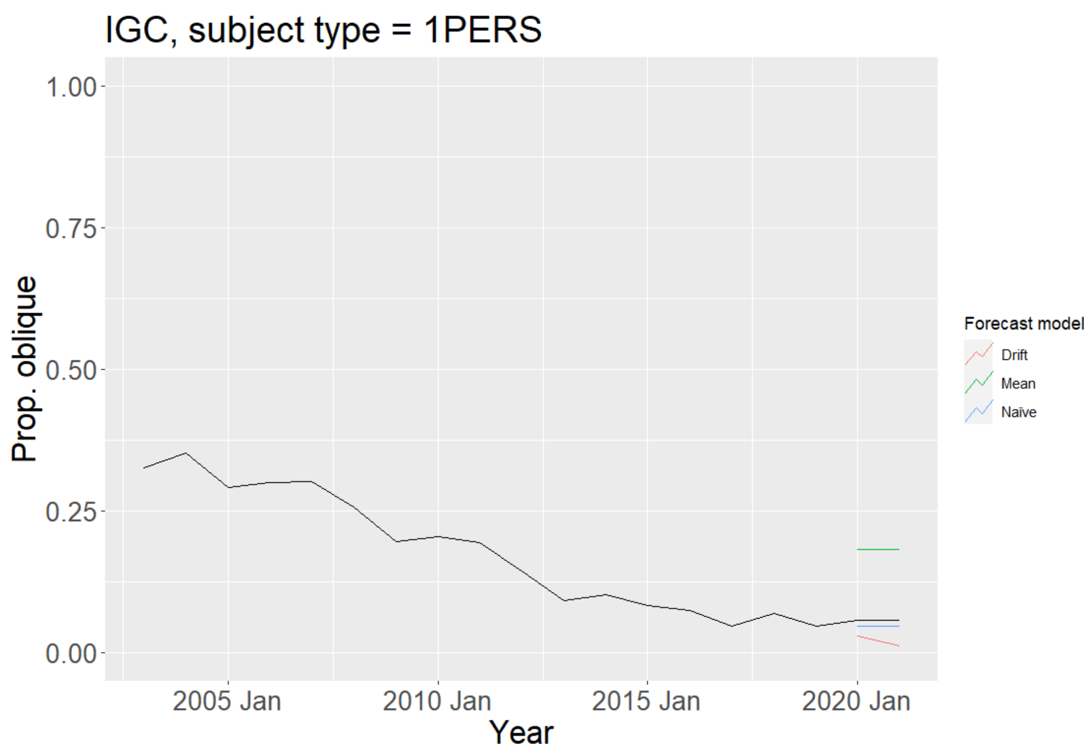


Figure 9.25. Three simple forecasting methods used to predict the data in the test set, i.e., the two-year period from 2020–2021. The Naïve model appears to have the most accurate predictions.

In addition to the three simple forecasting models mentioned above, an ETS and an ARIMA model were also fitted to the training data. Seeing that the time series tends towards zero, a log transformation was used to keep point predictions and prediction intervals within a positive range. The type of ETS and ARIMA model along with initial states was determined using the ETS() and ARIMA() functions from the fable package (O'Hara-Wild, Hyndman & Wang 2021). While the ETS() function selects a model by minimizing the AICc (Hyndman & Athanasopoulos 2021:254), the ARIMA() function relies on the Hyndman-Khandakar algorithm for automatic ARIMA modeling (Hyndman & Athanasopoulos 2021:286).

An automatic selection of an ETS model resulted in Holt's linear method with additive errors, ETS(A,A,N). The smoothing parameters were alpha = 0.3277967 and beta = 0.0001000001, with the initial states l = -0.8326587 and b = -0.1288876. The residuals are within acceptable limits with the mean of the innovation residuals being around zero, or 0.001187903 and plausibly interpreted as white noise (Ljung-Box statistics = 8.89 with p-value = 0.543, lag = 10). The autocorrelation coefficients are within the significance levels but the residuals are not fully normally distributed which might affect prediction intervals slightly.

The automatic selection of ARIMA resulted in ARIMA(0,1,1) with drift and coefficients ma1 = -0.6322 (s.e. 0.2618) and constant = -0.1286 (s.e. 0.0203). Point forecasts with prediction intervals are shown in Figure 9.26. Aside from the residuals not being fully normally distributed (slightly skewed on the positive side and containing a negative outlier) which might affect prediction intervals, they appear within acceptable limits. The residuals are plausibly interpreted as white noise (Ljung-Box statistics = 11.4 with p-value = 0.325, lag = 10) with mean 0.01552396 and correlation coefficients are within expected limits.

Even though different types of models are being used, the predictions for the training period are very similar. In both instances, the models suggest that the proportion of oblique subjects is diminishing over time and that it is likely to be below 4.5% in 2020-2021. Observed values are 5.6% (2021) and 5.7% (2021).

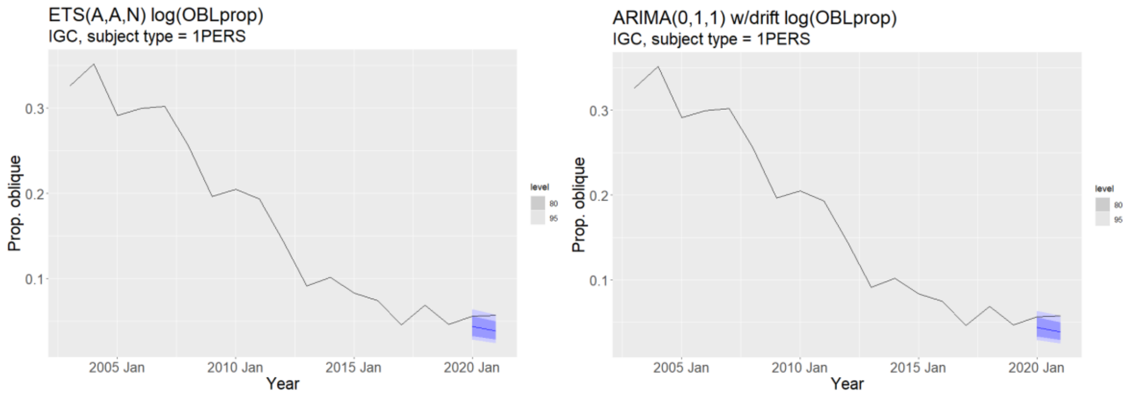


Figure 9.26. Prediction for the test period 2020-2021 from an ETS(A,A,N) model and ARIMA(0,1,1) w/drift. The series has been zoomed-in on.

A summary of the accuracy of the fit of the ETS(A,A,N) and ARIMA(0,1,1) model and the accuracy of the point forecasts for the test period is given in Table 9.17. Although the ETS(A,A,N) model appears to give slightly more accurate point predictions than ARIMA(0,1,1), it nevertheless performs worse than the naïve model discussed above. When evaluating the accuracy of the distributional forecasts using quantile scores, Winkler scores, Continuously Ranked Probability Scores and Scale-free comparison using skill scores, it turns out that no method performs better than the Naïve model when skill scores (based on crps) are considered. This suggests that the Naïve model might be the most appropriate for forecasting, although keeping in mind that the evaluations are based on a very small test set.

IGC, subject type = 1PERS							
Model	source	type	RMSE	MAE	MAPE	MASE	RMSSE
ETS(A,A,N) log(OBLprop)	IGC	Training	0.0302	0.0233	14.6	0.856	0.906
ARIMA(0,1,1) w/drift log(OBLprop)	IGC	Training	0.0291	0.0218	14.2	0.801	0.873
ETS(A,A,N) log(OBLprop)	IGC	Test	0.0157	0.0154	27.2	0.567	0.472
ARIMA(0,1,1) w/drift log(OBLprop)	IGC	Test	0.0159	0.0156	27.4	0.573	0.476

Table 9.17. A comparison of the accuracy of the fit and the point forecast of the ETS(A,A,N) and ARIMA(0,1,1) model. The ETS model appears to have a slightly worse fit but better prediction accuracy.

IGC, subject type = 1PERS						
Model	source	type	qs probs = 0.1	winkler (level 80) level 80	crps	skill
Mean	IGC	Test	0.00335	0.282	0.0764	-6.31
Naïve	IGC	Test	0.0124	0.103	0.0105	0
Drift	IGC	Test	0.0165	0.0925	0.022	-1.1
ETS(A,A,N) log(OBLprop)	IGC	Test	0.00519	0.0598	0.0112	-0.0756
ARIMA(0,1,1) w/drift log(OBLprop)	IGC	Test	0.00516	0.0645	0.0114	-0.0949

Table 9.18. A comparison of the accuracy of the forecast intervals of the three simple models and the ETS(A,A,N) and ARIMA(0,1,1) model. The skill scores suggest that the naïve method might be the most appropriate for forecasting. Note, however, that the statistics are based on a very small test set (2 periods).

Although the Naïve model appears to be the method best suited for forecasting the IGC time series for 1st person subjects, it must be admitted that the evaluation is based on a very small test set. Additionally, the quantile scores and the Winkler scores for the ETS(A,A,N) and ARIMA(0,1,1) models are lower than for the Naïve model, indicating a better estimate of the quantile and a narrower interval. It is, therefore, tempting to use these to generate forecasts, also keeping in mind that a forecast using the Naïve method will simply predict all future observations to be equal to the last observation. In the case of relying on the whole series, this would be 5.7%. In Figure 9.27 an ETS(A,A,N) and ARIMA(0,1,1) have been fitted to the whole series and used to produce a 20-step ahead forecast. The ETS(A,A,N) model predicts a diminished use of oblique first person subjects in IGC data, or from almost 4.6% in 2022 to ca. 0.6% in 2041. The ARIMA(0,1,1) model gives extremely similar results with oblique first person subjects making up around 4.6% (2022) to 0.65% (2041). Restricting the forecasts to positive predictions (above 0%) results in prediction intervals becoming increasingly smaller over time.[114]

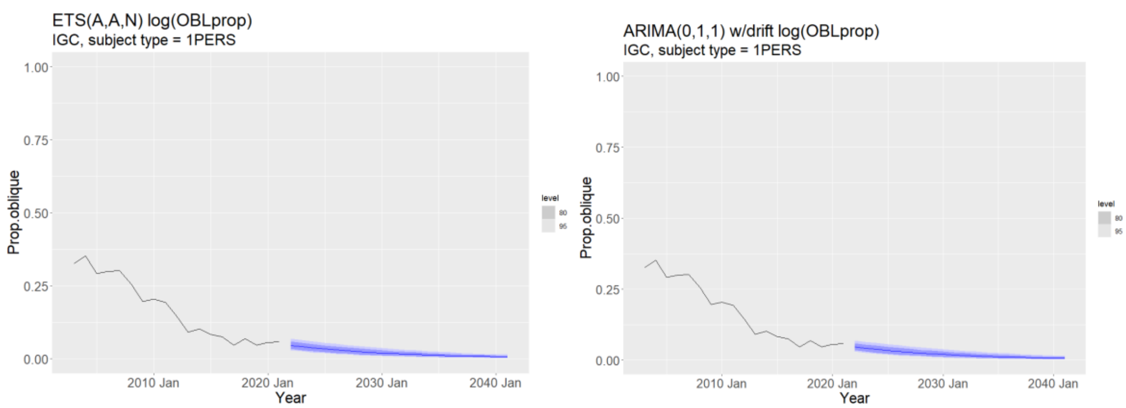


Figure 9.27. A 20-step ahead forecast using ETS(A,A,N) and ARIMA(0,1,1) with drift. Under both scenarios, a diminishing use of oblique 1st person subjects with *hlakka til* are expected.

In reality, producing a 20-step ahead forecast using time series models might not make much sense. The further in the future one uses a model to predict, the more inaccurate (and incorrect!) the predictions are likely to be, especially when there are sudden changes in the data like in the current time series. However, the predictions have a value in that they give rise to expectations. Under the current scenario, the proportion of 1st person oblique subjects is expected to continue to go down over time. The prediction intervals are quite narrow, staying below 5%.

[114] If the forecast had not been restricted to positive predictions, future observations would likely have been forecasted to show negative numbers.

The ETS model used above gave slightly better results than the ARIMA model when various figures were compared, cf. Table 9.17 and Table 9.18. The distribution of residuals were also slightly better. For the sake of experimenting, an ETS(A,A,N) model was fitted to the trend of the time series, which in turn was obtained through STL decomposition. Since there was very little noise in the series to begin with (see the remainder component in Figure 9.24), the results were very similar to fitting a model to the raw series. Figure 9.28 shows a 20-step ahead forecast for the trend. A potential problem with using the trend of the series rather than the raw time series is that we might be removing a part of the series that contains information about future development, i.e., the raw series was already quite smooth so the remainder component might contain information on what to expect of error terms etc.

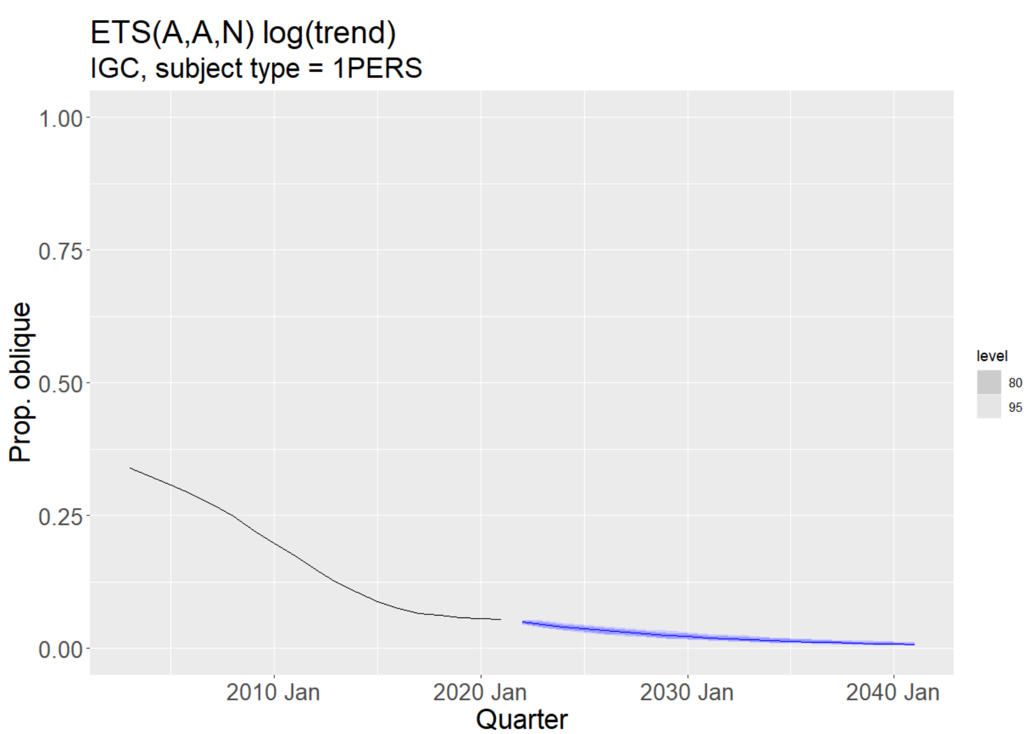


Figure 9.28. An ETS(A,A,N) model fitted to the trend of the complete time series decomposed with a STL method.

The overall picture that emerged from studying the proportion of first person oblique subjects with *hlakka til* in IGC data over time, is that the use of nominative is increasing. Extrapolating from the observed pattern, we may hypothesize that oblique subjects will make up between 5-6% of attested examples in a similar dataset in the next few years (according to the Naïve model), or that the proportion will decrease over time (the ETS and ARIMA models), potentially reaching almost zero around 2040. Whether or not these are realistic expectations for the future is discussed in Section 9.6.

9.5.3 IGC other types of subjects

The IGC time series containing information on case marking with subjects that are not a 1st person pronoun consists of 19 (yearly) observations from 2003 to 2021. Oblique subjects make up between 43.7% and 67% of attested examples throughout the period of 2003–2021. The series appears to have a an ever so slight negative trend (the strength of the trend being 0.740, the linearity -0.213 and the curvature 0.0425). The mean of the series is 0.5364725 and KPSS test (stat 0.536, p-value 0.0335) suggests that the series is not stationary and would have to be difference once to make it so. The autocorrelation function ACF, cf. Figure 9.29 shows the initial lag outside of the significance level. According to Ljung-Box statistics (statistics = 37.9, p-value = 0.0000392, lag = 10), the series is not a white noise series. An STL decomposition (see Chapter 7, Section 7.2.2) of the series (trend window = 13), shows how the series can be accounted for in terms of a small trend and a small remainder component which is between ca. -0.050 and 0.075, see Figure 9.29. Since the series contains yearly observations there is no seasonality. The gray boxes in each of the panels in Figure 9.29 are of the same scale and show that the remainder component is of a similar size as the trend component.

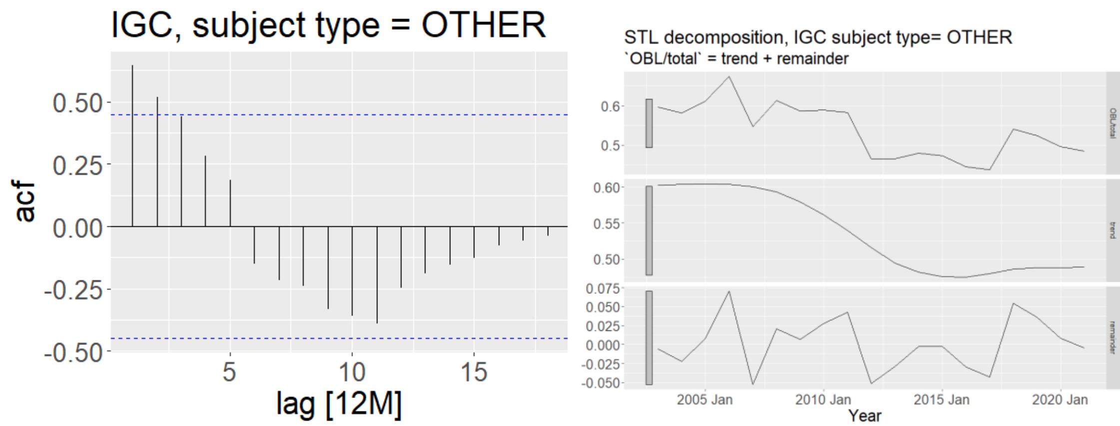


Figure 9.29. ACF and STL of the IGC series for subjects not in the first person. The series does not have a very clear trend and there is some amount of random noise (between 0.077 and -0.050).

Before choosing and fighting forecasting models, the series was divided into a training and test set. As noted earlier (Section 9.5.1), the training set contained 17 observations from 2003–2019, and the test set 2 observations from 2020 – 2021.

Three simple models were fitted to the training data, i.e., i) a Naïve model which assumes that all future observations will be the same as the last recorded observation in the series, ii) a Mean model which assumes that all future values will be equal to the mean of the whole series, and iii) a Drift model where changes over time are assumed to be average change in the historical data (see further in Chapter 7). The point-predictions of each of the models for the test period are shown in Figure 9.30 along with the whole series.

The three models, Naïve, Mean and Drift, yield similar predictions for the test period. They all suggest that the proportion of non-first person oblique subjects in 2020–2021 should be around 51%–54%. The actual observations fall between 48.4% and 49.6%.

A summary of the fit of each of the models to the training data as well as how accurate the point forecasts are, is provided in Table 9.19. The Drift model appears to have given the most accurate point prediction of the three models.

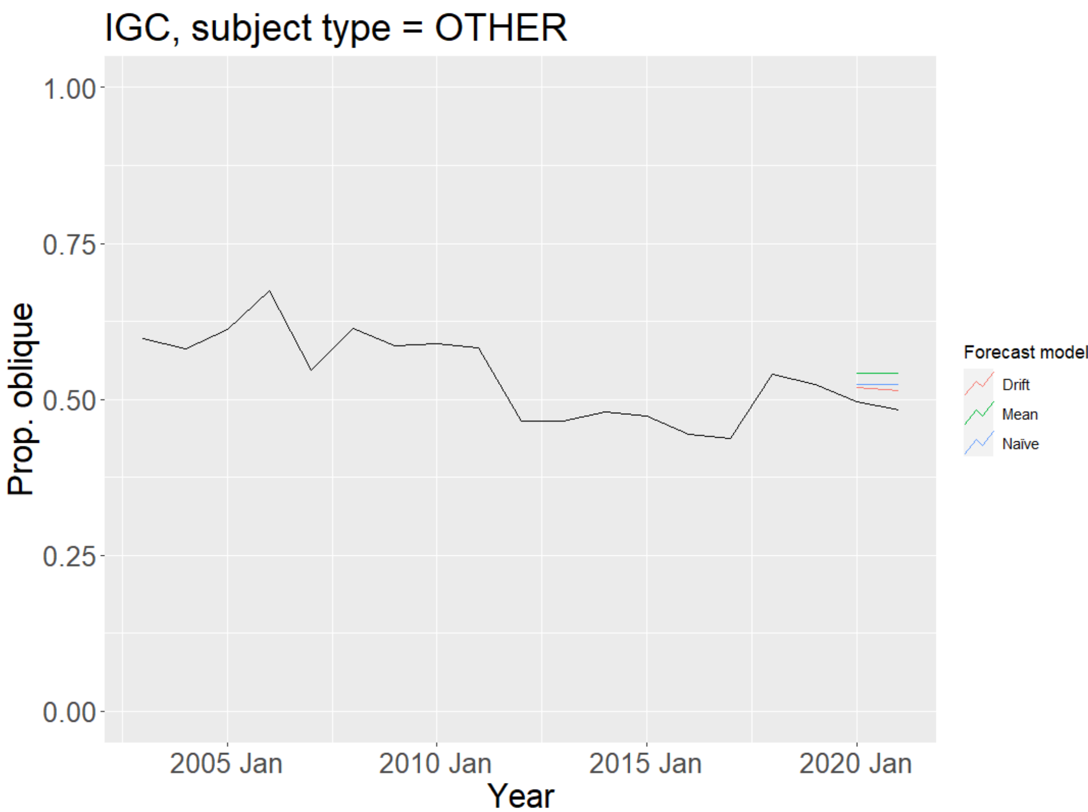


Figure 9.30. Three simple forecasting methods used to predict the data in the test set, i.e., the two-year period from 2020–2021. The three models give very similar results.

IGC, subject type = OTHER							
Model	source	type	RMSE	MAE	MAPE	MASE	RMSSE
Mean	IGC	Training	0.0684	0.0596	11.3	1.5	1.19
Naïve	IGC	Training	0.0572	0.0397	7.35	1	1
Drift	IGC	Training	0.0571	0.0391	7.2	0.986	0.997
Drift	IGC	Test	0.0273	0.027	5.52	0.68	0.476
Mean	IGC	Test	0.0522	0.0518	10.6	1.31	0.912
Naïve	IGC	Test	0.0344	0.0338	6.92	0.853	0.601

Table 9.19. A summary of the fit of the three simple models along with the accuracy of the point forecasts for the period 2020–2021.

In addition to the three simple forecasting models mentioned above, an ETS and an ARIMA model were also fitted to the training data. The type of ETS and ARIMA model along with initial states was determined using the ETS() and ARIMA() functions from the fable package (O'Hara-Wild, Hyndman & Wang 2021). While the ETS() function selects a model by minimizing the AICc (Hyndman & Athanasopoulos 2021:254), the ARIMA() function relies on the Hyndman-Khandakar algorithm for automatic ARIMA modeling (Hyndman & Athanasopoulos 2021:286).

An automatic selection of an ETS model suggested simple exponential smoothing with multiplicative errors, or ETS(M,N,N). The smoothing parameter was alpha = 0.6508007, and the initial state l = 0.5905189. The residuals of the model are acceptable (mean = -0.008986954) and plausibly interpreted as white noise (Ljung-Box statistics = 7.94 with p-value 0.635).

The automatic selection of ARIMA resulted in ARIMA(0,1,0). The residuals are not fully normally distributed (they have a slight tail on the positive side and an outlier on the negative site) which might affect forecast distributions. Otherwise the residuals look acceptable with mean -0.004276208 and they can be analyzed as white noise (Ljung-Box statistics = 10.5 with p-value 0.397, lag = 10). Point forecasts with prediction intervals are shown in Figure 9.31. Even though different types of models are being used, the point forecasts for the test period are very similar. Both models give a "flat" point forecast with ETS(M,N,N) predicting 51.8% oblique subjects and the ARIMA(0,1,0) 52.4%. Observed values are slightly lower or ca. 39.4% (2020) and 38.2% (2021) but are nevertheless inside the 80% confidence interval.

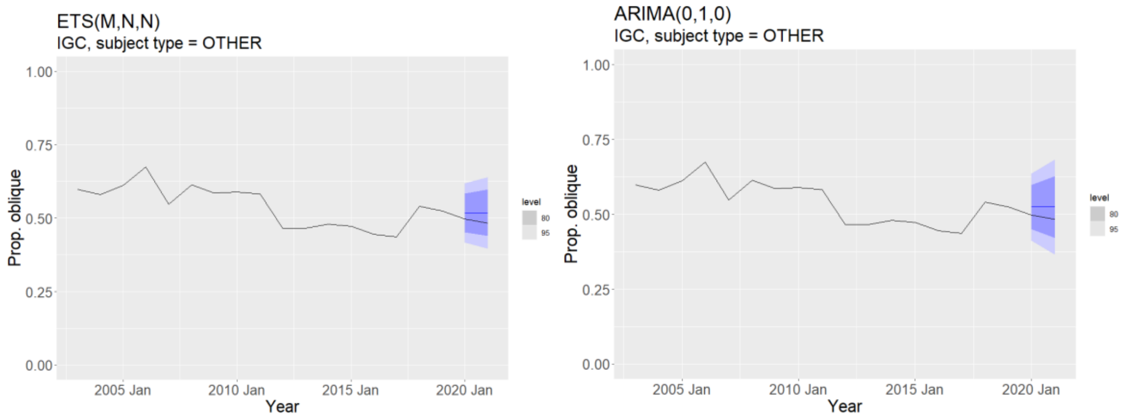


Figure 9.31. The ETS(M,N,N) model and AIMA(0,1,0) produce similar forecasts for the test period.

A summary of the accuracy of the fit of the ETS(M,N,N) and ARIMA(0,1,0) model with a constant along with the accuracy of the point forecasts for the test period is provided in Table 9.20. The ETS(M,N,N) model appears to give slightly more accurate point predictions than both the ARIMA(0,1,0) but slightly worse than the Drift model discussed above. An evaluation of the forecasts distribution using quantile scores, Winkler scores, Continuously Ranked Probability Scores and Scale-free comparison using skill scores, suggests that the ETS(M,N,N) performs 17.6% better than the Naïve method when skill scores (based on crps) are considered. This suggests that the ETS method might be the most appropriate for forecasting, although keeping in mind that the evaluation is based on a very small (2 steps ahead) test set. The Drift model also appears relatively promising, performing about 1.8% better than the Naïve model.

IGC, subject type = OTHER							
Model	source	type	RMSE	MAE	MAPE	MASE	RMSSE
ETS(M,N,N) OBLprop	IGC	Training	0.0514	0.036	6.79	0.906	0.899
ARIMA(0,1,0) OBLprop	IGC	Training	0.0555	0.0374	6.92	0.942	0.97
ETS(M,N,N) OBLprop	IGC	Test	0.0284	0.0277	5.68	0.699	0.497
ARIMA(0,1,0) OBLprop	IGC	Test	0.0344	0.0338	6.92	0.853	0.601

Table 9.20. A comparison of the accuracy of the fit and the point forecast of the ETS(M,N,N) and ARIMA(1,0,0) model. The ETS model appears to have a slightly better prediction accuracy.

IGC, subject type = OTHER						
Model	source	type	qs probs = 0.1	winkler (level 80) level 80	crps	skill
Mean OBLprop	IGC	Test	0.00821	0.186	0.0313	-0.383
Naïve OBLprop	IGC	Test	0.0109	0.177	0.0226	0
Drift OBLprop	IGC	Test	0.0132	0.186	0.0209	0.0769
ETS(M,N,N) OBLprop	IGC	Test	0.00898	0.145	0.0186	0.176
ARIMA(0,1,0) OBLprop	IGC	Test	0.0109	0.177	0.0226	-1.5E-06

Table 9.21. Based on quantile, Winkler, crps and skill scores, it looks like the ETS(M,N,N) has the best forecast distribution for the test period 2020–2021.

Since the ETS(M,N,N) model produced the best results for the test data, this type of model was fitted to the whole series and used to generate a 20-step ahead forecast, i.e., from 2019 to 2041 (Figure 9.32). Naturally, the forecast intervals become larger the further into the future the predictions reach. Note also that the method simply provides a “flat” future with the mean proportion of oblique being 49.1% all the way into 2041. The 80% confidence interval suggests oblique subjects are likely to make up anything from ca. 30% to up to just above. 60% of attested examples in 2041.

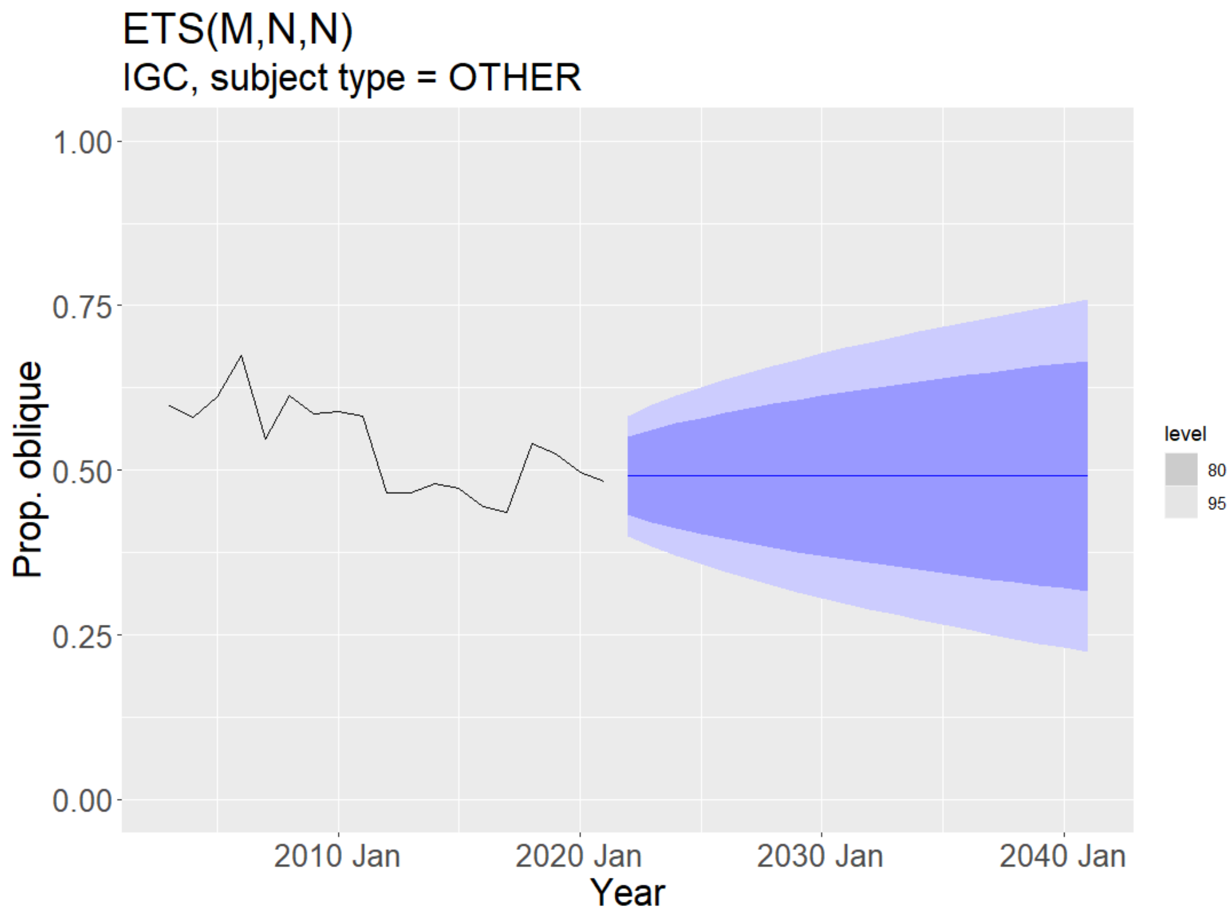


Figure 9.32. The ETS(M,N,N) model was used to predict 20 steps into the future or until 2041.

As noted above, an STL decomposition of the time series suggested that there was a considerable amount of noise (remainder component) in the series. Provided one claims that the "noise" is not meaningful and is simply due to randomness when measurements are conducted, it is possible to generate a forecast relying only on the trend component. This has been done in Figure 9.33 where an ETS(M,N,N) model was fitted to the trend of the whole time series and used to produce a forecast 20 steps into the future or until 2041. With the remainder component removed, the prediction intervals are narrower and the 80%

confidence interval now suggests the proportion of oblique subjects that are not 1st person pronouns will very likely be above 40% and below 60% in 2041.

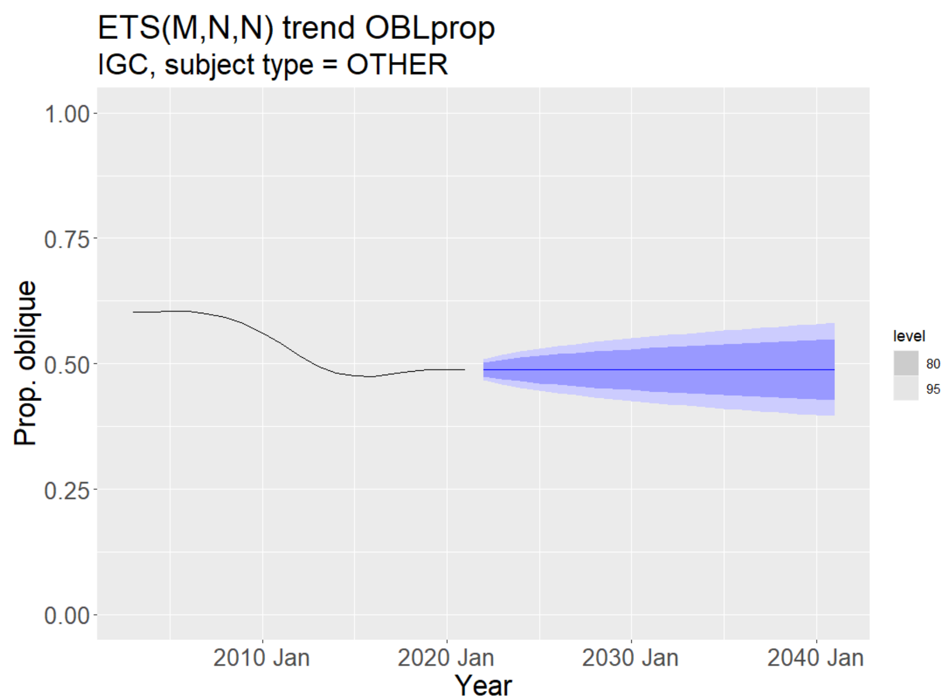


Figure 9.33. An ETS(M,N,N) model fitted to the trend of the STL-decomposed time series. Note that the 80% and 90% confidence intervals are narrower than when the raw series is used.

The general picture that emerges from studying the case marking of subjects of *hlakka til* that are not first person pronouns, is that the proportion of oblique remains relatively stable over time although there is some noise in the data (the series is a white noise series). Oblique subjects make up between 43.7% and 67% of attested examples throughout the period of 2003–2021. Using an ETS(M,N,N) model on the raw time series, it appears that the proportion of oblique subjects will be around 50% for the next several years, and will unlikely go below 25% or above or above 75% when taking the 80% and 95% confidence levels into consideration. If the remainder component is removed from the raw series an

ETS(M,N,N) model is applied to the trend part of the series, the 80% and 95% confidence intervals become narrower, suggesting that the proportion of oblique subjects will stay closer to the 48.8% point forecast than otherwise. Essentially, we are looking at a situation where the use of oblique subjects with *hlakka til* (when the subject is not a 1st person pronoun) will remain relatively stable in semi-formal and informal written material for the next several years. However, keep in mind that the number of examples behind each observation for this time series is less than that behind the time series focusing on 1st person subjects. The expectations generated from the forecasts are discussed further in Section 9.6.

9.5.4 Twitter 1st person subjects

The Twitter time series containing information on case marking with 1st person pronoun subjects consists of 44 (quarterly) observations from 2012 to 2022. The series, which has a mean = 0.1157312, is not stationary and would have to be differenced once to make it so (KPSS statistics = 1.02, p-value = 0.01). The series appears to show a slight negative trend, with oblique subjects making up ca. 17%–18% of the examples in Q1–Q2 2012, and around 6%–8.5% in Q3–Q4 2022. The strength of the trend is reported as 0.881, the linearity -0.2561735 and the curvature 0.08616929. There are no obvious outliers in the series. As suggested by the autocorrelation function (ACF), cf. Figure 9.34, several of the coefficients lie outside the expected limits, and the Ljung-Box statistic (statistics = 167., p-value = 0, lag = 10) the series is not a white noise series

Given that the series consists of quarterly observations, seasonality might be detected. This is in fact the case. An STL decomposition (see Chapter 7, Section 7.2.2) of the series (trend window = 13), where each observation is assumed to consist of a trend component, a seasonal component and a remainder component shows clearly the negative trend observed in the series. The decomposition also shows the small random noise, which lies somewhere between -0.04 and 0.06. The gray boxes in each of the panels in Figure 9.34 show the same scale i.e., they are the same size.

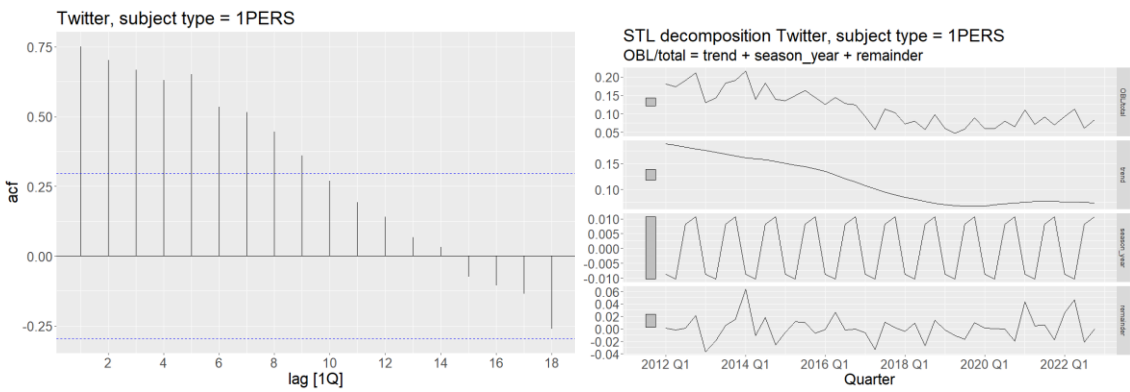


Figure 9.34. ACF and STL decomposition of the Twitter series for the proportion of oblique first person subjects with *hlakka til*. The series has a clear trend, some random noise (remainder) and extremely small seasonal component.

The series was divided into a training and test set. As noted earlier (Section 9.5.1), the training set contained 40 observations from Q1 2012–Q4 2021, and the test set 4 observations from Q1–Q4 2022.

Four simple models were fitted to the Twitter first person training data: i) a Naïve model which assumes that all future observations will be the same as the last recorded observation in the series, ii) a Seasonal naïve model which assumes each quarter in the future will be the same as last reworded quarter of the same type, iii) a Mean model which assumes that all future values will be equal to the mean of the whole series, and iv) a Drift model where changes over time are assumed to be average change in the historical data (see further in Chapter 7, Section 7.2.3). The point-predictions of each of the models for the test period are shown in Figure 9.35 along with the whole series. The model relying on the mean of the training data appears to give the least accurate prediction, suggesting dative subjects should make up almost 12% of examples through Q1 2022–Q2 2022 when the observed values are between ca 6%–11%, depending on the quarter. The Naïve and the Seasonal naïve models perform the best. A summary of the fit of each of the models to the training data as well as how accurate the point forecasts are, is provided in Table 9.22.

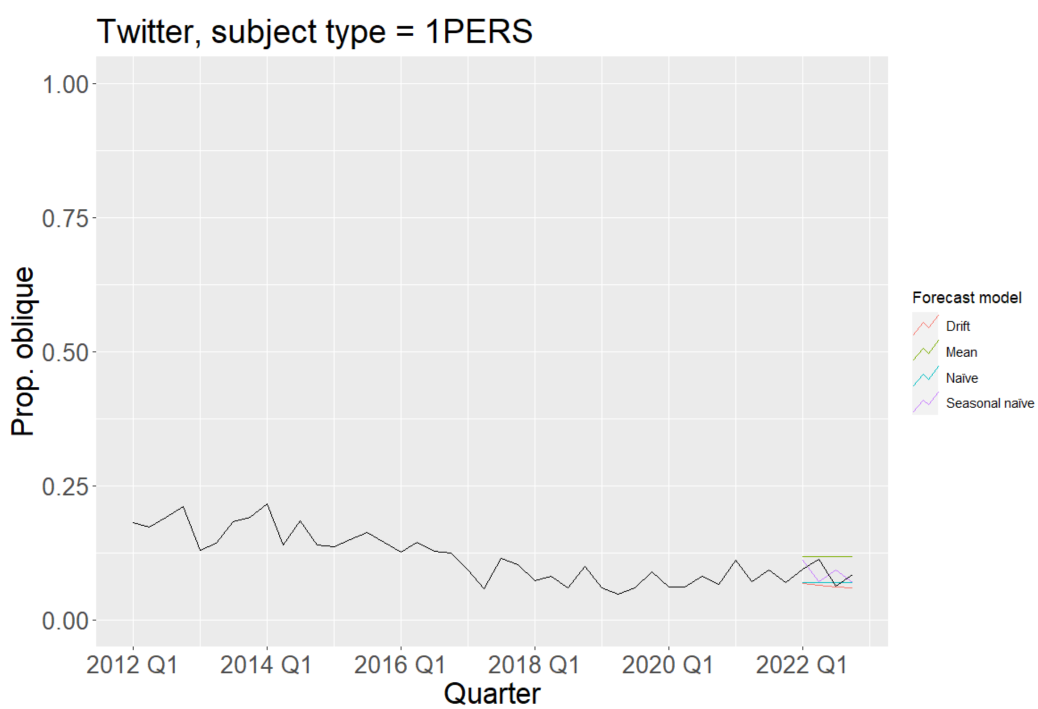


Figure 9.35. Four different models were fitted to the training data of the Twitter time series containing information on *hlakka til* with a 1st person subject. These models were: a Drift model, a model that picks out the mean of the whole series, a Naïve model and a Seasonal naïve model. The fitted model was used to forecast Q1 – Q4 2022.

Twitter, subject type = 1PERS							
Model	source	type	RMSE	MAE	MAPE	MASE	RMSSE
Mean OBLprop	Twitter	Training	0.0478	0.0413	43.1	1.68	1.41
Naïve OBLprop	Twitter	Training	0.0314	0.0256	25.4	1.04	0.928
Seasonal naïve OBLprop	Twitter	Training	0.0339	0.0246	25.8	1	1
Drift OBLprop	Twitter	Training	0.0313	0.0256	24.9	1.04	0.924
Drift OBLprop	Twitter	Test	0.0304	0.0254	25.8	1.03	0.898
Mean OBLprop	Twitter	Test	0.0352	0.0303	40.6	1.23	1.04
Naïve OBLprop	Twitter	Test	0.0257	0.0222	23.3	0.902	0.759
Seasonal naïve OBLprop	Twitter	Test	0.0278	0.0257	30.1	1.05	0.822

Table 9.22. Report on the fit of the simple models to the training data and the accuracy of the point predictions for the test period Q1–Q4 2022. Of the four models, the Naïve model appears to be the most accurate.

The four simple models above serve as a benchmark when fitting and choosing other models. Both ETS and ARIMA models were fitted to the training data in the hope that they would perform better than the models above. Since the data shows a negative trend, with values dipping below 10%, a log transformation was used to keep point predictions and prediction intervals within a positive range. The type of ETS and ARIMA model along with initial states was determined using the ETS() and ARIMA() functions from the fable package (O'Hara-Wild, Hyndman & Wang 2021). While the ETS() function selects a model by minimizing the AICc (Hyndman & Athanasopoulos 2021:254), the ARIMA() function relies on the Hyndman-Khandakar algorithm for automatic ARIMA modeling (Hyndman & Athanasopoulos 2021:286).

An automatic selection of an ETS model resulted in a simple exponential smoothing with additive errors, ETS(A,N,N). The smoothing parameters were alpha = 0.382763, with the initial states l = -1.724998. The residuals of the model are within desirable limits (mean = -0.05422317) and can be interpreted as white noise (Ljung-Box statistics = 8.13 with p-value 0.616).

The automatic selection of ARIMA resulted in ARIMA(0,1,1) with drift. The coefficient was ma1 = -0.6949 (s.e. 0.1119) and constant = -0.0239 (s.e. 0.0124). The residuals from this model are also within desirable limits (mean 0.001649338) and can be interpreted as white noise (Ljung-Box statistics = 10.8 with p-value 0.374, lag = 10). Note that neither of these models explicitly model the extremely small seasonal component found in the time series. Point forecasts with prediction intervals are shown in Figure 9.36. A summary of the accuracy of the point forecasts is provided in Table 9.23. The ETS(A,N,N) appears to generate more accurate predictions than the ARIMA(0,1,1). Both of the models perform better than the benchmark models discussed above.

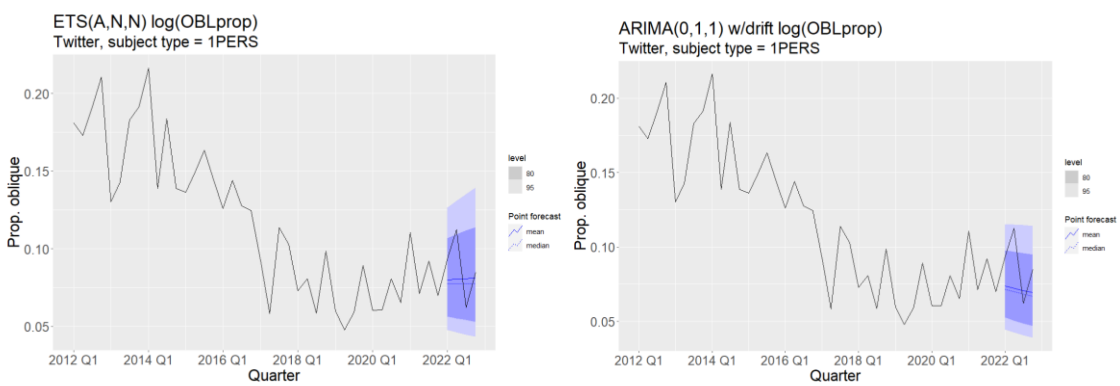


Figure 9.36. An ETS(A,N,N) and ARIMA(0,1,1) were fitted to the training data. A log transformation was used to ensure the forecast intervals and point forecast would stay within a positive range.

Twitter, subject type = 1PERS								
Model	source	type	RMSE	MAE	MAPE	MASE	RMSSE
ETS(A,N,N) log(OBLprop)	Twitter	Training	0.0258	0.0211	20.9	0.857	0.761
ARIMA(0,1,1) w/drift log(OBLprop)	Twitter	Training	0.0246	0.0194	18.3	0.788	0.727
ETS(A,N,N) log(OBLprop)	Twitter	Test	0.0198	0.0169	19.3	0.689	0.584
ARIMA(0,1,1) w/drift log(OBLprop)	Twitter	Test	0.0241	0.0211	22.3	0.856	0.713

Table 9.23. Report on the accuracy of the fit of the ETS(A,N,N) model and ARIMA(0,1,1) w/drift and accuracy of point forecasts for Q1–Q4 2022.

Even though both the ETS(A,N,N) model and the ARIMA(0,1,1) appear to have reasonably good predictions, they produce slightly different point forecasts and forecast distributions. The point forecast (going by the mean of the forecast distribution) of the ETS(A,N,N) suggests that the proportion of oblique first person subjects should be rising ever so slightly from 7.37% in Q1 2022 to 6.9% in Q4 2022. The ARIMA(0,1,1), on the other hand, suggests a minor downward trajectory from 7.37% in Q1 2022 to 6.9% in Q4 2022. The actual values lie between ca. 6%–11%, which are mostly within the 80% confidence interval of the ETS(A,N,N) forecast and the 95% confidence interval of the ARIMA(0,1,1) forecast.

An evaluation of the forecasts distribution of all the methods discussed above using quantile scores, Winkler scores, Continuously Ranked Probability Scores and Scale-free comparison using skill scores, suggests that the ETS(A,N,N) performs almost 27% better than the Seasonal naïve method when skill scores (based on crps) are considered. The ARIMA(0,1,1) with drift is only estimated to be around 8.7% better than the Seasonal naïve method. This suggests that the ETS method might be the most appropriate for forecasting, although keeping in mind that the evaluation is based on a very small (4 steps ahead) test set.

Twitter, subject type = 1PERS						
Model	source	type	qs probs = 0.1	winkler (level 80) level 80	crps	skill
Mean OBLprop	Twitter	Test	0.00651	0.126	0.0208	-0.276
Naïve OBLprop	Twitter	Test	0.016	0.124	0.0172	-0.0536
Seasonal naïve OBLprop	Twitter	Test	0.00914	0.0868	0.0163	0
Drift OBLprop	Twitter	Test	0.0178	0.128	0.0194	-0.189
ETS(A,N,N) log(OBLprop)	Twitter	Test	0.00671	0.0636	0.0119	0.269
ARIMA(0,1,1) w/drift log(OBLprop)	Twitter	Test	0.00773	0.0857	0.0149	0.0872

Table 9.24. Based on quantile, Winkler, crps and skill scores, it appears the ETS(A,N,N) has the best forecast distribution for the test period Q1 2022–Q4 2022.

Since both the ETS(A,N,N) and ARIMA(0,1,1) with drift performed better than the Seasonal näive method, these were fitted to the whole series and used to generate a 20-step ahead forecast, i.e., from Q1 2022 to Q4 2026 (Figure 9.37). The point prediction suggests is relatively flat. Only a minor increase in oblique over time is noted for the ETS(A,N,N) model, or from around 8% in Q1 2022 to almost 8.7% in Q4 2027. The 80% confidence interval suggests the proportion of oblique first person subjects will stay below 25% in Q4 2027. The ARIMA(0,1,1) model gives slightly narrower prediction intervals, suggesting that the proportion of oblique will likely not go above ca. 20%. The ARIMA(0,1,1) model also shows a slight increase in the use of oblique subjects over time. Going by the mean of the forecast distribution, the predictions go from 8.4% in Q1 2022 to 9.1% in Q4 2026.

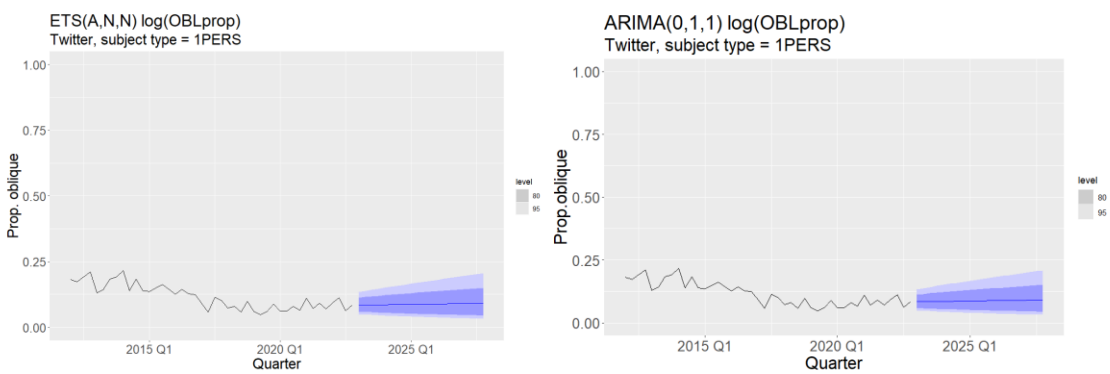


Figure 9.37. Models fitted to the training data and used to predict the test data and future periods until Q4 2026.

As noted above, an STL decomposition of the time series suggested that there was a considerable amount of noise (remainder component) in the series. If the "noise" is claimed to not be meaningful and simply the result of randomness in language use and how measurements are conducted, it is possible to generate a forecast relying on the trend component only. This has been done in Figure 9.38 where an ETS(A,N,N) model was fitted to the trend of the whole time series and used to produce a forecast 30 steps into the future or until Q4 2029. With the seasonal and remainder components removed, the prediction intervals are narrower. The point forecast suggests a very minor increase in oblique subjects over time, or from 7.44% in Q1 2023 to ca. 7.56% in Q2 2030. Note that these values are slightly lower than what was produced using the non-decomposed time series.

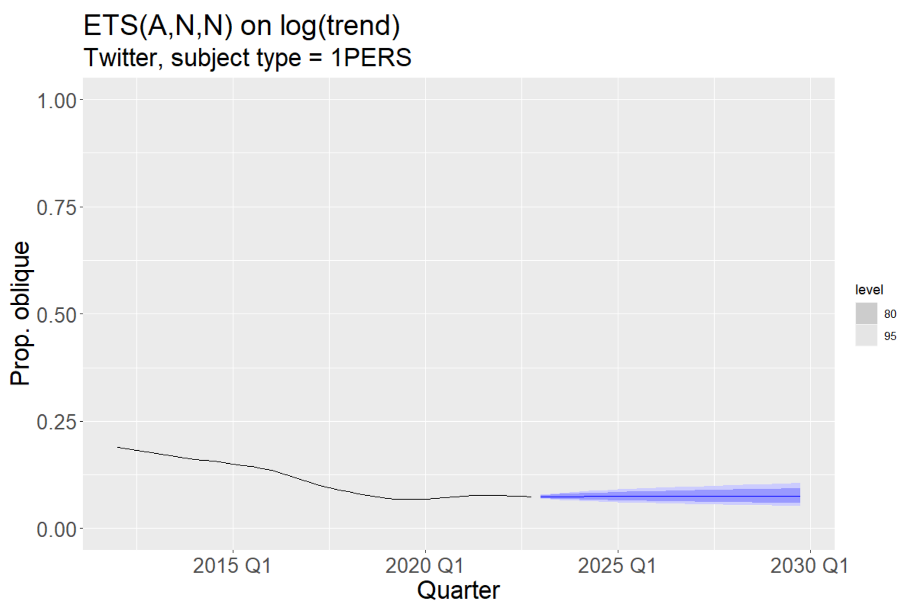


Figure 9.38. Simple exponential smoothing method with additive errors, ETS(A,N,N) fitted to the trend of the whole series and used to predict the proportion of oblique first person subjects from Q1 2023 to Q4 2029.

The general picture that emerges from studying the case marking of first person subjects of *hlakka til* in data from Twitter, is that the proportion of oblique subjects is going down over time. In Q1 2012 they make up about 18% of attested examples on Twitter and about 8.5% in Q4 2022. In addition to four simple models (naïve, seasonal naïve, drift and mean model), an ETS(A,N,N) and ARIMA(0,1,1) with drift were fitted to the training data. Both of these produced a better fit and better forecasts for the test data than the simple models (with the ETS model performing slightly better) so they were used to generate a 20-step forecast into the future, or until Q4 2026. The forecasts from the two models give rise to similar expectations for the direction of the propagation of oblique subjects. While both predict a low proportion of oblique subjects for the next several years, they do suggest a minor increase in their use of oblique subjects. Taking the 80% confidence interval into consideration, oblique subjects are not expected to make up less than 25% of attested examples on Twitter in 2026. If the remainder component and the very small seasonal component are removed from the raw series, an ETS(A,N,N) model is applied to the trend component, the 80% and 95% confidence intervals become narrower, suggesting that the proportion of oblique subjects will be very close to 7.5% in the future. Summarizing these predictions, we are looking at a situation where the use of oblique first person subjects with *hlakka til* will likely stay below 10% in Twitter material for the next several years. Expectations generated from the forecasts in this section are discussed further in Section 9.6.

9.5.5 Twitter other types of subjects

The Twitter time series containing information on case marking with subjects that are not a first person pronoun consists of 44 (quarterly) observations from 2012 to 2022. The series is based on only 1,177 examples in total, resulting in there being relatively few examples behind each observation in the series. The frequent ups and downs, or the large amount of noise in the series, might also be at least partially attributed to it being based on a few examples. Despite the noise, the series is not stationary (KPSS statistics 0.687, p-value 0.0147) and would have to be differenced once to make it so. The series, which has a mean of ca. 50.76%, appears to show a downwards trend from 2012 when proportion of oblique is around 75% until around ca. 2020 when it is down to around 30%. At the end of 2022 the proportion of oblique is just below 50%, suggesting an upwards trend. The strength of the trend is reported as 0.76659, the linearity -0.4961527 and the curvature 0.3572001. Some observations might be described as outliers, in particular those in Q4 2015 and Q1 2016. As shown on the correlogram in Figure 9.39, several of the coefficients lie outside the expected limits so this is not a white noise series (Ljung-Box statistics = 43.7, p-value = 0.0000037, lag = 10).

Given that the series consists of quarterly observations, seasonality might be detected. This is in fact the case. An STL decomposition (see Chapter 7, Section 7.2.2) of the series (trend window = 13), where each observation is assumed to consist of a trend component, a seasonal component and a remainder component also shows clearly the negative trend observed in the series. The window for the trend component was specified as 13 to minimize the amount of noise The decomposition also shows the amount of noise in the remainder component, which lies somewhere between -0.2 and 0.2. The gray boxes in each of the panels in Figure 9.39 show the same scale i.e., they are the same size.

The series was divided into a training and test set. As noted earlier (Section 9.5.1), the training set contained 40 observations from Q1 2012 – Q4 2021, and the test set 4 observations from Q1 – Q4 2022.

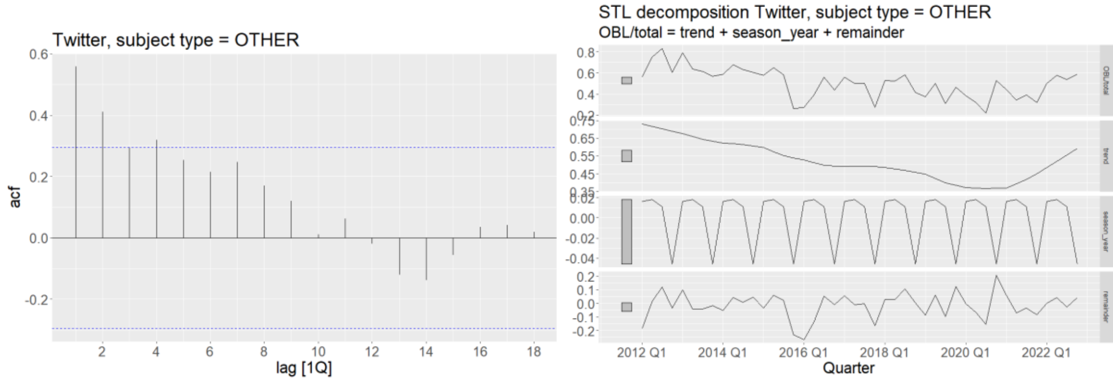


Figure 9.39. Some features of the Twitter time series that contains information on the proportion of oblique subjects with *hlakka til* that are not first person pronouns. Both the ACF and the STL decomposition suggest the series is trended. ACF shows some lags outside the significance line.

Four simple models were fitted to the Twitter first person training data: i) a Naïve model which assumes that all future observations will be the same as the last recorded observation in the series, ii) a Seasonal naïve model which assumes each quarter in the future will be the same as last reworded quarter of the same type, iii) a Mean model which assumes that all future values will be equal to the mean of the whole series, and iv) a Drift model where changes over time are assumed to be average change in the historical data (see further in Chapter 7, Section 7.2.3). The point-predictions of each of the models for the test period are shown in Figure 9.40 along with the whole series. None of the models appear particularly good. The model relying on the mean of the training data appears to give the most accurate prediction, suggesting dative subjects should make up around 50% of examples in Q1 2022–Q2 2022 when the observed values lie between ca 50%–59%. The second best model is the Seasonal naïve model, followed by the Naïve and Drift models which perform quite poorly. A summary of the fit of each of the models to the training data as well as how accurate the point forecasts are, is provided in Table 9.25.

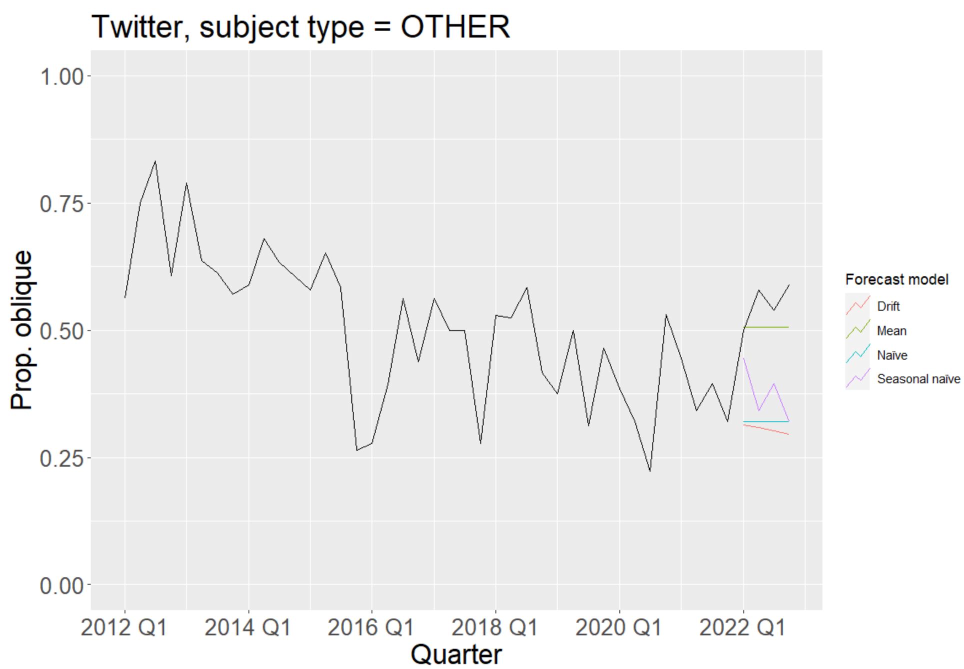


Figure 9.40. Four simple models, a Drift model, a Naïve model, a Seasonal naïve model and a Mean model, were fitted to the training data of the Twitter time series containing information on *hlakk til* with subjects other than first person pronouns. The model was used to forecast Q1–Q4 2022 with the mean model performing the best.

Twitter, subject type = OTHER							
Model	source	type	RMSE	MAE	MAPE	MASE	RMSSE
Mean	Twitter	Training	0.148	0.125	28.1	1.05	0.969
Naïve	Twitter	Training	0.135	0.109	24.9	0.918	0.887
Seasonal naïve	Twitter	Training	0.153	0.119	30.1	1	1
Drift	Twitter	Training	0.135	0.108	24.6	0.912	0.886
Drift	Twitter	Test	0.249	0.246	44.3	2.07	1.63
Mean	Twitter	Test	0.0579	0.0487	8.51	0.41	0.379
Naïve	Twitter	Test	0.233	0.231	41.6	1.94	1.53
Seasonal naïve	Twitter	Test	0.194	0.176	31	1.48	1.27

Table 9.25. Report on the fit of the simple models to the training data and the accuracy of the models given how they did in predicting Q1–Q4 2022.

In addition to the four simple forecasting models mentioned above, an ETS and an ARIMA model were also fitted to the training data. The type of ETS and ARIMA model along with initial states was determined using the ETS() and ARIMA() functions from the fable package (O'Hara-Wild, Hyndman & Wang 2021). While the ETS() function selects a model by minimizing the AICc (Hyndman & Athanasopoulos 2021:254), the ARIMA() function relies on the Hyndman-Khandakar algorithm for automatic ARIMA modeling (Hyndman & Athanasopoulos 2021:286).

An automatic selection of an ETS model suggested Holt's linear method with additive errors, or ETS(A,A,N). The smoothing parameters were alpha = 0.0001000142 and beta = 0.0001000002, and the initial states l = 0.6785951 and b = -0.008721004. While the residuals, with mean of 0.003066345, were plausibly analyzed as white noise (Ljung-Box statistics 4.34, p-value 0.931, lag = 10) and autocorrelation of residuals was within expected limits and, the distribution of the residuals was somewhat skewed, i.e., they were not fully normally distributed. However, a better model was not found.

An automatic selection of ARIMA resulted in ARIMA(0,1,1) with ma1 = -0.6955 (s.e. 0.1189). The ACF is within acceptable limits and residuals, with mean of -0.02126559, can be regarded as white noise (Ljung-Box statistics 5.55, p-value 0.852, lag = 10) even though they are not fully normally distributed which might have a slight effect on forecast distribution. An attempt was made to find a better model where $d = 1$ (in order to make the series stationary it needs to be differenced once), $p = 0–2$ (order of autoregressive part) and $q = 0 – 2$ (order of moving average part). However, based on the fit to the training data, relying on AIC, AICc and BIC, no better model was found. A model that came close to ARIMA(0,1,1) was ARIMA(0,1,2). Since the distribution of the residuals of these looked quite similar, we go ahead and use the automatically selected ARIMA(0,1,1).

Point forecasts for the ETS(A,A,N) and ARIMA(0,1,1) with prediction intervals are shown in Figure 9.41. Neither model performs well on the test data. The ETS(A,A,N) does not capture the observed values for the period Q1–Q4 2022. The point forecast suggests that oblique subjects should fall between 29.5% and 32.1% while observed values fall between 50%–60% which is outside of the 80% and 95% prediction intervals. It appears that the ETS(A,A,N) model is primarily picking up on the downward trend in the data that lasts until Q4 2021 so the shift in level of the series after that time is not captured. The ARIMA(0,1,1) model appears to perform slightly better at predicting the test data. The point forecast suggests a proportion of oblique subjects of 37% with observed values falling within the range of the 95% confidence interval of the forecast distribution. The accuracy of the fit of the two models and the accuracy of the point forecasts are summarized in Table 9.26.

Neither the ETS(A,A,N) nor the ARIMA(0,1,1) model appear particularly good at predicting the test data. An evaluation of the forecasts distribution of all the methods discussed above using quantile scores, Winkler scores, Continuously Ranked Probability Scores and Scale-free comparison using skill scores (cf. Table 9.27), suggests the ETS(A,A,N) performs almost 55% worse than the Seasonal naïve model when skill scores (based on crps) are considered. The ARIMA(0,1,1) is only estimated to be around 1.2% worse than the Seasonal näive model. The best forecast, both point forecast and forecast distribution, seems to come from the mean model that assumes all future values will be equal to the mean of the whole series. Note that the point forecast of the ARIMA(0,1,1) model also turns into a straight-line prediction, cf. Figure 9.41.

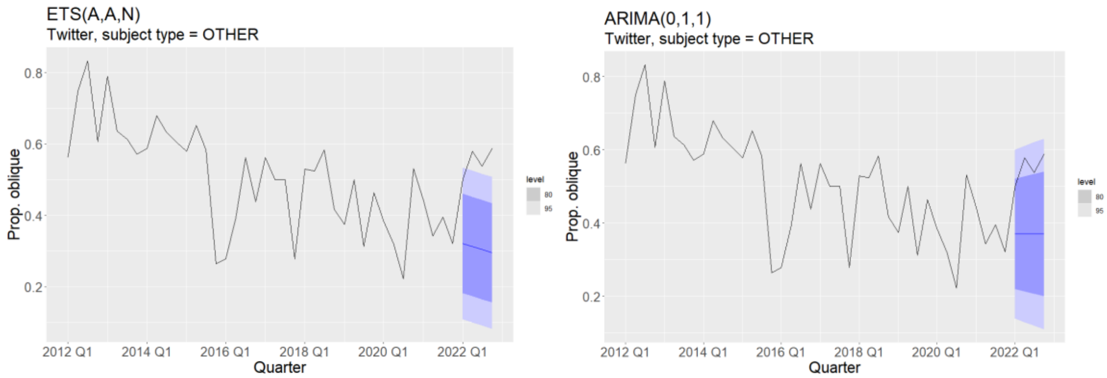


Figure 9.41. ETS(A,A,N) and ARIMA(0,1,1) were fitted to the training data and used to predict the test data (Q1–Q4 2022). Observed values lie outside of the 95% confidence interval of the ETS(A,A,N) model.

Twitter, subject type = OTHER							
Model	source	type	RMSE	MAE	MAPE	MASE	RMSSE
ETS(A,A,N) OBLprop	Twitter	Training	0.103	0.0788	19.2	0.662	0.674
ARIMA(0,1,1) OBLprop	Twitter	Training	0.115	0.0904	22.4	0.761	0.751
ETS(A,A,N) OBLprop	Twitter	Test	0.247	0.243	43.8	2.04	1.62
ARIMA(0,1,1)	Twitter	Test	0.185	0.181	32.6	1.52	1.21

Table 9.26. Accuracy of the fit of the ETS(A,A,N) and ARIMA(0,1,1) models and the accuracy of the point forecast for the test data. While these models fit the training data better than the simple models above, they produce worse point forecasts.

Twitter, subject type = OTHER						
Model	source	type	qs probs = 0.1	winkler (level 80) level 80	crps	skill
Mean OBLprop	Twitter	Test	0.0468	0.409	0.0516	0.564
Naïve OBLprop	Twitter	Test	0.0995	0.579	0.143	-0.207
Seasonal naïve OBLprop	Twitter	Test	0.0742	0.672	0.118	0
Drift OBLprop	Twitter	Test	0.104	0.624	0.153	-0.294
ETS(A,A,N) OBLprop	Twitter	Test	0.0764	1.32	0.183	-0.549
ARIMA(0,1,1) OBLprop	Twitter	Test	0.0684	0.579	0.12	-0.0124

Table 9.27. Comparing the forecast distribution of the simple models and that of ETS(A,A,N) and ARIMA(0,1,1).

The Mean model and ARIMA(0,1,1) were fitted to the whole dataset and used to generate a 15-step ahead forecast, i.e., from Q1 2022 to Q3 2025 (Figure 9.42).[115] Due to observations adjacent in times being very different, the forecast interval is quite large for both models. For the ARIMA(0,1,1) the interval becomes larger the further into the future predictions are made. The point forecasts of both models assume variation around a constant mean. According to the ARIMA(0,1,1) the proportion of oblique subjects will be right around 52.8%, whereas relying on the mean of the training period assumes a proportion around 50.9%.

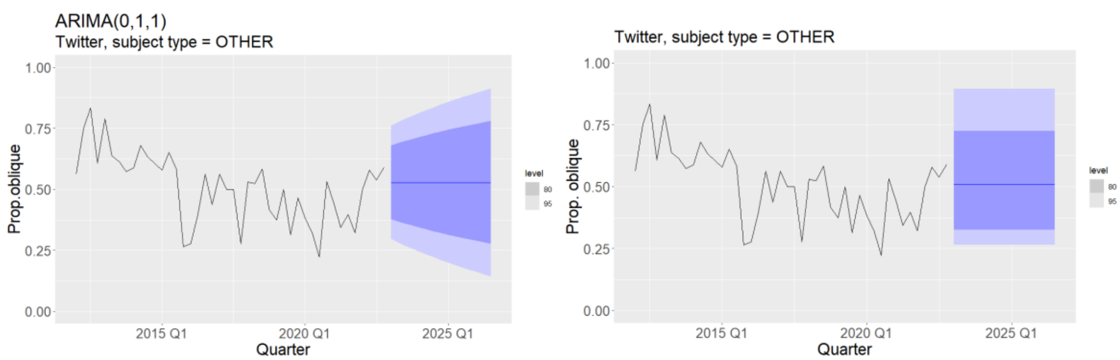


Figure 9.42. A model simply taking into account the mean of the whole series provides predictions where the level of the series is higher than when the ARIMA(0,1,1) model is used.

As noted above, an STL decomposition of the time series suggested that the remainder component was quite large. Assuming the remainder component reflects non-meaningful "noise", it is possible to generate a forecast relying on the trend component only. This has been done in Figure 9.43 where an ARIMA(0,1,1) model was fitted to the trend of the whole time series and used to produce a forecast 20 steps into the future or until Q4 2027.

[115] Forecasts for other series typically included 20-step ahead predictions. However, with such wide forecast distribution as the models show for the Twitter data with non-first person subject, this is not feasible.

With the seasonal and remainder components removed, the prediction intervals are initially much narrower. The point forecast suggests a very minor increase in oblique subjects over time, or from around 62% in Q1 2023 to ca. 71% in Q4 2027. Note that these values are considerably higher than what was produced using the non-decomposed time series. An odd feature of the STL decomposition is that the trend is assumed to show quite a strong upward trajectory around 2022.

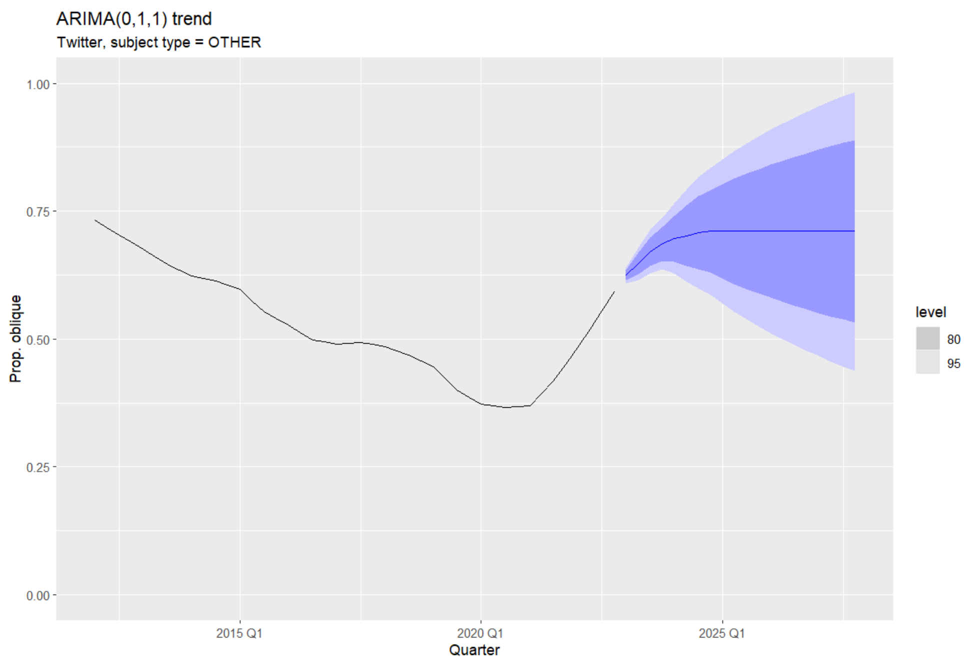


Figure 9.43. An ARIMA(0,1,1) model fitted to the trend of the whole series and used to predict the proportion of oblique first person subjects from Q1 2023 to Q4 2027.

The Twitter time series containing information on the proportion of oblique subjects that are not first person pronouns with *hlakka til* is a difficult series to work with. The whole series is based on only 1,177 examples which means that each of the 44 observations (from Q1 2012 to Q4 2022) is based on few data points. This is very likely the reason for considerable changes between observations close in time.

In general, the proportion of oblique subjects that are not first person pronouns shows a downward trend from Q1 2012 (from ca. 75%) to mid-2020 (around 30%). From around mid 2020 there appears to be an upwards trend to the end of 2022 when the proportion of oblique is just below 50%. In addition to four simple models (Naïve, Seasonal naïve, Drift and Mean model), an ETS(A,A,N) and ARIMA(0,1,1) were fitted to the training data. These did generally not perform very well on making predictions for the test set. In fact, the model relying on the mean of the whole series appeared to produce the best forecast, with ARIMA(0,1,1) being slightly worse than a seasonal naïve model. Depending on which of these two, the ARIMA(0,1,1) or the mean mode, are used, future observations are expected to show the proportion of oblique subjects being somewhere around 40%–50% with some chances of it dipping below 25% according to the 80% and 95% confidence interval for the forecast distribution of the ARIMA model, or reaching above 80% according to the 95% confidence interval of the mean model. If the trend component is isolated from the rest of the components of the series with STL decomposition (with trend window = 13) and used to generate a prediction with an ARIMA(0,1,1) model 20-steps ahead into the future, a slightly different picture emerges. Under this scenario, the proportion of oblique will rise from ca. 62% in Q1 2023 to 71% in Q4 2024 and then stay flat for the rest of the period until Q4 2027. Note that the reason for choosing ARIMA(0,1,1) instead of a mean model is that this is an attempt to capture the upwards trend in the data instead of simply assuming that all future values will be equal to the mean of the whole series. Summarizing these predictions, it looks like the use of oblique subjects with *hlakka til* (that are not first person subjects) will vary a lot in Twitter material for the next several years. Quite possibly, they will make up somewhere close to (or slightly above) 60% of attested examples, relying on results from the mean model which was fitted to the raw training set and the ARIMA(0,1,1) used on the trend of the whole series. Expectations generated from the forecasts in this section are discussed further in Section 6.

9.5.6 Interim summary

Summarizing the time series and the forecasts from Sections 9.5.1 – 9.5.5, it can be noted that the data from IGC and Twitter show a clear difference in the usage of oblique based on whether the subject is a first person pronoun or of a different kind.

When the subject of *hlakka til* is a first person pronoun, the time series from both IGC and Twitter show a rather low proportion of oblique (IGC around 35% in 2004, Twitter around 20% in 2012) that goes down over time (IGC around 6% in 2021, Twitter below 10% in 2022). When fitting ETS and ARIMA models to the time series, the models pick up on the downward trajectory in the series and predict a continuing decline in usage of oblique subjects. As an example, an ETS(A,A,N) model used for the IGC series, used on the trend with log transformation to keep prediction intervals positive, suggests that oblique first person subjects with *hlakka til* will be virtually nonexistent in 2040. For this particular series, a Naïve model actually turned out to provide the best predictions for the training period. Using a Naïve model instead of the ETS(A,A,N) suggests that oblique subjects with *hlakka til* in material of the type of semi-formal and informal language on IGC will be around 6% for all coming years. As for the Twitter series, the expectations are that the proportion of oblique will continue to be low in coming years, somewhere below 10%, but it will not go completely down to zero.

Turning to subjects that are not first person pronouns, it is important to note that the number of examples behind each observation in the series is way lower than the number of examples behind observation of series with first person subjects. Presumably, this reflects that individuals like to talk about themselves. The IGC series is based on a total of 6,376 examples and the Twitter series a total of 1,177 examples. Since there are so few examples behind each observation, it comes as no surprise that there is quite a lot of noise in the data i.e., adjacent observations are often quite different. This means that model fitting and projection into the future are somewhat uncertain and contain fairly large prediction intervals. Nevertheless, a few things may still be noted about the series. For instance, the proportion of oblique case marking with subjects that are not first person pronouns, is generally higher than with first person subjects. In the case of the IGC-series the proportion lies somewhere between 50 and 60% with little change over time. In the case of the Twitter-series, it is anywhere from ca. 20% to just above 80%. This is a considerably higher proportion of oblique than either of the series showed for first person subjects. Projecting the IGC series into the future, the expectations are that the proportion of oblique will continue to be somewhere around 50% although the prediction intervals are quite large. The Twitter series behaves slightly differently as it appears to show an initial decline in use of oblique subjects, followed by a rise. Predictions based on the trend in the series, suggests that the proportion of oblique subjects with non-first person pronouns on Twitter may increase slightly in the coming years. Also here are the prediction intervals quite large, especially after 2025. Again, taking into account the number of examples behind each observation, the results and predictions based on these series are not as well founded as those based on the first person time series.

9.6 Summary and discussion

Icelandic is known for having non-nominative subjects which can be either accusative, dative or genitive. While some oblique-subject predicates are old in the language, others have emerged at recent times. The predicate *hlakka til* 'look forward to' is an example of a recent oblique-subject predicate. It originally appeared with a nominative subject, as in example (9.1), but in Modern Icelandic it is often encountered with an oblique subject which is either in accusative or dative (9.2).

(9.1)	***Ég***	*hlakka*	*til*	*sumarsins*
	I-NOM	look.forward	to	the.summer
	'I look forward to the summer.'			

(9.2)	a.	***Mig***	*hlakkar*	*til*	*sumarsins*
		me-ACC	llook.forward	to	the.summer
	b.	***Mér***	*hlakkar*	*til*	*sumarsins*
		me-DAT	look.forward	to	the.summer
		'I look forward to the summer.'			

The innovative case marking on the subject (9.2) has been well documented with respect to its origins and its propagation through the language community. The current expectation regarding the change is that the use of oblique subjects will increase with the predicate over time. The expectations are based on several facts, for instance that previous documentation of the change has shown an increase in the use of oblique cases over time. In a survey from the early 80s (Svavarsdóttir 1982), the proportion of oblique case marking among 11-year-old children was 80.7% for 3rd person subjects. Documentation in the 2000s shows the proportion of oblique may have increased slightly to 81.2%, also for children aged ca. 11–12. These measurements are based on subjects that are not first-person pronouns. Another fact that contributes towards the expectation that oblique subjects with *hlakka til* should increase over time is that experiencer predicates often appear with oblique subjects (Jónsson 2003). It has even been argued that semantic bootstrapping may be used when determining the case of subjects, i.e., children pick up on semantic information such as thematic role when acquiring case marking (Nowenstein 2023).

The change from nominative to oblique with *hlakka til* goes against The Case Directionality Hypothesis, which suggests that lexical case becomes structural case. If the CDH were to hold at all times, oblique subjects would be expected to decrease over time. However, this is contrary to what is attested; there are instances of new oblique subjects emerging with various predicates at all times. A potential other reason for why oblique with *hlakka til* might be expected to decrease over time is that prescriptivism for the past 60-70 years has focused heavily on eliminating oblique subjects with the predicate. Despite systematic attempts to root out the oblique, oblique subjects continue to appear with *hlakka til*, suggesting that there is something about the structure that attracts oblique. For these reasons, oblique subjects might continue to be used and potentially increase. A note about expectations regarding first-person subjects is that prescriptive grammar in school focuses heavily on these. Therefore, it is possible that individuals may consciously learn to use nominative when the subject is a first person pronoun, even though they may use a different case for other kinds of subjects.

For testing expectations toward propagation of the change and for making predictions about the future, four time series were constructed, two from each data source, i.e., the Icelandic Gigaword Corpus and Twitter. One of the time series from each source contained information about the proportion of oblique first person subjects. The second series contained information about the proportion of oblique non-first person subjects. The series covered about 11 years (Twitter, projected into quarterly time series) to about 19 years (IGC, projected into yearly time series). Of course, a number of decisions were made when constructing the time series, such as what type of data to use, which examples to include in the data set, how to annotate the data, what the frequency of the series should be, how to split each series into a training and test set etc.

Interestingly, the time series that were constructed (Section 9.5.1) do not fully conform to expectations towards propagation of the change. The proportion of oblique subjects found in the data from IGC and Twitter is lower than what has been documented in surveys in the past. This applies to usage with first person subjects as well as other types of subjects. The documentation, which admittedly targets younger language users (see Section 9.3.2 for discussion), suggested about 80% dative with non-first person subjects (e.g., Svavarsdóttir 1982, Jónsson & Eythórsson 2003) and about 52% dative with first person subjects (Svavarsdóttir 1982). The proportion of dative with *hlakka til* in semi-formal and informal material on IGC and material on Twitter is slightly lower than expected, and it appears to go down over time. For first person subjects, the proportion is around 18% on Twitter in 2012 and dips below 10% in 2022, consistently going down over the period. For Twitter, first person subjects appear about 32% of the time with oblique in 2003, but in 2021 the proportion is around 6%, also consistently going down over time. Non-first person subjects appear more frequently in oblique than first person subjects. In semi-formal and informal data on IGC, the proportion of oblique is around 60% for 2003 but goes down to slightly less than 49% in 2021 (the downwards trajectory appears consistent over time). On Twitter, the proportion is quite varied, with observations being somewhere between 80% and 22% for the whole period 2012–2022.

When applying forecasting models, such as those from the ETS and ARIMA family (see Chapter 7, Section 7.24), to the four time series (see relevant sections in Chapter 9.5), the models pick up on (hopefully) informative patterns in the series. In cases when the series show a downward trajectory, the models predict a continued decrease of oblique subjects over time. The IGC series for first person pronouns is particularly interesting as models suggest that oblique first person subjects will be virtually non-existent around 2040. For non-first person subjects, predictions for both the IGC and Twitter series tend to be flat and suggest future observations will fall somewhere on either side of the 50% proportion. This is likely due to the large variation in the data which, in turn, can be traced to the fact that there are relatively few examples behind each observation. Even when a slight downward trajectory appears to be observed (IGC series, Section 9.5.3), predictions for the future are still fat. An exception from this was when the trend in the Twitter series was isolated from the large amount of noise and used to predict the future. In this case, a rise in the use of oblique was suggested although this would very unlikely exceed 80% when taking into account prediction intervals.

Summarizing the results briefly, documentation of oblique subjects with *hlakka til* along with forecasting suggests a split in the behavior of first person subjects versus other subjects. While the use of oblique subjects is expected to decrease with first person subjects in sources like IGC and Twitter, other types of subjects are expected to continue to be attested in the oblique case. The predictions are based on patterns attested for the past 12–20 years. The question is thus why the data and the forecast do not fully confirm with the expectation that oblique subjects should generally be on the rise or oblique subjects are attested around 20% less than they should be. There are a few things to consider.

One of the issues that come to mind. Is it we are dealing with a written language source that is all the examples based on convenient E-language data in written format. The reason for mentioning this explicitly is that there might be a gap between what written language shows versus what the situation is actually like in the larger language community. Written language may be different from spoken language simply because individuals may be more conscious about language and grammar when writing than when speaking. In writing, there is always the possibility of reviewing and editing the text.

A second reason for the data not fully conforming to expectations is that oblique subjects with *hlakka til* have been the focus of prescriptive grammar and normative pressure for over 60 years. Every individual who goes through the Icelandic school system is bound to have come across a phrase that tells them that “nominative is correct, oblique is incorrect” with this particular predicate. Due to this, it is not unlikely that many individuals may be particularly conscious about using the predicate. It might even be the case that the heavy focus on subject case marking with *hlakka til* has started to affect how the predicate is acquired. This might in turn be reflected in the diminished oblique with first person subjects. Those who favor prescriptivism might take the results and predictions as a sign that the constant discussion about the predicate for more than half a century is finally having an effect. This view might be premature, and things might be slightly more complex than that.

A third factor worth noting is that the study in this chapter is based on different kinds of data than has generally been used to test case marking with *hlakka til*. Earlier studies largely focus on young children (Svavarsdóttir 1982, Eythórsson & Jónsson 2003, Thráinsson et al. 2015, Nowenstein 2023) although some also surveyed older individuals (Thráinsson et al. 2015). The studies were based on questionnaires where individuals typically had one chance of using the predicate. The present study, on the other hand, is based on attested examples of *hlakka til* in written material. In other words, it looks at the proportion of traditional versus innovative variante where each speaker may have contributed more than one example to the data. The way the change is measured may therefore have an affect on how it is perceived to play out in the time series.

Related to the third point and how the same individual may have contributed multiple examples to the dataset one might ask the question: who is contributing towards the data? There is a possibility that the data is skewed in favor of individuals who conform more to language standards and have been able to consciously acquire nominatives with *hlakka til*, at least in writing. Perhaps there is a correlation between those who conform less to standards, write less on the internet, write less in general and those who use oblique subjects with *hlakka til*. If this is the case, the results from relying on written language might reflect some facts about Icelandic society, potentially the fear of being stigmatized. This would need to be investigated further.

A final thing to note about the present study is how many years the time series actually cover. The answer is, about 19–20 years at the most so an interesting question is whether a real change should be expected to be observed within that period. This is, no doubt, partially a theoretical question, i.e., how do we separate “noise” that arises from sampling data from meaningful signals? I believe this is a question that can only be answered by practicing forecasting systematically. For now, it is worth noting that both of the time series containing information about the proportion of oblique first person subjects appeared to consistently show a downward trend over time. The negative trend became even more noticeable when the series had been decomposed. How reliable this information is, and how to properly balance separating expected variation in observations (noise) and meaningful trajectory/change, is a task for the future.

10 Conclusion

10.1 Recapping the aim and scope of the dissertation

This dissertation set out to deal with questions related to language forecasting. In particular it was concerned with whether it is possible to predict a future state of a language, a task that has by many (Entwistle 1953:41; Keller 1994:72; Labov 1994:10; Croft 2000:3), although not all (Sanches-Stockhammer 2015; Van de Velde, 2017; Nevalainen 2015), been deemed unattainable. A claim was made that predicting the future of a language is indeed possible – even desirable – as it offers a new way to study language change (Chapter 2) and may even give rise to new expectations towards certain changes. Several chapters (Chapters 3–6) were dedicated to laying out important aspects of language forecasting, how it might be approached, and which issues might arise. The forecasting task includes identifying which questions need to be answered and what the context of those questions is (this is the forecasting problem), gathering appropriate data, deciding on forecasting methods and eventually producing a forecast, cf. Figure 10.1. Of course, any method chosen for a forecasting task needs to be compatible with available data.

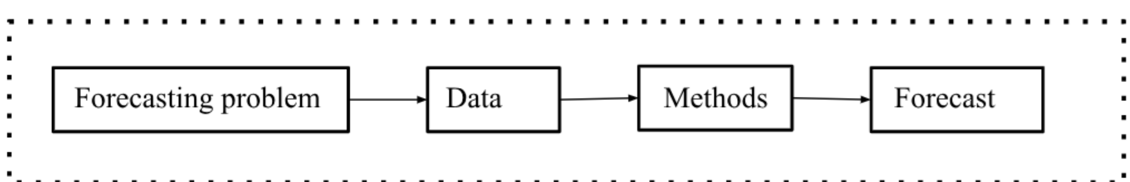


Figure 10.1. An overview of the major steps in a forecasting task. Chapters 2–6 in the dissertation explored topics related to the first two steps (the forecasting problem and data) while Chapters 7–9 applied some common forecasting methods relying on time series to produce forecasts.

In addition to providing a general background on language forecasting, two Icelandic case studies were presented (Chapters 8–9) with context of the forecasting and description of methods that were adopted (Chapter 7). The first study concerned changes in the complex prepositions *á bak við* 'behind' and *við hliðina* á 'next to'. The second study focused on changes in subject case marking with the predicate *hlakka til* 'look forward to'. Although changes in the complex prepositions *á bak við* 'behind' and *við hliðina* á 'next to' have been pointed out previously (Friðjónsson 2004, 2007 and Rögnvaldsson 2021), they have not been thoroughly discussed, nor have they been systematically documented. In contrast, changes in subject case marking with the predicate *hlakka til* 'look forward to' have been widely documented and discussed in the literature (e.g., Svavarsdóttir 1982; Jónsson & Eythórsson 2003). The documentation presented in Chapters 8 and 9 is novel in that it presents data in the form of regular time series that show the trajectory of the changes over roughly 11 to 19 years. The time series were used to produce forecasts for future time periods. The forecasts have given rise to expectations for these changes in upcoming years.

10.2 Summary and contribution

The present dissertation contributes both towards the understanding of language forecasting, including definition of forecasting problems, required data and available methods, and towards understanding of selected changes within Icelandic.

Language forecasting presents an interesting (and novel) way to study language change. It allows for the future to be used as a "playfield" for testing various hypotheses regarding expectations towards language change, including the understanding of propagation of novel variants through language communities. As such, it represents a new methodology for studying change over time, and new methodologies have the potential to alter the view on already known changes, generate new expectations and invite new questions.

Naturally, multiple different questions can be asked about language in the future, some which are easier to answer than others. For instance, attempting to foresee the introduction of a hitherto unattested linguistic variant might prove more challenging than predicting the propagation of a known linguistic variant through a community of speakers. Since language forecasting is in its infancy, it is appropriate to set up the task in such a way that makes it doable and feasible. This means that initial attempts might benefit from focusing on propagation of language change rather than predicting a hitherto unattested variant. Furthermore, it is practical to make predictions that rely on and are about E-language data in any form. This is because E-language data is more readily available and more easily checkable in the future than any prediction solely focusing on abstract grammars. Note, however, that any prediction about attested data also involves predicting what a future grammar can do. Focusing on these points does not exclude other types of forecasting questions and approaches, it simply tries to make the forecasting task as doable as possible in respect to defining how predictions are made and in respect to how easily predictions can be verified.

There are any approaches for generating predictions about the future, ranging from intuitive forecasting to complex statistical models (overview in e.g., Castle, Clements & Hendry 2019; Hyndman & Athanasopoulos 2021). For the studies presented in Chapter 8 and 9, a decision was made to rely on time series analysis and forecasting as this type of approach has many benefits, including being i) systematic, ii) requiring a well-defined type of data (time series), and iii) being widely used in forecasting. It should be noted that regular time series impose certain constraints on the data that is used in a forecasting model. Observations need to be made at regular intervals over an extended period of time, with one observation per time period. Additionally, there need to be enough observations to pick out important patterns that may emerge in the data. While there is no magic number for how many observations are required, it appears reasonable to use *at least* more than 15 observations for language data, although this may depend on how far into the future forecasts are to be made. This has implications for the documentation of language change. For instance, there is a need to determine how frequently changes need to be observed (every week, every month, every quarter, every year), keeping in mind that the chosen frequency affects how far into the future on-step-ahead forecast project. If a yearly time series is used, one-step-ahead forecast predicts the situation one year into the future. If a quarterly time series is used, one-step-ahead forecast predicts the situation one quarter into the future. This means that one may need to think about language data in a different way than is sometimes done as it suddenly matters whether data was generated in 2022 or 2023, or even if it was generated in spring 2023 or fall 2023. Thus, practicing language forecasting using regular time series data can be viewed as an argument in favor of more systematic documentation of selected changes, using smaller (quarterly and yearly) time-scales.

A further point made in the dissertation is that the future is not a single undefined period. Rather, it consists of adjacent points in time (Figure 10.2) and one needs to be explicit about which points (or which periods) in the future forecasts are to be made about.

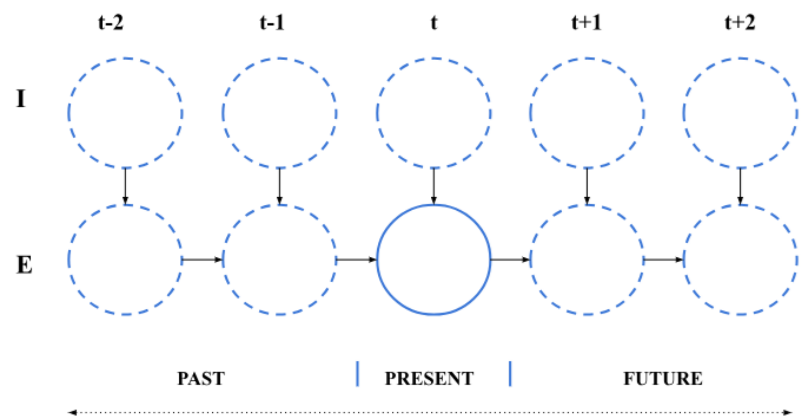


Figure 10.2. Snapshots of a language at various points in time. The past and future are not single undefined periods but can be broken up into smaller units that are ordered in time.

Language forecasts, just like any other forecasts, can be of short-, medium- and long-range. Although the exact length of these remains to be properly determined, it was proposed here that short-range forecasts might be concerned with predictions 1–15 years into the future, medium-range forecasting with 15–30 years in the future, and long-range forecasting with anything more than 30 years into the future. The definition of these was partially based on how far into the future predictions may be reasonably made relying on yearly time series data and loosely based on the idea of ‘generations’ which may cover a span of 15–30 years, although this depends on the exact definition (see Chapter 5). A claim was made that language forecasting should prioritize short- to medium-range forecasts as these can be properly checked and evaluated earlier than long-range forecasts, thus providing more rapid feedback on suitability of methods and data.

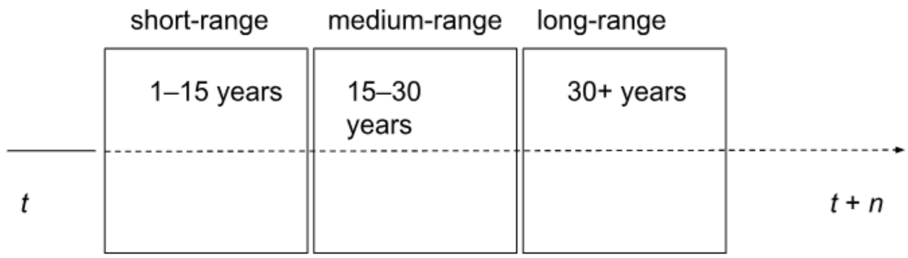


Figure 10.3. Language forecasting, like any other forecasting, can be concerned with producing short-range, medium-range or long-range forecasts. Although the exact length of these remains to be properly determined, some suggestions were made.

It should be emphasized that forecasts are always made within a certain context and with a particular audience in mind. Keeping this in mind, it was argued that a holistic approach to language forecasting needs to be adopted. This involves providing proper background information on the changes under investigation, along with a commentary on the data that is used and the forecasts that are generated. Without such reports, it may be difficult to obtain meaningful results from forecasts or properly evaluate them.

As noted earlier, the present dissertation does not only deal with language forecasting in a general way, but also includes two case studies which provide novel documentation of two types of changes within Icelandic, along with forecasts that show how the propagation of these changes may unfold in the next few years.

The first case study (Chapter 8) focused on the complex prepositions *á bak við* 'behind' and *við hliðina á* 'next to' which are frequently encountered in a simplified form where the initiation preposition (**P$_1$**) has been dropped, i.e., *bakvið* 'behind' and *hliðiná* 'next to'. The study represents the first serious documentation of variation in these phrases. The changes are viewed in light of grammaticalization, which arguably consists of semantic bleaching, syntactic reanalysis and phonological reduction. The dropping of the initial $\mathbf{P_1}$ was subsumed under phonological reduction and, in light of grammaticalization, it was expected to increase over time.

The documentation of *á bak við* 'behind' and *við hliðina á* 'next to' revealed various things. First, it suggested that *á bak við* was further along in the grammaticalization process than *við hliðina á*. This first and foremost reflected in *á bak við* showing more frequent dropping of $\mathbf{P_1}$ than *við hliðina á*. These results are not necessarily surprising as documentation of these changes at older stages of Icelandic show that grammaticalization appears to start earlier in *á bak við* than *við hliðina á* (Chapter 8, Sections 8.3.4–8.3.5). If the phrases are following the same kind of trajectory at similar speed (although, that is not at all guaranteed), it makes sense that the *á bak við* should be further along in the grammaticalization process than *við hliðina á.*

Second, the results show that these changes behave differently in different types of data, namely that the initial $\mathbf{P_1}$ is dropped slightly more in data from Twitter than in data from IGC. This has some implications for forecasting as it emphasizes that one needs to be aware of the type of data used to document changes and make forecasts. Most importantly, it needs to be consistent over time.

Third, although time-series documentation of the two complex prepositions shows in general that dropping of $\mathbf{P_1}$ remains relatively stable over time, one series shows a consistent downward trajectory which is picked up by forecasting models and assumed to continue into the future. This is the series for *á bak við*, based on data from IGC between 2003 and 2021. The downward trajectory is unexpected if dropping of $\mathbf{P_1}$ is understood as a part of grammaticalization, i.e., one would expect an increase in dropping of $\mathbf{P_1}$ not a decrease. The unexpected directionality may potentially be attributed to the nature of IGC data, i.e., some of it (the semi-formal data) is from online news media such as *Fréttablaðið.is* and *Mbl.is*. The change has been discussed by both Friðjónsson (2004, 2007) and Rögnvaldsson (2021). It is possible that awareness of this change among speakers may have caused a decline in use of the innovative variant over time. Another explanation would be that $\mathbf{P_1}$ is dropped less over time because Icelandic generally likes multi-word prepositions (Berthele et al. 2015). Under this view, the unexpected directionality might be taken to reflect some instability in the prepositional system that remains to be resolved. For further information on these results, the reader is directed to Chapter 8.

The second case study (Chapter 9) focused on variation in subject case marking with the predicate *hlakka til* 'look forward to'. The predicate originally appeared with a nominative subject, but is now frequently attested with either an accusative or dative subject. The change has been documented in some detail, although not at consistent time intervals. The first large-scale study is from the early 1980s (Svavarsdóttir 1982) with subsequent large-scale and smaller-scale studies being conducted in the 2000s (e.g., Jónsson & Eythórsson 2003; Thráinsson 2013). The present study adds a novel type of documentation through time series data. The time series consist of yearly and quarterly observations over the period 2003–2021 (IGC, yearly observations) and 2012–2022 (Twitter, quarterly observations). The results indicate that individuals who use an overt subject *hlakka til*, use first person subjects more frequently than other types of subjects. Thus, first person subjects appear 22,198 times in the IGC data and 9,629 times in the Twitter data. Other types of subjects are found 6,376 and 1,177 times in IGC and Twitter respectively.

The results also confirm a split in case marking with first person subjects versus other types of subjects. A difference in behavior of these was observed in the study Svavarsdóttir (1982) conducted in the 80s, where 11-year-old children used nominative more frequently for first person subjects than for third person subjects. In the present study, the split between first person subjects and other subjects is observed in two ways. First, non-first person subjects generally appear more frequently with an oblique subject than first person subjects. Second, the use of oblique appears to decline over time when the subject is a first person pronoun. This is shown both in the IGC data and to a lesser extent in Twitter data. A forecast based on the IGC data suggests that first person oblique subjects may virtually be non-existent in 2040. A forecast based on Twitter data shows that the proportion of first person oblique subjects will likely be less than 10 percent in 2040, although the proportion is not predicted to go down to zero.

Data for other types of subjects appear to show more stability over time (although this is not completely clear), with the proportion of oblique subjects being close to 50% and no drastic changes being predicted. It should be noted that data for non-first person subjects is based on relatively few examples and consequently the forecasts are less certain and should be taken with a grain of salt. For further information on results related to subject case marking with *hlakka til* ‘look forward to’, the reader is directed to Chapter 9.

Taken together, the Icelandic case studies provide valuable insight into language forecasting and the documentation of changes in the form of regular time series. In the case of the latter, results suggest that observable changes may occur in as little as 11–19 years in written language data. While these this needs to be investigated further, for instance in respect to whether these represent actual changes or only normal variation and whether this is due to individuals converging on variants over time or to increase in individuals who use particular variants, it does lend support to the idea that quarterly and yearly time series are appropriate for documenting language change.

It was hypothesized that seasonality, i.e., fluctuation in time series data at consistent and fixed (seasonal) intervals, would not play a role in the documentation and predictions of language change. This assumption seems reasonable given that the use of two or more linguistic variants in direct competition is not expected to depend on time of day, week, month or year.[116] However, in the case of quarterly time series, seasonality did emerge in both case studies presented in this dissertation. Nota bene, in all cases the seasonality was extremely small. In fact, it was virtually nonexistent and likely emerged only due to the chosen sampling frequency (quarterly observations) and due to the nature of the time series decomposition shown, which explicitly splits observations at any given time into a trend-cycle component, seasonal component and a remainder component. Thus, the claim still stands that seasonal variation is not meaningful in the context of language change.

In general, the models that were deemed most appropriate for forecasting each of the time series were rather simple in nature and mostly focused on modeling major trends (or stability) in the time series. This fact may indicate that complex forecasting models may not be necessary to analyze and predict the propagation of language change. Instead, it may be enough to rely on methods that are designed to model general directions in time series. Figure 10.1 shows the ARIMA and ETS models that were used for the time series in Chapter 8 and 9. Note how the same models were used for more than one series. A

[116] It is possible that the *use* of certain linguistic structures may show seasonal or cyclic behavior. However, it is not expected that the proportion of one variant (as opposed to another) should show a meaningful seasonal pattern. They might show cyclic patterns.

description of each of these models is provided in Table 10.2, showing that these either pick out trends (linear methods and exponential smoothing) or produce white noise based on the input data.

		ARIMA	**ETS**
IGC	*á bak við*	0,1,0	AAN
	við hliðina á	0,0,0	ANN
	1PERS	0,1,1	AAN
	OTHER	0,1,0	MNN
Twitter	*á bak við*	0,1,1	ANN
	við hliðina á	0,1,1	MNN
	1PERS	0,1,1	ANN
	OTHER	0,1,1	AAN

Table 10.1 An overview of ARIMA and ETS models that were used.

The ETS and ARIMA models that were used		
ETS	AAN	Holt's linear method w/additive errors
	ANN	simple exponential smoothing w/additive errors
	MNN	simple exponential smoothing w/multiplicative errors
ARIMA	0,1,1	equivalent to simple exponential smoothing
	0,1,0	random walk
	0,0,0	white noise

Table 10.2 The most common ETS model was ETS(A,A,N) which is Holt's linear method with additive errors. The most common ARIMA model was ARIMA(0,1,1) which is essentially a simple exponential smoothing model.

These results can be claimed to be in line with an observation made by Van de Velde and Petré (2020:349), namely that "ARIMA modelling is definitely an overkill, but it can be done in principle". As a minor object to the overkill-comment, it should be noted that ARIMA models are versatile in nature and there is no reason not to use them. They can deal with both stationary and non-stationary time series and model seasonal and non-seasonal data. Additionally, some ARIMA models are actually quite simple and may be regarded as equivalent to some linear exponential smoothing models (see discussion in Hyndman & Athanasopoulos 2021:306–307). In short, ARIMA models offer a lot of flexibility without demanding that all possible parameters need to be included in the modeling and forecasting, although using them is almost like relying on a graphing calculator to compute simple addition and subtractions. However, the question still remains as to when ARIMA or ETS are appropriate and when more simple models such as a Naïve, Seasonal naïve, Mean or Drift model might be used. In some cases, a simplistic model like the Naïve model (Twitter, *á bak við*) or the Mean model (Twitter, *við hliðina á*) gave more accurate predictions for the test period in the case studies in Chapter 8 and 9. While this may be simply due to chance (the test period was very short), it might also be that simple models like the Naïve or the Mean model might perform well for changes that show stability over time. Such stability might, for instance, be expected in the initial phase of a change (recall the starting tail of an S-curve) or in the final phase of a change (recall the end-tail of an S-curve). Note that using the models would require that the time series only covered a stable phase as the Naïve and Mean models restrict what the future might have in store. If a series shows a mild-life S-curve, these models would probably not perform well. The Naïve model would assume the propagation would come to a halt (all future observations would equal the last observation in the series) while the Mean model might indicate "diminished" propagation of the relevant linguistic feature, because future observations would be hypothesized to simply be the mean of the observed series. For these reasons, one might think that models predicting stability would *not* be appropriate for changes that are in the "quick quick" phase (Denison 2002:56) or, in the terms of

Nevalainen and Raumolin-Brunberg (2017:54–55) who use terminology from Labov (1994:67, 79 – 83), being new and vigorous (15 and 35 per cent), mid-range (36 – 65 per cent) or nearing completion (66 and 85 per cent). However, as shown in the case study on non-first person oblique subjects with *hlakka til* in IGC and on Twitter, a stable future might still be predicted when a novel variant makes up about 50% of attested example. Thus, the type of models that are deemed appropriate does not map directly onto where a change is at in terms of an S-curve propagation. Model choice depends on signals from the time series that is used.

It is clear that a lot of things remain to be done in the realm of language forecasting. As a final contribution, I wish to note that I believe that language forecasting may, in the future, offer an insight into how to deal with the actuation problem (Weinreich et al. 1968). The insight may be in the form of more detailed documentation of various changes over time, or through observations and predictions about changes in the level of time series over time. What has been presented here is only the beginning of language forecasting.

10.3 Remaining questions

A single dissertation is not capable of answering all questions relating to language forecasting. While an attempt has been made to define what a language forecasting task might look like and provide some examples of how it might be carried out, it is the view of the author that the present work has given rise to multiple new questions regarding language change and language forecasting. Questions that are still open pertain can be placed in two categories, i) those that relate to the Icelandic case studies presented in

Chapters 8 and 9, and i) those that pertain to various aspects of the whole forecasting process.

In the context of the Icelandic case studies, one might question the observed directionality in relation to some changes. As noted in the discussion in Chapters 8 and 9, expected directionality did not always emerge. In some cases (e.g., for 1st person subjects with *hlakka til*), a diminished use of the novel variant was observed. This raises questions about the material that was used to create the time series, e.g., who are the individuals contributing to the convenient E-language data and are we observing the same individuals over time or different individuals?

Another question raised by the case studies is when to treat a phenomenon as a single change or as multiple changes. This was particularly noticeable in the case *of hlakka til* where a split in case marking is observed between first person subjects and other subjects. It would have been possible to treat all subjects together, but prior studies on the change suggested that first person subjects might behave differently from other subjects. This, of course, raises questions about where the data might be split up in other ways as well, for instance whether examples in the present tense show the same behavior as examples in the past tense.

There are good chances that prescriptivism may play a role in the propagation of some changes. How big of a role it plays remains to be studied further. It would, for instance, be a worthwhile task to attempt to quantify in some way the effect of prescriptivism, normative pressures and copy editing. Perhaps this might be done with two or more time series that contain information about the same change but are hypothesized to be of a different nature. If one had a series based on data where copy-editing and prescriptivism was thought to play a large role, and another series where at least one of those things was thought to be absent, it might be possible to quantify the role of copy-editing and prescriptivism over time. This, however, is a task for the future.

For questions that pertain to various aspects of the forecasting process, these are listed below and include:

DATA Which type of data is the most appropriate for language forecasting? If it includes convenient E-language data, what kind of sources are most suitable w.r.t. language style etc.? Additionally, in what way is it most appropriate to measure changes over time to give good forecasting results? Is there a meaningful difference between measuring the overall proportion of innovative variants in attested data versus measuring how many individuals use particular variants at any given time?

FORECAST LENGTH How far into the future is it realistic to make predictions about the propagation of language change? How much does this depend on the frequency of the time series that is used? How does forecast accuracy interact with short-, medium- and long-range forecasts? Definitions of the length of these will likely need to be revised in the future.

ACCURACY It is unrealistic to expect predictions to be 100% spot-on all the time. Some variation must be allowed, if only to account for random variation that may arise due to sampling and how changes are measured. The question is then, what is a good accuracy for predictions about language in the future? How large should forecast intervals be?

METHODS The forecasts produced here mostly relied on raw time series data, although some forecasts were also made where the trend-cycle component of the series served as an input for a forecasting model instead. While using the trend-cycle component resulted in narrower forecast intervals, the question remains whether isolating the trend-cycle component should be generally practiced or not. If the trend-cycle component is used to predict future observations, choices need to be made as to how the component is extracted from the raw series. This can be done using traditional moving averages of various orders or by relying on decomposition methods that do not shorten the time series. A second choice that needs to be made is how smooth the trend-cycle component should be. In the case studies in Chapters 8 and 9, the trend-cycle window was set to 13 to make the component extremely smooth over time. It may be the case that better results would be obtained by using a smaller trend-cycle window. If a time series is smoothed too much, the risk is that some meaningful signal in the data will be lost. If it is not smoothed at all, a lot of unnecessary noise may be included in the prediction and that may affect how wide the forecast intervals turn out to be. Essentially, the question is how to separate signals pointing to language change from components that are more related to language use. In a way, this is like asking how to extract information on climate by observing everyday weather. Another question relating to methodology is what the optimal frequency is for observing changes over time and how long a time series should be.

EXPECTATIONS TOWARD CHANGES Using time series raises the question of how frequently the level of the series is expected to change over time. From the results of the case studies in Chapters 8 and 9, it appears that changes may be detected in as little as 11–19 years. However, the question is still up to what extent these are meaningful changes and up to what extent they represent random variation. Again, this gets to the question of the difference between predicting weather and predicting climate. When is the prediction about variation in language use and when is it about language change?

RELATIONSHIP BETWEEN TIME SERIES Finally, a considerable amount of work remains to be done in investigating the relationships between two or more time series. There are multiple reasons for wanting to do this. One such reason is to explore the interaction of two changes or two grammatical variants over time. If two or more grammatical features are hypothesized to interact synchronically, it is expected that these should also show some kind of relationship over time. In order to investigate such relationships, one would need two or more time series based on comparable sources. This work remains for the future.

While I do not pretend to have answers to the questions proposed here, I hope I have provided sufficient insight into language forecasting and pointed towards a direction in which the solutions to these questions may be sought in the future. In the end, forecasting does not have to only be about producing 100% accurate predictions, but rather about what can be learned from making predictions. – *Vituð ér enn eða hvat?*

Appendix A

Time series[117]

The IGC time series consist of data from semi-formal and informal sources (see Chapters 8 and 9).

Time series for *hlakka til* are based on oblique-biased annotation (see Chapter 9).

[117] These time series took ages to create. I do not know how many hours I spent in front of the computer just annotating examples, identifying features I was interested in. Of course, there is a possibility that my weary eyes may have made an occasional mistake. However, I do not believe that these would be large enough to affect the overall analysis and results.

IGC *á bak við*					
DATE	**PNP**	**a**	**NOPE**	**Total**	**Source**
1/1/2000	bak	310	85	395	IGC
1/1/2001	bak	646	276	922	IGC
1/1/2002	bak	687	323	1010	IGC
1/1/2003	bak	1166	610	1776	IGC
1/1/2004	bak	2533	1166	3699	IGC
1/1/2005	bak	4137	1855	5992	IGC
1/1/2006	bak	4100	1772	5872	IGC
1/1/2007	bak	4028	1762	5790	IGC
1/1/2008	bak	4544	1784	6328	IGC
1/1/2009	bak	4498	1577	6075	IGC
1/1/2010	bak	3805	1372	5177	IGC
1/1/2011	bak	3883	1211	5094	IGC
1/1/2012	bak	3739	1029	4768	IGC
1/1/2013	bak	3366	842	4208	IGC
1/1/2014	bak	3323	795	4118	IGC
1/1/2015	bak	3143	783	3926	IGC
1/1/2016	bak	3023	665	3688	IGC
1/1/2017	bak	3107	707	3814	IGC
1/1/2018	bak	3253	757	4010	IGC
1/1/2019	bak	3579	737	4316	IGC
1/1/2020	bak	3105	679	3784	IGC
1/1/2021	bak	3326	619	3945	IGC

IGC *við hliðina á*					
DATE	**PNP**	**NOPE**	**vid**	**Total**	**Source**
1/1/2000	hlid	5	420	425	IGC
1/1/2001	hlid	37	177	214	IGC
1/1/2002	hlid	62	272	334	IGC
1/1/2003	hlid	92	702	794	IGC
1/1/2004	hlid	120	1756	1876	IGC
1/1/2005	hlid	270	2561	2831	IGC
1/1/2006	hlid	287	2544	2831	IGC
1/1/2007	hlid	264	2660	2924	IGC
1/1/2008	hlid	333	2620	2953	IGC
1/1/2009	hlid	223	1933	2156	IGC
1/1/2010	hlid	279	4367	4646	IGC
1/1/2011	hlid	186	1709	1895	IGC
1/1/2012	hlid	156	1334	1490	IGC
1/1/2013	hlid	88	901	989	IGC
1/1/2014	hlid	41	558	599	IGC
1/1/2015	hlid	32	332	364	IGC
1/1/2016	hlid	16	151	167	IGC
1/1/2017	hlid	6	68	74	IGC
1/1/2018	hlid	3	44	47	IGC
1/1/2019	hlid	5	49	54	IGC
1/1/2020	hlid	87	3232	3319	IGC
1/1/2021	hlid	0	27	27	IGC

IGC *hlakka til* 1pers subjects					
DATE	**SUBJ**	**OBL**	**NOM**	**Total**	**Source**
1999-1	1PERS	1	60	61	IGC
2000-1	1PERS	3	67	70	IGC
2001-1	1PERS	9	105	114	IGC
2002-1	1PERS	18	89	107	IGC
2003-1	1PERS	148	306	454	IGC
2004-1	1PERS	517	952	1469	IGC
2005-1	1PERS	548	1332	1880	IGC
2006-1	1PERS	501	1173	1674	IGC
2007-1	1PERS	503	1164	1667	IGC
2008-1	1PERS	422	1229	1651	IGC
2009-1	1PERS	244	999	1243	IGC
2010-1	1PERS	261	1014	1275	IGC
2011-1	1PERS	279	1165	1444	IGC
2012-1	1PERS	190	1125	1315	IGC
2013-1	1PERS	109	1084	1193	IGC
2014-1	1PERS	115	1013	1128	IGC
2015-1	1PERS	76	837	913	IGC
2016-1	1PERS	71	881	952	IGC
2017-1	1PERS	44	914	958	IGC
2018-1	1PERS	59	798	857	IGC
2019-1	1PERS	58	1187	1245	IGC
2020-1	1PERS	23	385	408	IGC
2021-1	1PERS	27	445	472	IGC

IGC *hlakka til* other subjects					
DATE	**SUBJ**	**NOM**	**OBL**	**Total**	**Source**
1/1/1999	OTHER	22	31	53	IGC
1/1/2000	OTHER	17	41	58	IGC
1/1/2001	OTHER	21	49	70	IGC
1/1/2002	OTHER	33	36	69	IGC
1/1/2003	OTHER	58	86	144	IGC
1/1/2004	OTHER	134	186	320	IGC
1/1/2005	OTHER	152	240	392	IGC
1/1/2006	OTHER	125	259	384	IGC
1/1/2007	OTHER	206	249	455	IGC
1/1/2008	OTHER	161	256	417	IGC
1/1/2009	OTHER	146	207	353	IGC
1/1/2010	OTHER	152	218	370	IGC
1/1/2011	OTHER	155	216	371	IGC
1/1/2012	OTHER	190	165	355	IGC
1/1/2013	OTHER	195	170	365	IGC
1/1/2014	OTHER	201	185	386	IGC
1/1/2015	OTHER	176	158	334	IGC
1/1/2016	OTHER	216	173	389	IGC
1/1/2017	OTHER	187	145	332	IGC
1/1/2018	OTHER	148	174	322	IGC
1/1/2019	OTHER	189	208	397	IGC
1/1/2020	OTHER	68	67	135	IGC
1/1/2021	OTHER	80	75	155	IGC

Twitter *á bak við*					
DATE	**PNP**	**NOPE**	**a**	**Total**	**Source**
1/1/2009	bak	6	7	13	Twitter
4/1/2009	bak	4	14	18	Twitter
7/1/2009	bak	13	29	42	Twitter
10/1/2009	bak	8	23	31	Twitter
1/1/2010	bak	20	31	51	Twitter
4/1/2010	bak	15	27	42	Twitter
7/1/2010	bak	13	27	40	Twitter
10/1/2010	bak	17	36	53	Twitter
1/1/2011	bak	19	51	70	Twitter
4/1/2011	bak	38	65	103	Twitter
7/1/2011	bak	25	58	83	Twitter
10/1/2011	bak	27	90	117	Twitter
1/1/2012	bak	51	109	160	Twitter
4/1/2012	bak	65	117	182	Twitter
7/1/2012	bak	48	70	118	Twitter
10/1/2012	bak	48	96	144	Twitter
1/1/2013	bak	60	100	160	Twitter
4/1/2013	bak	70	102	172	Twitter
7/1/2013	bak	71	110	181	Twitter
10/1/2013	bak	111	163	274	Twitter
1/1/2014	bak	118	183	301	Twitter
4/1/2014	bak	126	176	302	Twitter
7/1/2014	bak	102	143	245	Twitter
10/1/2014	bak	96	146	242	Twitter
1/1/2015	bak	117	202	319	Twitter
4/1/2015	bak	164	229	393	Twitter
7/1/2015	bak	128	221	349	Twitter
10/1/2015	bak	146	237	383	Twitter
1/1/2016	bak	130	228	358	Twitter
4/1/2016	bak	172	312	484	Twitter
7/1/2016	bak	171	260	431	Twitter
10/1/2016	bak	120	236	356	Twitter
1/1/2017	bak	135	214	349	Twitter
4/1/2017	bak	101	175	276	Twitter
7/1/2017	bak	98	160	258	Twitter
10/1/2017	bak	115	225	340	Twitter
1/1/2018	bak	126	200	326	Twitter
4/1/2018	bak	140	268	408	Twitter
7/1/2018	bak	87	198	285	Twitter
10/1/2018	bak	111	225	336	Twitter
1/1/2019	bak	115	311	426	Twitter
4/1/2019	bak	110	244	354	Twitter
7/1/2019	bak	114	180	294	Twitter
10/1/2019	bak	113	202	315	Twitter
1/1/2020	bak	116	211	327	Twitter
4/1/2020	bak	116	229	345	Twitter
7/1/2020	bak	120	207	327	Twitter
10/1/2020	bak	138	300	438	Twitter
1/1/2021	bak	127	297	424	Twitter
4/1/2021	bak	157	303	460	Twitter
7/1/2021	bak	211	445	656	Twitter
10/1/2021	bak	183	467	650	Twitter
1/1/2022	bak	166	435	601	Twitter
4/1/2022	bak	112	356	468	Twitter
7/1/2022	bak	150	273	423	Twitter
10/1/2022	bak	163	341	504	Twitter

Twitter *við hliðina á*					
DATE	**PNP**	**NOPE**	**vid**	**Total**	**Source**
4/1/2009	hlid	1	3	4	Twitter
10/1/2009	hlid	3	7	10	Twitter
7/1/2010	hlid	3	4	7	Twitter
10/1/2010	hlid	2	8	10	Twitter
1/1/2011	hlid	1	13	14	Twitter
4/1/2011	hlid	4	31	35	Twitter
7/1/2011	hlid	6	32	38	Twitter
10/1/2011	hlid	4	31	35	Twitter
1/1/2012	hlid	14	39	53	Twitter
4/1/2012	hlid	12	56	68	Twitter
7/1/2012	hlid	6	52	58	Twitter
10/1/2012	hlid	8	62	70	Twitter
1/1/2013	hlid	5	70	75	Twitter
4/1/2013	hlid	12	57	69	Twitter
7/1/2013	hlid	9	74	83	Twitter
10/1/2013	hlid	20	125	145	Twitter
1/1/2014	hlid	32	152	184	Twitter
4/1/2014	hlid	35	128	163	Twitter
7/1/2014	hlid	18	114	132	Twitter
10/1/2014	hlid	31	151	182	Twitter
1/1/2015	hlid	30	160	190	Twitter
4/1/2015	hlid	33	207	240	Twitter
7/1/2015	hlid	42	195	237	Twitter
10/1/2015	hlid	28	210	238	Twitter
1/1/2016	hlid	32	193	225	Twitter
4/1/2016	hlid	40	262	302	Twitter
7/1/2016	hlid	38	220	258	Twitter
10/1/2016	hlid	29	205	234	Twitter
1/1/2017	hlid	18	219	237	Twitter
4/1/2017	hlid	16	136	152	Twitter
7/1/2017	hlid	16	153	169	Twitter
10/1/2017	hlid	22	180	202	Twitter
1/1/2018	hlid	18	162	180	Twitter
4/1/2018	hlid	13	159	172	Twitter
7/1/2018	hlid	13	154	167	Twitter
10/1/2018	hlid	16	125	141	Twitter
1/1/2019	hlid	14	161	175	Twitter
4/1/2019	hlid	13	153	166	Twitter
7/1/2019	hlid	11	132	143	Twitter
10/1/2019	hlid	11	110	121	Twitter
1/1/2020	hlid	17	136	153	Twitter
4/1/2020	hlid	8	121	129	Twitter
7/1/2020	hlid	21	140	161	Twitter
10/1/2020	hlid	18	143	161	Twitter
1/1/2021	hlid	22	166	188	Twitter
4/1/2021	hlid	27	196	223	Twitter
7/1/2021	hlid	33	254	287	Twitter
10/1/2021	hlid	26	208	234	Twitter
1/1/2022	hlid	19	197	216	Twitter
4/1/2022	hlid	29	178	207	Twitter
7/1/2022	hlid	34	170	204	Twitter
10/1/2022	hlid	25	152	177	Twitter
1/1/2009	hlid	0	4	4	Twitter
7/1/2009	hlid	0	8	8	Twitter
1/1/2010	hlid	0	6	6	Twitter
4/1/2010	hlid	0	8	8	Twitter

Twitter *hlakka til* 1pers subjects					
DATE	**SUBJ**	**OBL**	**NOM**	**Total**	**Source**
2009 Q1	1PERS	2	8	10	Twitter
2009 Q2	1PERS	1	13	14	Twitter
2009 Q3	1PERS	1	8	9	Twitter
2009 Q4	1PERS	1	14	15	Twitter
2010 Q1	1PERS	3	15	18	Twitter
2010 Q3	1PERS	3	17	20	Twitter
2010 Q4	1PERS	3	28	31	Twitter
2011 Q1	1PERS	5	25	30	Twitter
2011 Q2	1PERS	13	64	77	Twitter
2011 Q3	1PERS	12	47	59	Twitter
2011 Q4	1PERS	16	92	108	Twitter
2012 Q1	1PERS	19	86	105	Twitter
2012 Q2	1PERS	23	110	133	Twitter
2012 Q3	1PERS	22	93	115	Twitter
2012 Q4	1PERS	32	120	152	Twitter
2013 Q1	1PERS	16	107	123	Twitter
2013 Q2	1PERS	20	120	140	Twitter
2013 Q3	1PERS	32	143	175	Twitter
2013 Q4	1PERS	57	241	298	Twitter
2014 Q1	1PERS	61	221	282	Twitter
2014 Q2	1PERS	44	273	317	Twitter
2014 Q3	1PERS	45	200	245	Twitter
2014 Q4	1PERS	34	211	245	Twitter
2015 Q1	1PERS	38	241	279	Twitter
2015 Q2	1PERS	50	286	336	Twitter
2015 Q3	1PERS	40	205	245	Twitter
2015 Q4	1PERS	43	256	299	Twitter
2016 Q1	1PERS	32	222	254	Twitter
2016 Q2	1PERS	47	279	326	Twitter
2016 Q3	1PERS	34	232	266	Twitter
2016 Q4	1PERS	37	260	297	Twitter
2017 Q1	1PERS	22	216	238	Twitter
2017 Q2	1PERS	10	162	172	Twitter
2017 Q3	1PERS	20	156	176	Twitter
2017 Q4	1PERS	27	236	263	Twitter
2018 Q1	1PERS	14	178	192	Twitter
2018 Q2	1PERS	17	194	211	Twitter
2018 Q3	1PERS	11	177	188	Twitter
2018 Q4	1PERS	20	183	203	Twitter
2019 Q1	1PERS	10	157	167	Twitter
2019 Q2	1PERS	9	179	188	Twitter
2019 Q3	1PERS	9	143	152	Twitter
2019 Q4	1PERS	14	143	157	Twitter
2020 Q1	1PERS	11	171	182	Twitter
2020 Q2	1PERS	12	186	198	Twitter
2020 Q3	1PERS	15	171	186	Twitter
2020 Q4	1PERS	20	286	306	Twitter
2021 Q1	1PERS	25	201	226	Twitter
2021 Q2	1PERS	18	235	253	Twitter
2021 Q3	1PERS	24	237	261	Twitter
2021 Q4	1PERS	24	319	343	Twitter
2022 Q1	1PERS	19	184	203	Twitter
2022 Q2	1PERS	19	150	169	Twitter
2022 Q3	1PERS	11	166	177	Twitter
2022 Q4	1PERS	15	162	177	Twitter
2010 Q2	1PERS	0	9	9	Twitter

Twitter *hlakka til* other subjects					
DATE	**SUBJ**	**OBL**	**NOM**	**Total**	**Source**
2009 Q1	OTHER	1	0	1	Twitter
2009 Q2	OTHER	3	0	3	Twitter
2009 Q3	OTHER	1	0	1	Twitter
2009 Q4	OTHER	3	2	5	Twitter
2010 Q1	OTHER	2	1	3	Twitter
2010 Q2	OTHER	3	0	3	Twitter
2010 Q3	OTHER	2	1	3	Twitter
2010 Q4	OTHER	2	1	3	Twitter
2011 Q1	OTHER	3	2	5	Twitter
2011 Q2	OTHER	12	4	16	Twitter
2011 Q3	OTHER	9	5	14	Twitter
2011 Q4	OTHER	14	12	26	Twitter
2012 Q1	OTHER	9	7	16	Twitter
2012 Q2	OTHER	21	7	28	Twitter
2012 Q3	OTHER	20	4	24	Twitter
2012 Q4	OTHER	17	11	28	Twitter
2013 Q1	OTHER	15	4	19	Twitter
2013 Q2	OTHER	21	12	33	Twitter
2013 Q3	OTHER	19	12	31	Twitter
2013 Q4	OTHER	20	15	35	Twitter
2014 Q1	OTHER	20	14	34	Twitter
2014 Q2	OTHER	17	8	25	Twitter
2014 Q3	OTHER	19	11	30	Twitter
2014 Q4	OTHER	23	15	38	Twitter
2015 Q1	OTHER	22	16	38	Twitter
2015 Q2	OTHER	15	8	23	Twitter
2015 Q3	OTHER	21	15	36	Twitter
2015 Q4	OTHER	14	39	53	Twitter
2016 Q1	OTHER	5	13	18	Twitter
2016 Q2	OTHER	13	20	33	Twitter
2016 Q3	OTHER	18	14	32	Twitter
2016 Q4	OTHER	14	18	32	Twitter
2017 Q1	OTHER	9	7	16	Twitter
2017 Q2	OTHER	17	17	34	Twitter
2017 Q3	OTHER	6	6	12	Twitter
2017 Q4	OTHER	5	13	18	Twitter
2018 Q1	OTHER	9	8	17	Twitter
2018 Q2	OTHER	11	10	21	Twitter
2018 Q3	OTHER	7	5	12	Twitter
2018 Q4	OTHER	10	14	24	Twitter
2019 Q1	OTHER	12	20	32	Twitter
2019 Q2	OTHER	9	9	18	Twitter
2019 Q3	OTHER	5	11	16	Twitter
2019 Q4	OTHER	13	15	28	Twitter
2020 Q1	OTHER	5	8	13	Twitter
2020 Q2	OTHER	8	17	25	Twitter
2020 Q3	OTHER	4	14	18	Twitter
2020 Q4	OTHER	17	15	32	Twitter
2021 Q1	OTHER	8	10	18	Twitter
2021 Q2	OTHER	12	23	35	Twitter
2021 Q3	OTHER	17	26	43	Twitter
2021 Q4	OTHER	17	36	53	Twitter
2022 Q1	OTHER	12	12	24	Twitter
2022 Q2	OTHER	11	8	19	Twitter
2022 Q3	OTHER	14	12	26	Twitter
2022 Q4	OTHER	10	7	17	Twitter

References

Ahrens, Donald C. 2007. *Meteorology today: an introduction to weather, climate, and the environment*. Belmont, CA: Thomson/Brooks/Cole.

Ammon, Ulrich. 2015. “On the social forces that determine what is standard in a language – with a look at the norms of non-standard language varieties.” *Bulletin VALS-ALSA, n° spécial, tome 3 / Bulletin suisse de linguistique appliquée*, 53–67.

Andersen, Henning. 1973. “Abductive and Deductive Change.” *Language, 49*(4):765–95, doi:10.2307/412063.

Andersen, Henning. 1990. “The structure of drift”. In Henning Andersen & Konrad Koerner (Eds.) *Historical Linguistics 1987: Papers from the 8th International Conference on Historical Linguistics* (pp. 1–20). Amsterdam: John Benjamins.

Anderson, Stephen R. 2016. “Synchronic versus diachronic explanation and the nature of the Language Faculty.” *Annual Review of Linguistics, 2*(1):11–31. doi:10.1146/annurev-linguistics-011415-040735.

Andrews, Avery D. 1976. “The VP-complement analysis in Modern Icelandic.” *Proceedings of the North East Linguistic Society, 6*:1–21 [Reprinted 1990 with small revisions in Maling & Zaenen (Eds.), pp. 165–185.]

Andrews, Avery D. 1982. “The representation of Case in Modern Icelandic.” In Joan Bresnan (Ed.), *The Mental Representation of Grammatical Relations*, pp. 427–503. Cambridge, Massachusetts: The MIT Press.

Angantýsson, Ásgrímur. 2011. *The syntax of embedded clauses in Icelandic and related languages* [Ph.D. dissertation, University of Iceland].

Arends, Jacques. 1986. “Genesis and development of the equative copula in Sranan.” In Pieter Muysken and Norval Smith (Eds.), *Substrata versus universals in Creole genesis* (pp. 103–27). Amsterdam and Philadelphia: John Benjamins.

Aristar, Anthony Rodrigues. 1991. “On Diachronic Sources and Synchronic Pattern: An Investigation into the Origin of Linguistic Universals.” *Language, 67*(1):1–33. doi: 10.2307/415537.

Árnason, Kristján. 1994–1995. “Tilraun til greiningar á íslensku tónfalli.” *Íslenskt mál og almenn málfræði* 16–17:99–131.

Árnason, Kristján. 2002. “Upptök íslensks ritmáls.” *Íslenskt mál og almenn málfræði, 24*:157–193.

Árnason, Kristján. 2003. “Language planning and the structure of Icelandic”. In Kristján Árnason (Ed.), *Útnorður: West-Nordic standardisation and variation* (pp. 193–218) Reykjavík: Iceland University Press.

Árnason, Kristján. 2011. *The phonology of Icelandic and Faroese*. Oxford: Oxford University Press.

Asbury, Anna, Berit Gherke, Henk van Riemsdijk, & Joost Zwarts. 2008. “Introduction”. In Anna Asbury, Jakub Dotlačil, Berit Gehrke & Rick Nouwen (Eds.), *Syntax and semantics of spatial P* (pp. 1–34). Amsterdam: John Benjamins.

Axelsdóttir, Katrín. 2002. “Hvarf eignarfornafnanna okkarr, ykkarr og yð(v)arr.” *Íslenskt mál og almenn málfræði, 24*:107–156.

Bailey, Charles-James N. 1973. *Variation and Linguistic Theory*. Washington, DC: Center for Applied Linguistics.

Bailey, Guy, Tom Wikle, Jan Tillery, and Lori Sand. 1991. “The Apparent Time Construct.” *Language Variation and Change, 3*:241–64.

Barðdal, Jóhanna. 2001. *Case in Icelandic: A synchronic, diachronic and comparative approach* [Ph.D. dissertation, Department of Scandinavian Languages, Lund University.

Barðdal, Jóhanna. 2002. “Oblique subjects in Icelandic and German.” *Working Papers in Scandinavian Syntax, 70*:61–99.

Barðdal, Jóhanna & Thórhallur Eythórsson. 2003. “The change that never happened: the story of oblique subjects.” *Journal of Linguistics, 39*(3):439–472.

Barðdal, Jóhanna & Thórhallur Eythórsson. 2009. “The origin of the oblique-subject construction: An Indo-European Comparison.” In Vit Bubenic, John Hewson & Sarah Rose (Eds.), *Grammatical Changes in Indo-European Languages: Papers presented at the workshop on Indo-European Linguistics at the XVIIIth International Conference on Historical Linguistics, Montreal, 2007* (pp. 179–194). Amsterdam: John Benjamins.

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

Barrie, Christopher & Justin Chun-ting Ho. 2021. “academictwitteR: an R package to access the Twitter Academic Research Product Track v2 API endpoint.” *Journal of Open Source Software, 6*(62):3272. doi:10.21105/joss.03272

Bauer, Laurie. 1994. *Watching English change: An introduction to the study of linguistic change in Standard Englishes in the twentieth century*. London: Longman.

Benediktsson, Hreinn. 1976. “Ísl. vera að + nafnh.: Aldur og uppruni.” In Lars Svensson, Anne Marie Wieselgren & Åke Hansson (Eds.), *Nordiska studier i filologi och lingvistik. Festskrift tillägnad Gösta Holm på 60 årsdagen 8 juli 1976* (pp. 25–47). Lund: Carl Bloms Boktryckeri.

Benediktsson, Hreinn. 2002. “Relational Sound change: vá > vo in Icelandic.” In Guðrún Þórhallsdóttir, Höskuldur Þráinsson, Jón G. Friðjónsson og Kjartan Ottosson (Eds.), *Linguistic studies historical and comparative* (pp. 227–242). Reykjavík: Institute of Linguistics.

Berdicevski, Aleksandrs, Evie Coussé, Alexander Kopeling & Yvonne Adesam. 2024. “To drop or not to drop? Predicting the omission of the infinitival marker in Swedish future construction.” *Corpus Linguistics and Linguistic Theory, 20*(1):219–261. doi: https://doi.org/10.1515/cllt-2022-0101

Bergs, Alexander. 2012. “The Uniformitarian Principle and the risk of anachronisms in language and social history.” In Juan Manuel Hernández-Campoy & Juan Camilo Conde-Silvestre (Eds.), *The handbook of historical sociolinguistics* (pp. 80–98). Oxford: Wiley-Blackwell.

Bernharðsson, Haraldur. 2016. *Icelandic. A historical linguistic companion. 5th draft. Handouts on phonology, morphology and related matters*. [Printed by Háskólaprent on Suðurgata.]

Berthele, Raphael, Matthew Whelpton, Åshild Næss, Pieter & Duijff. 2015. “Static spatial descriptions in five Germanic languages”. *Language Sciences, 49*:82-101. doi: http://doi.org/10.1016/j.langsci.2014.07.006

Bjarnadóttir, Kristín. 1989. *Dativus sympatheticus*. [Course essay, University of Iceland, Reykjavík].

Blevins, Juliette. 2006. “A theoretical synopsis of Evolutionary Phonology.” *Theoretical Linguistics, 32*(2):117–166. doi: https://doi.org/10.1515/TL.2006.009

Box, George E. P., Gwilym M. Jenkins & Gregory C. Reinsel. 2008. *Time series analysis: Forecasting and control* (Fourth edition). Hoboken, New Jersey: Wiley.

Box, George E. P., Gwilym M. Jenkins, Gregory C. Reinsel & Greta M. Ljung. 2016. *Time series analysis: Forecasting and control* (Fifth edition). Hoboken, New Jersey: Wiley.

Booth, Hannah. 2018. *Expletives and clause structure: Syntactic change in Icelandic*. [Ph.D. dissertation, University of Manchester].

Booth, Hannah. 2019. “Cataphora, expletives and impersonal constructions in the history of Icelandic.” *Nordic Journal of Linguistics, 42*(2):139–164.

Booth, Hannah. 2020. “Expletives in Icelandic.” In Bridget Drinka (Ed.), *Historical linguistics 2017: Selected papers from the 23rd international conference on historical linguistics, San Antonio, Texas 31 July*–4 August 2017 (pp. 363–384). doi: https://doi.org/10.1075/cilt.350.17boo

Bowie, David. 2005. “Language Change over the Lifespan: A test of the Apparent Time construct.” *University of Pennsylvania Working Papers in Linguistics, 11*(2):45–58.

Brenzinger, Matthias, Arienne M. Dwyer, Tjeerd de Graaf, Michael Krauss, Osahito Miyaoka, Nicholas Ostler, Osamu Sakiyama, María E. Villalón, Akira Y. Yamamoto & Ofelia Zepeda. 2003. “Language vitality and endangerment.” Document adopted by the International expert meeting on UNESCO programme safeguarding of endangered languages. Paris, 10–12 March 2003: UNESCO Ad Hoc Expert Group on Endangered Languages. CLT/CEI/DCE/ELP/PI/2003/1

Bresnan, Joan. 2007a. “Is syntactic knowledge probabilistic? Experiments with the English dative alternation.” In Sam Featherston and Wolfgang Sternefeld (Eds.) *Roots: Linguistics in search of its evidential base*, *96*:77–96. Berlin: Mouton de Gruyter

Bresnan, Joan 2007b. “A few lessons from typology.” *Linguistic Typology, 11*:297–306.

Brown, Robert G. 1959. *Statistical forecasting for inventory control*. New York: McGraw/Hill.

Bugge, Sophus. 1867. *Norræn fornkvæði. Islandsk samling af folkelige oldtidsdigte om Nordens guder og heroer almindelig kaldet Sæmundar edda hins fróða*. Christiania: P. T. Mallings Forlagsboghandel.

Burgess, Jean & Nancy K. Baym. 2020. *Twitter: A biography*. New York: New York University Press.

Bybee, Joan L. 1988. “The diachronic dimension in explanation.” In John A. Hawkins (Ed.), *Explaining language universals* (pp. 350–379). Oxford: Blackwell.

Campbell, Lyle. 2003. *Historical linguistics: An introduction* (2nd edition). Cambridge, Massachusetts: MIT Press.

Campbell, Lyle. 2020. *Historical linguistics: An introduction* (4th edition). Cambridge, Massachusetts: MIT Press.

Castle, Jennifer L., Michael P. Clements & David F. Hendry. 2019. *Forecasting: An essential introduction*. New Haven: Yale University Press.

Chambers, J.K. & Peter Trudgill. 1980. *Dialectology*. London: Cambridge University Press.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, Massachusetts: The MIT Press.

Chomsky, Noam. 1986. *Knowledge of language: Its nature, origin and use*. New York: Praeger.

Chomsky, Noam & Andrea Moro. 2022. *The secrets of words*. Cambridge, Massachusetts: The MIT Press.

Cleveland, Robert B., William S. Cleveland, Jean E. McRae & Irma Terpenning. 1990. "STL: A seasonal-trend decomposition procedure based on loess." *Journal of Official Statistics, 6*(1):3–33.

Corbett, Greville G. 2006. Agreement. Cambridge: Cambridge University Press.

Croft, William. 2000. *Explaining language change: An evolutionary approach.* London: Longman.

Cukor-Avila, Patricia & Guy Bailey. 2013. "Real Time and Apparent time." In J.K. Chambers & Natalie Schilling (Eds.), *The Handbook of Language Variation and Change* (2nd edition) (pp. 239–262). Wiley-Blackwell.

Curzan, Anne. 2009. "Historical corpus linguistics and evidence of language change." In Anke Lüdeling & Merja Kytö (Eds.), *Corpus linguistics. An international handbook* (Vol. 2, pp. 1091–1109). Berlin: Mouton de Gruyter

Déchaine, Rose-Marie. 2005. "Grammar at the borderline. A case study of P as a lexical category." In John Alderete, Chung-hye Han & Alexei Kochetov (Eds.), *Proceedings of the 24th West Coast Conference on Formal Linguistics* (pp. 1–18). Somerville, MA: Cascadilla Proceedings Project.

Denison, David. 2003. "Log(ist)ic and simplistic S-curves." In Raymond Hickey (Ed.), *Motives for Language Change* (pp. 54–70). Cambridge: Cambridge University Press.

den Dikken, Marcel. 2010. "On the functional structure of locative and directional PPs." In Guglielmo Cinque & Luigi Rizzi (Eds.), *Mapping spatial PPs,* Cartography of syntactic structures (Vol. 6, pp. 74–127). Oxford: Oxford University Press.

De Smet, Hendrik & Freek Van de Velde. 2017. “Experimenting on the past: a case study on changing analysability in English ly-adverbs.” *English Language and Linguistics*, *21*(2):317-340.

De Smet, Isabeau & Freek Van de Velde. 2019. “Reassessing the evolution of West Germanic preterite inflection.” *Diachronica, 36*(2):139–179.

Dronke, Ursula. 1997. The Poetic Edda (Vol. II Mythological Poems). Oxford: Clarendon Press.

Entwistle, William J. 1953. *Aspects of language*. London: Faber and Faber.

Eythórsson, Thórhallur. 2002. “Changes in subject case-marking in Icelandic.” In David W. Lightfoot (Ed.), *Syntactic effects of morphological change* (pp. 196–212). Oxford: Oxford University Press.

Eythórsson, Thórhallur. 2015a. “Aukafallshneigð.” *Ástumál kveðin Ástu Svavarsdóttur sextugri* (pp. 94–96). Reykjavík: Stofnun Árna Magnússonar í íslenskum fræðum.

Eythórsson, Thórhallur. 2015b. “The Insular Nordic experimental kitchen: Changes in case marking in Icelandic and Faroese.” In Matthew Whelpton, Guðrún Björk Guðsteinsdóttir, Birna Arnbjörnsdóttir & Martin Regal (Eds.), *An Intimacy of Words / Innileiki orðanna*. *Essays in Honour of Pétur Knútsson / Festschrift til heiðurs Pétri Knútssyni* (pp. 328–352). Reykjavík: Stofnun Vigdísar Finnbogadóttur í erlendum tungumálum / Háskólaútgáfan.

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

Eythórsson, Thórhallur & Jóhanna Barðdal. 2016. “Syntactic reconstruction in Indo-European: State of the art.” *VELEIA, 33*:82–102.

Eythórsson, Thórhallur & Sigríður Sæunn Sigurðardóttir. 2016. “A Brief History of Icelandic Weather Verbs: Syntax, Semantics and Argument Structure.” *Working Papers in Scandinavian Syntax, 96*:91-125.

Eythórsson, Thórhallur & Höskuldur Thráinsson. 2017. “Variation in oblique subject construction in Insular Scandinavian.” In Höskuldur Thráinsson, Caroline Heycock, Hjalmar P. Petersen & Zakaris Svabo Hansen (Eds.), *Syntactic variation in Insular Scandinavian* (pp. 53–90). Amsterdam: John Benjamins Publishing Company.

Faarlund, Jan Terje. 2004. *The syntax of Old Norse: With a survey of the inflectional morphology and complete bibliography*. Oxford: Oxford University Press.

Faarlund, Jan Terje. 2015. “The Norwegian infinitive marker.” *Working papers in Scandinavian syntax*, *95*:1–10.

Falk, Cecilia. 1997. *Fornsvenska upplevarverb*. Lund: Lund University Press.

Ferguson, Charles A. 1966. "Assumptions about nasals: A sample study in phonological universals." In Joseph H. Greenberg (Ed.), *Universals of language* (2nd edition, pp. 53–60). Cambridge, Massachusetts: The MIT Press.

Finley, Alison. 2003. "Interpretation or over-interpretation: The dating of two Íslendingasögur." *Gripla, 14*:61–69.

Friðjónsson, Jón G. 1988. *Forsetningar í íslensku*. Reykjavík: Málvísindastofnun Háskóla Íslands.

Friðjónsson, Jón G. 2004. "Íslenskt mál 27. þáttur." *Morgunblaðið, 2004*(125):52.

Friðjónsson, Jón G. 2005. "Kerfisbundnar breytingar á notkun nokkurra forsetninga í íslensku. Samspil tíma og rúms." *Íslenskt mál og almenn málfræði, 27*:7–40.

Friðjónsson, Jón G. 2007. "Íslenskt mál 94. þáttur." *Morgunblaðið, 2007*(12):28.

Friðjónsson, Jón G. 2009. "Samband forsetninga og samsettra orða." *Orð og tunga, 11*:65–74.

van Gelderen, Elly. 2011. *The linguistic cycle: Language change and the Language Faculty*. Oxford: Oxford University Press.

Givón, Talmy. 1979. *On Understanding Grammar*. New York: Academic Press.

Gneiting, Tilmann, & Matthias Katzfuss. 2014. "Probabilistic forecasting." *Annual review of statistics and its application*, *1*(1):125–151.

Gold, Jana Willer, Boban Arsenijevi, Boban Arsenijević, Mia Batinić, Michael Becker, Nermina Čordalija, Marijana Kresić, Nedžad Leko, Franc Lanko Marušič, Tanja Milićev, Nataša Milićević, Ivana Mitić, Anita Peti-Stantić, Branimir Stanković, Tina Šuligoj,f Jelena Tušek, and Andrew Nevinsa. 2018. "When linearity prevails over hierarchy in syntax." *Proceedings of the National Academy of Sciences of the United States of America*, 115(3): 495–500. https://doi.org/10.1073/pnas.1712729115

Greenberg, Joseph H. 1966. "Synchronic and diachronic universals in phonology." *Language, 42*(2):508–517.

Greenberg, Joseph H. 1966. "Some universals of grammar with particular reference to the order of meaningful elements." In Joseph H. Greenberg (Ed.), *Universals of language* (2nd edition, pp. 73–113). Cambridge, Massachusetts: The MIT Press.

Greenberg, Joseph H. 1969. “Some methods of dynamic comparison in linguistics.” In Jan Puhvel (Ed.), *Substance and structure of language* (pp. 147–203). Berkeley: University of California Press.

Greenberg, Joseph H., Charles E. Osgood & James J. Jenkins. 1966. “Memorandum concerning language universals.” In Joseph H. Greenberg (Ed.), *Universals of language* (2nd edition, pp. xv–xxvii) . Cambridge, Massachusetts: The MIT Press.

Grestenberger, Laura. 2015. “Number marking in German measure phrases and the structure of pseudo-partitives.” *Journal of comparative German linguistics, 18*:93–138.

Guðmundsdóttir, Dagbjört, Sigríður Sigurjónsdóttir, & Iris Nowenstein. 2022. “Digital language contact between Icelandic and English.” In Kristiina Kumpulainen, Anu Kajamaa, Ola Erstad, Åsa Mäkitalo, Kirsten Drotner and Sólveig Jakobsdóttir (Eds.). *Nordic Childhoods in the Digital Age: Insights into Contemporary Research on Communication, Learning and Education* (pp. 79–91). Routledge, London.

Gunnlaugsson, Guðvarður Már. 2005. “Manuscripts and palaeography.” In Rory McTurk (Ed.), *A companion to Old Norse-Icelandic literature* (pp. 245–264). Malden, Massachusetts: Blackwell Publishing.

Hale, Mark. 2007. *Historical linguistics theory and method*. Malden, Massachusetts: Blackwell Publishing.

Hall, Robert A. Jr. 1954. “Linguistics Oxford style.” *American Speech, 29*(2):125–130.

Halldórsson, Halldór. 1982. “Um méranir: Drög að samtímalegri og sögulegri athugun.” *Íslenskt mál og almenn málfræði, 4*:159–89.

Harðarson, Gísli R. 2017. *Cycling through grammar: On compounds, noun phrases and domains* [Ph.D. dissertation, University of Connecticut].

Haspelmath, Martin. 2019. “Can cross-linguistic regularities be explained by constraints on change?” In Karsten Schmidtke-Bode, Natalia Levshina, Susanne Maria Michaelis & Ilja A. Serzant (Eds.), *Explanation in typology: Diachronic sources, functional motivations and the nature of the evidence* (pp. 1–23). Berlin: Language Science Press.

Hatton, Timothy J. & Jeffrey G. Williamson. 1998. *The age of mass migration: Causes and economic impact*. Oxford: Oxford University Press.

Haugen, Einar. 1966. *Language conflict and language planning: The case of Modern Norwegian.* Cambridge, Massachusetts: Harvard University Press.

Heine, Bernd. 1995. “Conceptual grammaticalization and prediction.” In John R. Taylor, Robert E. MacLaury (Eds.), *Language and the Cognitive Construal of the World* (pp. 119–136). Berlin, New York: De Gruyter Mouton.

Heine, Bernd, Ulrike Claudi & Friederike Hünnemeyer. 1991. *Grammaticalization. A Conceptual Framework*. Chicago: University of Chicago Press.

Heine, Bernd & Tania Kuteva. 2002. *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.

Hewson, John & Vit Bubenik. 2006. *From case to adposition: The development of configurational syntax in IndoEuropean languages*. Amsterdam: John Benjamins.

Hilmarsson-Dunn, Amanda & Ari Páll Kristinsson. 2010. “The language situation in Iceland.” *Current issues in language planning*, *11*(3):207–276.

Hockett, Charles F. 1958. *A course in modern linguistics*. New York, NY: Macmillan.

Hoff, John C. 1983. A *practical guide to Box-Jenkins forecasting*. Belmont, California: Lifetime learning publications.

Hoffmann, Sebastian. 2004. “Are low-frequency complex prepositions grammaticalized? On the limits of corpus data – and the importance of intuition.” In Hans Lindquist & Christian Mair (Eds.), *Corpus approaches to grammaticalization in English* (pp. 171–209). Amsterdam: John Benjamins.

Van Hofwegen, Janneke & Walt Wolfram. 2010. “Coming of age in African American English: A longitudinal study.” *Journal of Sociolinguistics*, *14:*427–455.

Holt, Charles C. 1957. *Forecasting Seasonals and Trends by Exponentially Weighted Averages*. Pittsburgh: Carnegie Institute of Technology.

Honeybone, Patrick. (2016). Are there impossible changes? θ > f but f ≯ θ. *Papers in Historical Phonology*, *1*(0), 316. https://doi.org/10.2218/pihph.1.2016.1705

Hopper, Paul J. 1991. “On some principles of grammaticalizaiton.” In Elizabeth Closs Traugott & Bernd Heine (Eds.), *Approaches to grammaticalization I* (pp. 17–35). Amsterdam: John Benjamins.

Hopper, Paul J. & Elizabeth Closs Traugott. 2003. *Grammaticalization* (2nd edition). Cambridge: Cambridge University Press. (Reprint of the original 1993 edition).

Hyndman, Rob J. 2023. *fpp3: Data for "Forecasting: Principles and Practice" (3rd edition)*. R package version 0.5. https://CRAN.R-project.org/package=fpp3

Hyndman, Rob J. & George Athanasopoulos. 2021. *Forecasting: Principles and practice* (3rd edition). Melbourne, Australia: Otext. https://otexts.com/fpp3/.

Hyndman, Rob J. & Anne B. Koehler. 2006. "Another look at measures of forecast accuracy." *International Journal of Forecasting, 22*(4):679–688.

Ingason, Anton Karl, Julie Anne Legate & Charles Yang. 2012. "The evolutionary trajectory of the Icelandic New Passive." *U. Penn working papers in linguistics, 18*(2):91–100.

ÍT = *Íslenskt Textasafn*, Stofnun Árna Magnússonar í íslenskum fræðum (Ed.). Accessed 2021–2023 at: http://corpus.arnastofnun.is/

Jackendoff, Ray. 1973. "The base rules for prepositional phrases." In Stephen Anderson, & Paul Kiparsky (Eds.), *A Festschrift for Morris Halle* (pp. 345–56). New York: Holt, Rinehart & Winston.

Janda, Richard D. 2001. "Beyond 'pathways' and 'unidirectionality': On the discontinuity of language transmission and the counterability of grammaticalization." *Language Science, 22*(2–3):65-340.

Janda, R., Gil, D., 1980, August 1-3. *Beyond `gradualness': abrupt, stepwise, and continuous language-change and spread of change*. [Paper presentation]. 42nd Summer Meeting of the Linguistic Society of America; Albuquerque, NM, United States.

Jóhannsdóttir, Kristín M. 2011. *Aspects of the progressive in English and Icelandic*. [Ph.D. dissertation, University of British Columbia].

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

Jónsson, Jóhannes Gísli. 2003. "Not so Quirky: On Subject Case in Icelandic." In Ellen Brandner & Heike Zinsmeister (Eds.). *New Perspectives in Case Theory* (pp. 129–165). Stanford: CSLI Publications.

Jónsson, Jóhannes Gísli. 2017. "Avoiding genitive in Icelandic." In Höskuldur Thráinsson, Caroline Heycock, Hjalmar P. Petersen & Zakaris Svabo Hansen (Eds.), *Syntactic Variation in Insular Scandinavian* (pp. 141–163). Amsterdam: John Benjamins Publishing Company.

Jónsson, Jóhannes Gísli & Thórhallur Eythórsson. 2003. "Breytingar á frumlagsfalli í íslensku." *Íslenskt mál og almenn málfræði, 25*:7–40.

Jónsson, Jóhannes Gísli & Thórhallur Eythórsson. 2005. "Variation and change in subject case marking in Insular Scandinavian." *Nordic Journal of Linguistics, 28*(2):223–245.

Jónsson, Jóhannes Gísli & Thórhallur Eythórsson. 2011. "Structured exceptions and case selection in Insular Scandinavian." In Heike Wiese & Horst Simon (Eds.), *Expecting the unexpected: Exceptions in grammar* (pp. 213–241). Berlin: Mouton de Gruyter.

Karlsson, Stefán. 2000. "Tungan." In *Stafkrókar: Safn ritgerða Stefáns Karlsson* (pp. 19–75). Reykjavík: Stofnun Árna Magnússonar.

Keller, Rudi. 1994. *On language change: The invisible hand in language*. London: Routledge.

Kikusawa, Ritsuko. 2012. "Standardization as language loss: Potentially endangered Malagasy languages and their linguistic features." *People and culture in Oceania, 28*:23–44.

Kiparsky, Paul. 2019. *Linguistics then and now: The view from NELS*. [Plenary talk]. 50th Annual Meeting of the North East Linguistic Society, Cambridge, MA, United States.

Kootstra, Gerrit Jan & Pieter Muysken. 2019. "Structural priming, level of awareness, and agency in contact-induced language change." *Languages, 4*(3):65. doi:10.3390/languages4030065

Kress, Bruno. 1982. *Isländische Grammatik*. Leipzig: VEB Verlag Enzyklopädie.

Kristinsson, Ari Páll. 2006. "Um málstefnu." *Hrafnaþing, 3*:47–63.

Kristinsson, Ari Páll. 2007. "Málræktunarfræði." *Íslenskt mál og almenn málfræði, 29*:99–124.

Kristinsson, Ari Páll. 2019. "Um greiningu á málstöðlun og málstefnu: Haugen, Ammon og Spolsky í íslensku samhengi." *Orð og tunga, 21*:129–151.

Kristinsson, Ari Páll. 2023. "Fössarar of fantasíur." *Árnastonfun, pistlar, 18*. January 2023. https://arnastofnun.is/is/utgafa-og-gagnasofn/pistlar/fossarar-og-fantasiur

Kroch, Anthony S. 1989. "Reflexes of grammar in patterns of language change." *Language Variation and Change, 1*:199–244.

Kuryłowicz, Jerzy. 1975. "The evolution of grammatical categories." *Esquisses Linguistiqes, 2,* 38-54. (Reprinted from "The evolution of grammatical categories," 1965, *Diogenes, 51*:55–71.)

Kuteva, Tania, Bernd Heine, Bo Hong, Haiping Long, Heiko Narrog & Seongha Rhee. 2019. *World lexicon of grammaticalization* (2nd edition). Cambridge: Cambridge University Press.

Kwiatkowski, Denis, Peter C. B. Phillips, Peter Schmidt, & Yongcheol Shin. 1992. "Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?" *Journal of Econometrics, 54*(1-3):159–178.

Labov, William. 1963. "The social motivation of sound change." *Word, 19*:273–309.

Labov, William 1966. *The Social Stratification of English in New York City.* Washington, D.C.: Center for Applied Linguistics

Labov, William. 1994. *Principles of Linguistic Change*: *Internal Factors*. Oxford: Blackwell.

Langacker, Ronald W. 1977. "Syntactic reanalysis." In Charles N. Li (Ed.), *Mechanisms of syntactic change* (pp. 57–139). Austin: University of Texas Press.

Lass, Roger. 1997. *Historical linguistics and language change*. Cambridge: Cambridge University Press.

Lavidas, Nikolaos. 2021. *The diachrony of written language contact*. Leiden: Brill.

Lehmann, Christian. 1991. "Grammaticalization and related changes in contemporary German." In Elizabeth Closs Traugott and Bernd Heine (Eds.), *Approaches to Grammaticalization*, (Vol. II, 439–535). Amsterdam: Benjamins.

Lehmann, Christian. 2002. "New reflections on grammaticalization and lexicalization." In Ilse Wischer & Gabriele Diewald (Eds.), *New reflections on grammaticalization* (pp. 1–18). Amsterdam: John Benjamins.

Lehmann, Christian. 2015. Thoughts on grammaticalization (3rd edition). Berlin: Language Science Press. (Original work published 1982).

Leonard, Stephen Pax & Kristján Árnason. 2011. "Language ideology and standardisation in Iceland." In Tore Kristiansen & Nikolas Coupland (Eds.), *Standard languages and language standards in a changing Europe* (pp. 91-96). Oslo: Novus Press.

Lewis, M. Paul & Gary F Simons. 2010. "Assessing endangerment: Expanding Fishman's GIDS." *Revue Roumaine de Linguistique*, *55*(2):103–120.

Lieberman, Erez, Jean-Baptiste Michel, Joel Jackson, Tina Tang & Martina A. Nowak. 2007. "Quantifying the evolutionary dynamics of language." *Nature, 449*:713–716.

Lightfoot, David. 1979. *Principles of diachronic syntax*. Cambridge: Cambridge University Press.

Lightfoot, Davi W. 2002. "Myths and the prehistory of grammars." *Journal of Linguistics, 38*(1):113–136. doi: https://doi.org/10.1017/S0022226701001268

Lohndal, Terje. 2009. "The copula cycle." In Elly van Gelderen (Ed.) *Cyclical change* (pp. 209–243). Amsterdam: John Benjamins.

Longobardi, Giuseppe. 2001. "Formal syntax, diachronic Minimalism, and etymology: The history of French, Chez." *Linguistic Inquiry, 32*(2):275–302. doi: https://doi.org/10.1162/00243890152001771

Magnússon, Ásgeir Blöndal. 1989. *Íslensk orðsifjabók*. Reykjavík: Orðabók Háskólans.

Makridakis, Spyros & Steven C. Wheelwright. 1978. *Forecasting: Methods and applications*. New York: John Wiley & Sons.

Makridakis, Spyros, Steven C. Wheelwright, & Rob J. Hyndman. 2018. *Forecasting methods and applications*. India: Wiley. (Original work published 2005).

Matras, Yaron & Jeanette Sakel. 2007. "Investigating the mechanisms of pattern replication in language convergence." *Studies in language, 31*(4):829–865.

Microsoft Corporation. 2022. *Microsoft Excel*, version 2308. Retrieved from https://office.miscrosoft.com/excel

Milroy, James & Lesley Milroy. 1985. "Linguistic change, social network and speaker innovation." J*ournal of linguistics, 21*(2):339–384.

Narrog, Heike. "Exaptation in Japanese and beyond." In Muriel Norde & Freek Van de Velde (Eds.), *Exaptation and language change* (pp. 93–121). Amsterdam: John Benjamins.

Nativi, Stefano & Max Craglia. 2021. *Destination Earth: Ecosystem architecture description*, *EUR 30646 EN*. Luxembourg: Publications office of the European Union. doi: doi:10.2760/08093.

Newmeyer, Frederick. 1998. *Language form and language function*. Cambridge, Massachusetts: MIT Press.

Nevalainen, Terttu. 2015. "Descriptive adequacy of the S-curve model in diachronic studies of language change." In Christina Sanchez-Stockhammer (Ed.), *Can We Predict Linguistic Change? (Studies in Variation, Contacts and Change in English 16)*. Helsinki: VARIENG. https://urn.fi/URN:NBN:fi:varieng:series-16-4

Nevalainen, Terttu & Helena Raumolin-Brunberg. 2017. *Historical sociolinguistics: Language change in Tudor and Stuart England* (2nd edition). London: Routledge.

Nijs, Julie & Freek Van de Velde. Under review. "How predictable is language chang?"

Noreen, Adolf. 1923. *Altnordische Grammatik I. Altisländische und altnorwegische Grammatik (Laut- und Flexionslehre)*. Vierte vollständig umgearbeitete Auflage. Halle (Saale): Verlag von Max Niemeyer.

Norde, Muriel. 2009. *Degrammaticalizaiton*. Oxford: Oxford University Press.

Norde, Muriel. 2011. "Degrammaticalization." In Heiko Narrog & Bernd Heine (Eds.), *The Oxford Handbook of Grammaticalization* (pp. 475-487). Oxford: Oxford University Press.

Nowenstein, Iris Edda. 2014a. "Intra-speaker Variation in Subject Case: Icelandic." *University of Pennsylvania Working Papers in Linguistics, 20*(1):28.

Nowenstein, Iris Edda. 2014b. *Tilbrigði í frumlagsfalli á máltökuskeiði. Þágufallshneigð og innri breytileiki*. [M.A. thesis, University of Iceland].

Nowenstein, Iris Edda. 2017. "Determining the nature of intra-speaker variation." In Höskuldur Thráinsson, Caroline Heycock, Hjalmar P. Petersen & Zakaris Svabo Hansen (Eds.), *Syntactic Variation in Insular Scandinavian* (pp. 92–112). Amsterdam: John Benjamins.

Nowenstein, Iris Edda. 2023. *Building yourself a variable case system: The acquisition of Icelandic datives*. [Ph.D. dissertation, University of Iceland].

Nowenstein, Iris Edda & Sigríður Sigurjónsdóttir. 2021. "Stafrænt málsambýli íslensku og ensku. Áhrif ensks ílags og málnotkunar á málfærni íslenskra barna." *Ritið 3*:11–56.

Numismatic collection of the Central Bank and National Museum of Iceland. 2002. "Opinber gjaldmiðill á Íslandi. Útgáfa og auðkenni íslenskra seðla og myntar." *Myntrit 3*. Reykjavík: Myntsafn Seðlabanka og Þjóðminjasafns.

O'Hara-Wild, Mitchell, Rob Hyndman, & Earo Wang. 2021. *fable: Forecasting Models for Tidy Time Series*. R package version 0.3.1. https://CRAN.R-project.org/package=fable

Óladóttir, Hanna. 2017. *Skólamálfræði: Hver er hún og hver ætti hún að vera? Um markmið og áhrif málfræðikennslu á unglingastigi grunnskóla*. [Ph.D. dissertation, University of Iceland].

Ólason, Vésteinn. 2006. Íslendingasögur og þættir. In Vésteinn Ólason (Ed.) *Íslensk bókmenntasaga* (Vol. II, pp. 25-161). Reykjavík: Mál og Menning.

ONP = *Ordbog over det norrøne prosasprog*. Den arnamagnæanske kommission (Ed.), Copenhagen. Retrieved from: http://onp.ku.dk/

Osgood, Charles & Sebeok, Thomas. 1954. "Psycholinguistics: a survey of theory and research problems." *Journal of Abnormal and Social Psychology, 49*:1–203.

Osthoff, Herman. and Brugmann, Karl. 1878. *Morphologische Untersuchungen auf dem Gebiete der indogermanischen Sprachen*. Leipzig: S. Hirzel.

Ottósson, Kjartan G. 1992. *The Icelandic middle voice: The morphological and phonological development*. [Ph.D dissertation, Lund University].

Ottosson, Kjartan. 2008. "The Old Norse Middle voice in the pre-literary period: Questions of grammaticalization and cliticization." In Folke Josephson & Ingmar Söhrman (Eds.), *Interdependence of diachronic and synchronic analyses* (pp. 185–219). Amsterdam: John Benjamins.

Pálsson, Gísli. 1979. "Vont mál og vond málfræði." *Skírnir, 153*:175–201.

Paul, Hermann. 1886. *Principien der Sprachgeschichte* (Second printing). Max Niemeyer.

Pfaff, Alexander. 2023, March 24. *Fallmörkun og hlutamerking*. [Invited talk, Málvísindakaffi Reykjavík].

Pintzuk, Susan. 1999. *Phrase structures in competition: Variation and change in Old English clause Structure*. New York: Garland.

Pintzuk, Susan & Ann Taylor. 2006. "The loss of OV order in the history of English", in Ans van Kemenade & Bettelou Los (Eds.), *The handbook of the history of English* (pp. 249–78). Malden, Massachusetts: Blackwell.

Pintzuk, Susan, Ann Taylor & Anthony Warner. 2017. "Corpora and quantitative methods." In Adam Ledgeway & Ian Roberts (Eds.), *The Cambridge handbook of historical syntax* (pp. 218–240). Cambridge: Cambridge University Press.

Plank, Frans. 2015. "Time for change." In Carlotta Viti (Ed.), *Perspectives on historical syntax* (pp. 61–91). John Benjamins Publishing Company.

Posit team (2023). RStudio: Integrated Development Environment for R. Posit Software, PBC, Boston, MA. URL http://www.posit.co/.

Postma, Gertjan. 2010. "The impact of failed changes." In Anne Breitbarth, Christopher Lucas, Sheila Watts, David Willis (Eds.), *Continuity and change in grammar* (pp. 269–302). Amsterdam: John Benjamins.

Quirk, Randolph & Joan Mulholland. 1964. Complex prepositions and related sequences. *English Studies, 45*(1–6):64–73.

R Core Team. 2021. R: *A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Version 2023.6.0.421, https://www.R-project.org/.

Rask, Rasmus K. 1813, August 30. [78. Letter to Bjarni Thorsteinsson]. Printed in Louis Hjelmslev (Ed.), *Breve fra og til Rask I 1805–1819* (pp. 162–170). Copenhagen: Ejnar Mungsaards forlag.

Rask, Rasmus K. 1818. *Undersögelse om det gamle nordiske eller islandske sprogs oprindelse*. Copenhagen: Gyldendal.

Raumolin-Brunberg, Helena 2005. "Language change in adulthood: historical letters as evidence." *European Journal of English Studies, 9*(1):37–51.

Reykjahólabók I–II. 1969–1970. Agnete Loth (Ed.). Copenhagen: Munksgaard

Ritreglur. *Auglýsingar mennta- og menningarmálaráðuneytis nr. 695/2016 og 800/2018 með leiðréttingum*. Stofnun Árna Magnússonar í íslenskum fræð-um (Ed.), accessed January 2024 at: http://ritreglur.arnastofnun.is/

Roberts, Ian. 2017. "Uniformitarianism." In Adam Ledgeway & Ian Roberts (Eds.), *The Cambridge handbook of historical syntax* (pp. 338–359). Cambridge: Cambridge University Press.

Robins, Robert Henry. 1997. *A short history of linguistics* (4th edition). London: Routledge.

Robins, Robert Henry. 1979. *A short history of linguistics* (2nd edition). London: Routledge.

Rögnvaldsson, Eiríkur.1997. "Frumlag og fall að fornu." *Íslenskt mál og almenn málfræði, 18*:37–69

Rögnvaldsson, Eiríkur. 2002. "ÞAÐ í fornu máli og síðar." *Íslenskt mál og almenn málfræði, 24*:7–30.

Rögnvaldsson, Eiríkur. 2020, February 19. "Íslenska á öllum sviðum." *Málfarspistill 95*. Accessed at: https://uni.hi.is/eirikur/ritaskra/malfarspistlar/

Rögnvaldsson, Eiríkur 2021a, January 9. “hliðiná.” *Málfarspistill 236*. Accessed at: https://uni.hi.is/eirikur/ritaskra/malfarspistlar/

Rögnvaldsson, Eiríkur. 2021b, February 4. “Fyrir bak við.” *Málfarspistill 249*. Accessed at: https://uni.hi.is/eirikur/ritaskra/malfarspistlar/

Rögnvaldsson, Eiríkur. 2023a, May 12. “Verður enska aðalsamskiptamálið á Íslandi árið 2050?” [Will English be the main language of communication in 2050?]. *Málfarspistill 627*. Accessed at: https://uni.hi.is/eirikur/ritaskra/malfarspistlar/

Rögnvaldsson, Eiríkur. 2023b, September 4. “Fyrir bakvið hús.” *Málfarspistill* 719. Accessed at: https://uni.hi.is/eirikur/ritaskra/malfarspistlar/

Rögnvaldsson, Eiríkur. 2023c, October 18. “Hvernig aukum við íslenskan orðaforða unglinga?” [How do we improve the vocabulary knowledge of Icelandic youth?]. *Málfarspistill 768*. Accessed at: https://uni.hi.is/eirikur/ritaskra/malfarspistlar/

Rögnvaldsson, Eiríkur, Anton Karl Ingason, Einar Freyr Sigurðsson & Joel Wallenberg. 2012. “The Icelandic Parsed Historical Corpus (IcePaHC).” *Gripla* XXIII:1977–1984.

ROH = *Ritmálssafn Orðabókar Háskólans*. Stofnun Árna Magnússonar í íslenskum fræðum (Ed.). Accessed 2021–2023: https://ritmalssafn.arnastofnun.is/

Rosen, Aliza. 2017, September 26. “Giving you more characters to express yourself.” Blog posted on September 26th. Accessed at: https://blog.twitter.com/official/en_us/topics/ product/2017/Giving-you-more-characters-to-express-yourself.html

Rosenkvist, Henrik. 2023. “Structural ambiguity and reanalysis: The case of Swedish *fortsatt*.” [Conference Presentation]. ICHL, Heidelberg, Germany.

Roy, Isabelle, and Svenonius, Peter. 2009. “Complex prepositions.” In Jacques François, Eric Gilbert, Claude Guimier & Maxi Krause (Eds.), *Autour de la préposition* (pp. 105–116). Caen: Presses Universitaires de Caen.

Sanchez-Stockhammer, Christina. 2015. “Can we predict linguistic change? An introduction.” *Studies in Variation, Contacts and Change in English, 16*:15.

Sankoff, Gillian. 2013. “Longitudinal studies.” In Robert Bayley, Richard Cameron & Ceil Lucas (Eds.), *The Oxford handbook of sociolinguistics* (pp. 261–279). Oxford: Oxford University Press.

Sapir, Edward. 1941. *Language: An introduction to the Study of Speech.* Harcourt, Braze and Company, New York. (Original work published 1921).

de Saussure 1959. Course in general linguistics. (Perry Meisel & Haun Saussy Eds., Wade baskin Trans.). New York: Columbia University Press. (Original work published 1916)

Schneider, Edgar W. 2018. "Christina Sanchez-Stockhammer (Ed.), Can we predict linguistic change? (Studies in Variation, Contacts and Change in English 16). 2015. www.helsinki.fi/varieng/series/volumes/16/ (2 March 2017)." [Review of the eSeries S*tudies in Variation, Contacts and Change in English*, Christina Christina Sanchez-Stockhammer (Ed.),]. *English Language and Linguistics, 22*(3):531–536. doi: https://doi.org/10/ghbc8m

Searle, John. 1969. *Speech Acts: An essay in the philosophy of language*. Cambridge: Cambridge University Press.

Seppänen, Aimo, Rhonwen Bowen & Joe Trotta. 1994. "On the so-called Complex Prepositions." *Studia Anglica Posnaniensia, 29*:3–29.

Sigurðardóttir, Sigríður Sæunn. 2019–2020. "'Haf góðan dag' Um uppkomu nýrrar kveðju út frá hugmyndum um talgjörðir." *Íslenskt mál og almenn málfræði, 41–42*:77–94.

Sigurðardóttir, Sigríður Sæunn & Thórhallur Eythórsson. 2022. "The emergence of oblique subjects: Oblique-Case Substitution and Shift in Anticausative Strategy in Modern Icelandic." *Working papers in Scandinavian syntax, 107*:47–82.

Sigurðardóttir & Eythórsson. To appear. "The Emergence of Oblique Subjects: Identifiable Processes in the History of Icelandic." In Holly Kennard, Emily Lindsay-Smith, Aditi Lahiri and Martin Maiden (Eds.). *Historical Linguistics 2022. Selected papers from the 25th ICHL. Oxford, 1–5 August 2022*.

Sigurðsson, Einar Freyr. 2006. *Tölvan hjá mér er biluð. Notkun forsetningarinnar* hjá *í eignarmerkingu*. [B.A. thesis, University of Iceland].

Sigurðsson, Einar Freyr. 2017a. *Deriving case, agreement and Voice Phenomena in Syntax*. [Ph.D. dissertation, University of Pennsylvania].

Sigurðsson, Einar Freyr. 2017b. "Eignarfall og eignarliðir." In Höskuldur Thráinsson, Ásgrímur Angantýsson & Einar Freyr Sigurðsson (Eds.), *Tilbrigði í íslenskri setningagerð III. Sérathuganir* (pp. 61–97). Reykjavík: Málvísindastofnun Háskóla Íslands.

Sigurðsson, Halldór Ármann. 1989. Verbal Syntax and Case in Icelandic: In a Comparative GB Approach. [Ph.D dissertation, Lund University]. [Reprinted by the Linguistics Institute, University of Iceland, Reykjavík, 1992].

Sigurðsson, Halldór Ármann. 1997. “Öðruvísi frumlög.” In Anna Agnarsdóttir, Pétur Pétursson and Torfi Tulinius (Eds.), *Milli himins og jarðar* (pp. 299–306). Reykjavík: Háskólaútgáfan.

Sigurðsson, Halldór Ármann. 2002a. “Agree and Agreement: evidence from Germanic.” *Working Papers in Scandinavian Syntax, 70*:101–56.

Sigurðsson, Halldór Ármann. 2002b. “To be an Oblique Subject: Russian vs. Icelandic.” *Natural Language & Linguistic Theory, 20*(4):691-724.

Sigurjónsdóttir, Katrín Ásta. 2023, December 5. “Ísland hrapar í nýrri PISA-könnun.” *Heimildin*, https://heimildin.is/grein/19913/

Sigurjónsdóttir, Sigríður, & Eiríkur Rögnvaldsson. 2018. “Stafrænt sambýli íslensku og ensku.” *Netla - Veftímarit um uppeldi og menntun: Sérrit 2018 - Menntavika 2018*. https://doi.org/10.24270/serritnetla.2019.29

Sigurjónsdóttir, Sigríður. & Nowenstein, Iris. 2021. “Language acquisition in the digital age: L2 English input effects on children’s L1 Icelandic.” S*econd Language Research, 37*(4):697–723.

Sigurjónsson, Sveinbjörn. 1960. “Meðferð talaðs orðs í íslenzkum skólum.” *Menntamál, 1960*(2):97–102.

Sigtryggsson, Jóhannes B. 2022. “Íslensk réttritun – nýtt rit um stafsetningu.” *Orð og tunga 24*:123–126.

Skard, Vemund. 1951. *Dativstudien: Dativus Sympatheticus und Dativus Comparationis in der norrönen Sprache*. Oslo: Jacob Dybwad.

Smith, Adam. 1970. *The Wealth of Nations* (Vol. I). London and New York. (Originally published 1776).

Sóskuthy, Márton. 2015. “Understanding change through stability: A computational study of sound change actuation.” *Lingua, 163*:40–60. doi: https://doi.org/10/ggf3dp

Sprouse, Jon. 2007a. *A program for experimental syntax: Finding the relationship between acceptability and grammatical knowledge*. [Ph.D. dissertation, University of Maryland].

Sprouse, Jon. 2007b. “Continuous acceptability, categorical grammaticality, and experimental syntax.” *Biolinguistics, 1*:123–134.

Stefanowitsch, Anatol, Elena Smirnova & Matthias Hüning. 2020. “Complex adpositions in three West Germanic Languages: German, Dutch, and English.” In Benjamin

Fagard, José Pinto de Lima, Dejan Stosic & Elena Smirnova (Eds.), *Complex adpositions in European languages: a micro-typological approach to complex nominal relators* (pp. 65–138). Berlin: De Gruyter.

Steingrímsson, Steinþór, Sigrún Helgadóttir, Eiríkur Rögnvaldsson, Starkaður Barkarson and Jón Guðnason. 2018. “Risamálheild: A Very Large Icelandic Text Corpus.” *Proceedings of LREC 2018* (pp. 4361–4366). Japan.

Svavarsdóttir, Ásta. 1982. “Þágufallssýki.” *Íslenskt mál og almenn málfræði*, *4*:19–62.

Svavarsdóttir, Ásta. 2013. “Þágufallshneigð í sjón og raun. Niðurstöður spurningakannana í samanburði við málnotkun.” In Höskuldur Thráinsson, Ásgrímur Angantýsson & Einar Freyr Sigurðsson (Eds.), *Tilbrigði í íslenskri setningagerð I. Markmið, aðferðir og efniviður* (pp. 83–110). Reykjavík: Málvísindastofnun.

Svavarsóttir, Hrefna. 2023. *Fallfesta og fallglötun í frumlagslyftingu. Breytileiki í fallmörkun í frumlægri og andlægri frumlagslyftingu með tal- og álitssögnum í íslensku.* [B.A. thesis, University of Iceland].

Svenonius, Peter. 2004. *Adpositions, particles and the arguments they introduce*. [Ms, Tromsø University].

Svenonius, Peter. 2004. *Spatial P in English*. [Ms, Tromsø University].

Svenonius, Peter. 2006. “The emergence of axial parts.” *Tromsø Working Papers in Linguistics*, *33*(1):50–71.

Svenonius, Peter. 2007. “Adpositions, Particles, and the Arguments They Introduce.” In Eric Reuland, Tanmoy Bhattacharya, and Giorgos Spathas (Eds.), *Argument Structure* (pp. 63–104). Amsterdam: Benjamins.

Svenonius, Peter. 2008. “Projections of P.” In Anna Asbury, Jakub Dotlačil, Berit Gehrke & Rick Nouwen (Eds.), *Syntax and semantics of spatial P* (pp. 63–84). Amsterdam: John Benjamins.

Sæmundsson, Stefán Þór. 2022, December 31. “Gullaldaríslenskan deyr út.” In Orri Páll Ormarsson (Ed.), Online news article with an interview on Mbl.is https://www.mbl.is/frettir/innlent/2022/12/31/gullaldarislenskan_deyr_ut/

Talmy, Leonard. 1978. “Figure and ground in complex sentences.” In Joseph H. Greenberg (Ed.), *Universals of Human Language* (Vol. 4, pp. 625–649). Stanford, CA: Stanford University Press.

Talmy, Leonard. 2000. *Toward a Cognitive Semantics: Concept Structuring Systems* (Vol. I). Cambridge, MA: MIT Press.

Thomas, George. 1991. *Linguistic purism. Studies in language and linguistics*. New York: Longman.

Thomason, Sarah Grey. 2003. “Contact as a source of language change.” In Brian D. Joseph & Richard D. Janda (Eds.), *The handbook of historical linguistics* (pp. 687–712). Malden, MA: Blackwell Publishing.

Thráinsson, Höskuldur. 1979. *On Complementation in Icelandic*. [Ph.D. dissertation, Harvard University]. New York: Garland.

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

Thráinsson, Höskuldur. 2005. *Setningar: Handbók um setningafræði*. [Editor and main author Höskuldur Thráinsson, co-authors Eiríkur Rögnvaldsson, Jóhannes Gísli Jónsson, Sigríður Magnúsdóttir, Sigríður Sigurjónsdóttir and Thórunn Blöndal.*]* *Íslensk tunga III*. Reykjavík: Almenna bókafélagið.

Thráinsson, Höskuldur. 2013. “Formáli.” In Höskuldur Thráinsson, Ásgrímur Angantýsson & Einar Freyr Sigurðsson (Eds.), *Tilbrigði í íslenskri setningagerð I. Markmið, aðferðir og efniviður* (pp. 7–9). Reykjavík: Málvísindastofnun.

Thráinsson, Höskuldur, Ásgrímur Angantýsson, Einar Freyr Sigurðsson, Sigrún Steingrímsdóttir & Thórhallur Eythórsson. 2013. “Efnissöfnun og aðferðafræði.” In Höskuldur Thráinsson, Ásgrímur Angantýsson & Einar Freyr Sigurðsson (Eds.), *Tilbrigði í íslenskri setningagerð I. Markmið, aðferðir og efniviður* (pp. 19–68). Reykjavík: Málvísindastofnun.

Thráinsson, Höskuldur, Thórhallur Eythórsson, Ásta Svavarsdóttir & Thórunn Blöndal 2015. “Fallmörkun.” In Höskuldur Thráinsson, Ásgrímur Angantýsson & Einar Freyr Sigurðsson (Eds.), *Tilbrigði í íslenskri setningagerð II. Helstu niðurstöður, tölfræðilegt yfirlit* (pp. 33–76). Reykjavík: Málvísindastofnun.

Thráinsson, Höskuldur, Ásta Svavarsdóttir, Eiríkur Rögnvaldsson, Jóhannes Gísli Jónsson, Sigríður Sigurjónsdóttir & Thórunn Blöndal. 2013. “Markmið.” In Höskuldur Thráinsson, Ásgrímur Angantýsson & Einar Freyr Sigurðsson (Eds.), *Tilbrigði í íslenskri setningagerð I. Markmið, aðferðir og efniviður* (pp. 11–18). Reykjavík: Málvísindastofnun.

Thráinsson, Höskuldur & Sten Vikner. 1995. “Modals and double modals in the Scandinavian languages.” *Working Papers in Scandinavian Syntax, 55*:51–88.

Ussery, Cherlon. 2013. “Variability in Icelandic agreement: An interaction of DP licensing and Multiple Agree.” In Seda Kan, Claire-Moore-Cantwell & Robert Staubs (Eds.), *Proceedings of NELS 40* (pp. 211–226). Amherst, MA: Graduate Linguistic Student Association, University of Massachusetts, Amherst.

Van de Velde, Freek. 2017. *Retro-predicting language change with binomial regression analysis*. [Conference Presentation]. ICHL, Austin, Texas, United States.

Van de Velde, Freek & Muriel Norde. 2016. “Exaptation. Taking stock of a controversial notion in linguistics.” In Muriel Norde & Freek Van de Velde (Eds.), *Exaptation and language change* (1–35). Amsterdam: John Benjamins.

Van de Velde, Freek & Peter Petré. 2020. “Historical linguistics.” In Svenja Adolphs & Dawn Knight (Eds.), *The Routledge handbook of English language and digital humanities* (pp. 328–359). London: Routledge.

Van de Velde, Freek, Hendrik De Smet & Lobke Ghesquiére. 2013. “On multiple source constructions in language change.” *Studies in Language, 37*(3):473–489.

Ventura, Rafael, Joshua B. Plotkins & Gareth Roberts. 2022. “Drift as a driver for language change: An artificial language experiment.” *Cognitive science, 46*(9):e13197. doi: 10.1111/cogs.13197

Vincent, Nigel. 2020. “Complex versus compound prepositions: Evidence from Gallo-Romance.” In Sam Wolfe & Martin Maiden (Eds.), *Variation and change in Gallo-Romance grammar* (pp. 347–363). Oxford: Oxford University Press.

Vincent, Nigel. 1999. “The evolution of c-structure: prepositions and PPs from Indo-European to Romance.” *Linguistics, 37*(6):1111–1153.

Walkden, George. 2014. *Syntactic Reconstruction and Proto-Germanic*. Oxford: Oxford University Press.

Walkden, George. 2019. “The many faces of uniformitarianism in linguistics.” *Glossa: a journal of general linguistics, 4*(1):52.1–17.

Wallenberg, Joel. 2009. *Antisymmetry and the conservation of c-command: Scrambling and phrase structure in synchronic and diachronic perspective*. [Ph.D. dissertation, University of Pennsylvania].

Wallenberg, Joel. 2019. “A variational theory of specialization in acquisition and diachrony.” In Anne Breitbarth, Miriam Bouzouita, Lieven Danckaert & Melissa Farasyn (Eds.), *The determinants of diachronic stability* (245–262). Amsterdam: John Benjamins.

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

Watkins, Calvert. “Towards Proto-Indo-European syntax: Problems and pseudo-problems.” In Sanford B. Steever, Carol A. Walker & Salikoko S. Mufwene (Eds.), *Papers from the Parasession on Diachronic Syntax* (pp. 306–326). Chicago: Chicago Linguistic Society.

Wedel, Andrew. 2015. “Simulation as an Investigative Tool in Historical Phonology.” In Patrick Honeybone and Joe Salmons (Eds.), *The Oxford handbook of historical phonology* (pp. 149–163). Oxford: Oxford University Press.

Weinreich, Uriel, William Labov, & Marvin I. Herzog. 1968. “Empirical foundations for a theory of language change.” In Winfred P. Lehmann & Yakov Malkiel (Eds.), *Directions for historical linguistics* (pp. 95–195). Austin: University of Texas Press.

Wickham, Hadley. 2016. *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag.

Winkler, Robert L. 1972. “A decision-theoretic approach to interval estimation.” *Journal of the American Statistical Association, 67*(337):187–191.

Winters, Peter R. 1960. “Forecasting sales by exponentially weighted moving averages.” *Management Science, 6*(3):324–342.

Wood, Jim. 2015. *Icelandic Morphosyntax and Argument Structure*. Dordrecht: Springer.

Wood, Jim. 2022. “Reflexive datives and argument structure.” Ms., Yale University.

Wood, Jim & Einar Freyr Sigurðsson. 2011. “Icelandic verbal agreement and pronoun-antecedent relations.” *Working papers in Scandinavian syntax,* 88:81–130.

Yang, Charles. 2002. *Knowledge and learning in natural language*. Oxford: Oxford University Press.

Yip, Moira, Joan Maling & Ray Jackendoff. 1987. “Case in Tiers.” *Language*, *63*:217–250.

Zaenen, Annie, Joan Maling and Höskuldur Thráinsson. 1985. “Case and grammatical functions: The Icelandic Passive.” *NLLT* 3: 441–83. [Reprinted 1990 in Joan Maling & Anne Zaenen (Eds.), Modern Icelandic Syntax, pp. 165–185. San Diego: Academic Press].

Zwart, Jan-Wouter. 2011. *The syntax of Dutch*. Cambridge: Cambridge University Press.