

# New Data on Language Change: Compositional Definiteness in American Norwegian

Yvonne van Baal

Postdoctoral Fellow, Department of Language and Literature, Norwegian University of Science and Technology, Trondheim, Norway  
yvonne.van.baal@ntnu.no

## Abstract

This study investigates language change in American Norwegian (AmNo), a heritage language with a long migration and research history. Specifically, the paper examines the co-occurrence of a prenominal determiner and a suffixed article in modified definite phrases, known as compositional (or double) definiteness (CD). Recordings from three different points in time, recently made available in the *Corpus of American Nordic Speech* (Johannessen, 2015), are used to track historical developments. The data show that previous generations of speakers use CD in a homeland-like manner. Present-day speakers, on the other hand, frequently omit the determiner, which reflects a language change in this generation. I suggest that this innovation is preceded by a shift in the distributional frequency of exceptions to CD, which is intertwined with changes in language maintenance in the communities. Together, these factors contribute to the timing of the language change.

**Keywords:** language change; heritage language; American Norwegian; definiteness; compositional definiteness; corpus linguistics.

## 1 Introduction

The Germanic heritage languages<sup>1</sup> in the United States all have a long migration history, typically with a period of large migration during the 19<sup>th</sup> and early 20<sup>th</sup> century (see, for example, Johannessen & Salmons, 2015 for an overview). These languages have been minority languages for several generations, and the current speakers tend to be elderly third- to fifth-generation immigrants (cf. Putnam et al., 2018). This is also the case for American Norwegian (AmNo), the language investigated in this paper.

The long migration history can be a challenge in studying these heritage languages, as it complicates the establishment of a proper baseline. Patterns observed for current speakers of the language might reflect recent innovations in the heritage language, or they might be innovations from previous generations of speakers. Despite this challenge, the long migration history of the Germanic heritage languages can also provide an interesting window on language change in these languages, provided that linguistic data from previous generations are available. For AmNo, speech recordings from 1931, 1942, and the period 1987-1992 have recently been added to the *Corpus of American Nordic Speech* (Johannessen, 2015),<sup>2</sup> enabling historical studies of the language.

In this paper, I present a study of language change in AmNo modified definite phrases, using the data from the corpus. In Norwegian, modified definite phrases obligatorily

---

<sup>1</sup> I adopt the definition in Rothman (2009): “A language qualifies as a *heritage language* if it is a language spoken at home or otherwise readily available to young children, and crucially this language is not a dominant language of the larger (national) society” (p. 156).

<sup>2</sup> The corpus was established as the *Corpus of American Norwegian Speech* (CANS) in 2015 (Johannessen, 2015). After the addition of American Swedish recordings, the name has been changed to *Nordic*. In the fall of 2019, the historical AmNo material was the most recent extension of CANS. This data serves as the empirical basis for the present paper.

contain both a prenominal determiner and a suffixed article. This phenomenon is known as compositional (or double) definiteness (CD). An example of CD is given in (1);<sup>3</sup> the phenomenon is discussed further in Section 3.1.

- (1) *den stor-e bil-en*  
DEF.SG large-DEF car-DEF.M.SG  
'the large car'

Previous studies have found that CD in AmNo diverges from CD in homeland Norwegian (Anderssen et al., 2018; van Baal, 2018, 2020). The studies so far are exclusively based on present-day speakers of AmNo. Their behavior might be the result of an earlier language change affecting their input (“incipient changes in the input” in Polinsky, 2018, pp. 33-35; “missing input competence divergence” in Pires & Rothman, 2009), and historical data are therefore needed to investigate the timing of innovative use of CD. The central question of the present paper is at which point in time CD has changed in American Norwegian. In addition, the paper discusses some of the factors that have contributed to the timing of the change.

The paper is structured as follows. In Section 2, the migration and research history of American Norwegian is introduced. Compositional definiteness and previous research on CD in AmNo are discussed in Section 3. Section 4 discusses the corpus search, and the results are presented in Section 5. In Section 6, I discuss the language internal and sociolinguistic factors that have likely contributed to the language change in AmNo and its timing. Section 7 concludes.

## 2 The Migration and Research History of American Norwegian

Norwegian immigration to the United States began almost two centuries ago with the arrival of the first Norwegian immigrants in New York in 1825. A period of mass migration from Norway to the United States started a few decades later, concluding again following the enactment of stricter immigration laws in the 1920s. Between 1850 and the 1920s, around 850,000 Norwegians moved to North America (Haugen, 1953, p. 29).<sup>4</sup> Most Norwegians settled in the Midwest and Pacific Northwest, where they built Norwegian American communities in which the Norwegian language was preserved as a community language, in churches and schools, and in Norwegian American newspapers (see Haugen, 1953; Lovoll, 1999 for more on the Norwegian American history).<sup>5</sup>

Most of the current speakers of AmNo are descendants of the 19<sup>th</sup> and early 20<sup>th</sup> century immigrants, and they primarily live in the states of Wisconsin, Minnesota, Iowa, and North and South Dakota. The speakers are third- to fifth-generation immigrants, and virtually all are of advanced age. They are the final generation that acquired Norwegian during childhood; their children are mostly monolingual speakers of English. AmNo can thus be classified as a moribund heritage language. Although most current speakers no longer use the

---

<sup>3</sup> The following glosses are used in the paper: DEF = definite, F = feminine, M = masculine, N = neuter, PL = plural, SG = singular. Note that some Norwegian dialects have apocope and therefore no definiteness inflection on the adjective.

<sup>4</sup> Most Norwegian immigrants moved to the United States, but around 40,000 of them settled in Canada. In this period, Norway had the second largest emigration rate (after Ireland) from Europe to the United States (Haugen, 1953, p. 29).

<sup>5</sup> Norwegian immigrants also settled in other U.S. states. However, mainly the communities from the Midwest have been recorded and studied, both historically and in current research on AmNo.

language for daily communication, they are still relatively fluent in AmNo (Johannessen & Salmons, 2012, p. 140).

The long history of AmNo poses some challenges to our research. First, it complicates the notion of “baseline.” It has been argued that the baseline should be the input to the speakers rather than the standard homeland variety of the language (Benmamoun et al., 2013, p. 134; Pascual y Cabo and Rothman, 2012; Polinsky, 2018, pp. 11-16; Polinsky and Kagan, 2007, p. 8, inter alia). For a population like the AmNo speakers, the establishment of a baseline is difficult. It can be hard to decide what to compare the speakers to when heritage speakers are several generations removed from migration from the homeland (see also D’Alessandro et al., 2021). Related to this challenge is the issue of language change. Innovations and restructurings observed in present-day speakers possibly occurred in an earlier generation of speakers, who were heritage speakers of the language themselves.<sup>6</sup>

An example of changes in the input to heritage speakers is presented in Montrul & Sánchez-Walker (2013). They found that both heritage speakers and immigrant speakers of Spanish in the United States showed tendencies to omit the obligatory differential object marking (DOM) on animate, definite objects. In other words, the language of the immigrant speakers (who grew up monolingually Spanish) had undergone a change that was transmitted to the next generation, the heritage speakers, to whom they provide the input.

A way to mitigate the baseline and language change challenges would be to study data from previous generations of heritage speakers of the language in question. However, for this to be possible, recordings or descriptions of previous generations need to be available. In the case of AmNo, we are fortunate that the language not only has a long migration history but also a long research history, with recordings from several points in time. Recently, a subset of these recordings has been transcribed, linguistically tagged, and added to the already existing, and much used, *Corpus of American Nordic Speech* (CANS) (Johannessen, 2015). This makes it possible to use these recordings in our research without first having to conduct the time-consuming transcription.<sup>7</sup>

Already at the start of the 20<sup>th</sup> century, George Flom and Nils Flaten studied the American Norwegian language and published small papers based on their fieldnotes (Flaten, 1900; Flom, 1900, 1903, 1926). The first large-scale data collection that also included audio recordings was conducted in 1931 by Didrik A. Seip and Ernst W. Selmer. Unfortunately, their recordings were not well preserved after returning to Norway, and large parts of the audio collection became damaged or lost (Haugen, 1992). The recordings that remained are quite hard to work with, but some of them have been transcribed and are now part of CANS.

During the 1930s and 40s, the American scholar Einar Haugen collected a large body of recordings, resulting in his seminal work, *The Norwegian Language in America* (1953). The book is based on fieldwork in the period 1936-1948, during which Haugen interviewed 260 AmNo speakers. Recordings from the year 1942 have been added to CANS. The next period of research took place in the 1980s, when Arnstein Hjelde conducted fieldwork on the dialects of Trøndelag (a region in central Norway), as well as with speakers from other dialectal backgrounds (Hjelde, 1992, 1996, 2015). The most recent collection of recordings of AmNo started in 2010 with the “Norwegian in America” (NorAmDiaSyn) project led by

---

<sup>6</sup> Based on their study of the dative case in heritage varieties of German, Yager et al. (2015) argue that observed changes in heritage languages are best described in terms of grammatical restructuring rather than the traditional narrative of loss and simplification.

<sup>7</sup> Many of the older recordings have been available at the website of the Tekstlaboratory at the University of Oslo (<http://tekstlab.uio.no/norskiamerika/opptak.html>). This means that, in principle, historical research on AmNo has been possible even before these recordings were transcribed, tagged, and added to CANS (e.g., Hjelde, 2015). However, the process of transcription is extremely time-consuming, and the addition of the older recordings to CANS have without a doubt made historical investigations on AmNo much more feasible.

Janne Bondi Johannessen. In the period 2010-2016, several trips were made to the United States and Canada to make both audio and video recordings of AmNo.

Large parts of the material collected in these different periods are now available in CANS, and more material may be added in the future. At the moment of data collection for this study (fall 2020), the corpus contained recordings of 205 American Norwegians who together utter somewhat over 700,000 tokens. The majority of the speakers are from the most recent period of data collection (N=152), but the addition of the older recordings is without doubt a great resource for studies on AmNo and heritage languages in general, as it enables us to investigate change in the heritage language. This paper examines the historical development of compositional definiteness (CD) in AmNo.

### 3 Compositional Definiteness

#### 3.1 *CD in Homeland Norwegian*

Like the other Scandinavian languages, Norwegian expresses definiteness with a suffixed article on the noun, as in (2). When the definite noun is modified by an adjective or numeral, the phrase contains both a pronominal determiner and the suffixed article, see (3). This is known as compositional definiteness (CD) (Anderssen, 2012).<sup>8</sup> Apart from some exceptions discussed below, modified definite phrases are only grammatical when both the pronominal determiner and the suffixed article are present.

(2) *hest-en*  
horse-DEF.M.SG  
'the horse'

(3) *den hvit-e hest-en*  
DEF.SG white-DEF horse-DEF.M.SG  
'the white horse'

As can be seen in the glosses in the examples above, both definiteness morphemes inflect for gender and number. This investigation is concerned with the use of the determiner and the suffix in American Norwegian. In other words, the focus is on the presence or absence of the morphemes rather than their morphological form. The gender and number inflection will therefore not be discussed further (for studies on gender in AmNo, see Johannessen & Larsson, 2015; Lohndal & Westergaard, 2016; Rødvang, 2017).

Although CD is generally obligatory in Norwegian, the pronominal determiner may be left out with a restricted set of adjectives. An example is given in (4), with the adjective *andre* 'other' that allows for omission of the determiner. Other exceptions are superlatives and *venstre* 'left' and *høyre* 'right', among others (see also van Baal, 2020, pp. 36-39, and Dahl, 2015, pp. 124-125 on Swedish). There is inter-speaker variation as to preferences for inclusion or omission of the determiner with these adjectives, but in general we can say that they constitute exceptions to CD, and I refer to them as "exceptional adjectives" throughout this paper.

(4) (*den*) *andre sid-a*  
DEF.SG other.DEF side-DEF.F.SG  
'the other side'

---

<sup>8</sup> Norwegian, Swedish, and Faroese have compositional definiteness in modified definite phrases. Only Norwegian is discussed in the present paper.

Current syntactic analyses of CD assume that Norwegian DPs contain two determiner-like projections, a high one (above the adjective) where the prenominal determiner is inserted, and a lower one (below the adjective, above the noun) where the definite suffixed article is located (e.g., Anderssen, 2006, 2012; Julien, 2002, 2005). Under this analysis, CD is a rather complex syntactic structure, and it is at least more complex than unmodified phrases like those in (2) (cf. Anderssen et al., 2018).

In addition to its syntactic complexity, CD is a rather infrequent phenomenon, as pointed out by both Dahl (2015, p. 121) and Anderssen et al. (2018). The definite suffix is used much more frequently than the prenominal determiner, since the former also appears in unmodified phrases, whereas the latter only appears in modified phrases.<sup>9</sup> As a result of the low frequency and high complexity, CD is acquired relatively late by monolingual children (Anderssen, 2006, 2012; see Bohnacker, 2003 for similar findings in Swedish). The fact that the prenominal determiner is unstressed is argued to play a large role in its late acquisition (Anderssen, 2006; Bohnacker, 2003). The bilingual Norwegian-English child studied in Anderssen & Bentzen (2013) shows both slower development and a different error pattern compared to her monolingual peers. In a study of bilingual Norwegian-English speakers, such as the American Norwegian heritage speakers, it has to be kept in mind that CD is unique to Norwegian. English only uses prenominal determiners in definite phrases, modified and unmodified phrases alike.

Given the factors complexity, low frequency, cross-linguistic differences, and late acquisition in monolinguals, it could be expected that CD is vulnerable for change in American Norwegian. This is indeed what has been found in previous research on CD in present-day AmNo speakers, as discussed in Section 3.3.

### 3.2 *The Alternative: Adjective Incorporation*

An alternative to CD that is particularly frequent in some Norwegian dialects is so-called “adjective incorporation.” In these instances, the uninflected adjective is incorporated to the definite noun and there is no prenominal determiner, e.g., *hvit-hest-en* ‘the white horse’ (cf. with (3)). Some of the present-day speakers in van Baal (2018, 2020) use adjective incorporation, while Anderssen et al. (2018) do not observe it in their data. In the data for the present paper, hardly any adjective incorporation was found either. However, such constructions are hard to find with the search queries used in these corpus studies. In CANS, adjective incorporations (and other compounds) are transcribed as single words. A search query to adjectives followed by nouns (see Section 4 for details) would thus not provide such phrases.<sup>10</sup> In other words, the apparent lack of adjective incorporation in AmNo corpus data is likely related to the corpus searches and not to the lack of such phrases from the language.

As pointed out by Natvig & van Baal (2020), it is not always easy to establish whether a phrase is an adjective incorporation or rather consists of a juxtaposed adjective and noun. Adjective incorporations come with a typical prosody for compounds, but individual listeners might interpret these prosodic patterns differently. Acoustic analyses could shed light on this matter, but the quality of the older recordings does not always allow for such an analysis. Moreover, even if acoustic analysis were possible, one may not be able to draw conclusions

---

<sup>9</sup> The distal demonstrative is homophonous with the prenominal determiner and this form can occur in phrases without an adjective: *den hesten* ‘that horse’. However, the demonstrative can carry stress, whereas the prenominal determiner cannot (Faarlund et al., 1997, p. 327), and the two should therefore be considered separate lexemes (Anderssen, 2006, p. 118).

<sup>10</sup> The one example of adjective incorporation in the historical data was found in one of the other sentences that was part of the results, i.e., not as a direct result of the search query.

from it. There are hardly any studies on intonation or prosody in AmNo,<sup>11</sup> and it is therefore not clear whether the (older) heritage speakers use the same prosody as present-day homeland speakers (also noted by Natvig & van Baal, 2020).

In the present paper, the data from CANS are used. This corpus search may miss some instances of adjective incorporation, but this is unfortunately the consequence of the transcription conventions and search possibilities of the corpus. The CANS data is manually checked (see Section 4), but not analyzed acoustically for the reasons stated above. In other words, I relied on the transcriptions in CANS and if they consider an adjective non-incorporated to the definite noun, I followed that. This approach furthermore facilitates replicability.

### 3.3 Previous Research on CD in AmNo

Two studies have been conducted on the use of compositional definiteness by the current speakers of American Norwegian. Anderssen et al. (2018) investigated the data in CANS, which consisted of 50 speakers of the current generation at that time. Van Baal (2018, 2020) studied CD with the help of two elicited production tasks with 20 participants.<sup>12</sup> While these studies make use of different types of data (corpus vs. elicitation), they find rather similar results.

In both studies, the speakers produced many modified definite phrases without CD, in contexts where it would be obligatory in homeland Norwegian. In the corpus data, 39% of the modified definite phrases had CD,<sup>13</sup> while 28% of the elicited data showed CD. As is expected with a population of heritage speakers, there is a certain degree of inter-individual variation, but all speakers produce a reasonable number of phrases without CD. Some speakers produce very few instances of CD, and there are even a couple of speakers who never use it at all.

The most typical modified definite phrase for AmNo speakers is one without a prenominal determiner. Two examples are given in (5), with an example from the corpus data in (5a) and an instance from the elicited production data in (5b). All speakers in van Baal (2020) produce phrases like this without the determiner, and most do so (very) frequently.

(5a.) *norsk-e*                      *ordbok-a*  
 Norwegian-DEF                dictionary-DEF.F.SG  
 ‘the Norwegian dictionary’  
 (westby\_WI\_05gm, CANS, Anderssen et al., 2018, p. 755)  
 Homeland Norwegian: *den norske ordboka*

(5b.) *stor-e*                      *båt-en*  
 large-DEF                      boat-DEF.M.SG  
 ‘the large boat’  
 (flom\_MN\_01gm, elicited production)  
 Homeland Norwegian: *den store båten*

<sup>11</sup> Haugen (1941) discusses AmNo intonation, but focuses on the intonation of English loanwords.

<sup>12</sup> Seven of van Baal’s (2020) participants were also part of the corpus study by Anderssen et al. (2018). The elicitation consisted of a translation task and a picture-based elicitation. In both, modified definite phrases were elicited in the context of a sentence, but the picture-based task also provided phrases in isolation. See van Baal (2020, pp. 67-84) for details on the elicitation tasks.

<sup>13</sup> The former percentage is calculated using the following raw data in Anderssen et al. (2018, pp. 760-761): 93 phrases contained CD, out of the 237 phrases which required CD (i.e., 39.24%). The latter percentage is calculated using the following raw data in van Baal (2020), excluding the phrases that contained a demonstrative rather than determiner: 190 out of 687 phrases contained CD (i.e., 27.66%).

In addition to the frequent omission of the determiner, the studies found that a subgroup of the speakers omits the suffixed article. They do this to varying degrees; most omit the suffix only occasionally, while a small group does so frequently. Two examples from the corpus data and the elicitation data are given in (6). Omission of the suffix in modified definite phrases has been found to correlate with a low(er) proficiency in the heritage language, as measured by speech rate and vocabulary knowledge (van Baal, 2019), or by homeland-like gender marking (Anderssen et al., 2018).

(6a.) *den best-e gang*  
 DEF.SG best-DEF time  
 ‘the best time’  
 (chicago\_IL\_01gk, CANS)  
 Homeland Norwegian: *den beste gangen*

(6b.) *den grønn-e bil*  
 DEF.SG green-DEF car  
 ‘the green car’  
 (sunburg\_MN\_11gk, elicited production)  
 Homeland Norwegian: *den grønne bilen*

Investigations of CD in present-day AmNo found that CD is used in a different manner than in homeland varieties. Van Baal (2020) argues that phrases without the determiner have become a part of the current AmNo language, since they are highly frequent and used by all speakers.

Given the long history of AmNo, however, questions about the timing of this language change naturally arise. Are the current speakers the first generation for whom the typical modified definite phrase lacks the determiner; or, is this rather something that happened earlier in the history of the heritage language? It is important to keep in mind that the current speakers received input from people who were heritage speakers of Norwegian themselves, so it is quite possible that the omission of CD, and of the determiner in particular, already happened in the (grand-) parents’ generation. Thus, the central question of this paper is at which point in time modified definite phrases without the determiner became part of American Norwegian.

#### 4 The Corpus and Data Collection

In order to answer the research question outlined above, I investigated the historical data in the *Corpus of American Nordic Speech* (CANS). CANS was established first in 2015 (Johannessen, 2015). At that time, it contained recordings of present-day AmNo speakers that were made in the period 2010-2012. The corpus has been expanded several times with material from present-day speakers. Recordings from present-day American Swedish speakers were added in 2017 (Andréasson et al., 2017; see Larsson et al., 2012 for details on the recordings). In 2019, the corpus was expanded again, this time with transcribed recordings from earlier points in time. This means it is now possible to conduct historical research on AmNo, as is done in the present paper.

This study is based on CANS-v.3.<sup>14</sup> The corpus was searched for strings that contained an adjective followed by a noun. After removal of a small proportion of noise, the

---

<sup>14</sup> The corpus can be accessed at <https://tekstlab.uio.no/glossa2/cans3>. Since January 27, 2021, version 3.1 is available, which contains 45 more recordings from Einar Haugen. These are not included in the present study.

results were divided into indefinite phrases, which are excluded as they never contain CD, and definite phrases. Demonstratives (e.g., *disse små land-a* ‘these small countries’, CANS-1942) and phrases expressing dates (e.g., *den femte september* ‘September 5th’, CANS-1942) were excluded. The remaining definite phrases were split into those that require CD, and those that contained an exceptional adjective and no prenominal determiner. As described above (Section 3.1), such phrases do not require CD. For all phrases that require CD, I checked whether both the prenominal determiner and the suffixed article were present. This way of categorizing the data allowed me to investigate not only how frequent CD is, but also which proportion of the modified definite phrases requires CD. As we will see, this turns out to be relevant for understanding the historical development.

The recordings from 1931 unfortunately provided too little data for analysis, so they are not included here. This paper thus investigates two time periods, 1942 and 1987-1992, and compares them with the data from present-day speakers. An overview of the data is given in Table 1, while Figure 1 shows a map of the locations of recording. Einar Haugen’s recordings from 1942 contain utterances of 31 speakers from four places across Wisconsin and Minnesota. The age range of these speakers is large, indicating that at that point in time, American Norwegian was still spoken across all generations. The recordings from 1987-1992 by Arnstein Hjelde contain utterances produced by five individuals from three locations in Wisconsin and Minnesota. They were all elderly at the time of recording.

TABLE 1. Overview of historical American Norwegian data used in the study.

	Haugen recordings (1942)	Hjelde recordings (1987-1992)
N of mod. def. phrases	96	43
N of speakers	31	5
Places of recording	Blair (WI), Coon Valley (WI), Spring Grove (MN), Westby (WI)	Appleton (MN), Coon Valley (WI), La Crosse (WI)
Year of birth	1850 – 1927	1909 – 1919
Age at recording	14 – 92	73 – 83

<< **FIGURE 1 HERE** >>

The data from these two periods are compared with each other and with the current speakers. As discussed above, the data of present-day AmNo consist of 50 speakers in CANS (Anderssen et al., 2018) and elicitation data of 20 speakers (van Baal, 2020). There is some overlap in the speakers in these two studies. However, there is no overlap between the three different time points, and no speakers from the Haugen recordings are also part of the Hjelde recordings. As a result, the present paper investigates the historical development of the AmNo language, rather than the development across the lives of individual speakers.

## 5 Results

### 5.1 General Results

In this section, I compare the use of compositional definiteness at three time points; subsequent sections discuss the data in more detail. Table 2 presents the number of modified definite phrases at each time point: 1942 (the Haugen recordings), 1987-1992 (the Hjelde recordings), the data from Anderssen et al. (2018), and the data from van Baal (2020). The latter two are present-day American Norwegian. As noted above, there are adjectives that do not require CD. Table 2 also presents the number of exceptions, the number of phrases that require CD, and the number of phrases that have CD. Since van Baal (2020) did not elicit



exceptions, the number of exceptions in her data is zero. In the table, the number of phrases with CD includes adjective corporation, and thus presents the number and percentage of homeland-like modified definite phrases. The use of CD is presented in Figure 2, where the bars indicate the use of CD as a proportion of the phrases that require it, corresponding to the bottom row in Table 2.

TABLE 2. Number of modified definite phrases, exceptions, phrases that require CD and phrases with CD at different points in time: 1942 (Haugen), 1987-1992 (Hjelde), present-day in Anderssen et al. (2018), and present-day in van Baal (2020).

	1942	1987-1992	ALW 2018	van Baal 2020
Modified definite phrases	96	43	422	687
Exceptions (no determiner)	52	33	185	0
Phrases that require CD	44	10	237	687
Phrases that contain CD	31 (70%)	8 (80%)	93 (39%)	190 (28%)

<<FIGURE 2 HERE>>

As can be seen in Table 2 and Figure 2, speakers in all subsamples produced phrases with CD, but the proportion of CD varies. In (7)-(10), an example of a modified definite phrase with CD is given for each time point. In all examples, both the prenominal determiner (*det* or *den*) and the suffixed article (*-et* or *-en*) are present.

(7) *det norsk-e språk-et i Amerika*  
 DEF.N.SG Norwegian-DEF language-DEF.N.SG in America  
 ‘the Norwegian language in America’  
 (spring\_grove\_MN\_19gm, CANS-1942)

(8) *det billigst-e land-et*  
 DEF.N.SG cheapest-DEF land-DEF.N.SG  
 ‘the cheapest land’  
 (la\_crosse\_WI\_02gm, CANS-1992)

(9) *det stor-e hus-et*  
 DEF.N.SG big-DEF house-DEF.N.SG  
 ‘the big house’  
 (blair\_WI\_04gk, 2010, Anderssen et al., 2018, p. 755)

(10) *den rød-e blomm-en*<sup>15</sup>  
 DEF.SG red-DEF flower-DEF.M.SG  
 ‘the red flower’  
 (coon\_valley\_WI\_06gm, 2016, van Baal, 2020, p. 120)

While CD is used by all generations, they do not all use it to the same extent. In the recordings from 1942, 70% of the modified definite phrases that require CD contain it (31 out of 44, see Table 2). For the recordings from 1987-1990, this score is 80% (8 out of 10). These scores are quite high, but somewhat lower than what would be expected in (present-day)

<sup>15</sup> The word for flower is *blomster* in standard Bokmål Norwegian; the variants *blom* and *blome* are allowed in Nynorsk Norwegian. They are used in many dialects and by several AmNo speakers as well.

homeland Norwegian where CD is obligatory. This could be evidence of a change in progress.

The picture for present-day speakers is rather different. In Anderssen et al. (2018), CD is used in 39% of the required contexts and in van Baal (2020), this score is only 28%. One might wonder why these scores are different since they are both from present-day speakers, but we should keep in mind that the studies are based on different types of data, viz. corpus and elicitation. In addition, the studies do not include the exact same speakers, although there is some overlap (7 speakers). Putting aside the precise reasons for this difference, the speakers in Anderssen et al. (2018) and in van Baal (2020) clearly use CD to a much lower extent than speakers of previous generations. There is a large decline in the use of CD between the recordings from 1987-1992 (80%) to the current recordings (39-28%).

One may want to analyze the data by year of birth of the speakers, rather than by year of recording as done above. Since CD is quite infrequent (see Section 3.1), the number of phrases per speaker is typically low. Unfortunately, it is therefore not very informative to plot the data by year of birth. If the data are arranged by immigrant generation, a similar problem arises. Most groups produce only few phrases that require CD, and a comparison is not informative; therefore, Figure 2 instead plots by date of recording.

It should be noted that the speakers recorded in 1987-1992 as well as those recorded from 2010 onwards are all elderly speakers. These groups are rather homogenous with respect to their years of birth. The recordings from 1942 (Haugen’s material), on the other hand, may be divided into two groups based on their age at the time of recording. Speakers are classified as elderly in CANS when they are above 50. The 1942 subsample contains 19 elderly speakers (born 1850-1892), and 72% of their modified definite phrases contains CD. For the remaining 12 younger speakers (born 1893-1927), the proportion of CD is 67%. In other words, the two age groups show rather similar behavior, despite the large variation in year of birth.

I conclude from the data presented in this section that the language change in AmNo has happened in the current generation of speakers (recorded since 2010), rather than earlier in the AmNo history (recorded in 1942 or 1987-1992). The next section takes a closer look at the data from phrases that require CD.

## 5.2 *Phrases That Require Compositional Definiteness*

The previous section showed that a large proportion of the modified definite phrases in the older recordings contains CD. At the same time, there is also a small number of phrases that require CD but does not contain it. Table 3 shows the types of structures used in modified definite phrases in 1942 and 1987-1992, as well as in present-day speakers based on Anderssen et al. (2018) and van Baal 2020.

TABLE 3. Types of modified definite phrases used at different time points: 1942 (Haugen), 1987-1992 (Hjelde), present-day in Anderssen et al. (2018), and present-day in van Baal (2020).

	1942	1987-1992	ALW 2018	van Baal 2020
Compositional definiteness	31 (70%)	7 (70%)	93 (39%)	143 (21%)
Adjective incorporation	-	1 (10%)	-	47 (7%)
Without determiner	5 (11%)	0	113 (48%)	339 (49%)
Without suffixed article	7 (16%)	0	31 (13%)	35 (5%)
Bare phrase	1 (2%)	2 (20%)	0 (0%)	123 (18%)
Total phrases	44	10	237	687

In 1942, 11% of the phrases that requires CD lacks the prenominal determiner, as in the example in (11a). Somewhat more frequent is the omission of the suffixed article (16%), as in (11b). Some of those phrases include a kinship term (like *bror* ‘brother’). In the homeland Norwegian dialects of the ancestors of the AmNo speakers, bare kinship terms are typically allowed, a pattern which has been found to be more frequent in AmNo (Kinn, 2021). Phrases such as the one in (11b) could be considered homeland-like. Most important for present purposes, however, is the fact that these phrases do not contain CD and contribute to the variation in the data. Finally, one phrase from the 1942 data lacks both determiner and suffix. As is clear from Tables 2 and 3, phrases like those in (11) are not very frequent in this period, and the (vast) majority of the phrases contains CD in these recordings.

In the period 1987-1992, no phrases without the determiner or without the suffixed article were found. There are two instances of bare phrases, in which neither determiner nor suffix are present while the context is definite. An example is given in (12). The two bare phrases were uttered by the same individual and contain a kinship term (*søster* ‘sister’ or *datter* ‘daughter’).

(11a.) *på hard-e bakk-en*  
 on hard-DEF ground-DEF.M.SG  
 ‘on the hard ground’  
 (blair\_WI\_23um, CANS-1942)  
 Homeland Norwegian: *på den harde bakken*<sup>16</sup>

(11b.) *den eldst-e bror min*  
 DEF.SG oldest-DEF brother my  
 ‘my oldest brother’  
 (westby\_WI\_24gm, CANS-1942)  
 Homeland Norwegian: *den eldste bror(en) min*

(12) *yngst-e søster hans far min*<sup>17</sup>  
 youngest-DEF sister his father my  
 ‘my father’s youngest sister’  
 (coon\_valley\_WI\_17gm, CANS-1992)  
 Homeland Norwegian: *(den) yngste søstera hans*

Present-day speakers, i.e., the two rightmost columns in Table 3, look rather different. While CD is the most frequent pattern in the data from 1942 and 1987-1990, it is not in the data from Anderssen et al. (2018) and van Baal (2020). In these data, the omission of the determiner is most frequent, and found in almost 50% of the modified definite phrases that require CD. Phrases without the suffix and bare phrases are also found in present-day speakers, but these are subject to more variation (see Section 3.3).

The data presented in this section show two things. First, CD was still frequently used at previous points in time, but it has become quite infrequent in the present generation. This was also observed in the previous section. Second, there is variation with respect to the utterances divergent from homeland Norwegian, that is, utterances that do not contain CD. Three types of non-CD phrases were found across the different time points: omission of the

<sup>16</sup> When the phrase has a generic (non-specific) reading, the omission of the determiner would be homeland-like.

<sup>17</sup> An anonymous reviewer would transcribe the phrase as *yngste søster åt far min* with the preposition *åt* ‘to, of’. I chose to follow the transcription from CANS (cf. Section 5.2), but it’s worth noting that in both interpretations, the prenominal determiner and suffixed article are missing from this phrase.

determiner, omission of the suffix, and bare phrases. In this respect, there also is a change: present-day speakers typically use modified definite phrases without a determiner, while this pattern is not frequent at all in the language of previous generations. This raises a question about the historical development: why did AmNo change towards the frequent omission of determiners, given that this pattern at first glance does not appear much in their input? I argue in the next section that the answer to this question lies in the use of exceptions to CD.

### 5.3 *The Exceptions to CD*

Recall from Section 3.1 that the prenominal determiner may be omitted, even in homeland Norwegian, with a specific set of adjectives. These “exceptional adjectives” are superlatives (including *første* ‘first’ and *siste* ‘last’), ordinal numbers (including *andre* ‘second, other’), and directional terms (e.g., *venstre* ‘left’ and *høyre* ‘right’). With the adjective *hele* ‘whole’, the determiner is always omitted in Norwegian.<sup>18</sup> Examples of modified definite phrases with an exceptional adjective and no prenominal determiner can be found in the old recordings of AmNo as well. A few examples are given in (13) and (14), including some context to show that they do not contain a prenominal determiner.

(13a.) *der var eldst-e jent-a vår født*  
 there was oldest-DEF girl-DEF.F.SG our born  
 ‘there, our eldest daughter was born’  
 (coon\_valley\_WI\_45gk, CANS-1942)

(13b.) *og da brant hele by-en*  
 and then burned whole city-DEF.M.SG  
 ‘and then the whole city burned down’  
 (blair\_WI\_25gm, CANS-1942)

(14a.) *det var sist-e plass-en*  
 it was last-DEF place-DEF.M.SG  
 ‘it was the last place’  
 (coon\_valley\_WI\_17gm, CANS-1992)

(14b.) *jeg hadde masse gris-er hele tid-a*  
 I had many pig-PL whole time-DEF.F.SG  
 ‘I had a lot of pigs the whole time’  
 (appleton\_MN\_01gm, CANS-1987)

Phrases like those in (13) and (14), with an exceptional adjective and no prenominal determiner, are very frequent in the old recordings. Table 4 presents the frequency of exceptions and phrases that require CD at the three time points in AmNo. Since van Baal (2020) did not elicit exceptional adjectives (cf. Section 5.1), only Anderssen et al. (2018) is included for present-day AmNo.

TABLE 4. Frequency of modified definite phrases that require CD and those that are exceptions and occur without the determiner at different time points: 1942 (Haugen), 1987-1992 (Hjelde), and present-day in Anderssen et al. (2018).

<sup>18</sup> In this respect, *hele* ‘whole’ is an atypical exception: with the other exceptional adjectives, the prenominal determiner can be omitted, but in the case of *hele* the determiner must be omitted. In these phrases, *hele* could be classified as a strong quantifier rather than an adjective. For the present study on AmNo, it suffices to classify *hele* as an exception to the use of CD.

	1942	1987-1992	ALW 2018
Modified definite phrases	96	43	422
Require CD	44 (46%)	10 (23%)	237 (56%)
Exceptions	52 (54%)	33 (77%)	185 (44%)

Table 4 shows that the exceptions are very frequent in historical AmNo. In 1942, slightly over half of the modified definite phrases did not require CD and did not contain the pronominal determiner (54%). For the recordings from 1987-1992, this percentage is even higher (77%, although based on a small data set). At this point in time, less than a quarter of the modified definite phrases requires CD, because most of them contain an exceptional adjective. In other words, there seems to have been a rise in the frequency of the exceptions.

In fact, they are so frequent that one would start to wonder what actually the “exceptions” are in these data. In the recordings from 1987-1992, the use of CD seems to be the exception, as it is very low frequent and most of the definite phrases with an adjective do not contain a pronominal determiner. This pattern can be considered the input to the present-day speakers. In this light, it is not surprising that they use CD to a low extend, and particularly omit the determiner: They hardly received evidence for the obligatoriness of this determiner in their input.

The present-day speakers in Anderssen et al. (2018) use less exceptions than the generations before them. However, the exceptions make up a considerable part of the modified definite phrases even in this group. Between 1942 and 2010-2016, the relative frequency of exceptional adjectives fluctuates somewhat. One might wonder what the reasons for these fluctuations could be, but this question is left for future research. For the present paper, the important observation is that the phrases with a determiner were very infrequent in the input of present-day speakers, which likely contributed to the observed change with respect to CD.

At the same time, it is important to point out that phrases which do not require a determiner are frequent in homeland Norwegian as well (Anderssen et al., 2018, p. 751). Present-day homeland Norwegian should not be used as a baseline for AmNo, especially not for the old recordings (see the references and discussion in Section 2). Yet, a comparison with homeland Norwegian can make clear that the frequent use of exceptional adjectives is not a specific trait of older AmNo. For this comparison—and *not* as a baseline or target—data from three homeland Norwegian corpora are provided in Table 5: the Norwegian part of the Nordic Dialect Corpus (NDC; Johannessen et al., 2009), the Language Infrastructure made Accessible corpus (LIA),<sup>19</sup> and the Norwegian Web as Corpus (NoWaC; Guevara, 2010). NDC is a corpus of spoken present-day Norwegian, LIA contains recordings of older dialects (including some from the period of migration to the United States), and NoWaC is a present-day corpus of written Norwegian. These corpora are much larger than CANS, and they have therefore not been checked manually.<sup>20</sup>

TABLE 5. Frequency of modified definite phrases with CD and those that are exceptions and occur without the determiner in non-heritage Norwegian: NDC (spoken present-day language), LIA (older spoken language), and NoWaC (written present-day language).

	NDC	LIA	NoWaC
--	-----	-----	-------

<sup>19</sup> See <https://www.hf.uio.no/iln/english/research/projects/language-infrastructure-made-accessible/>.

<sup>20</sup> As discussed in Section 4, the AmNo data from CANS were manually checked and categorized. For the homeland Norwegian corpora, I searched for phrases that consist of one of the determiners followed by an adjective followed by a definite noun (i.e., phrases with CD), and phrases with an adjective and definite noun that are not preceded by a determiner. The latter are taken to be exceptional adjectives. Admittedly, this is a rough categorization, which was warranted by the size of the corpora.

Modified definite phrases	4560	7505	3,459,060
Phrases with CD	1049 (23 %)	2585 (34.4 %)	2,300,981 (66.5 %)
Exceptions	3511 (77 %)	4920 (65.6 %)	1,158,079 (33.5 %)

In both of the spoken non-heritage corpora (NDC and LIA), the exceptional adjectives are very frequent, and more than half of the modified definite phrases are of this type. These frequencies are comparable with those of the AmNo recordings from 1942 and 1987-1992 (see Table 4). In written Norwegian (NoWaC), however, CD is much more frequent than determiner omission. This difference between spoken and written homeland Norwegian is also pointed out by van Baal (2020, pp. 172-175). In this respect, it is relevant that present-day AmNo speakers received an almost exclusively spoken input (see below).

From this and the previous sections, a historical development in two steps arises. First, the distribution of the exceptional adjectives seems to have become more frequent between 1942 and 1987-1992. Second, and likely as a consequence of this first change, the use of CD has changed in the current generation of speakers. They now produce modified definite phrases without a determiner in contexts where homeland speakers and previous generations would not omit the determiner. Furthermore, the observed change exemplifies how two categories have merged. While previous generations made a distinction between regular adjectives (with CD) and exceptional adjectives (without the determiner), present-day speakers omit the determiner with all adjectives.

#### 5.4 A Note on Individual Results

Present-day speakers of AmNo show considerable inter- and intra-speaker variation with respect to the use of CD, as is typical for heritage speaker populations (e.g., Benmamoun et al., 2013, p. 133; Polinsky, 2018, p. 17; and many others). This variation occurs on two levels: the proportion of phrases that contain CD, and the types of non-CD phrases that are used. With respect to the latter, Section 3.3 outlined that present-day speakers more frequently omit the determiner than the suffixed article. Since numbers of non-homeland-like modified definite phrases are low in the previous generations, it is hard to analyze them on the level of individual speakers. This section therefore briefly discusses the first type of variation, viz. variation in the proportion of phrase that contain CD.

In van Baal (2020), the proportion of phrases that is homeland-like differs considerably from speaker to speaker. The individual scores range from as high as 70% to as low as 0%. In other words, a few of the present-day speakers in this study display scores that are close to the overall frequency of CD in the previous generations (see Section 5.1). There are five speakers in van Baal (2020, p. 239) who never use CD during the elicitation tasks.<sup>21</sup>

Since modified definite phrases are quite infrequent in spontaneous language, the number of phrases per speaker in Anderssen et al. (2018) is typically quite low, making it difficult to draw conclusions.<sup>22</sup> It can be observed that six of the speakers only produce homeland-like phrases, while as many as 13 did not use CD in any of their phrases.

Corpus data from the previous generations of AmNo speakers are difficult to analyze on the level of the individual for the same reason. Each speaker produces just a handful of modified definite phrases, and the number of exceptions (which do not require CD) tends to be high, as discussed in Section 5.3 above. From the 31 speakers recorded in 1942, 11 only

<sup>21</sup> Of these five speakers, three speakers produce a few instances of adjective incorporation and therefore score slightly above 0% homeland-like. The remaining two speakers never use homeland-like modified definite phrases.

<sup>22</sup> Three of the speakers in Anderssen et al. (2018, pp. 760-761) did not produce any modified definite phrases, and two more speakers only produced a few phrases with an exceptional adjective but no phrases which require CD.

produce phrases with an exceptional adjective and no phrases that require CD. There are furthermore 11 speakers that only produce homeland-like phrases, seven speakers that produce one non-homeland-like phrase, and two speakers who produce several non-homeland-like phrases. Four out of five speakers recorded in 1987-1992 produce only homeland-like phrases, and the other speaker produced two non-homeland-like phrases. As noted in Section 5.2, some of these utterances may be considered homeland-like despite their lack of CD.

In sum, most speakers in the two previous generations produce only phrases that would be found in homeland Norwegian as well. Those who do not, each produce just a few divergent phrases. This is in line with the conclusion from Section 5.1 that CD is still used to a homeland-like extent in 1942 and 1987-1992, although the small proportion of phrases different from the baseline may indicate a change in progress.

## 6 Discussion

Thus far, we have seen that the change in CD took place in the current generation of AmNo speakers. They frequently omit the prenominal determiner in contexts where homeland Norwegian speakers and previous generations of heritage speakers would use CD. A question that arises is why the change happened at this point in time, and not, for example, a generation earlier, when the speakers recorded in 1987-1992 were also heritage speakers.<sup>23</sup> In this section, I discuss the factors that can explain the timing of the observed language change.

There are several differences between the generations recorded in 1942 and 1987-1992 on the one hand, and the generation recorded from 2010 onwards on the other. First, the current speakers of AmNo have a different sociolinguistic profile compared to previous generations. The previous generations seem to have been more balanced bilinguals. They grew up during a time when Norwegian was still frequently used in their communities and most local churches, and new immigrants from Norway (i.e., monolingual speakers) were still arriving (until the 1920s). Haugen (1953; 1989, p. 64) describes how the use of Norwegian declined in formal and church contexts during the first half of the 20<sup>th</sup> century. The generations of speakers recorded by Haugen (1942) and Hjelde (1987-1992) used Norwegian to a high extent throughout their lives, while the current speakers switched to using English relatively early in their lives. During their youth, the local churches also switched towards English as the language for their services. As adults, the current speakers no longer speak Norwegian on a day-to-day basis, and some had not spoken Norwegian for years when they participated in the recordings for CANS. In other words, previous generations of speakers were part of a pre-language-shift heritage language community, while present-day speakers grew up and lived their lives in a post-shift community.

The shift from Norwegian to English as the community language also includes a change with respect to literacy. Previous generations likely had higher levels of (Norwegian) literacy than today's speakers. At that time, American Norwegian newspapers such as the *Decorah Posten* were still frequently read.<sup>24</sup> The AmNo speakers recorded by Haugen and Hjelde were familiar with the standard written language, unlike the present-day speakers

---

<sup>23</sup> Most, but not all, speakers recorded in 1942 were also heritage speakers. Five speakers were first-generation immigrants (born in Norway between 1850 and 1892) who emigrated as adults, and are therefore not considered heritage speakers according to Rothman's (2009) aforementioned definition.

<sup>24</sup> There were three large and long-existing American Norwegian newspapers: *Minneapolis Tidende* (published until 1935), *Skandinaven* (published until 1941), and *Decorah Posten* (published until 1972), and several hundred short-lived newspapers have existed (Lovoll, 1999, p. 181). Although these newspapers were no longer published during Hjelde's recordings in 1987-1992, they were published and read during a large part of these speakers' lives, and Hjelde (2015, p. 290) notes that these speakers were familiar with the written language through church services and confirmation, as well as Norwegian "summer schools."

(Hjelde, 2015, p. 290). Currently, most speakers cannot read or write Norwegian. This loss of literacy in present-day speakers is illustrative for the change in the use of the language in the communities.

As a result of these community changes, present-day speakers likely received less input than previous generations of speakers. In addition, the input itself also seems to have changed. As discussed in Section 5.3, phrases with exceptional adjectives that are not preceded by a determiner are very frequent in the older recordings. Although we do not have data on the exact input to present-day speakers, it is likely that this input resembled the earlier recordings. In other words, this input was heterogenous and contained large amounts of exceptions, with the consequence that the present-day speakers likely heard only few phrases with CD. The speakers recorded by Haugen and Hjelde, on the other hand, probably had an input that contained more instances of CD.

Present-day speakers are different from the previous generations in several respects. The latter were part of a pre-language-shift community; they were balanced bilinguals who were often literate in Norwegian, and they likely received higher rates of CD in their input. The current speakers, on the other hand, live in a post-shift community. As a consequence, they are English-dominant bilinguals with fewer opportunities to hear and speak Norwegian during their lives, they have no or limited reading skills, and they received an input in which CD was (much) less common. The outcome of the changes in the community and the input is that present-day speakers no longer distinguish between contexts where CD is obligatory and contexts where it is not. Unlike previous generations, they frequently omit the prenominal determiner from phrases that require CD.

Some caveats are in order in the interpretation of the data presented here. First, CANS contains only conversation data and interviews (i.e., semi-spontaneous speech), while the present-day data in van Baal (2020) consist of elicited speech data. This could potentially influence the results. However, we have seen that the data from Anderssen et al. (2018), which are corpus data from the current speakers, show much less frequent use of CD than the previous generations in CANS. In other words, the data in the two present-day studies of AmNo show similar results despite their different methods, and are both unlike the data from the older recordings.

Another caveat is the locations where the recordings were made, or the locations of the AmNo communities. There is some overlap between the different periods of recordings; there are recordings from the neighboring villages Westby and Coon Valley (WI) from all three time points, for example. On the other hand, there are also differences, and we do not have recordings in all three periods for all the locations where AmNo speakers live(d). This could potentially influence the results, as there might be or have been differences between the different communities. However, the data from Westby and Coon Valley conform to the general pattern described in Section 5.1: CD is used to a large extent in the previous generations, while there is a large drop in present-day speakers. In Westby and Coon Valley, the proportions of CD (out of phrase that require it) are as follows: 61% in 1942, 57% in 1987-1992, and only 32% in present-day in Anderssen et al. (2018).

It is worth noting that previous research has found similarities across the different locations where AmNo is spoken (Hjelde, 2015; Johannessen & Laake, 2017). There is thus no principled reason to believe that there are large differences between the communities with respect to CD. Hjelde (2015) argues there has been a process of dialect levelling (or koinéization). People from different dialectal backgrounds migrated to the United States, and the data collected at different time points also include people from various dialectal backgrounds. Given the findings in Hjelde (2015) and Johannessen & Laake (2017), and the fact that many speakers of Eastern Norwegian dialects are included in all samples, it is unlikely that dialect backgrounds have influenced the results.



While these caveats should be kept in mind, the available data show that the use of modified definite phrases with a determiner has been declining during the history of AmNo, and that the language change of determiner omission has taken place in the current generation of speakers, i.e., the speakers who have been recorded since 2010.

## 7 Conclusion

This article has presented a historical study on American Norwegian (AmNo) on the phenomenon of compositional definiteness (CD). Data from previous generations of AmNo speakers that have recently been made available in the *Corpus of American Nordic Speech* were used to study the change in the frequency of use of CD. These data show that the change has taken place in the present-day generation of AmNo speakers, as the previous generations (recorded in 1942 and in 1987-1992) still used CD to a (nearly) homeland-like extent. The current speakers, on the other hand, typically omit the prenominal determiner and produce phrases like *hvite hesten* ‘the white horse’ where homeland speakers as well as the previous generations heritage speakers would say *den hvite hesten*.

Language change does not happen overnight or with sudden jumps, but rather proceeds along an S-curve trajectory (Chambers & Trudgill, 1998, p. 163; Labov, 2001, p. 450; Pintzuk, 2003, p. 512). Both previous generations included in this study are at a low point of the S-curve, where the change is only in its initial stages. A small amount of the modified definite phrases in these time periods does not contain CD, as a subset of the speakers occasionally omit the determiner, the suffix, or both. The current generation, on the other hand, is at a point much further along the S-curve, at a point where the language change is almost completed, and most phrases no longer contain CD. In this generation, much individual variation is found, and some of the present-day speakers show patterns similar to those in previous generations. Most speakers, however, frequently omit definite morphemes, particularly the prenominal determiner.

The described change can be analyzed as the merging of two categories. While previous generations of AmNo speakers only omit the determiner in front of a subset of adjectives (the “exceptional” adjectives), present-day speakers omit the determiner regardless of the adjective used. The distinction between regular and exceptional adjectives found in homeland Norwegian and the previous generations AmNo speakers no longer exists in current AmNo. A pattern of variation found in the homeland—omission of determiners with certain adjectives—is thus generalized in the present-day heritage variety.

Data from this paper are drawn from the *Corpus of American Nordic Speech* (Johannessen, 2015), which has recently been extended with recordings of previous generations AmNo speakers, thereby enabling historical studies as the one presented here. This paper exemplifies how historical data can provide insights on change in heritage languages.

I suggest that the historical development of CD in AmNo has proceeded in two steps. First, phrases with exceptional adjectives have become more frequent over time. This is a change in the distributional pattern, rather than the use of non-homeland-like modified definite phrases. Similar changes in distributional patterns have been observed for AmNo in the use of pre- and post-nominal possessives (Anderssen et al., 2018) and non-subject initial declaratives, which are the context for verb-second (V2) word order (Westergaard & Lohndal, 2019; Westergaard et al., 2021). As a second step in the historical development, modified definite phrases without the prenominal determiner became a part of AmNo. This is a change in the expression of modified definite structures, as the typical AmNo modified definite phrase is unlike homeland Norwegian. This second historical change has taken place

in the present generation, which is now different from both the previous generations of speakers and the homeland variety.

Future historical studies on change in heritage languages can provide more insights into the two-stage process of change identified in this study. It is possible that more changes are preceded by a change in the frequency distribution of a given structure. Several studies have observed certain linguistic structures to be less frequent in heritage languages, without necessarily being used in a way divergent from the homeland. For example, Anderssen et al. (2018) found post-nominal possessives in AmNo to be more frequent than in homeland Norwegian, and speakers of heritage Swedish and heritage Norwegian use fewer relative clauses than monolingual speakers of these languages (Karstadt, 2003; Taranrød, 2011). For American Russian, Polinsky (2018) found a sharp decrease in the use of the genitive of negation, hypothesizing that this phenomenon will eventually disappear from the heritage variety (pp. 34-35).

The change in frequency distribution coincides with a change in the sociolinguistic profile of the speakers. The heritage speakers recorded in 1942 and 1987-1992 were part of a community where the heritage language was still frequently used. The present-day speakers recorded since 2010, on the other hand, are part of a post-language-shift community. This results in less-balanced bilingualism and decreased levels of literacy. In other words, the observed language change is likely caused by a combination of sociolinguistic changes at the level of the individual and the heritage communities, and the (quality of the) input. These two developments are interrelated.

It is striking that a syntactically complex and infrequent phenomenon as CD has not changed earlier in AmNo. Possibly, several facts contributed to the “survival” of CD: AmNo was still actively used in the communities, and the speakers were relatively balanced bilinguals. This changed in the post-shift community in which the present-day speakers grew up. They had less input in Norwegian, an input that contained a great number of phrases without the determiner (i.e., the exceptions), and fewer opportunities to use the language. As a consequence, the language changed.

American Norwegian is not the only heritage language that is currently moribund, having undergone changes with respect to the use of the language in communities and by individuals. Many Germanic heritage languages in the United States are in the same situation, as are moribund indigenous languages. There is an increased possibility for variation in communities where the heritage language is no longer the main language (Polinsky, 2018, p. 335). The findings from the present study suggest that this sociolinguistic context might contribute to language changes in the present-day speakers in these populations as well.

Furthermore, the finding that the change in American Norwegian has happened in the present-day generation may have implications for future studies on heritage languages with a long migration history. The establishment of a baseline may be somewhat less problematic, given that change seems more likely in the present-day speakers than in previous generations. Future studies can establish whether other morphosyntactic changes are indeed typical for the current generation that is part of a post-shift community.

The present study has pointed out how pre- and post-shift communities of AmNo are different in terms of their sociolinguistic profile and their language. We can wonder if there is a true comparison for present-day speakers: the generation of their (grand-) parents was very different in terms of language dominance and the status of the heritage language as a community language. The discussion of the baseline typically centers around the structural, linguistic baseline: the input the speakers received (e.g., Benmamoun et al., 2013, p. 134). In addition to this “narrow” baseline of input quality, we could formulate a “broad” baseline that includes the quantity of input and the opportunity to speak and hear the heritage language across the lifespan. In this study, interrelated changes in the broad and narrow baseline cause

the observed change in CD in American Norwegian. This underlines the importance of the broad baseline in addition to the narrow baseline.

It is important to note that none of the findings in this study are highly surprising in terms of what is known about language change. Although the sociolinguistic context might be specific for heritage languages, the change itself is not. This is yet another example that heritage languages are natural human languages, and that heritage language studies inform our understanding of human languages in general, as argued by Benmamoun et al. (2013), D'Alessandro et al. (2021), Lohndal et al. (2019), Polinsky (2018), and Scontras et al. (2015), among others.

In summary, the present study has described the two-step language change of compositional definiteness in American heritage Norwegian and discussed the sociolinguistic and language internal factors contributing to the change. Further studies on the historical development of heritage languages can help provide the answers to large questions of which linguistic phenomena change, how they change, and why they change. The present study has emphasized the joint effect of changes in the input and of language use and maintenance in communities and individuals. Future studies might be able to disentangle the respective shares of each factor, and identify other factors in heritage language change.

### **Acknowledgements**

Many thanks to the *Heritage Language Journal* editors and two anonymous reviewers, the audience at WILA11, David Natvig, and Alexander Lykke. I would also like to thank the Text Laboratory at the University of Oslo for their work on the *Corpus of American Nordic Speech* that made this study possible.

### **Notes on Contributor**

*Yvonne van Baal* is a postdoctoral researcher at the Norwegian University of Science and Technology (NTNU). Her research focuses on the Scandinavian languages, and her interests include heritage language bilingualism, dialectal variation, and nominal morphosyntax. <https://orcid.org/0000-0002-1766-2911>

### **References**

- Anderssen, M. (2006). *The acquisition of compositional definiteness in Norwegian* [Unpublished doctoral dissertation]. University of Tromsø, Norway.
- Anderssen, M. (2012). A spanning approach to the acquisition of double definiteness in Norwegian. *IBERIA: An International Journal of Theoretical Linguistics*, 4(1), 1–34.
- Anderssen, M., & Bentzen, K. (2013). Cross-linguistic influence outside the syntax-pragmatics interface: A case study of the acquisition of definiteness. *Studia Linguistica*, 67(1), p. 82–100. <https://doi.org/10.1111/stul.12011>
- Anderssen, M., Lundquist, B., & Westergaard, M. (2018). Crosslinguistic similarities and differences in bilingual acquisition and attrition: Possessives and double definiteness in Norwegian heritage language. *Bilingualism: Language and Cognition*, 21(4), 748–764. <https://doi.org/10.1017/S1366728918000330>

- Andréasson, M., Adamsson Eryd, H., Larsson, I., & Tingsell, S. (2017, October 12-14). *Do you speak Swedish? American Swedish in the corpus of American Nordic speech* [Conference presentation]. 8th Workshop on Immigrant Languages in the Americas (WILA), Copenhagen, Denmark.
- Benmamoun, E., Montrul, S., & Polinsky, M. (2013). Heritage languages and their speakers: Opportunities and challenges for linguistics. *Theoretical Linguistics*, 39(3-4), 129–181. <https://doi.org/10.1515/tl-2013-0009>
- Bohnacker, U. (2003). Nominal phrases. In G. Josefsson, C. Platzack, & G. Håkansson (Eds.), *The acquisition of Swedish grammar* (pp. 195–260). John Benjamins Publishing.
- Chambers, J. K., & Trudgill, P. (1998). *Dialectology* (2nd ed.). Cambridge University Press.
- Dahl, Ö. (2015). *Grammaticalization in the North: Noun phrase morphosyntax in Scandinavian vernaculars*. Language Science Press.
- D'Alessandro, R., Natvig, D., & Putnam, M. T. (2021). Addressing challenges in formal research on moribund heritage languages: A path forward. *Frontiers in Psychology*, 12, Article 700126. <https://doi.org/10.3389/fpsyg.2021.700126>
- Faarlund, J. T., Lie, S., & Vannebo, K. I. (1997). *Norsk referansegrammatikk* [Norwegian reference grammar]. Columbia University Press.
- Flaten, N. (1900). Notes on the American-Norwegian with vocabulary. *Dialect notes*, 2, 115–126.
- Flom, G. T. (1900). English elements in Norse dialects of Utica, Wisconsin. *Dialect notes*, 2, 257–268.
- Flom, G. T. (1903). The gender of English loan-nouns in Norse dialects in America; A contribution to the study of the development of grammatical gender. *The Journal of English and Germanic Philology*, 5(1), 1–31.
- Flom, G. T. (1926). English loanwords in American Norwegian as spoken in the Koshkonong settlement, Wisconsin. *American Speech*, 1(10), 541–558.
- Guevara, E. (2010). NoWaC: A large web-based corpus for Norwegian. In A. Kilgarriff & D. Lin (Eds.), *Proceedings of the NAACL HLT 2010 6<sup>th</sup> Web as Corpus Workshop* (pp. 1–7). Association for Computational Linguistics.
- Haugen, E. (1941). Intonation patterns in American Norwegian. *Language*, 17(1), 40–48. <https://doi.org/10.2307/409459>
- Haugen, E. (1953). *The Norwegian language in America: A study in bilingual behavior*. Indiana University Press.

- Haugen, E. (1989). The rise and fall of Norwegian in America. In N. Dorian (Ed.), *Investigating obsolescence. Studies in language contraction and death* (pp. 61–73). Cambridge University Press.
- Haugen, E. (1992). A language survey that failed: Seip and Selmer. *American Speech*, 67(3), 330–336. <https://doi.org/10.2307/455572>
- Hjelde, A. (1992). *Trøndsk talemål i Amerika* [The Trønder variety of Norwegian in America]. Tapir Forlag.
- Hjelde, A. (1996). The gender of English nouns used in American Norwegian. In P. S. Ureland & I. Clarkson (Eds.), *Language Contact across the North Atlantic* (pp. 297–312). De Gruyter.
- Hjelde, A. (2015). Changes in a Norwegian dialect in America. In J. B. Johannessen & J. Salmons (Eds.), *Germanic heritage languages in North America: Acquisition, attrition and change* (pp. 283–298). John Benjamins Publishing.
- Johannessen, J. B. (2015). The Corpus of American Norwegian Speech (CANS). In B. Megyesi (Ed.), *Proceedings of the 20th Nordic Conference of Computational Linguistics (NODALIDA 2015)* (pp. 297–300). Linköping University Electronic Press.
- Johannessen, J. B., & Laake, S. (2017). Norwegian in the American Midwest: A common dialect? *Journal of Language Contact*, 10(1), 5–21. <https://doi.org/10.1163/19552629-01001002>
- Johannessen, J. B., & Larsson, I. (2015). Complexity matters: On gender agreement in Heritage Scandinavian. *Frontiers in Psychology*, 6, Article 1842. <https://doi.org/10.3389/fpsyg.2015.01842>
- Johannessen, J. B., Priestley, J., Hagen, K., Åfarli, T. A., & Vangsnes, Ø. A. (2009). The Nordic Dialect Corpus—an advanced research tool. *NEALT Proceedings Series*, 4, 73–80.
- Johannessen, J. B., & Salmons, J. (2012). Innledning [Introduction]. *Norsk lingvistisk tidsskrift*, 30(2), 139–148.
- Johannessen, J. B., & Salmons, J. (2015). The study of Germanic heritage languages in the Americas. In J. B. Johannessen & J. Salmons (Eds.), *Germanic heritage languages in North America: Acquisition, attrition and change* (pp. 1–17). John Benjamins Publishing.
- Julien, M. (2002). Determiners and word order in Scandinavian DPs. *Studia Linguistica*, 56(3), 246–314. <https://doi.org/10.1111/1467-9582.00094>
- Julien, M. (2005). *Nominal phrases from a Scandinavian perspective*. John Benjamins Publishing.
- Karstad, A. H. (2003). *Tracking American-Swedish English. A longitudinal study of linguistic variation and identity*. University of Uppsala.

- Kinn, K. 2021. Split possession and definiteness marking in American Norwegian. *Nordic Journal of Linguistics*, 44(2), 182–219. <https://doi.org/10.1017/S0332586521000135>
- Labov, W. (2001). *Principles of linguistic change: Social factors* (Vol. 2). Blackwell.
- Larsson, I., Tingsell, S., Andréasson, M., Lyngfeldt, B., & Nilsson, J. (2012). Amerikasvenskan förr och nu [American Swedish then and now]. *Norsk lingvistisk tidsskrift*, 30(2), 263–286.
- Lohndal, T., Rothman, J., Kupisch, T., & Westergaard, M. (2019). Heritage language acquisition: What it reveals and why it is important for formal linguistic theories. *Language and Linguistic Compass*, 13(12), 1–19. <https://doi.org/10.1111/lnc3.12357>
- Lohndal, T., & Westergaard, M. (2016). Grammatical gender in American Norwegian heritage language: Stability or attrition? *Frontiers in Psychology*, 7, Article 344. <https://doi.org/10.3389/fpsyg.2016.00344>
- Lovoll, O. S. (1999). *The promise of America: A history of the Norwegian-American people* (Rev. ed.). University of Minnesota Press.
- Montrul, S., & Sánchez-Walker, N. (2013). Differential object marking in child and adult Spanish heritage speakers. *Language Acquisition*, 20(2), 109–132. <https://doi.org/10.1080/10489223.2013.766741>
- Natvig, D., & van Baal, Y. (2020). Crossing borders to enhance our understanding of variation in heritage languages. *Oslo Studies in Language*, 11(2), 319–334. <https://doi.org/10.5617/osla.8506>
- Pascual y Cabo, D., & Rothman, J. (2012). The (il)logical problem of heritage speaker bilingualism and incomplete acquisition. *Applied Linguistics*, 33(4), 450–455. <https://doi.org/10.1093/applin/ams037>
- Pintzuk, S. (2003). Variationist approaches to syntactic change. In B. D. Joseph & R. D. Janda (Eds.), *The handbook of historical linguistics* (pp. 509–528). Blackwell.
- Pires, A., & Rothman, J. (2009). Disentangling sources of incomplete acquisition: An explanation for competence divergence across heritage grammars. *International Journal of Bilingualism*, 13(2), 211–238. <https://doi.org/10.1177/1367006909339806>
- Polinsky, M. (2018). *Heritage languages and their speakers*. Cambridge University Press.
- Polinsky, M., & Kagan, O. (2007). Heritage languages: In the “wild” and in the classroom. *Language and Linguistic Compass*, 1(5), 368–395. <https://doi.org/10.1111/j.1749-818X.2007.00022.x>
- Putnam, M. T., Kupisch, T., & Pascual y Cabo, D. (2018). Different situations, similar outcomes. Heritage grammars across the lifespan. In D. Miller, F. Bayram, J.

- Rothman, & L. Serratrice (Eds.), *Bilingual cognition and language. The state of the science across its subfields* (pp. 251–279). John Benjamins Publishing.
- Rothman, J. (2009). Understanding the nature and outcomes of early bilingualism: Romance languages as heritage languages. *International Journal of Bilingualism*, 13(2), 155–163. <https://doi.org/10.1177%2F1367006909339814>
- Rødvang, L. I. S. (2017). *Empirical investigations of grammatical gender in American heritage Norwegian* [Unpublished master's thesis]. University of Oslo, Norway.
- Scontras, G., Fuchs, Z., & Polinsky, M. (2015). Heritage language and linguistic theory. *Frontiers in Psychology*, 6, Article 1545. <https://doi.org/10.3389/fpsyg.2015.01545>
- Taranrød, B. (2011). *Leddstillingen i relativsetninger i amerikansk* [Word order in relative clauses in American Norwegian] [Unpublished master's thesis]. University of Oslo, Norway.
- van Baal, Y. (2018). Compositional definiteness in heritage Norwegian: Production studied in a translation experiment. In K. Kühn & J. Heegård (Eds.), *Selected proceedings of the 8<sup>th</sup> Workshop on Immigrant Languages in the Americas, Copenhagen 2017*, 9–17. Cascadilla Proceedings Project.
- van Baal, Y. (2019, October 10-12). *The proficiency of American heritage Norwegian speakers* [Conference presentation]. 10<sup>th</sup> Workshop on Immigrant Languages in the Americas (WILA), Halden, Norway.
- van Baal, Y. (2020). *Compositional definiteness in American heritage Norwegian* [Unpublished doctoral dissertation]. University of Oslo, Norway.
- Westergaard, M., & Lohndal, T. (2019). Verb second word order in Norwegian heritage language: Syntax and pragmatics. In D. W. Lightfoot & J. Havenhill (Eds.), *Variable properties in language: Their nature and acquisition* (pp. 91–102). Georgetown University Press.
- Westergaard, M., Lohndal, T., & Lundquist, B. (2021). Variable V2 in Norwegian heritage language: An effect of crosslinguistic influence? *Linguistic Approaches to Bilingualism*. Advance online publication. <https://doi.org/10.1075/lab.20076.wes>
- Yager, L., Hellmond, N., Joo, H., Putnam, M. T., Rossi, E., Stafford, C., & Salmons, J. (2015). New structural patterns in moribund grammar: Case marking in heritage German. *Frontiers in Psychology*, 6, Article 1716. <https://doi.org/10.3389/fpsyg.2015.01716>

## FIGURES

FIGURE 1. Map of the locations in Minnesota and Wisconsin where recordings of American Norwegian were made in 1942 and 1987-1992.

FIGURE 2. The ratio of compositional definiteness in modified definite phrases in American Norwegian at different points in time: 1942 (Haugen), 1987-1992 (Hjelde), present-day in Anderssen et al. (ALW 2018), and present-day in van Baal (2020).