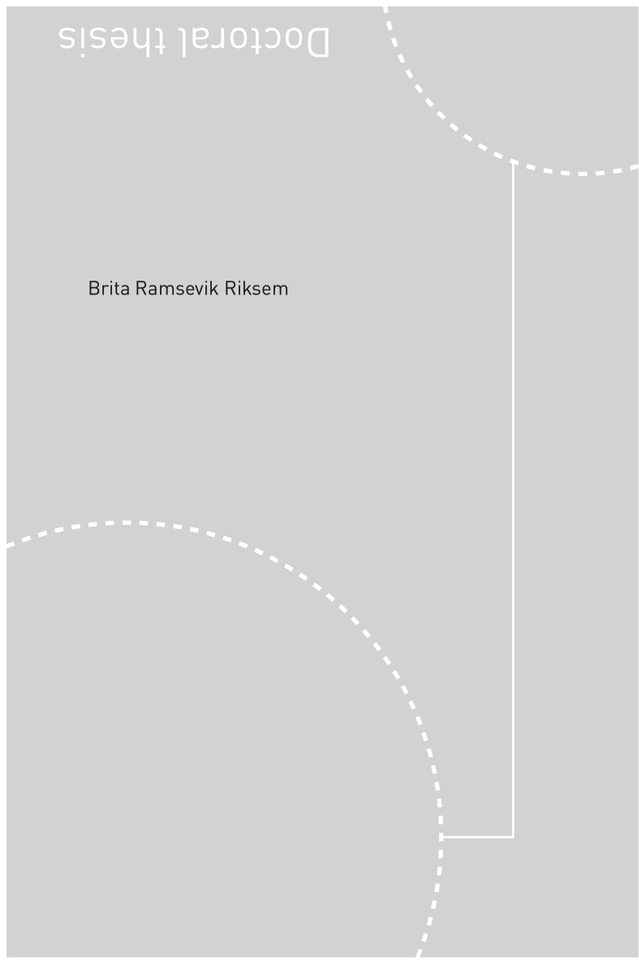
Doctoral theses at NTNU, 2018:7

Brita Ramsevik Riksem

Language mixing in American Norwegian noun phrases

An exoskeletal analysis of synchronic and diachronic patterns



ISBN 978-82-326-2826-1 (printed ver.)
ISBN 978-82-326-2827-8 (electronic ver.)
ISSN 1503-8181

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of Science and Technology
Thesis for the Degree of
Philosophiae Doctor
Faculty of Humanities
Department of Language and Literature

Doctoral theses at NTNU, 2018:7

NTNU
Norwegian University of
Science and Technology

Brita Ramsevik Riksem

Language mixing in American Norwegian noun phrases

An exoskeletal analysis of synchronic and diachronic patterns

Thesis for the Degree of Philosophiae Doctor

Trondheim, February 2018

Norwegian University of Science and Technology
Faculty of Humanities
Department of Language and Literature

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of Science and Technology

Thesis for the Degree of Philosophiae Doctor

Faculty of Humanities
Department of Language and Literature

ISBN 978-82-326-2826-1 (printed ver.)
ISBN 978-82-326-2827-8 (electronic ver.)
ISSN 1503-8181

Doctoral theses at NTNU, 2018:7

Printed by NTNU Grafisk senter

Abstract

The overall objective of the current dissertation is to investigate the morphosyntax of noun phrases in the heritage language American Norwegian (AmNo) that show mixing between English and Norwegian. AmNo was spoken by Norwegian immigrants to America in the years roughly from 1850 to 1920, and has been maintained by some of their descendants. Frequent usage of English items is characteristic of the language. Here I examine AmNo from both a synchronic and diachronic perspective, uncovering a generally consistent and systematic pattern of language mixing. A late-insertion exoskeletal model is employed to provide formal analyses of the observed patterns in the empirical material.

The dissertation contains three articles addressing the issue of language mixing in AmNo noun phrases from different perspectives. The first article presents empirical evidence that favors an exoskeletal model in analyzing language mixing, as compared to a mainstream lexicalist model. The exoskeletal model crucially separates syntactic structures from their phonological realizations. Moreover, realization of functional features is restricted by feature matching, whereas insertion of non-functional terminals is less restrictive. The second and third articles employ such a model in analyzing synchronic and diachronic patterns of language mixing in AmNo noun phrases, respectively. The former exploits data in the recently collected *Corpus of American Norwegian Speech* and finds a distinct pattern: English noun stems are incorporated into Norwegian structures and provided with Norwegian functional suffixes and determiners. The pattern is successfully analyzed in the late-insertion exoskeletal model. The third article conducts a diachronic investigation of the language mixing pattern in AmNo noun phrases by comparing the recently collected data with material collected by Einar Haugen in the 1930s and 1940s. The categories of number and definiteness are studied in detail and systematic changes are found: Norwegian functional suffixes are occasionally omitted or replaced by English alternatives. This is attributed to structural reanalysis of the AmNo grammar.

In combination, a thorough investigation of mixed AmNo noun phrases is provided, as well as a discussion of the mechanisms of language mixing in general. A late-insertion exoskeletal model is arguably well suited to account for language mixing without exploiting theoretical mechanisms other than those required for the analysis of monolingual speech. Thus, this model can be seen as a null theory of language mixing.

Acknowledgements

For the past three years I have had language mixing in American Norwegian on the brain. Investigating this phenomenon has meant diving into a pool of peculiar and enchanting data, and by listening to recordings and reading transcriptions, I feel that I have *watcha*[1] the speakers and gotten to know them, their *familyen*[2], their *kider*[3], and even their *barnebarnebarnebarn*[4]. I have taken a peek into their everyday routines of *mowe lawnen*[5] and having *disse lutefisk dinners*[6]. And this has been surprisingly exciting! Wrapping it all up does indicate the end of a journey, and I am both sad and relieved to say that *så it were ferdig*[7]. In retrospect, the data have at times been close to "a horrible jargon of mutilated English words",[8] but *det ble easyere om ei stund*[9], and there was definitely *stuff å ta care av*[10]. From my current perspective, I feel lucky for this opportunity, and grateful to the people who supported me during my exploration of "the gloriously mixed speech [that is] the Norwegian-American language".[11]

First and foremost, I would like to thank my two eminent supervisors, Professor Tor Anders Åfarli and Professor Terje Lohndal. Your linguistic expertise is (apparently) unlimited, and you constitute an inexhaustible source of inspiration and strangely accurate illustrations. You have taken great interest in my project and always been available for questions and discussion, for which I am grateful. Thank you also for insightful and extremely detailed comments on my work from the very first sketches to the final versions.

Tor and Terje put together the research group EXOGRAM (Exoskeletal Approaches to Grammar), where we are joined by my fellow PhD candidates, Maren Berg Grimstad and Ragnhild Eik. Our weekly discussions have been the incubator for most ideas pursued in the current project, and I thank you all for valuable comments, questions, and suggestions.

---

[1] watch-PST. All examples in italics are authentic and drawn from the *Corpus of American Norwegian Speech* (Johannessen, 2015a)
[2] family-DF
[3] kid-PL
[4] Literally *childchildchildchild*, and quite an innovative development of the Norwegian word for grandchildren *barnebarn.* However, in Norway we do have a word for this specific generation, *tippoldebarn* 'great-great-grandchild'.
[5] mow[e]-PRS lawn-DF
[6] these lutefisk dinners
[7] then it was finished
[8] Flaten, 1900: 119
[9] it got easy-COMP in a while
[10] stuff to take care of
[11] Haugen, 1992: 335

Furthermore, I would like to thank Professor Marit Julien for a warm welcome and many fruitful discussions during my time at Lund University from November 2016 to February 2017, and for acting as opponent in my mock defense in spring 2017. I have without a doubt benefited greatly from this experience. Moreover, I thank Christiane Müller and Anders Agebjörn for making my time in Lund about more than writing and knitting.

Most working days have, nevertheless, been spent at Dragvoll Campus, where the Department of Language and Literature has been a great place to work, and a fantastic collection of colleagues ensured a good social environment. Special thanks to Tina Louise Ringstad, Maria Boer Johannessen, Anne Mette Sunde, Agata Bochyńska, and Ragnhild Eik—the amazing NoSLiP 2016 crew—and Johanne Kristiansen for your advice and encouragement, not to mention all the nerdy (and not so nerdy) lunch breaks!

I would also like to thank Bridget Samuels for her excellent work proofreading the bulk of the dissertation. Still, if any mistakes should remain, I have no one to blame other than the person who put them there in the first place – myself.

My friends and family also deserve a warm thank you. You may not understand the intricate details of an exoskeletal analysis of grammar (or my fascination with it, for that matter), but you have understood my need for care and encouragement along the way. Distractions in the form of mountain hikes, board games, and sushi have been very much appreciated.

A final word of gratitude is reserved for my dear Lars – your patience, support, and, even though I hate to admit it, your dry sense of humor have surely contributed to the actual completion of this project, and made the journey a whole lot better.

Trondheim, August 2017
Brita Ramsevik Riksem

List of articles

1. Grimstad, Maren B.; Riksem, Brita R.; Lohndal, Terje, & Åfarli, Tor A. (in press). Lexicalist vs. exoskeletal approaches to language mixing. *The Linguistic Review.*

2. Riksem, Brita R. (in press). Language Mixing in American Norwegian Noun Phrases. *Journal of Language Contact, 12* (1).

3. Riksem, Brita R. (2017). Language Mixing and Diachronic Change: American Norwegian Noun Phrases Then and Now. *Languages,* 2 (2). doi: 10.3390/languages2020003

Note on Article 1

Article 1 in the current dissertation, entitled ”Lexicalist vs. exoskeletal approaches to language mixing”, is joint work with Maren B. Grimstad, Terje Lohndal, and Tor A. Åfarli. The article scrutinizes a proposed lexicalist analysis of language mixing between Spanish and English in DPs and shows how an alternative exoskeletal model fares better in analyzing these data, as well as DPs which show mixing between Norwegian and English. As the main component of the article deals with DPs, I have been particularly involved in the collection and analyses of data. However, I have also taken an active part, together with my co-authors, in developing the particular version of the exoskeletal model that is proposed. Moreover, in the process of writing, I have contributed especially to Sections 5 and 6, as well as the conclusion. I have also contributed to revisions after peer reviews.

List of abbreviations

AmNo	American Norwegian
AP	Adjective Phrase
αP	α Phrase
C_{HL}	Computational system for human language
CANS	Corpus of American Norwegian Speech
CardP	Cardinal Phrase
CASE	Case
COMP	Comparative
CS	Codeswitching
DEF	Definiteness
DF	Definite
DM	Distributed Morphology
DP	Determiner Phrase
DxP	Deixis Phrase
EL	Embedded Language
EPP	Extended Projection Principle
F	Feminine
FP	Functional Phrase
GEN	Gender
GenP	Gender Phrase
GFSH	Grammatical Features Spell-out Hypothesis
HL	Heritage Language
HS	Heritage Speakers
INDF	Indefinite
KP	Functional phrase associated with uniqueness
L1	First language
L2	Second language
LF	Logical Form
M	Masculine
ML	Matrix Language
MLF	Matrix Language Frame

MP	Minimalist Program
MSIH	Missing Surface Inflection Hypothesis
N	Neuter
NOM	Nominative
***n*P**	NP-related functional Phrase
NP	Noun Phrase
NUM	Number
NumP	Number Phrase
PERS	Person
PF	Phonetic Form
PL	Plural
PRS	Present tense
PST	Past tense
REFL	Reflexive
SG	Singular
TP	Tense Phrase
UG	Universal Grammar
VoiceP	Voice Phrase
vP	Verbalizer Phrase
VP	Verb Phrase
W	Weak adjectival inflection
WQP	Weak Quantifier Phrase

Table of contents

Abstract ..... **i**
Acknowledgements ..... **ii**
List of articles ..... **iv**
Note on Article 1 ..... **v**
List of abbreviations ..... **vi**
Table of contents ..... **viii**

Cover article

1 Introduction **1**
1.1. Empirical and theoretical points of departure 2
1.2. Objectives and research questions 3
1.3. Outline 4

2 Summary of the articles **6**
2.1. Article 1 6
2.2. Article 2 7
2.3. Article 3 8

3 Heritage languages and American Norwegian **10**
3.1. Heritage languages 10
3.2. American Norwegian 13
3.2.1. Historical background 14
3.2.2. Previous studies of American Norwegian 16

4 Theoretical foundation **20**
4.1. The structure of grammar 20
4.1.1. The Minimalist Program 21
4.1.2. The lexicalist approach to grammar 22
4.1.3. Critique of the lexicalist model 24
4.1.4. Exoskeletal approaches to grammar 26
4.1.5. Exoskeletal accounts for argument structure and categorization 29
4.1.6. A late-insertion exoskeletal model 30
4.1.7. Interpretation 39
4.1.8. Summary 41
4.2. Language mixing 42
4.2.1. Codeswitching and borrowing 44
4.2.2. Early approaches 47
4.2.3. Matrix Language Frame Model 48
4.2.4. A null theory or constraint-free approach to language mixing 49
4.3. Norwegian noun phrase structure 52

4.3.1. Early analyses of Scandinavian DPs 54
4.3.2. Julien (2005) 57
4.3.3. Norwegian versus English DP structure 60
4.3.4. Danish DPs and lack of double definiteness 62

5 Data and methodology 64
5.1. General considerations 64
5.2. Sources of AmNo data 66
5.2.1. Haugen (1953) 68
5.2.2. Corpus of American Norwegian Speech 70
5.3. Some methodological remarks 73
5.3.1. The observer's paradox 73
5.3.2. Composition of the groups of speakers 74
5.3.3. Tagging, transcribing, and transliterating in CANS 76
5.3.4. Metadata 77

6 Main findings and implications 79
6.1. Main findings 79
6.1.1. Empirical patterns of language mixing in American Norwegian 80
6.1.2. The late-insertion exoskeletal model as a null theory model of language mixing 83
6.2. Some implications and points of contention 85
6.2.1. Is it really mixing? 86
6.2.2. Analyses of grammatical gender 89
6.2.3. Gender assignment through translation? 93
6.2.4. FP 95
6.2.5. The road ahead 97

7 Final remarks 99

References 100

Article 1: Lexicalist vs. exoskeletal approaches to language mixing
Article 2: Language Mixing in American Norwegian Noun Phrases
Article 3: Language Mixing and Diachronic Change: American Norwegian Noun Phrases Then and Now

Cover Article

1 Introduction

What can language mixing in American Norwegian (AmNo) noun phrases tell us about the principles of language mixing, and subsequently about the structure of grammar in general? This has been the driving question for the current dissertation. AmNo is a Norwegian heritage language spoken by Norwegians who immigrated to and settled in America roughly from the mid-1800s and until the 1920s. Today, some of their descendants still speak the language. An apparent trait of AmNo is the frequent usage of English items. This is the phenomenon I will refer to as language mixing, which can be understood as utterances containing (substantial or functional) elements from two or more languages. Throughout the dissertation, I argue that AmNo noun phrases constitute a fruitful empirical domain for investigating language mixing, and that the insights gained here are conducive to establishing a model of grammar capable of analyzing these patterns.

The goal of the current dissertation is twofold. First, it aims to describe language mixing outcomes observed in AmNo noun phrases, and to provide formal analyses of these. A so-called late-insertion exoskeletal model is employed in these analyses, and the second goal of the dissertation is thus to demonstrate the suitability of such a model for analyzing language mixing. The dissertation contains three articles which approach these goals through theoretical investigations as well as by conducting both synchronic and diachronic empirical investigations. Together the articles shed light on the topics of interest: the field of language mixing in general, and AmNo in particular. Specifically, new insights into AmNo and the way that these speakers use English items in their Norwegian utterances are provided. Here a clear pattern is revealed: English items are typically incorporated into Norwegian grammatical structures. Moreover, diachronic investigations show that this pattern is consistent, though not immutable, and systematic changes are also discussed. Taking a broader perspective, the empirical material from AmNo shows that language mixing by and large follows a predictable pattern; content items from a secondary language are incorporated into the structure of a main language where they are assigned functional properties. These patterns are analyzed in a specific exoskeletal approach to grammar, a framework that separates syntactic structures and their phonological exponents. The latter are inserted “late” through a process which is crucially less restrictive for non-functional components than for functional ones. The results of the dissertation support an exoskeletal model as an excellent analytical tool for language mixing.

The current cover article provides a general review of the relevant literature and methodological considerations which form the base for the investigations conducted in all three articles. Moreover, overall findings and proposals are presented and discussed, combining the results from the individual studies.

1.1. Empirical and theoretical points of departure

AmNo can be characterized as a heritage language due to being situated in a community where a different language, in this case English, is the dominant language. Research on heritage languages in general is a flourishing field, especially in the American context, and prominent works investigating such languages are, among others, Polinsky (2006, 2011, 2016), Rothman (2007, 2009), Montrul (2008, 2016), and Benmamoun, Montrul, and Polinsky (2013). A primary objective of these investigations is to document the properties of various heritage languages and investigate potential linguistic differences as compared to the corresponding baseline variety (often a non-heritage variety). Moreover, studying the characteristics and competence of heritage speakers is beneficial in investigating linguistic competence in general and language change in a minority context.

AmNo is particularly interesting as a heritage language since data have been collected and studies conducted at different points of time since the large wave of immigration from Norway to America took place, primarily in the 19th century. This enables studies across a long time span. A significant door-opener for new investigations of AmNo is the recently established online *Corpus of American Norwegian Speech* (henceforth CANS) (Johannessen, 2015a). Studies of AmNo have flourished in the aftermath of establishing this corpus. Most contributions have focused on the Norwegian properties of the language. However, the influence from the dominant language, English, is clearly visible in AmNo through, among other things, frequent usage of English items. This language mixing has also been investigated to some extent by other scholars, mostly focusing on the verbal domain (see more in Section 3.2.2).

The current dissertation adds to the previous studies by investigating the patterns of language mixing in noun phrases.[1] The nominal domain proves to be an excellent area for

[1] Language mixing in the nominal domain has been investigated in various previous studies, e.g., Fuller & Lehnert (2000), Jake, Myers-Scotton, & Gross (2002), Cantone & Müller (2008), Cantone & MacSwan (2009), Herring, Deuchar, Parafita Couto, & Moro (2010), Carter, Deauchard, Davies, & Parafita Couto (2011), Pierantozzi (2012), Moro (2014), and Parafita Couto, Munarriz, Epelde, Deuchard & Oyharçabal (2015). These studies investigate different language pairs than the present dissertation, and they typically adopt a different theoretical framework. Space limitations prevent a comparison of the frameworks.

investigating mixing between Norwegian and English as the noun phrases of these two languages show certain structural differences. In addition to the synchronic investigation of language mixing based on CANS, the third article in this collection is the first to systematically investigate and analyze diachronic changes (in language mixing) in AmNo by comparing the recent data to previously collected data (Haugen, 1953). The overall results are significant both for expanding our knowledge of AmNo and the properties of language mixing.

Theoretically, a key objective for formal linguistics is to investigate the nature of grammatical representations. In this respect, studying language mixing is advantageous, as it is an expression of an individual's language competence, and therefore falls within the range of possible languages (González-Vilbazo et al., 2013). Furthermore, combinations of different languages might provide a window into the basic structure and mechanisms of grammar by unveiling characteristics that are difficult to access or observe in monolingual data. Treating language mixing and monolingual speech alike is an important goal for current language mixing research, known as a "null theory" or "constraint-free approach" to language mixing (e.g., MacSwan, 1999, 2014). This stands in opposition to several previously proposed restrictions on language mixing, which in one way or another presume that the language faculty is able to distinguish between languages and mix them using special mechanisms.

The analyses conducted in the current dissertation employ a late-insertion exoskeletal model for analyzing language mixing in AmNo and argue that this model is capable of analyzing the observed patterns without resorting to special mechanisms. This model is primarily motivated by monolingual data, and its success in analyzing language mixing data provides evidence in favor of such a model as a null theory of language mixing.

1.2. Objectives and research questions

Given these points of departure, the hypotheses and research questions of the dissertation have both an empirical and a theoretical angle. The aims of this work involve mapping systematic patterns in the empirical data both synchronically and diachronically, and in addition providing a formal analysis of these patterns by way of a late-insertion exoskeletal model. Thus, the core hypothesis of the dissertation is the following.

(1) *Mixing of English and Norwegian in American Norwegian noun phrases is systematic, and a late-insertion exoskeletal model is well suited to capture the empirical patterns.*

Based on this core hypothesis, specific research questions can be formulated. For the purpose of the current investigations, these may be separated into two groups. The first group, in (2), concerns the empirical, synchronic and diachronic patterns of language mixing in AmNo noun phrases, whereas the second group, in (3), covers theoretical aspects relating to the model of grammar.

(2) a. Does language mixing in American Norwegian noun phrases follow systematic patterns?
    b. Have the patterns of language mixing in American Norwegian noun phrases changed diachronically?

(3) a. How can the typical patterns of language mixing be formally analyzed in a late-insertion exoskeletal model?
    b. How can a late-insertion exoskeletal model account for the patterns of diachronic change?
    c. How do the analyses of language mixing in American Norwegian noun phrases support a late-insertion exoskeletal model as a null theory of language mixing?

Article 1 in this dissertation addresses in particular the general theoretical question in (3c), whereas Articles 2 and 3 are concerned with the questions of typical mixing patterns (questions 2a and 3a) and diachronic change (questions 2b and 3b), respectively. The overall objective is to provide a thorough account of language mixing and AmNo noun phrases.

1.3. Outline

This cover article is structured as follows. Section 2 provides a summary of the three articles that form the core of the dissertation, reporting the findings of the individual studies. Section 3 introduces the language under investigation, AmNo, with respect to its historical background and documentation through previous research. In addition, Section 3 places AmNo in the broader context of heritage languages in America. Section 4 provides a discussion of the theoretical foundation of the dissertation. This section is divided into three main subsections: Section 4.1 addresses the question of the structure of grammar and discusses the exoskeletal approach, as understood and employed in the current dissertation. Language mixing is the topic of Section 4.2, whereas Section 4.3 introduces Norwegian DP

structure. Section 5 offers an introduction to the corpora from which I obtained data, as well as a discussion of the methodology of the dissertation. Section 6 provides a discussion of the main findings and implications, and Section 7 contains some final remarks. Then the three articles follow in the same order as they are summarized in Section 2.

2 Summary of the articles

2.1. Article 1

The title of the first article is "Lexicalist vs. exoskeletal approaches to language mixing" and is the product of joint work with Maren Berg Grimstad, Terje Lohndal, and Tor Anders Åfarli. The article has been accepted for publication in *The Linguistic Review.*

This article is a critical review of a lexicalist feature-driven analysis of language mixing, an approach promoted especially by Jeff MacSwan (1999, 2000, 2005a, b, 2009, 2014). By scrutinizing one paper that builds specifically on MacSwan's proposals, namely Moro (2014), we uncover and discuss the shortcomings of such a framework as an analytical tool for language mixing. Moro studies a group of bilingual speakers from Gibraltar. We investigate the data provided by Moro (2014), and argue that both the presentation of the data and the analyses of them are inadequate. Studying English–Spanish mixing within a DP, Moro claims that the combination of a Spanish determiner and an English noun is well-formed, whereas the opposite, an English determiner with a Spanish noun, is ill-formed. However, we show that under standard lexicalist minimalist feature checking procedures, this is not the predicted pattern. Concerning the former pattern, English nouns would not be able to provide a gender feature that could check the unvalued gender feature of a Spanish determiner; hence, this pattern should be illicit. The allegedly ill-formed pattern, however, would not leave unchecked features behind, and should therefore converge. For these reasons, in order to obtain the attested pattern, Moro (2014) utilizes special principles for feature checking, which we argue are implausible.

We propose that an exoskeletal analysis is more promising for capturing this type of data. Exoskeletal approaches to grammar assume grammatical features to be part of the syntactic structure. The specific implementation adopted here moreover assumes that content items emerge from combining a neutral root with a categorizer, yielding a stem. Importantly, the root/stem does not provide functional features to syntax. Employing the exoskeletal approach, and the Subset Principle as a regulator of functional feature realization, we successfully analyze English–Norwegian mixing within a DP in the heritage language American Norwegian. First, we analyze the pattern where a Norwegian determiner occurs together with an English noun stem: The functional features in a Norwegian structure are best realized by Norwegian exponents, due to feature matching requirements, whereas an English stem easily is inserted into an available, less restricted position. The opposite pattern, an

English determiner with a Norwegian noun, is analyzed as a Norwegian stem inserted into an English structure. This is further supported by evidence from the verbal domain, supporting exoskeletal analyses of language mixing in general.

2.2. Article 2

Article 2 is entitled "Language Mixing in American Norwegian Noun Phrases" and is in press with *Journal of Language Contact.* This article provides a synchronic investigation of the morphosyntax of noun phrases in contemporary American Norwegian that show mixing between English and Norwegian. Data are extracted from the *Corpus of American Norwegian Speech* (Johannessen, 2015a), and the goals of the article are first to provide a detailed description of the observed mixing patterns, and secondly to show how a late-insertion exoskeletal model can be used in the analyses.

The results show that language mixing in American Norwegian noun phrases typically takes the form of an English noun stem being inserted into a Norwegian structure, and thereby being provided with functional suffixes and associated functional words. I propose that an exoskeletal model is able to capture the observed empirical patterns without adopting special mechanisms for language mixing, thus being a null theory of language mixing. The crucial factor in exoskeletal analyses is, as also addressed in Article 1, the separation between an abstract syntactic structure and its phonological realizations (exponents), and moreover, that the phonological realizations of functional and non-functional components are governed by different restrictions: Functional exponents are restricted by feature matching, but insertion of substantial exponents is less restricted. As it is a Norwegian heritage language, American Norwegian nominal structures are expected to contain the features definiteness, number, and gender. Consequently, functional components are typically realized by Norwegian exponents due to feature matching requirements, whereas English stems are inserted into the less restricted positions. Considering that the model is well supported by monolingual data, the objective of a null theory is maintained.

In addition to the most typical mixing patterns, the article further discusses a pattern that is unexpected from an exoskeletal perspective, namely the occurrence of an English functional suffix, the plural *-s*. The article proposes an analysis through which this

phenomenon may also be accounted for, where the plural *-s* is considered a gender-neutral alternative.[2]

2.3. Article 3

The third article is entitled "Language mixing and diachronic change: American Norwegian noun phrases then and now" and is published in *Languages* (Riksem, 2017). This article offers a diachronic perspective on language mixing within noun phrases in American Norwegian by comparing the recently collected material in the *Corpus of American Norwegian Speech* (Johannessen, 2015a) to Haugen's (1953) collections from the 1930s and 1940s. The overall patterns of language mixing, as described in Article 2, appear to be stable, but upon investigating the data more closely, certain systematic differences emerge.

The article is focused on the three nominal categories of gender, number, and definiteness and how these are realized by functional suffixes or determiners and demonstratives in mixed noun phrases. Concerning gender, the article is primarily focused on the distribution across Norwegian's three genders, and finds no clear diachronic development. Thus, number and definiteness are the main topics for discussion in the article.

Number is realized as suffixes in both English and Norwegian, but due to being a Norwegian heritage language, American Norwegian plural noun phrases typically have a Norwegian suffix even when the noun stem is English. This is predicted by the exoskeletal model and is also the main pattern in Haugen's (1953) material. However, two patterns of change are identified for this category. First, the English plural *-s* is used in a majority of the relevant plural phrases in the new material, as compared to only in a subset in Haugen (1953). Moreover, some plural phrases surface without a suffix at all, which is a pattern not discussed by Haugen.

Definiteness is a category realized in different ways in Norwegian and English, and is therefore a good area for investigating language mixing. In the older material, definiteness is realized following a Norwegian pattern, i.e., with definiteness as a functional suffix and double definiteness in relevant cases. In the new material, two patterns of change are again identified. First, the Norwegian definite suffix is omitted in a number of cases, and in addition, the English determiner *the* is occasionally used in an otherwise Norwegian context.

[2] Not mentioned in the article is the fact that some Norwegian dialects also make use of a single exponent for plurality, parallel to this pattern. Whether the AmNo speakers and their ancestors have such realizations in their dialect background (to the extent that we can say something certain about this) has not been investigated in the current dissertation.

A similar pattern is also found for indefinite phrases where the indefinite article in some cases is either omitted or replaced by an English one.

The article continues discussing potential changes in the underlying grammar that could cause the observed patterns. From the exoskeletal perspective, two potential scenarios may cause a change. One possibility is expressed as the Missing Surface Inflection Hypothesis (Lardiere, 2000) suggesting that the structure is intact, but the exponent and its conditions for insertion are reduced. A second possibility is that changes have occurred in the actual structure, altering the composition of feature bundles, which subsequently led to changes in their realizations. Both alternatives are discussed and potential evidence in the data is provided. Somewhat tentatively, the article argues in favor of the latter alternative. Finally, the article briefly discusses the nature of the change and possible trajectories like incomplete acquisition, attrition, and cross-linguistic influence.

3 Heritage languages and American Norwegian

3.1. Heritage languages

The language under investigation in the current dissertation, AmNo, is described as a heritage language, and thus belongs to a group of languages which are currently the subject of extensive research.[3] According to Polinsky and Kagan (2007), both a broad and a narrow conception of a heritage language can be identified. The former expresses a link between cultural and linguistic heritage, such that speakers may use their family heritage as motivation for learning the language. Despite their initial motivation, these speakers are considered L2 learners. The latter conception, however, emphasizes the linguistic heritage alone and the manner and order of acquisition; the heritage language in the narrow sense should be the first language, acquired naturally. The current dissertation, like the literature it relies on, follows this narrow, linguistically focused, conception of heritage languages.

A fundamental definition of a heritage language and its speakers, as understood here, is provided by Rothman (2009: 156):[4]

> A language qualifies as a *heritage language* if it is a language spoken at home or otherwise readily available for young children, and crucially this language is not a dominant language of the larger (national) society. […] From a purely linguistic point of view, we assume that an individual qualifies as a heritage speaker, if and only if he or she has some command of the heritage language acquired naturalistically. (Rothman, 2009: 156)

Important points to take away from this definition are (i) that a heritage language is not the main language of the society where it is situated, and (ii) that it is naturally acquired. Heritage speakers acquire the heritage language at home, but either simultaneously or at the time of starting school acquire the dominant language of the society, which due to its status and usage in the society typically becomes the dominant language of the heritage speaker. This places heritage speakers in the interesting position of being dominant in a language that is not their

[3] E.g., Polinsky (2006, 2011, 2016), Polinsky & Kagan (2007) Rothman (2007, 2009), Montrul (2008, 2012, 2016), Pascual y. Cabo & Rothman (2012), Benmamoun et al. (2013), Putnam & Sánchez (2013), Rothman & Treffers-Daller (2014), Scontras, Fuchs, & Polinsky (2015), Yager et al., (2015), Kupisch & Rothman (2016), and Goldrick, Putnam, & Schwarz (2016a, b).

[4] Other definitions can be found in, e.g., Polinsky & Kagan (2007) or Montrul (2008). The core properties of a heritage language and its speakers are nevertheless recognizable: A heritage language is a minority language in a community where another language is clearly dominant. Heritage speakers acquire this language naturally, as their L1, but crucially also acquire the dominant language from birth or early childhood, thus becoming simultaneous or early sequential bilinguals.

first acquired language. Consequently, this raises the question of what a native language is, and whether or not heritage speakers can be considered native speakers of the heritage language. The fact of the matter is that many heritage speakers, when compared to non-heritage speakers of the language in question, display significant differences in competence and performance. Rothman and Treffers-Daller (2014) point out that the basic premise for a native language is that it is acquired naturally in early childhood, thereby arguing that heritage speakers are, in fact, native speakers. Heritage speakers thus typically fall between labels used to categorize language competence: The heritage language is their first-acquired and native language, but nevertheless their weaker language. Instead, a second language, acquired simultaneously or later, constitutes their dominant language. Their competence in the heritage language thus often differs from that of native, dominant speakers of the corresponding non-heritage variety.

Despite their diverging competence in their heritage language, Scontras et al. (2015) emphasize the potential in studying heritage speakers. They argue that heritage speakers constitute "a unique testbed for issues of acquisition, maintenance, and transfer within linguistic theory" as they do not follow the traditional trajectory for acquisition and typically exhibit a non-native-like competence of their first language as adults (Scontras et al., 2015: 3). Benmamoun et al. (2013) also address the potential that lies in investigating heritage speakers and how this may enrich the field of grammatical theory. For instance, it sheds light on the role of input, especially when the input conditions are reduced in a bilingual environment. Data from heritage languages could then help reveal how vulnerable or resilient different areas of grammar are in such a process. In addition, the heritage language's co-existence with a different, dominant language may yield interesting insights into the processes of language mixing and contact-induced change. In short, studying these speakers and their production provides new perspectives on language structure and what the scope of human linguistic competence is.

A basic definition of a heritage speaker is captured in the quote by Rothman (2009) above. However, determining who is a heritage speaker in practice may not always be as straight-forward. For instance, Scontras et al. (2015) present some hypothetical individuals, who they argue are all heritage speakers, but who nevertheless differ considerably in proficiency, production, and age of exposure to the dominant language in the society. Arguably, the group of heritage speakers shows great variation, and their proficiency may vary widely between individuals, from near native to merely receptive knowledge (Benmamoun et al. 2013; Rothman, 2009). In many cases, the competence of adult heritage

speakers may resemble that of L2 learners, despite being native speakers. Benmamoun et al. (2013) discuss different aspects of the grammatical system in order to map the competence of heritage speakers. They find that phonological competence is well-preserved and that heritage speakers typically outperform L2 learners in this area. Moreover, core structural properties appear consistent, whereas inflectional morphology is more vulnerable. Thus, overall, heritage speakers show good speaking and listening abilities, but at the same time differ significantly from what is typically expected of full attainment of the language in some grammatical aspects (Rothman, 2009; Montrul, 2012; Benmamoun et al., 2013).

The divergence in competence that heritage speakers show is not easily explained: "It should come as no surprise, then, that the proposed trajectories to the competence of heritage speakers are at least as complex as the speakers and abilities they are meant to characterize" (Scontras et al., 2015: 3). Three possible, and widely used, trajectories that may describe this diverging competence are *incomplete acquisition* (Polinsky, 2006; Montrul, 2008), *attrition* (Rothman, 2007; Polinsky, 2011) [5] and *cross-linguistic influence.* Incomplete acquisition is especially debated, due to the term *incomplete.* In short, this implies that the acquisition of the heritage language was interrupted by exposure to the dominant language and thus not properly completed. It has been argued that, due to reduced and possibly also divergent input, the acquisition is not incomplete, just different (Pascual y. Cabo & Rothman, 2012; Putnam & Sánchez, 2013). Actually determining the significance and progress of these trajectories is complicated, as it requires an extensive collection of data over the lifespan of the individual speaker, including the input the speaker had, as well as the speaker's competence in the dominant language. A complete set of such data is rarely available, which limits the possible conclusions.

A related, and quite crucial, question when discussing the reduced competence of heritage speakers concerns the baseline for comparison. What is a proper baseline language for evaluating heritage speakers' competence? Benmamoun et al. (2013: 134) put it aptly when they say that the baseline language "is defined as the language of input for heritage speakers". They elaborate on this definition in a related footnote, saying that "[c]rucially, the baseline language is not the monolingual variety of that language" (Benmamoun et al., 2013: 134, footnote 2). However, when a speaker's input is not available in the data, the researcher will in many cases resort to the non-heritage variety of the language in question for comparison. This issue is especially prominent in the case of AmNo. As I will introduce

[5] See Sorace (2011) for a different definition of attrition.

below, most of today's speakers of AmNo are 3rd and 4th generation, suggesting that the contact with Norway, and the Norwegian language as spoken there, is rather weak, if not entirely non-existent. These speakers have instead been exposed to Norwegian as spoken by other heritage speakers one or two generations prior. However, as today's speakers of AmNo are already adults, analyzing their actual childhood input is not possible. Fortunately for AmNo studies, data from the 1930s and 1940s (Haugen, 1953) are available, and may be used as a compensation.

One final question in defining heritage speakers, which is also important for the present work, is whether or not 1st generation immigrants should be considered heritage speakers. Benmamoun et al. (2013) argue that heritage speakers are children and later descendants of the original immigrants, thus not counting the first generation (see also Montrul, 2012). The rationale for this divide is that the original immigrants are dominant in the non-heritage variety of the language, which they have acquired under different circumstances. Possible differences in their competence are presumably rather caused by L1 attrition. On the other hand, first generation immigrants are also speakers of a minority language in their new society and may experience changes due to the contact with and dominance of the majority language (Pascual y. Cabo & Rothman, 2012; Åfarli, 2015a). In the current dissertation, the issue emerges when data from the 1930s and 1940s are addressed, as this group of speakers includes first generation immigrants. I return to the questions concerning comparisons of informants in Section 5.3.2.

3.2. American Norwegian

AmNo is a Norwegian heritage language spoken by immigrants who came from Norway to America between 1825 and the 1920s. Today some of their descendants still speak the language. Although heavily influenced by English during its time of existence, the language is still unmistakably Norwegian. However, the majority of current AmNo speakers are already in their 70s or 80s, and considering that English has for most been the dominant language from school age, it is not surprising that individual competence varies. Some speakers have a Norwegian spouse or close Norwegian friends, inviting a more frequent usage of the language. Others might not have spoken the language since their parents passed away several years prior. These differences notwithstanding, all speakers are relatively fluent (Johannessen & Salmons, 2012; Johannessen & Laake, 2017).

In this section, I will introduce the heritage language AmNo. First, I give a brief introduction to the historical and sociolinguistic background of these speakers and their communities, and thereafter I turn to discussing a selection of research on AmNo conducted since the early 1900s. In this latter subsection, I limit the discussion to only introducing a selection of relevant studies on formal grammar and/or language mixing. The available data and methodology are discussed further in Section 5.

3.2.1. Historical background

The first Norwegian migrants left for America in 1825, and over the century that followed, Norwegians emigrated in large numbers. In the 1920s, new restrictions on immigration in America slowed down the flow of migrants. By then, more than 800,000 Norwegians had emigrated, a number nearly equal to the population of Norway in 1800 (Haugen, 1953: 28–29). The Great Depression in the 1930s, as well as World War II, caused additional declines in emigration, and even though Norwegians continued to immigrate to America in later years, the quantity was never close to the exodus leading up to 1920 (Lovoll, 1999). The speakers investigated in the articles contained in the current dissertation are the descendants of immigrants who came to America prior to 1920.

The migrants who left Norway in 1825 were a group of religious dissenters, but for the majority of the following migrants, the hope of social betterment was the chief motive. What this hope actually implied could vary from person to person; some sought economic advantages, others religious or political freedom, and still others were motivated by pure adventurousness (Haugen, 1953: 18). “America letters” from the pioneer immigrants were circulated widely back in Norway, providing knowledge about the promise of the new land, and thereby encouraging later immigrants to take the leap (Haugen, 1953; Lovoll, 1999). The migrants came from all over the country, but the majority came from the fjord districts in the west of Norway and the mountain valleys in the east (Lovoll, 1999). Hence, several different dialects were represented in the group.

Upon arriving in America, most immigrants continued inland. They went primarily to the Midwest, where they formed large Norwegian settlements in the area stretching from Illinois to North Dakota. Immigrating to a new country does mean turning everything you know upside down. However, by keeping together with people with a common heritage, language, and culture, these settlements postponed the necessity of reshaping one’s life (Haugen, 1967). As the settlements took shape, important institutions like churches, newspapers, social communities, colleges, and congregational schools where established.

These promoted both social life and the preservation of the Norwegian language. In some settlements, if one did not engage in political life or go to the big cities, one could actually go through life as a Norwegian monolingual and still have one's needs for social and religious life covered (Haugen, 1953: 45). In general, however, learning the English language was both necessary and encouraged, and children went to American schools where they were taught in English (Lovoll, 1999: 98).

The Lutheran church was an important institution in the settlements, and the first assembly houses and congregations were founded in the 1840s. A challenge was, however, getting competent pastors, which was a key motivation for the establishment of Luther College in 1861 and St. Olav's College some years later (Haugen, 1953; Lovoll, 1999). Congregational, or religious, schools were also initiated. These institutions would provide religious training for children as well as edification for adults, and they were also crucial for preservation of the Norwegian language.

Two additional institutions that were important for the continued usage of Norwegian were newspapers and a variety of social groups that were established in the settlements. The newspapers served the important dual functions of bringing news from Norway and helping the immigrants integrate into American life. As many as four hundred newspapers appeared in the Norwegian settlements, although many of these were short-lived. In the Midwest, three leading newspapers had long-lasting and large circulations: *Skandinaven* published in the years between 1866 and 1941, *Decora-Posten* 1874–1972 and *Minneapolis Tidende* 1887–1935 (Lovoll, 1999: 181). These newspapers provided news in Norwegian well into the 20th century. Moreover, newspapers functioned as a medium through which the immigrants could keep in touch with each other, and were used to announce meetings of various cultural or social societies. Voluntary associations characterized the Norwegian society and were established both for special political or charitable purposes, or simply as a forum for socializing (Lovoll, 1999). A unique and important group was the *bygdelag,* a social organization bringing together immigrants originating from the same *bygd* (rural community) or district in Norway. In these societies, the immigrants were able to preserve the ties to their Norwegian heritage, as well as the usage of the Norwegian language. Not until the 1950s did they start a gradual transition to English (Lovoll, 1999).

With the decline of the immigration after 1920, however, a decline in the use of the Norwegian language ensued. The cause of the change was probably complex, involving among other factors the "Americanization" following World War I and "the hysterical opposition to everything foreign" (Haugen, 1967: 30). Using the main language of the society

was in any case considered a natural and inescapable part of the immigrants' integration, and gradually, newspapers ceased publication, and religious services and instructions as well as social communities switched to using English. Due to the difficulties they experienced and stigma in schools, many parents also refrained from passing the language on to the next generations. Nevertheless, Norwegians did not abandon their ethnic and cultural qualities, which are still evident in, e.g., museums, shops, festivals, handcrafting, and folk music (Lovoll, 1999; Johannessen & Salmons, 2012). The extensive change to English notwithstanding, a considerable number of speakers did preserve and continue using the language and passed it on to following generations. Thus, Norwegian-speaking descendants of the immigrants from the mass emigration during the 1800s and early 1900s are still found today.

3.2.2. Previous studies of American Norwegian

The fact that Norwegian immigrants kept together in Norwegian settlements created a solid foundation for the continued usage of the Norwegian language. However, contact with English was eventually inescapable. In this section, I will introduce a selection of studies conducted on the AmNo language concerned with questions both about the viability of the Norwegian language and the mixing of Norwegian and English.

AmNo was the subject of research already around the beginning of the 20th century, when linguists Nils Flaten and George T. From published short articles, or notes, concerning the language. Flaten (1900) described a community where Norwegian was spoken almost exclusively, but where English words and idioms could freely be used. He continued to describe systematics in this mixing: "generally the root, or stem, is taken and Norse inflections are added as required by the rules of the language" (Flaten, 1900: 115). Flom (1900) described, in a similar manner, a language into which English elements were incorporated, constituting an essential part of it. In later publications, Flom (1903, 1926) elaborated on the question of gender assignment to English loans and provided a rich list of possible English word adoptions. The first large-scale data collections were conducted in 1931 by two linguists from the University of Oslo, Didrik A. Seip and Ernst W. Selmer. They wanted to learn if AmNo speakers had preserved old Norwegian dialect features or developed a new dialect through influencing each other. Discovering instead a language heavily

influenced by English, they did not pursue their study further, and most of their recordings were, unfortunately, broken or lost over the following decades (Haugen, 1992).[6]

One especially important and influential scholar when it comes to AmNo is Einar Haugen. His seminal work *The Norwegian Language in America* (1953) gives a thorough description of the AmNo society and language. The two volumes cover both the sociolinguistic surroundings from the first immigration to the time of his writing, as well as detailed descriptions and discussions of AmNo grammar, bilingualism, and the shift from Norwegian to English in the AmNo community. Another noteworthy contributor to the documentation and investigation of AmNo is Arnstein Hjelde. Hjelde collected data in the 1980s and studied especially the phonology and morphology of a specific Norwegian dialect, *trøndersk*[7]*,* in America (Hjelde, 1992).

After the initiative of collecting new data and establishing CANS in 2010, studies of AmNo have accelerated. Many of these studies are collected in the books *Germanic Heritage Languages in North America* (2015, edited by Johannessen and Salmons) and *Moribund Germanic Heritage Languages in North America* (2015, edited by Page and Putnam), as well as the special issue of *Norsk Lingvistisk Tidsskrift* 2012 (2) – *Norsk i Amerika [Norwegian in America]* (edited by Johannessen and Salmons). In accordance with the scope of this dissertation, I limit the following presentation to selected works dealing with the formal properties and development of AmNo.[8]

Most of the recent AmNo studies gathered in the publications mentioned above and elsewhere focus on the Norwegian properties of AmNo, discussing formal, historical, and sociolinguistic topics, to mention a few. One recurring question concerns whether or not AmNo is affected by either incomplete acquisition or attrition. These phenomena are difficult to distinguish empirically, and sparse information about the speakers' input and competence over time limits the possible conclusions. Nevertheless, studies show data supporting both trajectories as influential factors in AmNo. Larsson and Johannessen (2015) investigate word order in embedded clauses and argue that some observed patterns are not consistent with the input, so these patterns are interpreted as a result of incomplete acquisition. A different case is presented in Johannessen (2015b) where non-target-like production in certain verb categories is attributed to the process of attrition. This latter study also investigates nominal categories,

[6] Some remaining recordings can be found at http://www.tekstlab.uio.no/norskiamerika/opptak/seip-selmer.html [accessed: May 15, 2017], but the quality is rather poor.
[7] This is the dialect from *Trøndelag,* a region in the central part of Norway.
[8] Prominent scholars at the forefront of contemporary AmNo research are Tor A. Åfarli, Merete Anderssen, Kristin Melum Eide, Arnstein Hjelde, Janne B. Johannessen, Ida Larsson, Terje Lohndal, and Marit Westergaard, whose work can be found in the above-mentioned publications and in other journals.

where the data are less conclusive. In their study of possessive constructions, Westergaard and Anderssen (2015), on the other hand, reach the conclusion that high frequency serves as a protection against attrition. Possessive pronouns in Norwegian can be either prenominal or postnominal, the latter being most frequent. In their study, they expected that influence from English would have shifted the preference toward a prenominal possessive structure among the AmNo speakers, a hypothesis that was not borne out. Studies like these thus show that AmNo has preserved core Norwegian properties, but nevertheless shows some deviance from non-heritage European Norwegian.

A different topic of discussion that is relevant for the current dissertation is grammatical gender, specifically the outcome of a gender language like Norwegian encountering a dominant language like English, where nouns do not have gender. Hjelde (1996) discusses gender assignment to English nouns in AmNo and argues that it is possible to find morphological, semantic, and phonological rules which account for the observed patterns. More recently, Johannessen and Larsson (2015) investigate a selection of speakers in CANS and argue in favor of a general stability in gender assignment in AmNo. However, they do find an overgeneralization to the masculine, and show that the complexity of the phrase impacts the deviance in gender assignment. A similar study was conducted by Lohndal and Westergaard (2016), surprisingly with a different conclusion. Based on a cross-cutting investigation of all 50 speakers in CANS, they argue that gender is vulnerable due to the lack of transparency in the Norwegian gender system. They, too, find an overgeneralization to the masculine, but since both feminine and neuter are affected they interpret this as a process of attrition which may ultimately result in an erosion of the gender system. The contrast between these two studies is striking, and as Lohndal and Westergaard themselves point out, this might be due to their definition of gender: Whereas Johannessen and Larsson (2015) include the definite suffix as a gender marker, Lohndal and Westergaard (2016) argue that this suffix is rather a marker for declension class. Both studies find that the definite suffix in general is more persistent than other (unquestionable) gender markers such as the indefinite article (see also Rødvand, 2017, who reaches a similar conclusion).

The historical and sociolinguistic environment of AmNo, being a minority language in a territory heavily dominated by English, has promoted production showing a mix of the two languages. This was noticed already in the very first studies of AmNo, as mentioned above, and many of the recent works focus explicitly on language mixing (e.g., Grimstad, Lohndal, & Åfarli, 2014; Åfarli, 2015a, b; Alexiadou, Lohndal, Åfarli, & Grimstad, 2015; Grimstad, 2017; Riksem, Grimstad, Lohndal, & Åfarli, in press). These studies provide both a general

discussion of language mixing and how this phenomenon should be analyzed, as well as detailed analyses of AmNo data. Some of these studies investigate verb phases and tense inflection, finding that English verbs used in AmNo typically have Norwegian tense marking (Åfarli, 2015a; Grimstad et al., 2014; Grimstad, 2017). This is not expected if tense inflection is generated in the lexicon, which is assumed in Chomsky (1995) and later works within the lexicalist approach to the Minimalist Program, which I will discuss in Section 4.1. Their conclusion is instead that tense is syntactically assigned. A parallel pattern is indicated for the nominal domain, yielding English nouns with Norwegian functional suffixes (Grimstad et al., 2014; Alexiadou et al., 2015; Riksem et al., in press). The available studies on language mixing in AmNo thus suggest that it follows a systematic and predictable pattern where English lexical items are integrated into Norwegian functional structures.

The present study relates to and elaborates on these previous studies by investigating AmNo, and in particular language mixing in noun phrases, from a formal perspective. The results support the pattern of systematic language mixing suggested in the works mentioned above. Gender in AmNo is an interesting property of English nouns as well as Norwegian ones; English nouns are also assigned a gender when used in AmNo. The present study shows that all three genders are assigned to English nouns, arguing in favor of gender being an active, syntactic category in AmNo. Moreover, the present study is the first to take a diachronic perspective and investigate the stability of language mixing as well as document patterns of change. However, it also questions the foundation for arguing that the observed change is the result of incomplete acquisition or attrition.

4 Theoretical foundation

This section presents the relevant theoretical background and assumptions for the current dissertation. Broadly speaking, the section can be divided into three main components: the structure of grammar, language mixing, and Norwegian DP structure. These are discussed separately. The discussion of the structure of grammar is more extensive than that of the two others, since this lays the foundation for the theoretical model I have employed in the analyses. A comprehensive discussion of any of these topics is, nevertheless, beyond the limits of the current cover article, and therefore I focus on the main issues and questions relevant for the discussion in the three articles contained in the dissertation.

4.1. The structure of grammar

This dissertation adopts the Chomskyan generative approach to the study of language, with the key idea that language is anchored in the human mind and should be formalized to accurately characterize humans' tacit knowledge of language. This language faculty is what facilitates language acquisition and it is the mechanism behind the manifestations of language that we observe in speech or writing. The task of formal generative linguistic research is therefore to characterize the language faculty, with the goal of best understanding how it is structured in order to produce the observed outputs.

Broadly speaking, current formal generative linguistics can be separated into two main approaches, a lexicalist or endoskeletal approach, and a constructivist or exoskeletal approach (see e.g., Ramchand, 2008). A key issue distinguishing these approaches is the division of labor between syntax and the lexicon, and prominent questions concern, for instance, word formation and the source of formal features. The lexicalist approach has dominated formal linguistics over the last four to five decades, becoming especially prominent after the introduction of the Minimalist Program (MP) (cf. Chomsky, 1995). This framework relies on a model where words are built in the lexicon, potentially following procedures different than those that create phrases in the syntax. Moreover, formal features are considered properties of the individual words in the lexicon, and syntactic structures are assumed to project from these words when they are taken into the syntax and combined with other words (Adger, 2003).

The exoskeletal approach, on the other hand, assumes that syntactic structures are generated independently and that words are subsequently inserted into them. In such a model, no distinction is made between the creation of words and phrases; they are assumed to be

generated by the same computational system. Moreover, functional features are considered properties of the syntactic structures. The structure thus forms a grammatical frame of the sentence, where lexical items may be inserted into available positions (see, e.g., Borer, 2005a; Åfarli, 2007; Marantz, 2013; Lohndal, 2014).

The current dissertation takes a non-lexical, or exoskeletal, approach to linguistic structures, following the tradition of Borer (2003, 2005a, b, 2013) as well as the framework of Distributed Morphology (DM) (e.g., Halle & Marantz, 1993, 1994; Marantz, 1996, 1997, 2013; Harley & Noyer, 1999; Embick & Noyer, 2007; Embick, 2015). As I will discuss below, these traditions build on many similar arguments and hypotheses, but differ in the details of the technical implementation. The analyses conducted in the current dissertation can thus be described as DM analyses, assuming a version of DM compatible with some of the core ideas proposed in Borer's work. This will be elaborated upon in later subsections.

4.1.1. The Minimalist Program

MP is the most recent development within the generative tradition (Chomsky, 1995, 2000, 2001, 2005, 2008, 2013; see also Boeckx, 2006 for an introduction). The core of this approach is the hypothesis that Universal Grammar (UG) is perfectly designed in accordance with overarching principles of economy and simplicity. MP builds on a substantial theoretical foundation already established in the Principles and Parameters framework, which was key to important advances in the generative tradition. Through inquiry into a range of different languages, a rich selection of parameters as well as principled operations regulating different linguistic levels were proposed. Chomsky (2005) describes MP as an effort to sharpen the questions of linguistic research. The prominent issues now concern not only how structures look and what their components are, but also *why* grammar should look this way. With simplicity and economy as leading principles, MP pursues the question of "how far can we progress in showing that all such language-specific technology is reducible to principled explanation, thus isolating the core properties that are essential to the language faculty […]?" (Chomsky, 2005: 11). Principled explanations of language are thus limited to those motivated by the interfaces Phonetic Form (PF) and Logical Form (LF). In other words, in MP, emphasis is put on these interfaces, and linguistic expressions generated by syntax are well-designed to the extent that they are motivated and realized in accordance with the conditions set by the interfaces to the semantic and phonological systems (Chomsky, 1995, 2005).

As the name suggests, MP is a research program more than a new theory, with the hallmarks of asking broad questions, making room for multiple perspectives (Boeckx, 2006).

However, the lexicalist approach has enjoyed a prominent position in grammatical theory even prior to the introduction of MP and is furthermore the approach taken by Chomsky (1995) and most subsequent work within MP. Thus, a close connection between the two has been established, sometimes leading to the interpretation that the lexicalist tradition *is* the minimalist approach. In keeping with the programmatic nature of MP, however, I assume that an exoskeletal approach may also be considered minimalistic in that it pursues simple and economic analyses of the structure of grammar.

The main objective of the current discussion is to introduce the properties of the exoskeletal approach as they are understood and employed in the current dissertation. However, an important first question is why one would depart from the mainstream model, lexicalism. Therefore, the discussion in the following subsections will briefly introduce some main components of the lexicalist tradition, and subsequently point out some shortcomings of lexicalism motivating the break from this mainstream approach.

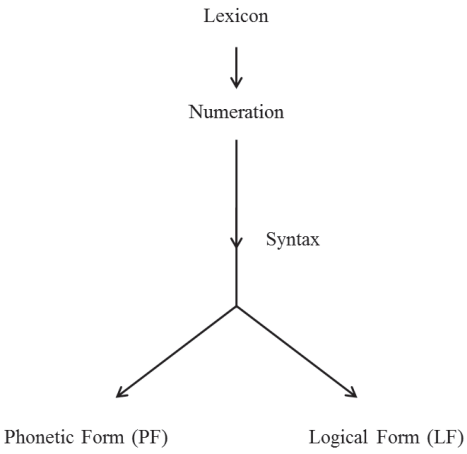
4.1.2. The lexicalist approach to grammar

The lexicalist approach promotes a model of grammar in which the lexical items listed in the lexicon serve a key role (see Adger, 2003, for a textbook introduction).[9] The assumption that lexical entries, or words, are the basic components of sentences is part of a traditional approach to grammar that predates the generative tradition. Within generative grammar, the emergence of the lexicalist tradition is typically attributed to Chomsky's (1970) work on nominalization. Subsequent developments (e.g., Stowell, 1981, Hale & Keyser, 1993), gradually expanded the lexicon and the range of possible information stored there. Lexical entries are now considered packages of (syntactic, semantic, and phonological) features necessary to project structures. These lexical entries, and the features they carry, are considered to be the basic building blocks of syntax, and their interplay with other features is taken to explain different linguistic phenomena. To illustrate this, Adger (2003) describes features as the atoms of language, and in a parallel description, he explains that lexical items constitute the molecules of language, comprising collections of features. Differences between languages thus emerge in the lexicon, as different languages will have different lexical items, or molecules, in their lexicon.

A syntactic derivation from the lexicalist perspective is illustrated in (4).

[9] Lexicalism is in itself an approach that comes in different guises, with differences concerning their assumptions about lexical representations, morphology, and syntax (see Ackerman, Stump, & Webelhuth, 2011 for discussion). For reasons of space, the current discussion is limited to the version of lexicalism put forth in Chomsky (1995) and Adger (2003).

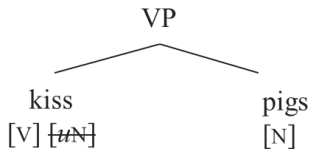
(4)



At the outset of the derivation, a selection of lexical items is drawn from the Lexicon and contained in the Numeration. In syntax, the operation *Merge* applies to these items, building structures which are further accommodated by *Move* and *Agree*, the latter establishing a connection between a probe and a goal ensuring agreement in different positions (Chomsky, 2000; Adger, 2003). The various features of a word are considered to be either interpretable/valued or uninterpretable/unvalued and the latter kind is inadmissible at the interfaces. Hence, the syntax is guided by a system of feature checking or feature valuation, assuring that the structures will converge when they are finally sent off to the interfaces. Lexical items and the functional features they provide thus play a crucial role in the syntactic derivation (Adger, 2003). In some lexicalist approaches operations may even take place in the lexicon (e.g., Hale & Keyser, 1993).

As a concrete example, category features are considered one of the most important sets of features relevant to syntax, and they serve to separate lexical entries into the traditional word classes. For instance, the noun *pig* will be equipped with the feature [N], and the verb *kiss* with the feature [V]. In addition to their category feature, a lexical item may have c(ategory)-selectional features, which will determine which categories that element is able to, or required to, merge with. An example is provided by Adger (2003: 86): "So a word like *kiss* has an interpretable [V] feature […] and an uninterpretable [*u*N] feature. If *kiss* merges with a noun bearing an interpretable [N] feature, then this Merge allows the checking of the uninterpretable [N] feature on the verb". A simple structural representation of this relation is presented below.

(5)



The [*u*N] feature of the verb in (5) requires it to merge with an element holding a corresponding [N] feature, i.e., a noun. These c-selectional features are furthermore commonly associated with Theta-roles, meaning that they instruct the argument structures that the verb will enter into (Adger, 2003: 87). This brief introduction thus demonstrates the crucial role played by lexical items in a lexicalist syntactic derivation. In general, these entries are assumed to provide the functional features required for, among other things, categorization, phrase building, agreement relations, and syntactic movement in this system.

4.1.3. Critique of the lexicalist model

In response to the increasing role played by the lexicon in mainstream approaches to grammar, a different approach emerged and has gained ground, motivated (among other things) by a different interpretation of the division of labor between the syntax and the lexicon. This approach is represented by the exoskeletal tradition, after Borer (2003, 2005a, b, 2013), and DM, both of which address questions like: How much information is stored, and where is it stored? What components and mechanisms does the language faculty consist of? These approaches argue that providing the lexicon with a heavy load of information does not necessarily serve the economic goals of MP. Instead they propose a shift in the work load from lexical items to syntax.

In this subsection, I introduce two core and recurring themes in the critique of the lexicalist approach, namely categorization and argument structure. Borer (2005a: 8) uses the following sentences from Clark and Clark (1979) to illustrate the problem:

(6) a. The factory horns sirened throughout the raid.
b. The factory horns sirened midday and everybody broke for lunch.
c. The police car sirened the Porsche to a stop.
d. The police car sirened up to the accident site.
e. The police car sirened the daylight out of me.

The sentences in (6) exemplify the problems of lexicalist analyses in at least two ways. First, the word *siren,* commonly known to be a noun, is here used as a verb, and moreover, the verb *siren* is used with different argument structures. This is surprising if one supposes that *siren* is listed in the lexicon with a category feature [N] and subsequently no argument structure at all. Nevertheless, a speaker of English will be able to interpret the sentences (6) without much trouble, and the core meaning, to produce sound, is maintained.

In addition, Borer (2017: 127–128) provides the examples in (7), demonstrating how a similar interpretation is conveyed by using different words in the syntactic configurations.

(7) a. The bells rang throughout the raid.
b. The factory signaled midday and everyone stopped for lunch (*e.g., by sirening*).
c. The police forced the Porsche to a stop (*e.g., through sirening*).
d. The police car rushed up to the accident (*e.g., while sirening*).
e. The police car scared the daylights out of me (*e.g., with its sirening*).

The instance of *siren* is not unique, and a wide variety of similar examples can be found (e.g., in Åfarli, 2007; Brøseth, 2007; Nygård, 2013; Grimstad et al., 2014; Lohndal, 2014; Frengstad, 2016; see also Rappaport Hovav & Levin, 1998). A selection of such examples is provided in (8)–(10):

(8) a. Pene jenter i TV2 *vær-er* solskinn og regn.
Nice girls in TV2 weather-PRS sunshine and rain.
'Nice girls on TV2 forecast sunshine and rain.'

b. En venn av meg *tesl-et* til Spania i sommer.
A friend of my tesl-PST to Spain in summer.
'A friend of mine drove to Spain in a Tesla this summer.'

c. Du skal vel bare *tante deg* i dag du.
You shall well only aunt you.REFL in day you.
'You are probably going to do nothing but be an aunt today, aren't you?'

(9) a. He ran *out* the door.
b. My son *outed* me to his preschool.
c. He was desperately looking for an *out*.

(10) a. Kim whistled.
b. Kim whistled a warning.
c. Kim whistled me a warning.

These examples are interesting and prominent in the critique of lexicalism as they show the flexibility of lexical items. Notice first the flexibility concerning word class. The examples in (6) have already shown that the noun *siren* can be used as a verb, and (8) provides additional, similar examples. Next, the examples in (9) show that word class flexibility does not only apply to using nouns as verbs, but seems to be more comprehensive. Here a lexical item typically recognized as a preposition (9a) is used as a verb (9b) and as a noun (9c). A second complication arises from sentences like (6) and (10), which show how *siren* and *whistle*, respectively, may occur with a variety of argument structures. In order to account for this flexibility of word classes and argument structures, lexicalist approaches would have to assume multiple listings of virtually identical lexical items in the lexicon, or to assume a high degree of optionality in a lexical item's specifications. Both alternatives appear uneconomical and reduce the explanatory power of the model (Harley & Noyer, 2000; Borer, 2005a; Åfarli, 2007; Marantz, 2013; Lohndal, 2014). Borer (2005a) and others use this as motivation to flip the perspective, asking "why *words* can mean so many different things, but structures cannot" (Borer 2005a: 3). This observation constitutes one of the core motivations behind exoskeletal approaches.

4.1.4. Exoskeletal approaches to grammar

The term "exoskeletal" originates in Borer (2003), describing a model with a severely impoverished lexicon.[10] The process of word formation, as well as all formal properties, are instead attributed to syntax and morphology: "I will call this view *exo-skeletal*, given its focus on the way in which the structure, rather than the listed item, determines not only grammatical properties, but also the ultimate fine-grained meaning of lexical items themselves" (Borer, 2003: 3). In this approach, the syntactic structure constitutes a skeleton, template, or frame (the precise label is not crucial) which determines the properties and overall shape of its content. Impoverished lexical items are fitted into this structure without any influence on its

[10] This is contrasted with the lexicalist framework, which Borer refers to as *endoskeletal*. Inspiration for the terms exoskeletal and endoskeletal is found in nature where certain animals, e.g., insects, have an external skeleton, whereas for others, e.g., humans, the skeleton is internal. In a parallel manner, exoskeletal approaches to grammar consider structures to be outside of the lexical items, as opposed to endoskeletal approaches that consider the syntactic structure to be embedded within the lexical items themselves.

properties. The structures are moreover considered to have well-defined properties which impose certain conditions on the interpretation of the frame (Borer, 2005a). In other words, structural positions, into which lexical items may be inserted, imply canonical interpretations. The full, actual meaning of a sentence arises from the combination of the canonical structure and the specific lexical items that are employed. As an illustrative example, the two sentences in (11) share the transitive structure, but their specific meanings differ radically due to the choice of lexical items.

(11) a. John ate cookies.
b. John killed Mary.

Although the term "exoskeletal" is commonly attributed specifically to the tradition of Borer, the core exoskeletal hypotheses can be used as an umbrella covering different models and implementations. The scope of the exoskeletal approach is pointed out by Borer:

> In what follows, I will continue to bring forth arguments that support a rich syntactic functional component, and a correspondingly impoverished lexical component. In turn, I will also propose a very specific syntactic functional structure for event structure, and will proceed to justify it as well. However, the validity of postulating an impoverished lexicon, in the sense employed here, is quite independent of the validity of any specific functional structure I will propose. In other words, it may very well turn out that the lexicon is every bit as impoverished as I suggest, but that the syntactic structure required in the presence of such an impoverished lexicon is different from that proposed below. (Borer, 2005b: 10)

As will become clear, this dissertation is anchored in the exoskeletal ideas and motivations proposed by Borer, but the specific analyses conducted in the three articles exploit a model developed within the DM framework.

The exoskeletal approach, broadly understood (i.e., also including DM), constitutes a family of approaches (e.g., Halle & Marantz, 1993, 1994; Marantz, 1996, 1997, 2013; van Hout, 1996; Harley & Noyer, 1999; Alexiadou, 2001; Borer, 2003, 2005a, b, 2013; Åfarli, 2007, 2015a; Embick & Noyer, 2007; Ramchand, 2008; Lohndal, 2012, 2014, Alexiadou, Anagnostopoulou, & Schäfer, 2015, and Embick, 2015). The details of these approaches and the models they propose are different, but crucially they share a common core, namely the notion of Separationism. This entails that abstract syntactic structures are separate and

independent from the phonological realizations that will come to realize them (Harley & Noyer, 1999).

Relatedly, in certain context, the exoskeletal approach is referred to as constructivist, or neo-constructivist. This establishes a link to Construction Grammar (e.g., Goldberg, 1995, 2006), in which constructions, i.e., conventionalized pairs of form and function, play an important role in grammar. Such constructions are combined in order to build larger structures. Without going into an elaborate discussion, core commonalities between the two approaches lie in the assumption that structures/constructions are separate components, and that the structures themselves carry meaning. Important differences, however, concern, for instance, the nature of the structures and how they are created. Whereas constructionist approaches assume that constructions are language-specific schemas stored in the lexicon alongside lexical items, generative exoskeletal approaches consider the structures to be generated by the computational system based on universal principles and fragments made available by UG (Borer, 2005a; Marantz, 1997; Ramchand, 2008). Still, a remaining question for generative approaches concerns whether (frequently used) structural templates may be stored, or if structures are built from scratch every time. Exploring these processes of generating and storing structural templates is interesting and essential to the development of syntactic theory, but nevertheless beyond the scope of the current work.

Two additional core components of the exoskeletal, or DM, framework are the assumptions that syntax works "all the way down" and Late Insertion (see e.g., Harley & Noyer, 1999). The former entails that morphology and syntax do not constitute separate processes, but that words and phrases are generated by the same computational mechanisms. Distinct from the lexicalist approach to grammar, the lexical item (or word) has no special or autonomous status in the exoskeletal framework. It is instead decomposed into more primitive components, crucially including a root (√) devoid of all grammatical categories. Hence, the lexicon is severely reduced, and all functional features, including word class categories, are instead determined by the structural configuration. The latter core component, Late Insertion, refers to the assumption that phonological realizations (known as Vocabulary Items or exponents) are inserted after generating the purely abstract, syntactic structure. I return to a more elaborate discussion of these issues in Section 4.1.6 below. In the following subsection, I will show how an exoskeletal analysis tackles the challenges of lexical flexibility in a more economical and adequate way than the lexicalist alternative.

4.1.5. Exoskeletal accounts for argument structure and categorization

In relation to the question of argument structure, Borer (2003: 2) turns the view within lexicalism upside down, proposing that "the entire burden of argument structure is shouldered by the syntax". Thus, instead of c-selectional features being contained within lexical entries, the structure determines the argument structure of the sentence, as well as providing all other relevant functional features. Lexical items, unable to influence the composition of the syntactic structure, rather function as "modifiers" of the structure. For instance, the item in a direct object position will be interpreted as a patient(-like) argument (Åfarli, 2007). Marantz (2013: 3) highlights this as a significant shift in the theory of grammar:

> [D]iscussion [has shifted] away from verb classes and verb-centered argument structures to the detailed analysis of the way that structure is used to convey meaning in languages, with verbs being integrated into the structure/meaning relations by contributing semantic content, mainly associated with their roots, to subparts of a structured meaning representation. (Marantz, 2013: 3)

Argument structure is, then, not something that a verb *has*, but something a verb *gets* by being inserted into certain structural frames. Subsequently, cases of argument structure flexibility, as in (6) and (10) above, may now be accounted for in a way that is more in line with the economic ideal of MP. Instead of assuming multiple listings of *siren* and *whistle* or a wide optionality in their argument structures, these lexical items receive the argument structure provided by the frame they are inserted into. For instance, in (10) *whistle* occurs in an intransitive, transitive, and ditransitive frame: (10a) *Kim whistled,* (10b) *Kim whistled a warning,* and (10c) *Kim whistled me a warning,* respectively. The verb is simply inserted into a given structure and consequently provided the corresponding argument structure.

The examples in (6), (8), and (9) posed yet another complication for the lexicalist approach, namely that of categorization. These examples showed, for instance, how items commonly recognized as nouns could easily instead occur as verbs in a sentence. If category features are properties of the individual word, it is problematic to account for such word class flexibility without compromising MP's economical ideal. As introduced in the previous subsection, exoskeletal approaches assume lexical items to be decomposable into more primitive components. The core, atomic element is the root (√), and crucially the root is devoid of all grammatical features, including word class category (see, e.g., Arad, 2005). In parallel to argument structures, word class category is also structurally defined in the

exoskeletal framework. The specific implementations of this procedure may vary across different versions of exoskeletal models (see Alexiadou & Lohndal, in press, for a discussion), and this is discussed further in Section 4.1.6 below. Crucially, this means that one root may be assigned different categories depending on its syntactic context. The nouns *siren, vær* 'weather', *Tesla,* and *tante* 'aunt', in (6) and (8), may therefore surface as verbs when their roots occur in a verbal position in the structure rather than a nominal one. The preposition *out,* in (9), may in a similar fashion surface as either a noun or verb by virtue of being inserted such structural contexts.

4.1.6. A late-insertion exoskeletal model

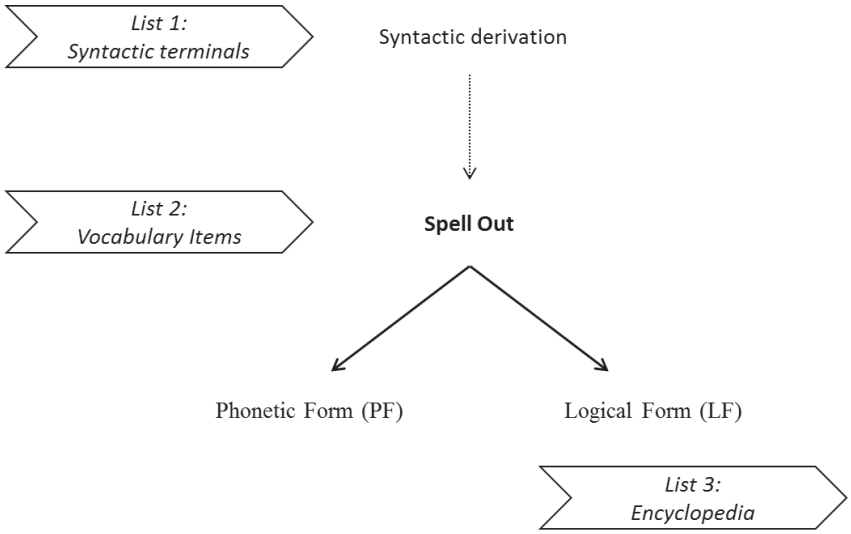
The specific model exploited for analyses in the articles contained in the current dissertation is dubbed a *late-insertion exoskeletal model*. The mechanisms in this model are proposed and developed within the framework of DM. Since DM is in itself a family of approaches differing in their specific implementations, a couple of clarifications are in order before proceeding. First, as already introduced, an important component of exoskeletal analyses is the root. Questions concerning how much information, if anything, can be attributed to the root, what the nature of the root is, and what role it plays in the derivation are nevertheless highly debated. I follow Borer's understanding of a severely impoverished lexicon, containing bare roots without any grammatical features, and the versions of DM compatible with this view. I return to a brief discussion of this later in the present subsection. Secondly, within DM a range of post-syntactic mechanisms have been proposed, e.g., Fission, Fusion, and Impoverishment (see, e.g., Harley & Noyer, 1999; Embick & Noyer, 2007; Embick, 2015). These mechanisms are not adopted in the current work, and will thus not be discussed in the following. However, I remain agnostic as to whether future research may or may not show that these mechanisms are relevant for the data under investigation here.

Essentially, the work within DM is a search for the primitive elements of grammar, the principles that construct complex objects from these primitives, and how these relate to phonological form. The primitive elements are described as morphemes, and complex items are derived by the rules of syntax (Embick, 2015). Still, DM does not refute the need to store information, but rather than assuming one comprehensive lexicon, distinct information is distributed across distinct lists.

DM proposes three such lists: first, a list of syntactic terminals comprising the basic building blocks of grammar (List 1); second, a list of vocabulary items, i.e., phonological realizations inserted post-syntactically (List 2 or the Vocabulary); and finally, the

Encyclopedia (List 3), comprising all additional conceptual knowledge. In the following, I will introduce and discuss these lists and how they, along with (some of) the mechanisms associated with them, constitute a model from which specific predictions can be made. I will focus on the two first lists, as these are crucial to the syntactic derivation and its phonological realization. List 3 , or the Encyclopedia, is assumed to store conceptual information (world-knowledge) about words and possibly phrases. A discussion of the Encyclopedia is beyond the scope of the current dissertation, and this list will not be considered throughout the main part of the discussion. A model of the distribution of the three lists is presented in (12).

(12)



This model illustrates the grammatical process from a DM perspective, and "when and where" the three different lists are accessed in the course of the derivation. The first list provides syntactic terminals, constituting the building blocks of syntax, and these are combined into abstract structural frames. The second list is accessed at Spell-Out, providing phonological exponents for the syntactic structure. I discuss these two lists separately in the following.

The syntactic terminals contained within List 1 are abstract pieces which are combined into complex structures during the syntactic derivation. These are separated into two types: functional features and roots (Embick & Noyer, 2007; Embick, 2015). In an exoskeletal context, formal morphological features constitute the relevant functional features. These are

considered properties of the abstract structures, where the features may occur separately or combined in bundles in the functional terminals. Moreover, these terminals are considered to show *Full Specification,* meaning that they contain all relevant functional features for the terminal in question. Crucially, they do not contain phonological features (Embick, 2015). A common assumption is that UG enables a complete set of such functional features, referred to as *The Universal Feature Inventory* (Embick, 2015; see also Cinque, 1999 and Cinque & Rizzi, 2016 for much discussion). A specific language is then identified based on its selection of features and how these are grouped together in various feature bundles. For instance, what we identify as a Norwegian structure is defined based on the features that are active there. However, even though scholars commonly assume that a universal selection of features is available, there is not yet a comprehensive theory about possible and impossible features. On the one hand, one can argue that this is an empirical question, and that the selection of features is as wide as the attested features found in the languages of the world. Adger and Svenonius (2011), on the other hand, stress the theoretical aspects over the descriptive capacity of features. They argue that a theory of features would enable us to, for instance, organize features into classes and to organize them according to their function. Constraining features would subsequently mean constraining the theory itself (Adger & Svenonius, 2011), thus constituting a central motivation for pursuing a comprehensive theory of features.

The second type of syntactic terminal in List 1 is that of roots.[11] As already mentioned above, the nature of roots and their contribution to the syntactic derivation is a much-debated issue. An exhaustive discussion is therefore beyond the scope of the current work and it would not be possible to do justice to all the contributions in the discussion here.[12] Instead, I will present some core hypotheses concerning roots, and discuss the notion of roots as adopted in the analyses in the current dissertation.

Like functional features, roots are basic building blocks of syntactic structures, but unlike the former, roots are devoid of grammatical properties. Roots are considered the atomic, non-decomposable core elements of what will surface as substantial items in an utterance (Arad, 2005). An important motivation for assuming roots is the potential relation between similar words: Is it just the result of incidental phonological similarities, or do they contain a common unit? Arad (2005) illustrates this with the examples in (13).

[11] These are referred to as listemes in Borer's early works (e.g., Borer, 2003, 2005a, b), but Borer also refers to roots in later publications (e.g., Borer, 2013, 2014).

[12] See, e.g., Arad (2005), Acquaviva (2009), Borer (2013), de Belder (2013), as well as articles compiled in Alexiadou, Borer, & Schäfer (Eds.) *The Syntax of Roots and the Roots of Syntax* (2014), and *Theoretical Linguistics* 40 (3/4) (2014).

(13) a. Lucy broke the glass.
b. The glass broke.
c. Lucy hammered the metal.
d. Lucy bought a hammer.

The relations between *broke* in (13a) and (13b) and between *hammer* in (13c) and (13d) are, according to Arad, crucial to the understanding of building blocks of grammar, as she states: "I believe that if we take causative and inchoative 'break' or the noun 'hammer' and the verb 'hammer' to be two different concepts, then we miss some generalization about what these words share" (Arad, 2005: 6). The idea that words are made of lexical kernels is a traditional assumption, but in the lexicalist tradition words were typically considered the basic elements, not available for further decomposition (Arad, 2005). In exoskeletal approaches, on the other hand, words are decomposed and the root constitutes the atomic unit accounting for the relation between similar words. Beyond that, however, a range of different approaches emerge, as pointed out by Borer:

> [A] central role is played not by a 'word' or a 'lexeme' in the traditional sense, but rather, by a 'root'. Within all of these approaches, there is a general understanding that roots are at the very least devoid of syntactic category as well as of any discernible morpho-phonological complexity. Beyond that, however, what 'roots' are, exactly, is by no means agreed upon, and as a consequence, there is little agreement on how, exactly, they interact with the syntax or, indeed, whether they are altogether necessary. (Borer, 2014: 343)

Although there is no single theory about roots, broadly speaking, two main approaches can be identified. On the one hand, one can assume that roots are totally bare, with no grammatically relevant properties at all, and on the other hand, one can assume that roots have certain features or selectional properties, enabling them to be somewhat more active in the syntactic derivation. A typology of different approaches to roots is provided by Gallego (2014: 192) and reproduced in (14).

(14)

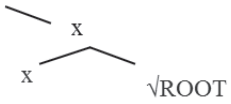
	Category	Argument structure	Semantic type	Conceptual content
Partially bare	No	Yes	Yes	Yes
	No	No	Yes	Yes
Totally bare	No	No	No	Yes

As can be read from the table, consensus is found in that roots are assumed to be category-neutral, i.e., not identified as a member of one specific word class, and in that roots are associated with conceptual content of some kind. Roots like √CAT and √SING thus represent concepts and are provided a category in the structure. This is elaborated upon below. A first difference among these approaches emerges in the question of whether roots have argument-taking properties, where some approaches assume that roots can take internal arguments (e.g., Harley, 2009, 2014). Moreover, some approaches assume roots to be bearers of a semantic type, for instance [event] or [state], which influences their structural distribution (see, e.g., Harley, 2005; Ramchand, 2008; Embick, 2015). These are all representatives of the partially bare approach to roots.

Borer (2005a, b, 2013, 2014) is an advocate for the approach considering roots to be totally bare, i.e., devoid of properties concerning category, argument structure, and semantic type. This is the view adopted in the current dissertation as well, and within the family of DM, I follow the proposals that are compatible with this approach (e.g., Marantz, 1997, 2013; Arad, 2005). Below I will return to the question of conceptual content, which may still constitute a difference among these approaches.

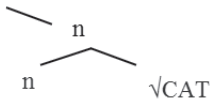
Concerning categorization, Borer and DM agree that this is syntactic, but in their specific implementations they take different approaches. Within DM, a designated category-defining head, called a categorizer, will merge with the root and provide it with a category (Marantz, 1997; Arad, 2005; Pylkkänen, 2008; Embick & Marantz, 2008; Embick, 2015). In what follows, the resulting complex is informally referred to as a stem. A structural representation of this is given in (15), where x can be (for instance) a verbal (v), nominal (n), or adjectival (a) categorizer.

(15)

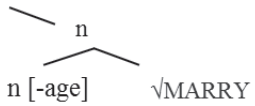


The categorizer may materialize as a derivational affix or have a phonologically empty (null) realization. Examples from the nominal domain are shown in (16) where the categorizer in (16a) *cat* is not phonologically realized, and (16b) where the categorizing n is overt, combining with the root to create *marriage* (Embick, 2015: 45–46).

(16) a.

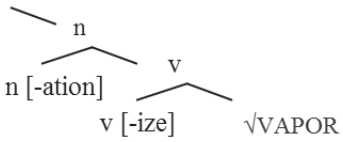


b.



Subsequent categorizing heads may also enter the structure, changing the category of a stem, for instance through adding another derivational affix (Embick, 2015). The example of *vaporization* is shown in (17).[13]

(17)



Borer (2005a, 2014), on the other hand, argues that null realizations, such as in (16a), are superfluous, and proposes a different process of categorization. In cases without an overt categorizer she assumes that the root is inserted directly into a designated complement position in the functional structure. For instance, a root inserted under a DP will be categorized as a noun. To account for derivational suffixes, as in (16b), she introduces a category-defining head similar to that in DM, called a C-functor in Borer's terms, which provides the category and materializes as a suffix. In the analyses in the current dissertation, I assume the analysis with category-defining heads, as proposed in DM.

As discussed above, there is a discussion in the literature concerning what properties, if any, are introduced by the root. In DM, it is typically assumed that, unlike functional morphemes, roots have phonological features (Embick, 2015). Nevertheless, uncategorized roots are considered unavailable for Spell-Out, and categorization, as discussed above, emerges as a requirement for a root to enter the syntax (Arad, 2005). A different question

[13] Notice that using multiple layers may make the theory less restrictive and potentially less falsifiable, especially if the layers have null realizations. See Borer (2014).

concerns whether roots have semantic properties. In her most recent work, Borer (2013, 2014) assumes roots to be pure phonological packages, thus not containing anything of semantic or conceptual value. Examples like those in (13) above, could, on the other hand, suggest that there is a conceptual as well as phonological similarity between these words. Thus, other works argue that roots carry conceptual properties, but that the extent of such properties may vary across languages (e.g., Arad, 2005; Harley, 2005; Levinson, 2007; Anagnostopoulou & Samioti, 2014; Alexiadou & Lohndal, 2017)

Arad (2005), for instance, assumes roots to have both core phonological and core conceptual properties, and this is the view I will briefly present in the following discussion. In Arad's work, roots are described as being underspecified potentialities, and their meaning largely depends on the environment into which they are inserted.

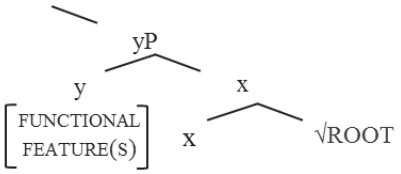
> I suggest a fundamental parallelism between phonology and semantics: in both, the root, in its pure form, is incomplete. Taken on its own, the root is neither phonologically pronounceable nor semantically fully specified. It may instead be represented as a potentiality for a range of sounds and meanings. (Arad, 2005: 60)

In other words, a root's semantic potential will take effect when combined with a categorizer in a specific environment. However, Arad (2005) highlights this as one of the main challenges for the root hypothesis, as it is not clear how this semantic content and the relationship between it and specific word-meanings should be characterized.[14]

So far we have established that abstract syntactic structures are generated based on the syntactic terminals in List 1 (i.e., functional features and roots) in DM. A simple representation of a syntactic structure is the following.

[14] The argument of roots being phonological and semantic potentialities may be more easily perceived in the case of Hebrew, the language Arad (2005) investigates, than for, e.g., English. The decomposition of roots is quite visible in Hebrew, where the root is made out of a string of consonants, and needs to be supplied with vowels in order to be pronounced. Thus, a relatively small selection of Hebrew roots may turn into numerous nouns and verbs. In English, on the other hand, the decomposition of roots is not equally evident, and an English root can only appear in a limited number of environments, thus requiring a larger selection of different roots. With this in mind, one can speculate whether English actually makes use of roots, at least in a similar fashion to Hebrew. Building on Arad (2003, 2005), Alexiadou and Lohndal (2017) propose a typology of languages depending on the role of the root, with Hebrew and English as potential terminal points on a scale. Hebrew roots have little independent meaning, and are dependent on functional vocabulary, whereas English roots carry substantial meaning, and word-making morphology is often not present. Different languages may then be placed on this scale depending on the semantic properties of their roots. See also Kastner (2017).

(18)



At the bottom of the structure, a root is combined with a category-defining head x. In the extended projection, relevant functional features, either represented separately or in bundles, are merged in one or more functional projections, yP. This yields the abstract syntactic structure subject to phonological realization. In an extended version, the structure will also comprise additional open positions, i.e., positions not containing functional features, such as specifier, complement, and potential adjunct positions, available for insertion of various constituents.[15] A further discussion of the structural template is, nevertheless, beyond the scope of the current dissertation, which primarily investigates the relation between a stem and associated functional features.

In the next step of the process, the Vocabulary (List 2) is accessed and the morphemes receive their phonological forms, governed by the operation Vocabulary Insertion. The Vocabulary is thus considered a collection of exponents, each consisting of, on the one hand, a phonological representation, and, on the other hand, specifications for functional features, which determine where in the structure the exponent may be inserted.

Employing an operation like Vocabulary Insertion adds an extra computation to the syntactic derivation, as compared to the lexicalist view where sound and meaning are connected already in the lexicon. Embick (2015) argues that this expansion is empirically motivated, especially by cases of syncretism. In cases where distinct functional terminals have the same phonological realization, Late Insertion allows generalizations across Vocabulary Items in a way that reduces the number of items needed to capture the observed variation. This implies *Underspecification*, which is an important assumption within DM. Vocabulary Items are assumed to be underspecified for functional features, meaning that the same item may spell out several positions. As a simple example, consider English verbal agreement in present tense where 3rd person singular requires the suffix *-s*, *John eats cake,* whereas 1st and 2nd person as well as plural cases all have null realizations, e.g., 1st person singular, *I eat cake.*

[15] Notice that in Article 2 of the current dissertation these positions are referred to as "open slots". This is intended to be a similar collective term covering the above-mentioned alternatives.

Thus, in the Vocabulary, only the exponent for 3rd person singular needs to be fully specified. The remaining cases can all be captured by one exponent specified only for present tense. In other words, Underspecification reduces the number of necessary exponents in the Vocabulary. The available Vocabulary Items, or exponents, for English verbal inflection can then be presented as in (19).

(19) a. [PRS, 3PERS, SG] ↔ *-s*
b. [PRS] ↔ -

When more than one exponent is a potential realization of the same terminal, insertion is competitive. This process is guided by the Subset Principle:

> The phonological exponent of a Vocabulary item is inserted into a morpheme in the terminal string if the item matches all or a subset of the grammatical features specified in the terminal morpheme. Insertion does not take place if the Vocabulary item contains features not present in the morpheme. Where several Vocabulary items meet the conditions for insertion, the item matching the greatest number of features specified in the terminal morpheme must be chosen. (Halle, 1997: 428)

In other words, the Subset Principle states that the exponent that best matches the functional terminal is preferred.[16] Moreover, an exponent cannot be inserted if it contains features that are not present in the structure. For English verbal inflection, the exponent *-s* will be the most appropriate exponent for feature bundles of present tense, 3rd person singular, but it is blocked from insertion in any other feature bundle composition. A key objective in DM is then to investigate the features of various exponents in order to make the most accurate predictions about their insertion. Concerning roots, many approaches within DM assume that they have phonological features, and are thus not subject to a competitive insertion process like the functional exponents (Embick, 2015). In sum, the process of Spell-Out in DM is piece-based, with phonological exponents for different terminals in the structure.

The process of Spell-Out represents an additional difference between DM and Borer's implementation. Borer (2013) argues in favor of Spell-Out of non-complex items.

[16] In addition to being the most appropriately specified alternative, many DM approaches also assume that phonological exponents may have contextual conditions governing their insertion (Embick, 2015). Such conditions are not assumed in the current dissertation and thus not discussed further.

> I will assume that inflection is, indeed, radically realizational, which is to say, I will assume that it is amorphous, and that e.g. *sang*, *walked*, and *dreamt* are all non-complex. As a consequence, I will also assume, as in Anderson (1982, 1992), but not as in Halle and Marantz (1993), that although there certainly is something that we may refer to as PAST, that *sang* or *walked* correspond to is not a combination of a stem (or a root) + PAST, but rather the spellout of a stem (or a root) marked as PST. (Borer, 2013: 22)

According to Borer (2013), some functional categories are dependent on their immediate environment for spell-out. The required phonological information is provided by an adjoining root, as these are assumed to be packages of (pure) phonological information. The Spell-Out of the root may also vary in different functional contexts. For instance, in the case of plurality in English, the root *tooth* must have a phonological package rich enough to contain information that whenever it is marked as plural it will be spelled out as *teeth.* In the absence of such specific information, Spell-Out will result in a default form, which for English plurals is the suffix *-s*. Concerning the empirical material under investigation in the current dissertation, dealing with language mixing, such an analysis of realization is problematic. For instance, the data investigated here show several cases in which an English root or stem occurs with a Norwegian plural suffix, and it is not expected that an English root should carry phonological information about Spell-Out in a Norwegian context. Moreover, as the English nouns appear with Norwegian suffixes varying according to gender, this cannot be accounted for as an instance of the default form. Due to these issues, a piece-based approach to Spell-Out, as in DM, is more promising and suitable for the data investigated here.[17]

4.1.7. Interpretation

The discussion so far has been focused on the syntactic derivation and the phonological realization of syntactic structures, i.e. the PF interface, which reflects the focus of the three

[17] Notice that language mixing also raises challenges for a theory of spanning (e.g., Svenonius, 2012, 2016). Spell-Out is, under this approach, assumed to recognize spans, i.e., sequences of heads in an extended projection. A critique similar to the one raised against Borer's (2013) take on Spell-Out can therefore be raised against spanning; it is problematic to account for Spell-Out of a combined sequence of heads when that sequence contains elements from more than one language. A particularly illustrative example from contemporary Norwegian is a case like *cardigan-s-ane* 'cardigan-PL-DF.PL.M', containing both an English plural *-s* as well as a Norwegian definite suffix. Such cases support a piece-based approach to Spell-Out, as it is hard to account for how a combined Spell-Out of such a sequence still would acknowledge the various contributions from each language.

On the other hand, cases of suppletion constitute a challenge for piece-based approaches, favoring instead spanning or other non-complex processes of Spell-Out. These questions are significant and important for further developments of syntactic theory, but since they do not arise from the data under investigation in the current dissertation, I will not elaborate upon them.

articles that follow. The second interface in the model of grammar, Logical Form (LF), is not prominent in the current dissertation and is therefore not elaborated upon (see, e.g., Lohndal, 2014 for an approach compatible with the present perspective). A few remarks should, nevertheless, be made concerning the interpretation of syntactic structures and especially their correspondence to the interpretation provided by the phonological exponents. As discussed above, a common assumption in exoskeletal approaches is that structural frames carry some meaning, and that an additional and concrete meaning is supplied when inserting lexical items. The meaning carried by the structure is assumed to be basic: Positions within the syntactic frame will impose a canonical interpretation concerning argument structure relations. The core semantics of the roots then provide an additional layer of meaning, and function as a "modifier" of the structure: "the insertion of lexical elements into the frame implies a semantic *enrichment* of the very rudimentary semantics of the frame itself" (Åfarli, 2007: 15).

Following from the exoskeletal approach to grammar, any root can in principle be inserted into a designated position in the frame. This, however, may in some cases result in sentences that are quite odd, subsequently raising interesting questions concerning the interaction between the interpretation of a root and the structure. An important difference to keep in mind in such a discussion is the difference between what is grammatical and ungrammatical, on the one hand, and what is acceptable and unacceptable, on the other. In many cases a sentence may be assessed as unacceptable due to its content rather than its actual grammatical components. In an exoskeletal approach, a sentence is considered grammatical insofar as it fulfils the necessary requirements set by the abstract structural frame and its realization of functional terminals, following the Subset Principle. The choice of roots will not directly affect the grammaticality of the sentence, but may cause an unexpected interpretation.

> From this perspective, the infelicity of [The police car fell up to the accident] emerges not from the grammatical properties of *fall* but from a clash between the basic meaning of FALL and the event interpretation that emerges from the syntax […]. This infelicity parallels that of juxtaposing *sleep* and *furiously* or *colorless* and *green*, as in Chomsky's (1957) famous example [*Colorless green* ideas *sleep furiously*]. (Borer, 2017: 128)

Key words in a discussion of interpretation and acceptability are harmony and convention (Borer, 2005a; Åfarli, 2007). In a syntactic derivation and its realization, both the structure

and the roots will provide fundamental semantics, and the proposed hypothesis is that the harmony between these two components will determine how "natural" or odd an utterance is perceived as being. A ditransitive frame will, for instance, provide an interpretation of a transaction, transferring *something* from *someone* to *someone else.* Roots with a similar transactional interpretation, like *give* or *send,* will harmonize nicely in the verbal position in such frames. In comparison, *whistle*, as used in the examples in (10) above, does not typically yield a transaction. Thus, the harmony with a ditransitive frame, as in (10c) *Kim whistled me a warning,* is not as ideal as with *give* or *send.* Borer (2005a) emphasizes such a match determining the felicity of the utterance:

> In the event of a mismatch, the grammar will always prevail. The interpretation put forth by the conceptual component can and will stretch, as much as possible within the confines of the concept under consideration, so as to match the rigid, absolute interpretational constrains circumscribed by the grammar. Indeed, one should never underestimate the stretching abilities of concepts. After all, even *square circles* can be assigned an interpretation. The more the conceptual system stretches, the more the utterance will appear odd, and at times, the oddity may be so extreme that it becomes difficult to distinguish from a straightforward case of ungrammaticality […]. (Borer 2005a: 11)

In other words, the conceptual content will stretch, and the more it stretches, the odder the sentence is. Consequently, the speaker will typically choose structures and roots that harmonize, and convention emerges as an important factor in what roots *usually* fill a particular structure (Åfarli, 2007; Marantz, 2013; Lohndal, 2014).

4.1.8. Summary

To summarize, this dissertation employs a late-insertion exoskeletal model in its analyses. The term "exoskeletal" is used in a broad sense here, emphasizing the separation of independently generated syntactic structures from phonological items and an impoverished lexicon, thus also including Distributed Morphology (DM). These exoskeletal approaches stand in contrast to the mainstream lexicalist approaches to grammar, and provide a more adequate analysis of argument structure and categorization than lexicalist approaches: In exoskeletal approaches, the syntactic structure/template/frame provides designated positions for inserting lexical items, and when an item is inserted into a "verbal position" in the

structure, it is interpreted as a verb, with argument structure provided by the structure. This means that lexical items themselves carry neither category nor argument structures.

In my specific implementation of the exoskeletal approach to grammar, I employ a model proposed and developed within DM. In this framework, the lexicon is distributed across three lists, accessed at different stages in the derivation. The elements in List 1, which are the syntactic terminals, i.e., functional features and roots, constitute the building blocks of the abstract structure; elements in List 2, or Vocabulary Items, are the phonological exponents inserted into the structure. Finally, the Encyclopedia provides a rich source of interpretations.

4.2. Language mixing

The phenomenon under investigation in the current dissertation is referred to as language mixing. This term is primarily employed as an informal, pre-theoretical account of the observed phenomenon: utterances containing (lexical or functional) elements from two or more languages. Related and frequently used terms for this phenomenon are codeswitching and borrowing, and I return to a discussion of the interpretation and relation between these terms in Section 4.2.1. As will emerge from that discussion, I adhere to a broad understanding of the term language mixing, covering various materializations of the co-occurrence of elements from different languages (see also Muysken, 2000; Lohndal, 2013).

Language mixing is a phenomenon prone to occur in any multilingual society. It may be especially prominent among balanced bilinguals or in immigrant communities, but also in historically multilingual societies, as well as among proficient L2 speakers.[18] In fact, today it seems more accurate to say that most speakers have, to some degree, knowledge or familiarity with more than one language, which subsequently enables mixing (Grosjean, 1998).

Nevertheless, language mixing, or multilingualism in general, has not always gained a lot of attention from linguists. Such studies emerged mainly in the 1960s and 1970s, although there were some even earlier (see Benson, 2001 for an overview). The majority of the early studies were conducted in the fields of sociolinguistics and pragmatics, whereas grammatical analyses have developed from these traditions, now constituting a growing area in linguistics (Gardner-Chloros, 2009: 9–10). In formal linguistic studies, the classical portrait of the ideal speaker/listener long constituted the core object of linguistic studies (cf. Chomsky, 1965), thus limiting the interest in phenomena like language mixing. This strategy was arguably fruitful in terms of developing the fundamentals of the theory, and even though much

[18] Notice that I use the term bilingual throughout the discussion, but that all parts of the discussion are naturally also relevant for multilingual speakers.

contemporary research continues in this tradition, attention is turning towards more complex linguistic phenomena.

From a formal point of view, a driving question concerns how free or how restricted language mixing actually is. In other words, there may be many (social) reasons for *why* and *when* people mix languages, but the interest of formal linguistics, and the current dissertation, is rather *how* mixing is produced by the language faculty and how it is restricted (see, e.g., Gardner-Chloros, 2009 for a more general introduction). Formal approaches to language mixing have therefore typically set out to identify general constraints regulating the mixing (e.g., Poplack 1980; Sankoff & Poplack, 1981; Woolford, 1983; Di Sciullo, Muysken, & Singh, 1986; Belazi, Rubin, & Toribio, 1994; Myers Scotton, 1993, 2002; MacSwan, 1999, 2000, 2005a, b, 2009, 2014; Muysken, 2000, 2015; Gónzalez-Vilbazo & López, 2011, 2012). I return to a discussion of some prominent proposals in the subsections that follow. In general, language mixing may constitute a potential window into our language capacity; the conditions and restrictions on language mixing can tell us which linguistic elements are possible to mix, and whether some are more available or resistant to mixing than others. Thus, studies of language mixing may refine and deepen our understanding of grammatical theory (Muysken, 2000; Gardner-Chloros, 2009, González-Vilbazo et al., 2013).

In the following subsections, I will discuss key terminology, as well as some important proposals and core goals for studies of language mixing. First, I will introduce a common, although much debated, pair of terms used in studies of language mixing, *codeswitching* and *borrowing,* and how these are understood and employed in different works (see Grimstad, 2017 for further discussion). In the subsequent subsections, I discuss a selection of formal analyses of language mixing. In general, one can identify two strategies among such analyses. On the one hand, one can argue that the ability to mix languages requires an additional setting or additional machinery in the language faculty of bilinguals. On the other hand, one can argue that language mixing is constrained by the same mechanisms and principles as monolingual speech. The proposals presented below will represent both sides. First, in Section 4.2.2, I discuss some constraints on language mixing proposed in early work. Then, I turn to the influential work by Myers-Scotton (1993, 2002) and her Matrix Language Frame Model in Section 4.2.3. These fall in under the former approach, proposing some special machinery for language mixing. In the final subsection, 4.2.4, I discuss the latter approach, which argues in favor of a uniform analysis of mixed and unmixed utterances, and introduce some important contributions.

4.2.1. Codeswitching and borrowing

The literature on language mixing comprises an extensive discussion of a wide variety of terms concerning utterances where elements from different languages occur together: codeswitching, code-mixing, language mixing, borrowing, and nonce borrowing, to mention the most frequent. The main objective of the current dissertation is not to bring new evidence or shed new light on the nuances between the different terms. Instead, as mentioned above, I use the term language mixing as a general (and neutral) term describing the observed phenomenon (see also Muysken, 2015; Grimstad, 2017). Given their prominence in the literature, it is still necessary to briefly address two of the terms.

At the core of the debate, we find the terms codeswitching (CS) and borrowing, and the question of whether or not these should be formally distinguished in linguistic theories. According to Poplack (2004), differentiating between these terms is uncontroversial. Borrowings are morphologically, syntactically, and phonologically integrated into the recipient language, making them available also for monolinguals; CS, on the other hand, means taking material from the donor-language "as is", which requires some bilingual competence. This difference between CS and borrowing is also stressed by MacSwan (1999, 2000, 2005a), who argues that borrowing only involves one lexicon into which the new element has been integrated, whereas CS results from drawing lexical items from two lexicons.

A main point of debate arises when encountering singly occurring foreign words, which potentially show word-internal mixing, for instance a lexical item from Language A occurring with morphology from Language B. Are these CS or borrowings? Some examples from AmNo are provided in (20), where the substantial item is English and the morphology is Norwegian.

(20)	a.	**company**-en	company-DF.SG.M	'the company'
	b.	**kid**-er	kid-INDF.PL.M/F	'kids'
	c.	**watch**-a	watch-PST	'watched'

This kind of mixing has, in a substantial part of the literature, been considered some kind of borrowing rather than CS (e.g., Poplack, 1980; MacSwan, 1999, 2000; MacSwan & Colina, 2014). Poplack's (1980) *Free Morpheme Constraint* for instance, states that switching may only occur between free morphemes, effectively ruling out (20) as instances of CS. I return to this constraint below. Poplack (2004), nevertheless, acknowledges that lone other-language

items have similarities to both borrowings and CS; they are lone lexical items potentially integrated into the recipient language, like borrowings, but they are not widespread and typically require bilingual competence, like CS. The term *nonce borrowing* was proposed to cover these cases, describing them as borrowed items, although borrowed only for the moment, the *nonce* (Poplack, Sankoff, & Miller, 1988; Poplack, 2004).

An important component of MacSwan's (1999, 2009) analysis is that PF conditions prevent word-internal CS. The argument, in brief terms, is that since CS presumably involves items being drawn directly from different lexicons, retaining their original phonology, word-internal CS is impossible: The phonological system would not be able to parse two types of phonology within one item (MacSwan, 2005a). Instead, the boundary for CS must found between heads. This is formulated as the PF Disjunction Theorem (revised as the PF Interface Condition in MacSwan, 2009; MacSwan & Colina, 2014). Studies by Poplack and her colleagues, however, have shown that phonology can vary greatly in the case of borrowing and CS, and thereby constitutes a weak criterion for distinguishing these phenomena (e.g., Poplack et al., 1988; Poplack & Dion, 2012).[19] Bandi-Rao and den Dikken (2014) take a DM approach to the alleged ban on word-internal CS and argue that this is restricted by the syntactic configuration rather than phonological systems. Thus, a phonological word composed of separate syntactic terminals, e.g., a root and a functional feature bundle, is realized by separate phonological exponents (cf. DM's piece-based approach to Spell-Out as discussed in Section 4.1.6), which may yield word-internal CS.

The term nonce borrowing, and the general exclusion of single items from CS, has been the subject of controversy. The combination of the terms *nonce* and *borrowing* itself is, for some researchers, a contradiction, hiding the diachronic development of a borrowed item. Single-word CS and borrowings are instead considered part of a continuum based on how established these items have become in the recipient language (Myers-Scotton, 1993; Treffers-Daller, 2005; Gardner-Chloros, 2009). Under this view, all borrowings start out as CS and are then gradually generalized among the speakers of the recipient language: "CS forms may become B [Borrowed] forms through an increase in their frequency and their adoption by monolinguals; nothing more is required" (Myers-Scotton, 1993: 182). Hence, a

[19] The conditions concerning phonology are not discussed specifically in this dissertation, largely because of Poplack and her colleagues' observation that it is a weak predictor in language mixing.

term like *nonce borrowing* is seen as superfluous, and CS and borrowed items are related in terms of frequency.[20]

Moreover, Jake et al. (2002) argue that by excluding singly occurring other-language items, we miss a generalization about language mixing, and they accuse MacSwan of taking the easy way out: "MacSwan avoids feature mismatches that would arise when singly occurring forms appear in mixed constituents by classifying all such forms as borrowings" (Jake et al. 2002: 71). They continue by pointing out that single items constitute the most frequent type of CS, and should be carefully considered in a theoretical model.

As mentioned above, the current dissertation uses the term language mixing to describe any instance where elements commonly associated with two (or more) languages occur in the same utterance. As far as the CS versus borrowing debate goes, this quote from Gardner-Chloros illustrates the core of the conflict, namely that many of the differences rest in the definitions we employ:

> CS is not an entity which exists out there in the objective world, but a construct which linguists have developed to help them describe their data. It is therefore pointless to argue about what CS *is*, because, to paraphrase Humpty Dumpty, the word CS can mean whatever we want it to mean. (Gardner-Chloros, 2009: 10–11)

The term language mixing is thus intended to function as an umbrella term for a phenomenon that may manifest itself in different ways, and which is described by a variety of terms. Specifically, this dissertation is primarily concerned with language mixing manifesting itself as singly occurring English items in a Norwegian utterance, and among these a considerable amount occur with Norwegian morphology. Hence, they are situated right at the center of the single-word or word-internal CS discussion (see also Muysken, 2000 and his discussion of cases of *insertion*). I believe the approach taking single-word CS and borrowings as part of the same continuum is on the right track; what separates single-word CS from borrowing is nothing other than how established the word has become in the recipient language (see also Grimstad, 2017). Consequently, all instances of language mixing should be accounted for by the same model. I return to this topic in Section 4.2.4 and later in Section 6.2.1. Below, I present some previously proposed approaches to analyzing language mixing.

[20] Frequency of occurrence is, nevertheless, quite intangible. It begs the question of how this frequency should be captured, and at what point should a codeswitched form be identified as a borrowed form. A discussion of this is, however, beyond the limits of the current dissertation, and I will not discuss it further. See, e.g., Myers-Scotton (1993) or Poplack & Dion (2012).

4.2.2. Early approaches

Early on it was recognized that language mixing follows certain rules, and previous formal studies of language mixing have focused on proposing different constraints on grammar in order to explain these rules. The limits of this cover article prevent an exhaustive discussion, but I will briefly introduce some of the most prominent proposals.

Two influential constraints were proposed by Poplack (1980), argued in combination to be general enough to account for all CS.

(21) **The Free Morpheme Constraint:** Codes may be switched after any constituent in discourse provided that constituent is not a bound morpheme.

(22) **The Equivalence Constraint**: Code-switches will tend to occur at points in discourse where juxtaposition of L_1 and L_2 elements does not violate a syntactic rule of either language, i.e. at points around which the surface structures of the two languages map onto each other.

(Poplack, 1980: 585–586)

As mentioned above, the Free Morpheme Constraint (21) effectively prevents CS from occurring within a word, whereas the Equivalence Constraint (22) prevents CS within a constituent that is only relevant for one of the languages. Subsequent empirical tests have shown that the latter constraint does not hold (Myers-Scotton, 1993; Mahoothian, 1993; MacSwan, 1999), but the former constraint remains controversial, as it is tied directly to the CS versus borrowing debate.

Other proposed constraints utilize general generative mechanisms, like government and f-selection, in accounting for language mixing. Two such examples are shown in (23) and (24).

(23) **The Government Constraint**

a. If Lq carrier has index q, then Yq^{max}

b. In a maximal projection Y^{max}, the Lq carrier is the lexical element that asymmetrically c-commands the other lexical elements or terminal phrase nodes dominated by Y^{max}.

(Di Sciullo et al., 1986: 6)

(24) **The Functional Head Constraint:** The language feature of the complement f-selected by a functional head, like all other relevant features, must match the corresponding feature of that functional head.

(Belazi et al., 1994: 228)

Both of these constraints, however, resort to a language feature/index in order to successfully account for mixing. Due to government constraints, either in a c-command relation (Di Sciullo et al., 1986) or from f-selection (Belazi et al., 1994), the items involved must match in terms of their language index or feature. However, neither of the two approaches discusses the motivation behind these language features/indexes to any extent. In fact, Di Sciullo et al. (1986: 4) simply state that "We take the notion of language index to be a basic one; it simply marks the words that are drawn from a particular lexicon". Belazi et al. (1994) similarly take the language feature to be on a par with any other feature. The language feature and the associated constraint are further held to be operative in all speech, though their effects will be apparent only in cases of mixing. If there is indeed something like a language feature, it must be present in all production, but as I will discuss below, specific references to the languages involved in mixing are questionable (MacSwan, 2014).

4.2.3. Matrix Language Frame Model

One of the most influential analyses of language mixing is the Matrix Language Frame Model (MLF model) proposed by Myers-Scotton (1993). According to this approach, it is impossible to analyze mixing phenomena without recognizing an asymmetry between the languages involved; one language, the Matrix Language (ML), enjoys a privileged status, as it is responsible both for word order and for providing the inflectional or functional morphemes. The other language(s), the Embedded Language(s) (EL), may only contribute content items. Thus, in the analyses of language mixing, two crucial distinctions are identified and should be accounted for. First, one must account for the hierarchical relation between the ML and one or more EL(s). Second, the model draws a distinction between system morphemes (i.e., functional morphemes) and content (substantial) morphemes, and claims that an EL can only contribute content morphemes. The MLF model accounts for these asymmetries by arguing that in a situation of language mixing, one language will function as the ML, providing a structural frame as well as all system morphemes. Content morphemes from the EL may

subsequently be fitted into this frame. This is formulated as the following two principles, upon which the MLF model rests.

(25) **The Morpheme-Order Principle:** In ML + EL constituents consisting of singly-occurring EL lexemes and any number of ML morphemes, surface morpheme order (reflecting surface syntactic relations) will be that of the ML.

(26) **The System Morpheme Principle:** In ML + EL constituents, all system morphemes which have grammatical relations external to their head constituent (i.e. which participate in the sentence's thematic role grid) will come from the ML.

(Myers-Scotton, 1993: 83)

The MLF model succeeds in capturing the crucial empirical observation about language mixing: One language, the ML, typically does determine word order and provide functional material. This is further supported by data from mixing of different language pairs, in Myers-Scotton's own work (1993, 2002), as well as other studies (e.g., Kamwangamalu, 1997; Muysken, 2000, Türker, 2000).

These empirical contributions notwithstanding, the MLF model shows theoretical weaknesses. First, a crucial shortcoming is that it takes ML and EL to be theoretical primitives. This means that the language faculty must be able to identify individual languages in order to provide them roles as either ML or EL in a mixing situation. Moreover, the proposed principles in (25) and (26) are more descriptions of the observed empirical patterns than actual theoretical principles (see, e.g., MacSwan, 2005a; Åfarli, 2015a; Åfarli & Subbarao, in press, for discussion). In sum, the MLF model highlights a crucial, empirical asymmetry in language mixing, but the model itself is not adequate for analyzing these patterns.

4.2.4. A null theory or constraint-free approach to language mixing

As mentioned in the introduction to this section, two main approaches to language mixing can be identified from a formal perspective: one assuming that language mixing requires some special machinery in our language competence, and one holding that language mixing is constrained by the same restrictions as unmixed speech. Within the field, there is currently a general consensus that one should pursue the latter approach, which is typically referred to as a null theory (Mahootian, 1993) or constraint-free approach to language mixing (MacSwan,

1999, 2000, 2005a, 2009, 2014; see also Lipski, 1985; Di Sciullo et al., 1986; Belazi et al., 1994; González-Vilbazo & López, 2011, 2012). A leading advocate for such an approach, Jeff MacSwan, maintains the view that "[n]othing constrains CS apart from the requirements of the mixed grammars" (MacSwan 2014: 18). In other words, CS is produced by exactly the same mechanisms as monolingual speech. González-Vilbazo and López (2011) support such an approach, arguing that language mixing is an expression of the linguistic competence, i.e., the I-language, of the bilingual speaker, just like unmixed speech. These approaches thus establish the goal of developing a theory and a model that can account for both mixed and unmixed language production using the same principles.

As also discussed above, different constraints have been proposed for language mixing, and according to MacSwan (2009, 2014) even some theories that claimed to be constraint-free struggled to propose analyses that do not rely on special machinery for mixing. In fact, MacSwan argues that although the early approaches apparently promoted a constraint-free approach, they made explicit reference to the languages involved or argued that the linguistic competence of bilinguals involves some mechanism instructing the mixing (MacSwan, 2014). He finds this problematic because "grammars are formally blind to the languages they generate", and thereby also blind to when something is CS or not (MacSwan, 2014: 4). The early approaches made use of different strategies for getting past this problem, for instance introducing a language feature or language index (cf. Di Sciullo et al., 1986; Belazi et al., 1994), but neither is easily motivated if not by cases of language mixing. Such CS-specific constraints were also the topic of the polemic between MacSwan (2000, 2005a, b) and Jake, Myers-Scotton, and Gross (2002, 2005), where MacSwan discarded the MLF model as an inadequate analysis of language mixing due to it not adhering to the constraint-free ideal.

MacSwan endorses a lexicalist approach within MP as a constraint-free analysis of language mixing. As discussed in Section 4.1, economy is a central idea in MP, which aims to eliminate all redundant mechanisms. This combined with the fact that a lexicalist approach is independently motivated are factors favoring such an analysis of language mixing. According to MacSwan (2000), parametric differences depend on individual lexical items (a tradition going back to Borer, 1984), and learning a language means learning its lexical items. The computational system, on the other hand, is invariant. A bilingual speaker is therefore able to draw lexical items from either lexicon, and the syntax will treat them no differently than if they were drawn from the same lexicon; the lexical items still have to satisfy the mechanisms of feature checking. Licit mixings, or CS in MacSwan's terms, are those that fulfil the interface requirement, despite being drawn from different lexicons.

The advantage of such an analysis is that it does not depend on language-specific or CS-specific mechanisms, thus being a constraint-free approach. Still, it depends on the lexicalist assumption that individual lexical items provide the relevant features, which is contested by the exoskeletal approaches to grammar (cf. the discussion in Section 4.1). Moro (2014) offers a concrete analysis of mixed DPs employing the lexicalist framework as proposed by MacSwan. This analysis is discussed thoroughly in Article 1 of this dissertation, which concludes that the proposal is highly questionable. Without going into the details, the analysis utilizes additional principles for the feature checking process, in order to generate the attested forms and avoid leaving unvalued features behind that would cause the derivation to crash.

González-Vilbazo and López (2011) also assume that no specific rules or mechanisms should be proposed for the language faculty in order to regulate language mixing. However, unlike other constraint-free approaches, they argue that the bilingual competence does not necessarily entail the union of the two monolingual grammars; they contend that "code-switchers may include features drawn directly from Universal Grammar which are absent in the component grammars" (González-Vilbazo & López, 2011: 833). This assumption is based on studies of mixing between Spanish and German containing light verbs that are not present in either monolingual grammar. In their analysis, they employ a DM framework, and as discussed in Section 4.1, in such a framework, the syntax is assumed to operate on abstract functional features made available by UG. Language acquisition is then considered an interaction between the input data and UG, and in order to accommodate a bilingual competence, the speaker may resort to UG ingredients not present in the input grammars, but otherwise available in UG.

The current dissertation continues the quest for a null theory analysis of language mixing, where this process is not constrained by principles different from those for monolingual speech. However, two challenges emerge at the outset. First, as the previous sections have made clear, I am not convinced that the lexical approach promoted by MacSwan is the right one (cf. the discussion in Section 4.1). Second, the empirical observations captured in the MLF model, i.e., that one language typically is more prominent in mixing than the other(s), seems to me to be too essential to be ignored. The crucial task then involves building a bridge between these two prominent accounts of language mixing. The current dissertation proposes that a late-insertion exoskeletal model is able to capture the insights from the MLF model and at the same time abstain from resorting to mixing-specific mechanisms. As discussed in Section 4.1, the specific implementation of this model falls within the DM framework, as does the account proposed by González-Vilbazo and López (2011). However,

the data under investigation here does not shed light on or require language-independent UG mechanisms, so their approach is therefore not discussed specifically. Instead, the prime objective of the current work is to combine the empirical insights from the MLF model with the null theory ideal. I return to a discussion of this in Section 6.

4.3. Norwegian noun phrase structure

The Norwegian DP structure forms the third component of the theoretical background of the current dissertation. The nominal domain constitutes a field of vast research within generative grammar, which flourished especially after the proposal of the DP hypothesis (Szabolcsi, 1983; Hellan, 1986; Abney, 1987). This drew attention to the functional properties of noun phrases and potential parallels with the verbal domain. In accordance with the data under investigation in the current dissertation, the following discussion is limited to the Norwegian DP structure.[21] I start by giving a brief introduction to the components of Norwegian noun phrases, before I turn to some prominent analyses within the generative framework. Towards the end, I briefly compare the Norwegian noun phrase with the English one.

Norwegian noun phrases can be quite complex, containing a determiner, weak quantifier, one or more adjectives, pre- or post-nominal possessive pronouns (the latter being the more frequent), and PPs. A simplified presentation of the available components of a Norwegian noun phrase is given in (27), here with the possessive pronoun realized post-nominally.

(27) [determiner [weak quantifier [adjective [noun [possessive pronoun [PP]]]]]]

The noun itself is inflected for definiteness, number, and gender. These functional features manifest themselves as functional suffixes on the stem, as well as through agreement relations with the other elements of the phrase. Norwegian has three genders, masculine, feminine, and neuter, and their assignment is non-transparent. In other words, one cannot predict a noun's gender purely from its semantic or phonological properties.[22] Number and definiteness, on the

[21] Norwegian DPs are in general similar to those of other Scandinavian languages. Some significant differences are nevertheless evident, for instance their realization of definiteness and case. These properties have been subject to comparative investigations (Delsing, 1993; Vangsnes, 1999; Julien, 2005). A comparison or comprehensive discussion of the Scandinavian systems is beyond the scope of the current discussion. The Danish realization of definiteness is briefly addressed in Section 4.3.4.

[22] Trosterud (2001) proposes a range of semantic, morphological, and phonological rules which he argues can account for 94% of gender assignment in Norwegian. However, in addition to being numerous (43 rules) and quite intricate, these rules are problematic as they do not account for the fact that gender assignment can actually

other hand, have binary values (singular/plural, indefinite/definite), and the interplay between these three features determines the realization of the functional exponents.

In indefinite singular phrases in Norwegian, the noun is accompanied by a pre-nominal article: the masculine *ein,* feminine *ei,* or neuter *eit.*[23] Plurality and definiteness are realized by suffixes on the noun stem, with the definiteness suffix commonly called the definite article. An overview of the available suffixes is provided in (28).[24]

(28)

	Singular	**Plural**	
	Definite	**Indefinite**	**Definite**
Masculine	-en	-ar	-ane
Feminine	-a	-er	-ene
Neuter	-et	-	-a

Definiteness, number, and gender are also reflected in the shape of associated words, i.e., determiners, adjectives, and possessive pronouns. Possessive pronouns are not addressed in the current dissertation, and I will therefore not elaborate upon them here (see Westergaard & Anderssen, 2015 for discussion). Attributive adjectives in Norwegian are typically described as displaying weak or strong inflection. The strong inflection is sensitive to number and gender and is found in indefinite phrases, whereas the weak inflection occurs in all definite phrases. An example is provided in (29), showing the inflection of *stor* 'big'.

(29)

	Strong inflection	**Weak inflection**
Masculine	stor-	stor-e
Feminine	stor-	stor-e
Neuter	stor-t	stor-e
Plural	stor-e	stor-e

Norwegian determiners are also sensitive to gender and number, where *den* is used for masculine and feminine, *det* for neuter, and *dei* for plural. Moreover, a characteristic of

vary across different dialects in Norway, meaning that the gender of a noun in one dialect need not be the same as in a different dialect (see, e.g., Enger, 2009). Rodina and Westergaard (2013, 2015a, b) also find the (abundant) collection of rules problematic from an acquisition perspective since many of the rules cover small groups of nouns that are typically infrequent in a child's input.

[23] Notice that I, here and in the following, use examples from *Nynorsk,* which is one of Norwegian's two written standards. The second written standard is *Bokmål.* See Venås (1993) or Vikør (1995) for discussion.

[24] In the literature, there is a discussion regarding whether the functional suffixes actually express gender or rather declension class (see, e.g., Enger, 2004; Lødrup, 2011). I assume that gender is a feature of the suffix.

Norwegian noun phrases is the phenomenon known as double definiteness, i.e., the simultaneous realization of definiteness in the determiner and the functional suffix. This occurs in cases involving an attributive adjective or a weak quantifier. Some examples are provided in (30).[25]

(30) a. den raude bil-en
the.M red.W car-DF.SG.M
'the red car'

b. det gule hus-et
the.N yellow.W house-DF.SG.N
'the yellow house'

c. dei to jakk-ene
the.PL two jacket-DF.PL.F
'the two jackets'

Double definiteness has been an important and recurring topic in the analyses of DPs in Norwegian (and Scandinavian in general), which I will discuss in the following sections.

4.3.1. Early analyses of Scandinavian DPs

Two early and central generative studies of Scandinavian noun phrases are those of Delsing (1993) and Vangsnes (1999). These works were both inspired by the relatively newly proposed DP analysis, and thus search for parallels between noun phrases and clause structure. As a consequence, they are primarily concerned with the higher levels of the DP structure, i.e., the role and position of determiners, adjectives, and quantifiers. Moreover, both studies were conducted at a time when the lexicalist approach to grammar dominated the field. Certain grammatical properties, such as gender, were therefore considered inherent features of the noun itself. Double definiteness is a key topic in both works, as the analysis of this phenomenon in Scandinavian has the potential to shed new light on the structure of noun phrases in general.

Delsing (1993) conducted a comparative study of Scandinavian DPs within the Principles and Parameters framework. A key proposal within his work is the head raising parameter, accounting for different realizations of definiteness:

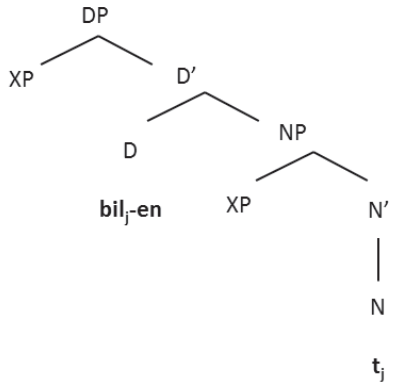
[25] A similar pattern without the attributive adjective or weak quantifier can be found when using a demonstrative, e.g., *den bil-en* 'that.M car-DF.SG.M', *det hus-et* 'that.N house-DF.SG.N'. The exponents appear identical, but importantly *den/det* are always stressed when occurring as demonstratives.

> I have argued that the suffixed article in Scandinavian implies that there can be head movement in the noun phrase, i.e. raising of N to D. I assume that the difference between languages with head raising (like Scandinavian) and languages without it (like English and German) is stated as a parameter, *the head raising parameter.* (Delsing, 1993: 77)

The structural representation of the singular, definite *bilen* 'the car' under this account is shown in (31), where N has raised to D and received the suffixed article.

(31) a. bil-en
car-DF.SG.M
'the car'

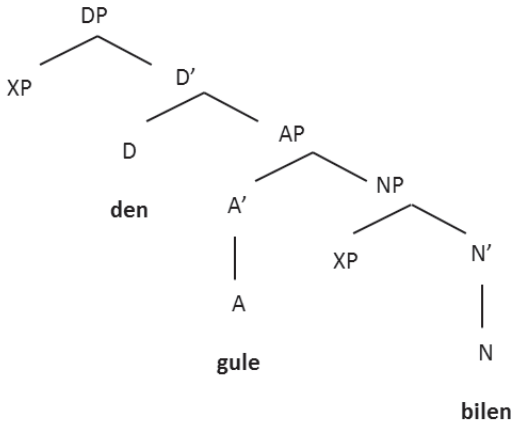
b.



Delsing analyzes attributive adjectives as heads in their own projection (AP), as proposed by Abney (1987). However, diverging from Abney's analysis, Delsing proposes that NP is generated as the right-hand specifier of the adjective (in order to account for attributive adjectives that take objects) (Delsing, 1993: 81). Important for the present discussion is the fact that an adjective in this position will block head raising, as in (31b), by virtue of the Head Movement Constraint (Travis, 1984). Instead, the separate realization of D by an article is required, yielding double definiteness. An example is presented in (32).

(32) a. den gule bil-en
the.M yellow.W car-DF.SG.M
'the yellow car'

b.



The outcome of such structures uncovers internal differences among the Scandinavian languages: Norwegian, Swedish, and Faroese have structures with double definiteness, as described above, whereas Danish and Icelandic do not have this phenomenon. The corresponding Danish phrase to (32a), for instance, would be *den gule bil.* Delsing's (1993: 127–130) proposal is that the suffixed article may be base-generated in N, and that this variation is parametric in nature. Danish, for instance, is assumed to be a language without a base-generated suffix. Definiteness is instead provided by D. In cases with an intervening adjective, the noun is unable to raise to D, and will therefore occur without the suffix, but with a determiner. Norwegian, Swedish, and Faroese, on the other hand, are assumed to have the suffix base-generated in N, and this is the source of double definiteness. In these languages, when an attributive adjective blocks raising, the base-generated suffix will be realized at the same time as D is lexicalized by an article. In cases without an intervening attributive adjective, on the other hand, the effect of this movement will not be "visible" as the N raises to D to lexicalize the D-position.

Vangsnes (1999) introduces a semantic principle of "Identification" as a necessary licensing mechanism for functional categories in the noun phrases, which also serves as a trigger for movement. This mechanism is parallel to feature checking in minimalism, requiring elements from lower in the structure to move up in order to identify functional projections. Vangsnes proceeds to propose three functional projections in the nominal

domain, each correlating with certain semantic classifications of determiners: KP introducing universally quantifying determiners, DP introducing definite determiners, and NumP introducing cardinal determiners (Vangsnes, 1999: 111). Regarding double definiteness, Vangsnes argues that the suffixed article has a dual status across Scandinavian: "I will argue that it is either a clitic element (heading Dx^0) or a part of the (pre-syntactic) inflectional system" (Vangsnes, 1999: 120).[26] Danish and Icelandic are languages with the former type of structure, whereas Norwegian, Swedish, and Faroese have the latter structure. Subsequently, Vangsnes uses this structural difference to account for double definiteness, arguing that DxP cannot have realizations of both its head (the clitic) and its specifier (a demonstrative). Consequently, double definiteness is not possible in the languages that make use of the clitic, i.e., Danish and Icelandic. In languages where the suffixed article is part of the noun's inflectional system, on the other hand, nothing prevents the additional realization of a demonstrative in Spec-DxP (Vangsnes, 1999: 123).

4.3.2. Julien (2005)

The most recent and extensive work on Scandinavian DP structure is that of Julien (2005). Her main focus is on Norwegian, but the proposed structure is arguably compatible with other Scandinavian languages and a cross-linguistic perspective is also included. The analysis incorporates certain ideas from the MP, e.g., that movement is triggered by features, but otherwise adopts a non-lexical approach to grammar. Julien (2005) argues that previous analyses of DP structure have not been able to accommodate all possible components of a Norwegian DP, which then constitutes a core motivation for her work. She takes as her starting point the following example, which shows the maximal expansion of the Norwegian noun phrase (Julien, 2005: 1):

(33) dei to gaml-e teikning-a-ne mine av byen
the.PL two old-W drawing-PL-DF my.PL of town-DF.M.SG
'my two old drawings of the town'

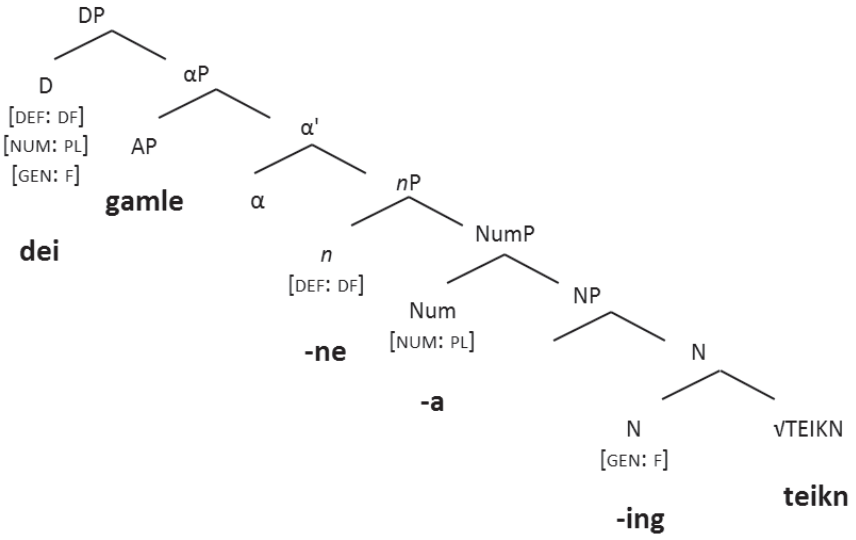
In this phrase, the noun *teikning* 'drawing' is accompanied by a determiner, weak quantifier, adjective, possessive pronoun, and a PP. Moreover, the noun itself is inflected for definiteness, number, and gender, which establishes an agreement relation with the other

[26] See Faarlund (2009) for a discussion of the historical development of definiteness markers in Norwegian, and their roles as grammatical words, clitics, or inflectional affixes.

items of the phrase, like the determiner, adjective, and possessive pronoun. My main interest lies in the noun itself, as well as the determiner, and the realization and agreement of definiteness, number, and gender. In what follows, I therefore limit the discussion to *dei gamle teikningane* 'the old drawings'. A structural representation for this phrase is the following:

(34) a. dei gaml-e teikning-a-ne
the.PL old-W drawing-PL.DF
'the old drawings'

b.



The representation in (34b) is based on Julien's (2005) proposed structure, including the labels she uses. Thus, this structure differs somewhat from the one employed in the analyses in the current dissertation. I return to a discussion of these differences in Section 6.2.4. Here, I address the structure in (34b), in which I have added the relevant features, in brackets, as well as the phonological exponents, in bold. Taking a non-lexical approach, the functional features are provided by the structure, and their phonological exponents are inserted through Vocabulary Insertion (cf. Section 4.1).

The derivation of this structure starts by building the nominal stem *teikning* from a root (√TEIKN) and a category-defining head, a nominalizer N, which in this case has an overt

realization. Grammatical gender is considered a feature introduced by the nominalizer, meaning that gender is a property of the stem, not of the root (Julien, 2005).[27]

Inflectional markers are added in the course of the derivation, and the fact that number and definiteness end up as suffixes on the nominal stem must mean that the noun (stem) moves up and combines with functional projections supplying it with these components (cf. The Mirror Principle; Baker, 1985). Thus, above the stem, Julien (2005) proposes two functional projections, NumP, providing a number feature, and *n*P, providing a definiteness feature. Whereas NumP, or another functional projection introducing number, is commonly assumed for various languages, the latter is a more novel proposal. The key motivation for *n*P is double definiteness, and by proposing a lower functional projection holding a definiteness feature in addition to D, Julien (2005) is able to account for this double representation of definiteness. I return to this presently.

Above *n*P, the structure has two additional functional projections, αP and DP, housing associated adjectives and determiners.[28] Concerning adjectives, Julien (2005) contests the previously proposed analysis with adjectives as heads in the extended nominal projection, as in Delsing's (1993) analysis in (32) above. This is problematic, she argues, considering that adjectives can have arguments and be modified by degree elements. Thus, Julien analyzes adjectives as heads of their own projection, AP, which, importantly, is merged in the specifier position of a designated functional projection, αP (cf. Cinque, 1994).[29] Adjectives are not discussed in detail in this dissertation, but constitute a relevant part of the DP concerning agreement. In general, I assume that adjectives are also built from a root and an adjectival categorizer. Moreover, this stem would be associated with a functional projection ensuring agreement through a probe–goal relation with valued features lower in the structure.[30] In (34b) above, AP is placed directly into Spec-αP without further elaboration for ease of exposition.

---

[27] Motivation for not placing gender higher in the structure, for instance as a feature of Num, comes from cases with deviant number marking. For instance, plurality on feminine nouns in Norwegian is typically realized with the suffix *-e*, whereas it in some cases, like with *teikning* above, is realized by the suffix *-a*, which is typical for masculine nouns. Despite such differences, feminine gender is reflected in associated elements. Thus, Julien (2005) concludes that gender must be fixed before Num is added.

[28] Julien (2005) proposes additional layers above DP to introduce strong quantifiers and demonstratives. These are not crucial to the current analyses and are therefore not discussed further. In the articles contained in this dissertation, a few cases involving a demonstrative occur, and the demonstratives in these cases are for convenience analyzed as realizations of D.

[29] Similar to adjectives, weak quantifiers are analyzed as heads of their own projection, WQP, merged in the specifier position of a functional projection in the extended nominal projection, CardP (Julien, 2005).

[30] This means that language mixing may also take the form of (e.g.) an English adjectival stem being incorporated into a Norwegian structure, and in fact, a couple such examples are attested and commented on briefly in Article 2. A thorough investigation of language mixing in the adjectival domain is, nevertheless, left for future research.

On top of the structure is a DP layer, headed by D, which contains an unvalued feature bundle of definiteness, number, and gender. These features receive valuation through an agree relation with the features lower in the structure. Moreover, the DP has some additional restrictions concerning its realization. Without going into an elaborate discussion, Julien (2005) finds that the Norwegian data essentially corroborate Longobardi's (2001) analysis, in which nominals that function as arguments (specifically referring, not existential or generic) are always introduced by an overtly realized D domain. In other words, there must be some overt material in D or Spec-DP to ensure that the DP is referential and thus can be an argument. In Norwegian, this requirement can be fulfilled in two ways: Either the *n*P moves to Spec-DP or a separate determiner is inserted in D. The former of these alternatives requires phrasal movement, and is common in definite noun phrases not modified by adjectives (or weak quantifiers). However, when the phrase contains a modifier, for instance an adjective, sitting in the specifier position of a functional projection, *n*P is blocked from moving to Spec-DP. In order to still fulfil the requirement for overt realization in the DP domain, D is realized by a separate determiner, yielding double definiteness.

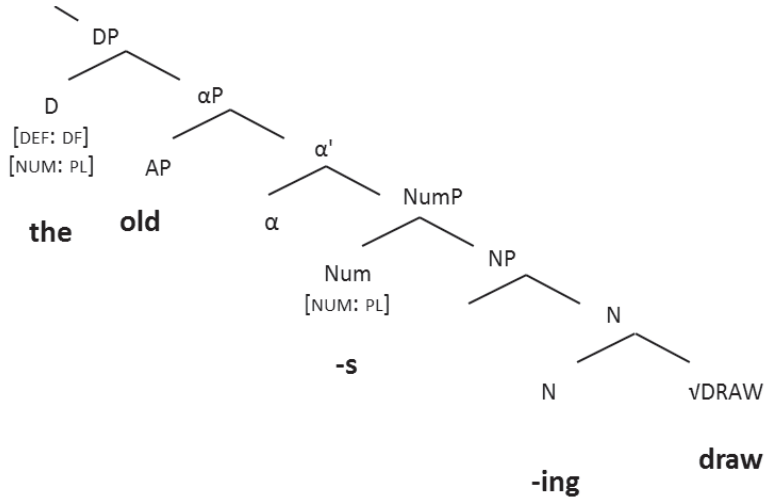
In the analyses conducted in the current dissertation, a modified version of Julien's (2005) model has been employed. Briefly summarized, the changes concern the arrangement of definiteness, number, and gender, which are compressed in a feature bundle contained in one functional projection, FP, above the stem. I will elaborate on this projection and its motivation in Section 6.2.4.

4.3.3. Norwegian versus English DP structure

The data under investigation in the current dissertation show how English and Norwegian are mixed within noun phrases. This has proven to be an advantageous domain in which to study the patterns of language mixing: English and Norwegian noun phrases are not radically different, but certain structural differences make the comparison fruitful. Of special interest in the current work are the categories of gender and definiteness. Whereas Norwegian has three genders, a suffixed definite article, and double definiteness, English has no grammatical gender on nouns, and realizes definiteness solely by a prenominal determiner. In this section, I will briefly introduce the English DP and its potential structure. Importantly, however, the limitations of the dissertation do not allow an extensive discussion of English, and I will focus on the comparison with the Norwegian DP as presented above (cf. Julien, 2005). In the interest of this comparison, I take as my starting point the phrase *the old drawings,* corresponding to the Norwegian phrase discussed above. A potential structure is presented

below. Again, this structure is based on Julien's (2005) proposal, as in (34b), and I have added valued features in brackets and phonological exponents in bold.

(35)



Similar to Norwegian, we can assume that the derivation starts by combining a root with a category-defining head, i.e., a nominalizer, which is realized overtly by *-ing*. An important difference from Julien's (2005) Norwegian structure is that the nominalizer in this case does not hold a gender feature, as we can assume that this feature is not active in English (cf. the discussion of feature selection from a universal assortment above).

Number is relevant for both Norwegian and English, and is furthermore realized as a suffix in both languages. Thus, a functional projection NumP is generated above the stem (see also Ritter, 1993; Alexiadou, Haegeman, & Stavrou, 2007), and its exponent, here *-s*, will be suffixed to the stem after movement. The additional functional projection *n*P in Norwegian is motivated primarily by double definiteness, and it introduces a definiteness feature below the DP layer and adjectives. As English does not have this construction, there should be no need to assume such an additional functional projection here.[31]

[31] Notice that the structure in (35) differs slightly from the structure used in analyses of English DPs in Article 1 in this dissertation. In Article 1, a functional projection FP is instead proposed above the stem, in parallel to the analyses of Norwegian DPs in the same article. The FP projection was mentioned in Section 4.3.2 and is discussed further in Section 6.2.4. What is important here is the fact that the FP in an English DP is assumed to contain only a number feature and not a definiteness feature, which would be present in a Norwegian structure. The difference between the English DP in Article 1 and in (35) may therefore be reduced to their labels: The functional projection is labeled NumP in (35) primarily to be directly comparable to Julien's (2005) proposed structure for Norwegian, as shown in (34b).

In the higher projections of the structure, similar to Norwegian, adjectives and weak quantifiers are presumably generated in specifier positions of functional projections. On top of the structure is the DP layer, which holds (at least) a definiteness feature and a number feature. The former seems uncontroversial, and the realization of D will typically be a definite or indefinite article. In relation to number, demonstratives and indefinite articles provide evidence for such a feature in D. According to Haegeman and Guéron (1999), indefinite articles (*a/the)*, possessive pronouns (*my, your,* etc.), as well as demonstratives (*this/that)* are in complementary distribution in English DPs. Possessive pronouns and the definite article are invariant with respect to number, thus requiring only a definiteness feature. Demonstratives, on the other hand, do vary according to number, and indefinite articles precede singular count nouns but are not realized in indefinite plural cases. This argues in favor of a number feature in D, which will be valued by the feature of Num through a probe–goal relation.

In general, then, we can say that Norwegian noun phrases appear to be more complex than English ones. Especially relevant for the current discussion, is the gender feature as well as the realization of definiteness. The differences here make noun phrases a good domain for investigating the "contribution" of each language in mixing between English and Norwegian.

4.3.4. Danish DPs and lack of double definiteness

The structure of Danish noun phrases, and more specifically their realization of definiteness, constitutes an interesting supplement to the current discussion. Like Norwegian, Danish realizes definiteness as a functional suffix on the noun, but in contrast to Norwegian, Danish does not have double definiteness. In phrases involving an attributive adjective or weak quantifier, Danish realizes definiteness only by a prenominal determiner, more like English (Lundskær-Nielsen & Holmes, 2010, see also Embick & Noyer, 2001, and Hankamer & Mikkelsen, 2002, 2005). Thus, concerning definiteness, Danish can be considered to occupy an intermediate position between English and Norwegian.

Omission of the functional suffix in contexts commonly requiring double definiteness is in fact possible in Norwegian too. Such structures were more common in older Norwegian and may today be conceived as archaic or as part of a specific stylistic choice. According to Faarlund, Lie, and Vannebo (1997), these forms are more common in certain contexts or fixed expressions, like *Det kvite hus* '(the) White House', and beyond that, phrases without definite suffixes are primarily found in written texts and considered to be quite formal. A potential reinforcing factor is the fact that Norway was in a union with Denmark for more than four

hundred years prior to 1814. During most of this time Danish was the official language in Norway, which presumably contributed to the persistence of these forms.

In relation to the AmNo data under consideration in the current dissertation, phrases where definiteness is only realized by a determiner constitute interesting cases. Should they be considered grammatical realizations following the stylistic pattern similar to Danish, or rather cases of attrition where a functional suffix is lost? On the one hand, the Norwegian that these speakers have acquired stems from the immigrants of the 1800s and has existed in surroundings quite shielded from the developments in contemporary non-heritage Norwegian. Thus, archaic forms, like the omission of the definite suffix, might enjoy better conditions for survival in the AmNo community. The church has moreover been crucial in preserving AmNo, and the liturgy typically displays a formal style. On the other hand, however, AmNo speakers are typically not literate in Norwegian, and as the grammatical omission of the definite suffix is mostly associated with written texts, it can hardly be considered an influence from this specific style. This topic is discussed briefly in Article 3.

5 Data and methodology

In the case of formal linguistics, the object of inquiry is not available for direct observation. In fact, the main object of inquiry in generative grammar, I-language, can only be studied through its manifestations, E-language (Chomsky, 1986). By investigating patterns in language production, linguists draw generalizations and propose models of the underlying capacity. This makes theoretical models necessarily hypothetical, proposed in order to answer the question of how the innate language capacity must be designed in order to characterize the observed phenomena.

The articles contained in the current dissertation exploit two specific corpora for obtaining data: Haugen (1953) and CANS (Johannessen, 2015a). These sources, like corpora in general, provide specimens of naturally occurring speech, which in the next step are used to test a theoretical model of grammar, the exoskeletal model. In addition, the data are used to demonstrate that this specific model is preferable to a different model of grammar, the lexicalist model. The current section contains a presentation and discussion of the methodology of the dissertation. I start out with a general, but brief, discussion of using corpora and theoretical models in linguistic research. Thereafter, I turn to the specific data and method used here. Finally, I discuss some methodological considerations concerning the available AmNo material.

5.1. General considerations

The methodology of three articles contained in the current dissertation all involve corpus data and testing of theoretical models of grammar. In general, corpora provide large collections of naturally occurring speech, and the large amounts of data they make available are advantageous for linguistic studies, enabling a range of different investigations (Gries & Newman, 2013). Corpora are typically intended to be representative of a population, and questions of interest may concern the relative frequency of specific linguistic phenomena in the population. In the current investigations, however, corpora are primarily used as a source of data, not to pin-point the frequency of the various linguistic phenomena. Rather, the main function is providing insight into how a linguistic phenomenon, in this case language mixing, appears in an actual context. In other words, corpora are here used to extract naturally occurring specimens of language mixing, whereas studying variation within the AmNo population as a whole is left for future research.

An essential benefit of using corpora is the large amount of natural data that they make available. A drawback, however, is the lack of negative data or information beyond the data itself, for instance acceptability judgments or the speakers' own evaluation of the material. Traditionally, these have been important sources of data in generative grammar, as gaining access to both acceptable and unacceptable utterances is useful in finding the limits of grammar (Schütze, 2011). The lack of such judgements thus makes it harder to infer conclusions concerning possible restrictions on language mixing. The use of acceptability judgements in cases of language mixing is, nevertheless, not unproblematic. Language mixing could in many communities be a stigmatized form of communication, and moreover, normative notions about what constitutes "proper language" could influence such judgements, potentially causing "negative over-reporting" (González-Vilbazo et al., 2013, see also Bauer, 2014). Thus, when using corpora, one cannot be conclusive concerning the limits of language mixing, but corpus data are still useful in studying language mixing, as they capture natural and spontaneous speech, which may be less influenced by social norms than data obtained through other means.

In the current dissertation, the cases of language mixing within noun phrases in the corpora have been used to provide a general description of the observed patterns and to support a specific model of grammar. In formal linguistic research, theoretical models and testing hypotheses are significant components. The models in question are meant to be explanatory and aimed at providing insight into the structure and mechanisms of the language faculty. Moreover, they are hypothetical and abstract descriptions of the structure of grammar, and the goal of testing is to show how one model is preferable to the alternatives (see, e.g., Beavers & Sells, 2013; Zuidema & de Boer, 2013). Continuous testing against new data will eventually corroborate or falsify theoretical models. For instance, the analyses of language mixing conducted here provide evidence in favor of an exoskeletal analysis of grammar.

As mentioned above, the common factor of the three articles in this dissertation is that they make use of corpus data to test theoretical models. Yet, they differ in the details of their methodology and their approaches to the available material. Article 1 is primarily an investigation of two theoretical models and how well they predict the observed patterns of language mixing. The article takes the predictions made by the different frameworks and examines how well they fare when tested against various types of language mixing.

Article 2 is a cross-sectional study of language mixing in contemporary AmNo. It conducts a synchronic investigation of language mixing patterns within noun phrases based on data available in CANS (which is discussed in detail below), and provides a descriptive

overview of the typical patterns. Subsequently, an exoskeletal model is used in analyses of these data.

Article 3 offers a diachronic perspective on AmNo, comparing early and recent material. The time dimension in linguistic studies can be captured in different ways (see, e.g., Blondeau, 2013). A direct method would be a longitudinal study, ideally following individual speakers over time. This is, however, not always possible. For instance, the current situation of AmNo, as a moribund language, makes a longitudinal study practically impossible. A pseudo-longitudinal study is an alternative, indirect approach where the apparent time dimension is captured by splitting the speakers of a cross-sectional, synchronic study into different age groups. Again, in the case of AmNo, this is problematic, as today's speakers all belong to the same age group. However, it is still possible to study AmNo diachronically. The third article in the dissertation offers a comparison of material collected at two different points of time, i.e., comparing material collected in the 1930s and 1940s (Haugen, 1953) to data collected after 2010 in CANS (Johannessen, 2015a). The connection between these two groups of speakers is that they are all immigrants or descendants of immigrants from Norway prior to 1920. Data from the two corpora I examine can therefore be considered two "snapshots" of the AmNo language at two different points of its development. This work therefore represents a "trend study" aimed at investigating language variation in a certain community over time (Blondeau, 2013). Such a methodology is commonly associated with sociolinguistic research, and ideally in such a case, the later study (i.e., CANS) would exactly replicate the original study. Although CANS is not a direct replication of Haugen (1953), it provides a significant follow-up perspective. Through returning to many of the same communities that Haugen investigated, the American Midwest, CANS provides data which can be said to represent a later step in the development of the AmNo language than the one documented by Haugen (1953).

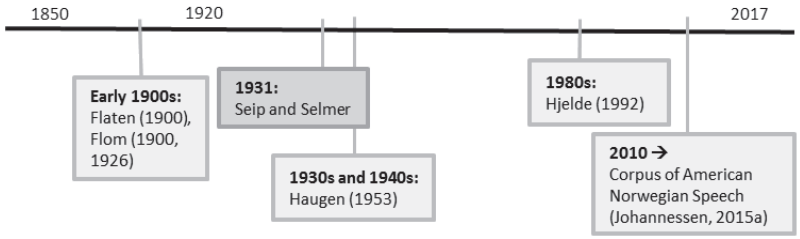
In the following subsections, I will present and discuss the different sources of AmNo data used in the current dissertation as well as how I have proceeded in working with them.

5.2. Sources of AmNo data

This section introduces the sources I have used for extracting AmNo data: what they contain, how the source material was collected, and how they are employed for the purposes of the current dissertation. AmNo has been a language of interest for researchers for more than a century, and data have been collected and the language studied at different points of time

during this period. The following timeline (reproduced from Article 3) provides an overview of the main data collection efforts.

(36)



As discussed in Section 3.2, articles containing lists of English words occurring in AmNo were published already in the early 1900s. These word lists are not proper corpora, but they still provide valuable data concerning the language mixing practices in early AmNo. The first major data collections were conducted by Seip and Selmer in 1931, but, unfortunately, a large proportion of these data has been lost, and what remains is rather incomprehensible. Later data have been collected through three main efforts. Einar Haugen collected data in the 1930s and 1940s, Arnstein Hjelde in the 1980s, and most recently new collections were initiated in 2010 and are still ongoing at the time of writing (summer 2017). These most recently collected data are gathered in CANS, created by the Text Laboratory at the University of Oslo.

In the current dissertation, Haugen (1953) and CANS have been the primary sources of data. CANS in particular has been the main source, exploited for data in all three articles. In addition, Flaten (1900), Flom (1900, 1926), and Hjelde (1992) have been consulted in relevant instances, especially in the diachronic study in Article 3, where this material contributed to a richer foundation for investigating the development of AmNo. In the following subsections, I introduce the two primary sources, Haugen (1953) and CANS, in order to show the kind of data these sources comprise and the methods by which they were collected. These subsections also include a discussion of how I have approached the two corpora in order to extract data for the current dissertation. The purpose of these discussions is to provide the most complete and transparent methodological description possible. Moreover, due to some methodological challenges, which I will return to in Section 5.3, the descriptions of working with CANS are quite detailed.

5.2.1. Haugen (1953)

Haugen conducted his field work during the years 1936 to 1948 and interviewed a total of 260 speakers. The informants were recruited mainly through local contacts, and the majority came from Wisconsin. Haugen primarily tried to find speakers born in the US, but his group of informants also included speakers who had themselves emigrated from Norway.

Haugen's main tool for collecting data was a rather extensive questionnaire. This questionnaire included a range of questions covering topics like home and family life, farming, travel and communication, and human relationships (Haugen, 1953: 323–324). Haugen also made sure to include concepts which usually would be expressed in English in order to elicit English words and thereby measure the influence of the language. These words were primarily selected from the word lists in Flaten (1900) and Flom (1926), but also based on Haugen's own experience of being an AmNo speaker (Haugen, 1953: 324). The complete questionnaire turned out to be lengthy, to say the least, and the dimensions of the task can be captured in Haugen's comment about one revision: "In all, 278 items were dropped, and 35 were added, which reduced the time required to complete interviews from 11 or 12 hours to about eight" (Haugen, 1953: 328). However, not all informants were subject to the full-length questionnaire, as Haugen separated them into groups of primary, secondary, and occasional informants, depending on the amounts of data they contributed. An example from the questionnaire is shown below.

(37)

MORPHOLOGICAL SYSTEM (1) NOUNS *Page 2.*

Name..... Dialect..... Amer. Community..... Date....

(1) *Nouns*		Singular			Plural		
		Indef-inite	Def-inite	Dative	Indef-inite	Def-inite	Dative
Masc. 1.-a	kalv hund, by kniv (rygg)						
2.-i	sau sekk(benk) vegg(legg) (gris)*						
3. cons.	fot, bror (nagl) (mann)						

Facsimile from Haugen (1953: 646)

As can be seen from the sample questionnaire, Haugen sought to identify full inflectional paradigms for a range of words, and in this way also to provide comparable material from various dialects. The interviews took place orally and were somewhat like a conversation. Moreover, they were conducted in Norwegian, which is described as the informants' preferred language.

In addition to the questionnaires, Haugen took notes and made recordings of most informants, since, in Haugen's (1953: 479) own words: "No study of a language is complete without texts to illustrate its living usage". These constituted part of Haugen's primary sources, but he also selected 31 specimens out of the various recordings to include in their entirety in his book (Haugen, 1953: 479ff.). These were selected in order to be representative of the dialects in the AmNo communities and authentic style of speech (Haugen, 1953), and they take form as, for example, stories told by the informant or introductions to social customs in the communities. These specimens are transcribed and supplied with English translations, and English items are indicated.

Haugen's material is available in his two-volume book *The Norwegian Language in America* (1953), which describes both the AmNo community and language. Some recordings are also available online, although they are not transcribed and not of the best sound quality.[32] However, the discussions in the book are detailed and thorough, thereby compensating to some extent for the shortage of recordings. Of special interest for studying language mixing, the volumes contain a detailed description concerning the grammar of registered English loan words. Moreover, a vocabulary of English loans and their grammatical properties is provided, as well as the above-mentioned specimens of the informants' free speech. Hence, Haugen's written material constitutes a solid source of AmNo data, and the recordings have therefore not been considered in this dissertation. An additional remark concerning the vocabulary should be made. Haugen (1953: 556) only includes 10% of the English loan words in the selected vocabulary. Moreover, as the conversations were not recorded in their entirety, the rest of the loans remain out of reach for further research. Thus, it is not possible for subsequent researchers to study the entirety of the material collected by Haugen. However, for the current work, Haugen's own descriptions of AmNo grammar, supplemented by a selection of examples, are considered a substantial and reliable source of information on the AmNo language and community, and sufficient for the purpose of the current work.

[32] See http://www.tekstlab.uio.no/norskiamerika/opptak/haugen.html [accessed: July 27, 2017].

5.2.2. Corpus of American Norwegian Speech

The most recent effort to collect AmNo data started in 2010 as an extension of an ongoing project collecting Norwegian dialect data (Johannessen & Salmons, 2012, 2015). The initial recruitment of speakers started in 2009 through ads in American Norwegian newspapers, which sought Norwegian-speaking Americans who were the descendants of immigrants who arrived in America prior to 1920, and who had acquired Norwegian at home. This resulted in 40 responses, and the first field work was conducted in the spring of 2010 ("Norsk i Amerika", 2017). Material was gathered through recorded interviews with the speakers as well as conversations between pairs of speakers, and in the subsequent work, these recordings were transcribed and published in an online corpus. In the years since 2010, even more informants have been recruited and more data have been gathered, and the corpus is expected to expand in the future.

At the time of gathering data for the purpose of the current dissertation, the corpus consisted of recordings of 50 individual speakers.[33] These speakers came from 22 different places in the US and Canada, primarily from areas in the American Midwest. As mentioned in Section 3.2, most speakers were between 70 and 100 years of age at the time of recording and they differed widely in their usage of AmNo. Some had used the language on a daily or weekly basis throughout their lives, whereas others had not spoken it for many years. Hence, their competence differed significantly, but all speakers were still relatively fluent (Johannessen & Salmons, 2012).

CANS is available for research as an online, searchable corpus. Sound and video files are provided, and the recordings are transcribed on two levels: a broad phonological transcription and an orthographical transcription into the Norwegian written standard *Bokmål.*[34] These transcriptions are tagged for different grammatical categories, thus enabling specific searches (see Kåsen, Olsen, Rødvand, & Tengesdal, 2016 for the procedures used to tag and transcribe the material). Of specific interest for the current dissertation is the tag "x", which labels all items that are not listed in the standard Norwegian *Bokmål* dictionary. This tag is relevant as there is no tag specifically targeting English words in the corpus.[35] I have instead captured these data manually sorting the English words from various dialect words, etc. found through searches for the tag "x".

[33] According to Khayitova (2016) the corpus contains approximately 182,000 words.
[34] As already mentioned in footnote 23, Norwegian has two written standards, *Bokmål* and *Nynorsk.* The former is used by the majority of Norwegians.
[35] A separate tag "+eng" is available, but this is used only to label larger English segments, consisting of more than one word (Kåsen et al., 2016).

My work collecting relevant data for the present dissertation has by and large involved extracting noun phrases containing an English noun from the corpus. One caveat is, however, in order before I proceed. Since it is an online corpus, CANS may be subject to readjustments, additions, and other updates, and in fact, as of early 2017 the search interface has been thoroughly updated, subsequent to the work presented here. I have conducted some random searches after this update, and these have shown that the results of specific searches might have been impacted. It is therefore necessary to stress that the data set investigated in the current dissertation was extracted on April 20th, 2016, and later updates have not been taken into account. Potential mismatches are thus unfortunate, but beyond my control. Due to these challenges, my procedures for identifying and collecting data will be presented step by step in what follows.

Since searching directly for English items is not an option, the tag "x" constituted the starting point for my purposes, providing all elements not listed in the Norwegian *Bokmål* dictionary.[36] This resulted in a collection of 6145 items which I subsequently exported to a separate spreadsheet. In the next step, these items were manually sorted in order to exclude Norwegian words, English words other than nouns, and other irrelevant items, e.g., various interjections. In cases of doubt, I consulted the sound files and contexts of the items in question. At this point, I had a list of 1610 potentially relevant items. From these, obvious proper names (e.g., *Chicago, Flom, Ford*) were excluded. In the case of compounds, I followed the general rule that the right-hand element constitutes the core of the compound (Williams, 1981). Thus, compounds with a Norwegian right-hand element were excluded even if their left-hand element was English (e.g., *powersag* 'power saw' (chainsaw)*, beerflasker* 'beer bottles'*, roadarbeid* 'road work'). Compounds that consisted of two English items both tagged "x" were counted as one item.

At this point, my collection comprised 1482 English nouns, which were subject to an even finer sorting. Fixed expressions (e.g., *oh boy, in fact*) were excluded, along with meta-comments, for instance where the speaker asks about the translation of an English noun, as in the example below.

[36] Any other approach would entail going through the corpus manually, logging all occurrences of English items in noun phrases. This would be quite time consuming, and frankly a bit strange considering the search options provided. Therefore, at least as a general rule, the tag "x" has been trusted to provide all English items in the corpus.

(38) hva du kall-er herring på norsk? (CANS; westby_WI_03gk).[37]
what you call-PRS herring on Norwegian?
'what do you call herring in Norwegian?'

I consider cases like the one in (38) to be citations, and not cases of language mixing proper. In cases where a word was produced twice, with one occurrence immediately following the other in the same utterance, only one occurrence was counted. An example is given in (39).

(39) han var i navy-en # navy-en (CANS; portland_ND_02gk)
he was in navy-DF.SG.M # navy-DF.SG.M
'he was in the navy # the navy'.

To ensure that this sorting is accurate, I have considered all 1482 sound files and contexts. The purpose of such an initiative was not primarily to capture phonological details, but rather to establish support for the selections. Moreover, as I will discuss in the following subsection, listening to all sound files has also been advantageous in uncovering discrepancies between the two levels of transcriptions. Consulting the sound files also enabled me to distinguish between some English and Norwegian nouns with an apparent lexical overlap (e.g., *school/skole, doctor/doktor*). Instead of excluding all such instances, and potentially throwing out many interesting items, the sound files provided me with clues as to whether these words should be considered English or Norwegian. For instance, in the case of the pair *school/skole* the Norwegian noun typically ends in a schwa, which the English noun does not. Such rudimentary phonological judgements have thus complemented the main sorting. After listening to all the sound files, the resulting collection was 1265 English nouns. These have subsequently been collected in a "mini corpus" and constitute the prime selection of data considered in the current dissertation.

A first relevant question concerns the context in which these 1265 English nouns occur. Seventy-five nouns occur without an immediate context, either due to being single-word utterances, or being surrounded by uninterpretable parts of an utterance. One hundred fifty-six nouns occur in an English context, i.e., an English phrase or larger piece of structure. The remaining 1034 items, constituting the majority of the selection, occur in otherwise Norwegian structures. The grammatical properties of these English items, more specifically the realization of associated grammatical components, have been the main interest of the

[37] The information provided in parentheses here and below is an informant code, identifying the individual speaker in CANS who produced these utterances.

current dissertation. This collection is thus the data source for the grammatical analyses as well as basic descriptive statistics in the three articles.

In addition to this main data selection, I conducted a complementary search (July 29th 2016) targeted specifically at English definite and indefinite articles. The objective of this search was to find whether these articles where used in combination with a Norwegian noun, or otherwise occurring in a Norwegian context.

5.3. Some methodological remarks

The previous subsection presented the main data sources for the current dissertation and how they have been approached. The corpora comprising AmNo provide rich sources of data, but some methodological considerations should also be made. In the following, I will discuss some concerns relating to the observer's paradox, the composition of the two groups of speakers, the available linguistic data, and also the available metadata. All these questions are connected to the process of collecting data, thus beyond the scope of my own influence, but they are worthwhile to comment on as they may have an impact on the data available and subsequently on the possible conclusions we draw from them.

5.3.1. The observer's paradox

A general consideration when collecting data is the observer's paradox (Labov, 1972): Is the speaker performing as naturally as he/she would without the researcher present? As discussed above, a large portion of Haugen's (1953) material was collected through questionnaires. Such a method could, as Haugen himself comments, come across as an examination, and some informants would even ask him whether or not they passed (Haugen, 1953: 336). A relevant question then concerns to what extent the informants' answers to the questionnaire correspond to their actual way of speaking. For instance, concerning the usage of English words, Haugen remarks that the informants were reluctant to admit to such usage in the questionnaires, but that this nevertheless became apparent in their free conversation. Moreover, Haugen reports that some informants became self-conscious when being recorded, and that this "mike fright" could have had an influence on their performance on tape. The specimens of free speech provided and transcribed in Haugen (1953: 479) were thus selected from those informants who seemed to show a more relaxed, everyday style of speech.

The collection of material for CANS took place in a way that one can assume would invite more natural speech: either conversations between two AmNo speakers or an interview

or conversation with the researcher. However, the interview situation could still produce some challenges. For instance, the fieldworkers have reported that the AmNo speakers were eager to meet and speak to Norwegian researchers (Johannessen & Salmons, 2012; Johannessen & Laake, 2017). In such a setting, one can therefore suspect that the speakers potentially wanted to demonstrate their skills in Norwegian, so they may have tried to speak as "correct" Norwegian as possible and mask their mixing practices. Relatedly, one can suspect that the speakers might have been more comfortable mixing languages when talking to other AmNo speakers. In such ways, the recording situation may influence practices like language mixing, and should be considered in the discussion of such data (see also González-Vilbazo et al. 2013).

A related concern is the efforts that the interviewers made in order to improve communication. One challenge that has been reported from the collections for CANS is that the AmNo speakers did not always understand some dialectal forms used by the interviewers, who consequently had to adapt to their interlocutors (Johannessen & Salmons, 2012; Johannessen & Laake, 2017). On other occasions the interviewers themselves would use mixed items in their questions, and such practices are likely to trigger mixing among the informants. Grosjean (1998) describes it as a dangerous strategy not to factor in the interviewers' bilingualism, as this may trigger or hinder language mixing. In CANS, the interviewers' utterances are not immediately accessible when doing simple searches. However, in the sound and video files, a specific clip of the conversation can be expanded, so that one can hear the questions and responses from the interviewer.

5.3.2. Composition of the groups of speakers

The individual speakers in both Haugen (1953) and in CANS differ greatly in their competence, background, and usage of Norwegian. In other words, the group of AmNo speakers is quite heterogeneous, and for research purposes, this comes with some challenges as well as some benefits. On the one hand, in a heterogeneous group it is difficult to find an overall pattern, and general conclusions cannot be drawn based on intensive studies of a few individuals. This was also pointed out as a methodological concern by Haugen (1953: 319–320). On the other hand, such a composition of speakers with a variety of competences will most likely disclose more possible patterns and constructions in the data.[38]

[38] See also Johannessen & Laake (2017), who discuss the development of a common variety across the present AmNo societies.

One significant difference between the two groups, also impacting their comparability, is a generational one. As noted in Section 3.1, a discussion in the heritage language literature concerns whether or not 1st generation immigrants should be included. These speakers have acquired the language in different surroundings, but come to be speakers of a minority language. For CANS, this is not a problem, as none of the speakers are 1st generation immigrants. Haugen's (1953) group of informants, on the other hand, does include 1st generation immigrants, some of whom immigrated as young children and others as adults.[39] The adult immigrants in particular presumably speak Norwegian without much influence from English. Thus, the range of competence within the group is quite broad. This is also noted by Haugen (1953: 45): "The degree of bilingualism in the Norwegian settlements varied greatly […]. Farmers who chose to restrict their contact with the English-speaking world to a minimum could live an entire life here as monolinguals". Consequently, it is not surprising that the speakers in Haugen's material deviate less from the typical Norwegian grammatical patterns than today's speakers.

When comparing these two groups of speakers, as has been done in the third article in the current dissertation, these differences in competence need to be taken into consideration. Haugen (1953) does not explicitly separate the 1st generation speakers from the rest in his discussions of grammatical components, for instance, and instead discusses the group as a whole. The competence of the 1st generation immigrants may therefore boost the results in a positive direction. However, in the list of all informants, provided in the final volume of his work, the generation to which each informant belongs is provided (Haugen, 1953: 618ff.), giving an overview of the members of the 1st generation immigrants: 65 informants belong to this group, and 48 of them were older than 14 when they immigrated. This tells us that in the group of 260 informants in total, the majority are in fact 2nd or later generation immigrants. Consequently, one can assume that Haugen's discussions represent this majority of speakers, who are unquestionably heritage speakers.

[39] In fact, Haugen also included a couple of speakers who were not of Norwegian heritage, but who still were considered proficient enough in Norwegian to be informants for his study.

5.3.3. Tagging, transcribing, and transliterating in CANS

Turning to CANS, some remarks should be made concerning the transcriptions and transliterations provided in the online corpus, the latter referring to the orthographic transcriptions. As mentioned in Section 5.2.2 above, the utterances in the corpus are tagged for various grammatical properties. A first concern here is, as already noted, the lack of a tag specifically targeting the English items in the corpus. Researchers interested in the English influence on present AmNo thus have to employ alternative searches. If one is interested in single items, as in the current dissertation, the tag "x" is the best option. Unfortunately, this search is not entirely accurate as random searches have shown that not all English items are tagged in such a manner. For the purpose of the current dissertation, the tag "x" alone has been used in order to make the methodological approach as transparent as possible, although there are more (untagged) English items in the corpus that have not been taken into consideration. However, as I have used the corpus primarily as a source of naturally occurring examples of language mixing, this remains unproblematic. If one pursues statistical studies of, e.g., frequency, on the other hand, a comprehensive collection of English items would be necessary.

A related tag is "+eng", which is used to indicate longer sequences of English. The instructions accompanying the use of this tag are the following: "Taggen +eng skal 'fjerne' lengre sekvenser av engelsk. Tanken er at korpuset skal være for amerika*norsk* og ikke engelsk", which translates to "the tag +eng should 'remove' longer English sequences. The idea is that the corpus should be for American *Norwegian,* not English" (Kåsen et al. 2016: 11, my translation). Consequently, the sequences in question do not appear unless specifically searching for "+eng". This could be seen as unfortunate since AmNo is in fact a language containing a lot of English items and sequences. Considering that the speakers in CANS are dominant in English, removing English segments from the corpus could mean removing a large portion of potentially interesting data both for formal and sociolinguistic investigations of the AmNo language and community.

A further remark about CANS concerns some of the transliteration practices. As mentioned in Section 5.2.2, the sound files and contexts of 1482 (potential) English items were carefully considered for the purpose of extracting relevant data for the current dissertation. During this process, I compared the two levels of transcription. Ideally, this should not be necessary, but it has proven to be crucial, as the phonological and orthographic transcriptions do not always sufficiently correspond, e.g., in terms of suffixes. Thus, in order

to confidently identify patterns and draw conclusions concerning the grammatical realizations for the current AmNo speakers, the phonological transcriptions need to be consulted.[40]

5.3.4. Metadata

Finally, a brief remark should be made concerning the available metadata in the two corpora. Haugen (1953) provides systematic information about the AmNo society and speakers. His (already extensive) questionnaire also included questions concerning basic demographic facts about the informants and their family, dialect background, and literacy, among other things. Thus, the background information both about speakers and the community in general is thoroughly described in Haugen (1953). A significant drawback is, nevertheless, that Haugen does not always make explicit the link between an individual speaker and a specific utterance. The descriptions of the grammar of English loanwords, for instance, are based on the group of informants as a whole. Studying individual differences is therefore not possible. This is partially compensated for by introducing the individual speakers in relation to the entries in the selected vocabulary of English loans as well as the transcribed recordings of free speech, even though this material only contains a fraction of the material collected. As the investigations conducted in the current dissertation are primarily concerned with the group level, these differences are not crucial for the results. Potential future studies of individual speakers, on the other hand, should give precedence to the individuals whose transcribed recordings are provided.

The metadata provided in CANS has other advantages and disadvantages. In general, the metadata provided for the AmNo communities and the individual speakers appear scarcer. The corpus makes available basic information about speakers, e.g., gender, year of birth, what language they used in school, and how much contact they have with Norway. Additional information is nevertheless not systematically collected. What would have been relevant additional information is, for instance, some remarks on the informants' actual usage of

[40] The origin of these differences can be traced to the instructions for transliteration, which state that this should be conducted in such a way that the result is most compatible with an ordinary Norwegian (*Bokmål*) text. The instructions further suggest that this entails, for instance, adding appropriate suffixes and changing wrong tempus marking on verbs or gender markings on nouns. The rationale for this instruction is that the transliterated version should have a better flow to it and that it improves searches into the material (Kåsen et al. 2016: 12). However, both the decision and its justification are unfortunate as they standardize utterances in a way that could potentially conceal valuable data. For instance, due to the processes of attrition, incomplete acquisition, and cross-linguistic influence, outcomes *not* conforming to the expected patterns are especially interesting, but will be hidden by such transliteration practices. Moreover, using *Bokmål* as a yardstick for the production of the AmNo speakers seems odd as their dialect background (insofar we can say something about it) is not especially close to this standard. Consequently, they will often produce differing forms. Thus, aiming for "ordinary Norwegian" does not ease search procedures.

Norwegian in their daily lives, what generation of immigrants they are, etc. This would be advantageous in the comparison to Haugen's (1953) informants, as well as when discussing differences of competence within the group. Nevertheless, there is one great advantage of CANS, namely that it provides a direct link between the individual speaker and each utterance by that speaker. This enables the researcher to study individual speakers, and thereby opens the door for more in-depth and detailed studies of the AmNo language.

6 Main findings and implications

The core hypothesis of the current dissertation was formulated as follows in the introduction of this cover article:

(1) *Mixing of English and Norwegian in American Norwegian noun phrases is systematic, and a late-insertion exoskeletal model is well suited to capture the empirical patterns.*

Building on this, a twofold goal has been pursued throughout the dissertation. First, I have aimed to describe the empirical patterns of language mixing observed in AmNo noun phrases, and provide grammatical analyses of these patterns. This has been conducted through a cross-sectional study of contemporary AmNo speech, made possible by CANS, as well as by a comparison of this material and data collected by Haugen (1953) in the 1930s and 1940s. In other words, both the synchronic and diachronic patterns of language mixing in AmNo have been investigated. Subsequently, these data were analyzed using an exoskeletal model of grammar. This introduces the second goal of the dissertation, namely to show that the proposed exoskeletal model constitutes an excellent analytical tool for language mixing in AmNo, and moreover constitutes a constraint-free, or null, theory of language mixing.

In the current section, I first summarize the overall findings of the dissertation in relation to the core hypothesis in general and the proposed research questions in particular. This is topic of Section 6.1 below, but as the results are thoroughly reported by the individual articles following this cover article, I will limit the discussion here to the general findings and indications of where a further discussion can be found in the articles. In Section 6.2 I turn to some further implications that can be drawn from the dissertation and finally some proposals for future research.

6.1. Main findings

The core hypothesis above formed the basis for formulating specific research questions. These were split into two groups: those concerning empirical patterns (2), both synchronic and diachronic, and those concerning more theoretical aspects (3).

(2) a. Does language mixing in American Norwegian noun phrases follow systematic patterns?

b. Have the patterns of language mixing in American Norwegian noun phrases changed diachronically?

(3) a. How can the typical patterns of language mixing be formally analyzed in a late-insertion exoskeletal model?
b. How can a late-insertion exoskeletal model account for the patterns of diachronic change?
c. How do the analyses of language mixing in American Norwegian noun phrases support a late-insertion exoskeletal model as a null theory of language mixing?

In the current subsection, I will discuss how these questions are answered by the articles to follow. The empirical questions in (2) are addressed in Section 6.1.1, whereas Section 6.1.2 deals with the theoretical issues concerning analyses of language mixing in general, (3c), and language mixing in AmNo noun phrases in particular, (3a) and (3b).

6.1.1. Empirical patterns of language mixing in American Norwegian

The short answer to the empirical questions in (2) is yes to both; the empirical investigations of the AmNo material show a systematic pattern of language mixing, and certain diachronic changes are also found. The first empirical question is directly addressed in Article 2. The synchronic investigations of the data available in CANS show that English content items are systematically integrated into a Norwegian structure. More specifically, English noun stems occur in AmNo with Norwegian determiners and Norwegian functional suffixes, in a Norwegian constituent order. Interestingly, English noun stems are also distributed across Norwegian's three genders, even though this is an alien category to English nouns. Article 2 provides a rich selection of data illustrating the various patterns, as well as a couple of cases involving adjectives which provide additional support for the general observation that English stems are integrated into the Norwegian structure. Moreover, one unexpected pattern is also uncovered, namely the frequent occurrence of the English plural *-s*. Two generalizations are discussed in relation to this latter pattern: First, it occurs primarily in indefinite phrases, and second, it appears almost exclusively with English noun stems.

The diachronic investigations of AmNo are addressed in Article 3, and two particularly interesting observations are made. First, the overall pattern of language mixing appears to be quite stable. Studies conducted in the early 1900s and throughout the 20th century show that English items occurring in AmNo are provided with Norwegian grammatical components, as

discussed in Article 2. However, a detailed investigation of the categories of number and definiteness reveals systematic diachronic changes in the data. First, certain functional exponents (which are otherwise obligatory in non-heritage Norwegian and in early AmNo, cf. Haugen, 1953) have come to be optional. This is especially apparent in definite contexts where the obligatory Norwegian definite suffix is not always realized where it should be. The pattern is also found in some plural phrases. However, the number of omitted plural suffixes is seemingly low, which might be due to the fact that the resulting bare forms are difficult to identify as plurals without an overt quantifier. In other words, the actual numbers here may be higher, but because these cases can be ambiguous, they are not easily identified. The second pattern of diachronic change is the use of English functional exponents. This concerns the English plural *-s,* as also discussed in Article 2, and the English determiners *a* and *the.* Haugen (1953) noted that the usage of the English plural suffix increased as time went by, and the results of Article 3 actually show that the English plural suffix is now more common than Norwegian plural suffixes. Using the determiner *the* was on the other hand not acceptable to the speakers Haugen consulted, and the usage of the indefinite article *a* was not mentioned in his work. Although these cases are not numerous, the newer material shows that these English determiners may be used together with a Norwegian noun in an otherwise Norwegian context. Both patterns are illustrated with examples in Article 3.

Beyond the concrete questions concerning AmNo, the results of the articles contribute to the study of heritage languages in general. Heritage languages are receiving increasing interest in linguistics, and the many immigrant languages in the US are especially prominent in the literature, cf. the discussion in Section 3.1. These previous studies have typically addressed heritage languages other than Norwegian, for instance heritage Spanish, heritage Russian, or heritage German (e.g., Polinsky, 2006, 2011, 2016; Rothman, 2007, 2009; Montrul, 2008, 2012; Putnam & Sánchez, 2013; Yager et al.,2015). Moreover, many studies have been concerned with general investigations of the more theoretical questions related to heritage language competence, comprehension, and production. More concrete studies have focused on topics like case realization, ellipsis, and grammatical gender, to mention a few (e.g., Yager et al., 2015; Scontras et al., 2015; Polinsky, 2016). The current study, then, adds to the already existing field of heritage language research by studying a different heritage language, AmNo, which in turn enables studying different structural components. For instance, double definiteness is a particularly interesting typological difference as it is characteristic of (some) Scandinavian languages.

One of the main interests in researching heritage languages is studying the outcomes of language contact across time, and as mentioned in Section 3.1, key objectives involve investigating the differences in competence that arise as a result of incomplete acquisition, attrition, or cross-linguistic influence. The time dimension is therefore crucial for such studies. AmNo is interesting in such a context since it is a heritage language that has been present in America for a long period of time, enabling diachronic studies on a larger scale. Importantly, this would not be possible if it were not for the pioneering work of scholars investigating and documenting the language during the 20th century. The work of Haugen (1953) is especially relevant in this respect, but the contributions of Flom (1900), Flaten (1900), and Hjelde (1992) are also notable. Combined, their work enables the unique possibility of studying the outcomes of prolonged language contact and diachronic change (see also, e.g., Roberts, 2007 on the relationship between contact and diachronic change).

Moreover, one obvious and interesting difference between the speakers of AmNo and the speakers investigated in many other heritage language studies is their age. AmNo speakers today are generally 70 years or older, and the fact that most of them have had English as their dominant language since they started school around the age of 6 makes them a particularly interesting group. AmNo has for these speakers been their non-dominant language for decades, and the contact with English has lasted significantly longer than for other groups. This enables a lifetime perspective which is not yet possible in studies of many other heritage languages. Relatedly, most speakers in other heritage language studies are 2nd generation immigrants. Thus, their input has often been shown to be fairly consistent with the non-heritage variety which their parents speak. AmNo speakers, on the other hand, represent even later generations and the input that they received was, importantly, not non-heritage Norwegian. Instead, their input was already a heritage language, AmNo spoken by 2nd or 3rd generation immigrants.

However, studying AmNo presently means studying the language in retrospective, which restricts the potential conclusions. Neither input data nor longitudinal data of individual speakers are systematically collected. The available corpora investigated in the current dissertation may provide indications of development, but conclusions should be drawn carefully. This insight might provide beneficial instructions for future collections of heritage language data, which should be careful to include input data when possible, and also attempt to study individual speakers across time.

6.1.2. The late-insertion exoskeletal model as a null theory model of language mixing

The theoretical questions raised in (3) are pursued throughout all three articles contained in this dissertation. In addition to mapping the empirical patterns of language mixing, a chief goal has been to show that a late-insertion exoskeletal model is capable of accounting for the observed patterns, and that this model constitutes a null theory of language mixing.

In Section 4.2, two challenges emerge from previous studies of language mixing. First, the empirical asymmetry observed by Myers-Scotton (1993) should be accounted for, and second, this should be done without proposing special machinery in the analyses (cf. MacSwan, 1999, 2014). MacSwan himself is critical of Myers-Scotton's MLF model and the notion of a grammatical frame at play in language mixing, stating that "if we are to believe that the 'language frame' is a principle of grammar and not a code-switching-specific constraint, then they must further show that it is a well-motivated construct for the analysis of monolingual data as well" (MacSwan, 2000: 42). This may be read as a challenge, and one objective for the current dissertation has been to demonstrate that the proposed late-insertion exoskeletal model is capable of uniting the two points of view.

The theoretical aspects of the dissertation are explicitly pursued in Article 1, which compares the late-insertion exoskeletal model to the mainstream lexicalist approach to grammar. By scrutinizing a concrete lexicalist analysis of Spanish–English mixing between a determiner and a noun, Article 1 shows that such an analysis is questionable, requiring principles beyond the common feature-matching mechanisms in order to account for the observed patterns. We argue that a late-insertion exoskeletal approach fares better in this respect, and is thereby favorable in analyzing language mixing. Articles 2 and 3 provide additional support for this proposal.

A first point to establish is that the late-insertion exoskeletal model is, as discussed in Section 4.1, motivated by monolingual data. Hence, its ability to analyze language mixing data by the same mechanisms will establish it as a null theory of mixing. The next concern is its ability to account for the observed asymmetrical contributions of the languages involved, where one language appears to have a more prominent role in providing both word order and (most) functional components. I argue that, in fact, the mechanisms of a late-insertion exoskeletal model predict such a pattern. The abstract structure will establish the word order prior to insertion of phonological exponents, and the functional features of the structure restrict their realization. Functional exponents will then typically be drawn from the same language that is associated with the overall structure, due to the feature-matching

requirements. Insertion of non-functional items is radically less restrictive, and they can be drawn from any language that the speaker knows.

This introduces a different issue, namely the notion of potentially language-specific structures, which has been criticized in prior analyses (see MacSwan, 2014). As discussed in Section 4.1.6, a general assumption in the exoskeletal approach is that UG makes available a complete set of features, and that a specific language is identified by the features activated in that language. Thus, referring to something as a Norwegian structure, as I do throughout this dissertation, means a structure comprising features typically associated with Norwegian. A bilingual speaker will then master one set of features associated with Language A and another set of features associated with Language B. Crucially, the usage of these sets of features depends on the speaker's language mode, i.e., the activation of a language in a given context (Grosjean, 1998, 2001, 2013). This means that when a speaker is in the mode of Language A, (s)he will make use of the set of features associated with this particular language. Even in a bilingual situation, one language will constitute the base or main language of the conversation. Relatedly, the activated main language will typically also be the source of most of the lexical items, although less activated language(s) still can be accessed if needed. Crucially, the term 'main language' is only used descriptively here and not as a theoretical primitive.

In this manner, an exoskeletal framework is able to account for the asymmetric contributions in language mixing: An abstract syntactic structure is generated prior to insertion of phonological exponents, based on the features of the main language. Then, the functional terminals are as a general rule realized by exponents associated with the same language, whereas substantial items are easily inserted into radically less restricted positions. Consequently, both the empirical observations pointed out by Myers-Scotton and the ideal of a null theory or constraint-free approach, as stressed by MacSwan, are preserved within an exoskeletal approach.

A concrete analysis of language mixing in AmNo can thus be summarized as follows. Since AmNo is a Norwegian heritage language, we can assume that Norwegian is the main language in these conversations and that the abstract syntactic structures will consist of functional features associated with Norwegian. In the nominal domain, this entails a structure comprising definiteness, number, and gender. When these functional terminals are subject to Spell-Out, Norwegian functional exponents are preferred over English alternatives, as the Norwegian ones provide a complete match to the features specified in the structures. English noun stems may still be inserted in the designated positions. Such analyses are provided in

Articles 1 and 2. The first article is primarily concerned with analyzing data showing a mix between the determiner and the noun. Mixing between Spanish and English is also discussed in order to propose alternative analyses to the data provided in Moro (2014). Article 2 in particular provides analyses of all the various mixing patterns in the nominal domain in AmNo, as discussed in Section 6.1.1 above. Additionally, this article provides an analysis of the usage of the English plural *-s* within the exoskeletal model.

The diachronic changes observed in AmNo are also accounted for in the exoskeletal model. This is explicitly addressed in Article 3, where both the omission of functional suffixes and the usage of English functional exponents are analyzed. In an exoskeletal model, these changes may be considered a result of either change in the phonological exponent or in the syntactic structure. Both alternatives will alter the conditions for insertion and may explain the observed change in the AmNo material. However, due to the way in which the change has taken place, combined with the dominant role of English in the AmNo communities and within the individual speakers, the changes are tentatively analyzed in Article 3 as structural reanalyses due to cross-linguistic influence from English (see, e.g., Putnam & Sánchez, 2013; Scontras et al., 2015; Polinsky, 2016).

The results of the current dissertation thus provide support for a late-insertion exoskeletal model as a good analytical tool for language mixing. By extension it also provides support for an exoskeletal model of grammar more generally. As mentioned in Section 4.2, González-Vilbazo et al. (2013), among others, point out that language mixing provides insight into components of language that are not easily detected in monolingual data. Language mixing between English and Norwegian in the nominal domain thus constitutes an interesting testing ground, as the noun phrases in these languages show certain crucial structural distinctions, e.g., grammatical gender and realization of definiteness. Details concerning these components can be quite opaque in monolingual Norwegian, whereas the different pieces might be more salient in mixed phrases. For instance, in Section 6.2.2, I discuss how language mixing data provide insight that can help to refine a syntactic analysis of grammatical gender.

6.2. Some implications and points of contention

In addition to the main findings discussed in the previous subsection, some additional implications and new questions arise from the data and results of the dissertation. Some of these issues may be points of contention and in need of further research. Here, I discuss a selection of these implications and questions and how they have been approached in the

current investigations. Subsections 6.2.2, 6.2.3, and 6.2.4 concern details of the specific analyses and the DP model employed in the current dissertation, and the final subsection proposes some issues to be considered in the future. First, however, I address a question concerning the basic understanding and interpretation of language mixing in the first place.

6.2.1. Is it really mixing?

AmNo has been a heritage language over a long period of time, and by investigating data from various time periods, one cannot escape noticing that certain English words keep recurring. For instance, Haugen (1953) refers to the Norwegian jurist Ole Munch Rædar, who visited America in 1847 and commented on the usage of English. Words like *field, river, fence,* and *lake* were already in use at that time, and these words are also found in the new material available in CANS. The emerging question is then whether or not we still can consider these instances of mixing, or if they are better described as items in the AmNo vocabulary. Furthermore, how does this affect the recent analyses of language mixing in AmNo, as conducted in the current dissertation?

Johannessen and Laake (2017) argue that a common language variety or dialect has developed in the AmNo society. They base this on, among other factors, the observation that many new words, i.e., English loans, have become established in the AmNo vocabulary. Moreover, they observe that these words are distributed across a large area, as well as over time (referring to attestations in Haugen, 1953, and in Hjelde, 1992), thus supporting the proposal of an emerging new variety. Still, a closer investigation of the English items in CANS reveal that they often occur after hesitation, which one would not expect for words truly established in a vocabulary. An example is shown in (40), where "#" marks a pause and "e" marks hesitation.

(40) og hadde # e **caren** i # **garagen** og # hadde ny olje på han
og had # e car.DF.SG.M in # garage.DF.SG.M and # had new oil on him
'and had the car in the garage in order to change the oil'

(CANS; westby_WI_06gm)

Both *car* and *garage* are listed in Johannessen and Laake (2017) as potential items of a new AmNo vocabulary. Both words are attested in the speech of their present informants, as well as Haugen (1953), which is presented as support for their proposal. The hesitations and pauses connected to these and similar words nevertheless suggest that they are not completely

integrated or easily accessed. Instead, it suggests that the speaker may still be searching for the proper Norwegian word, and resorts to English when this fails.

In a broader perspective, these data invite a discussion of what it really means to mix languages. This brings us back to the debate concerning CS and borrowing in Section 4.2.1 and the different analyses of these phenomena. From his lexicalist perspective, MacSwan (1999, 2009, 2014) argues in favor of language-specific lexicons. Consequently, for him, CS means drawing items from two different lexicons into syntax, whereas borrowing means copying a lexical item from Lexicon A to Lexicon B. From this perspective, the more established English items occurring in AmNo would be considered borrowings, not CS. Hence, as borrowed items, they are already established in the AmNo lexicon, and using them does not entail mixing any languages.

In the exoskeletal approach, as employed in the current work, this can be understood differently. As discussed in Section 4.1.6, the content of the traditional lexicon is here distributed across different lists, and functional features are separated from individual lexical items. The consequence of this separation is that the functional structure is not affected by the individual root or stem inserted: it might be Norwegian, English, or even made up, and the syntax will still provide it with the relevant functional components. In other words, our grammar remains unaffected; it handles functional features and roots as usual, and phonological exponents are inserted following the restrictions of the Subset Principle. When an utterance is phonologically realized, on the other hand, the speaker will obviously associate the various components with what we know as different languages. This is, from an exoskeletal perspective, assumed to be primarily encyclopedic knowledge and not a property of the individual items. The fact that we recognize an item as Norwegian is based our associations of that item with what we know to be properties of Norwegian vocabulary.[41] Put somewhat in the extreme, one might argue that what we consider to be language mixing is then principally based on our judgements of the product.

Concerning the relation between CS and borrowing, Myers-Scotton (1993: 174ff.), among others, emphasizes that these are not separate phenomena, but that CS and borrowing are two points on the same continuum, distinguished primarily by the frequency of the item in question. Importantly, any borrowed item will, according to this proposal, have entered the language initially as a spontaneous loan (i.e., CS), but as its frequency increases it becomes

[41] This invites a different question, namely concerning how roots are stored. Are they stored in separate (language-specific) lists or are they possibly stored in one comprehensive list? This is a crucial question for the development of syntactic theory and should be addressed in future research. Space prevents a comprehensive discussion here. See Grimstad (2017) for some discussion in the context of AmNo.

more and more established in the vocabulary of that language and is eventually considered a borrowed item. This is easily adapted to an exoskeletal approach where syntax makes no distinction between the first time a root or stem is used (i.e., what would be considered CS) and subsequent uses, when it is eventually classified as a borrowed item (see also Grimstad et al., 2014: Åfarli, 2015b; Grimstad, 2017). Hence, while MacSwan and other lexicalists must assume a process of copying an item between lexicons, the exoskeletal approach escapes this procedure: All items are subject to the same insertion restrictions, regardless of what vocabulary we consider that item a member of. Language identity is rather determined by our associations related to the outcomes, and these may change over time.

Finally, as discussed in Section 4.2, an important goal of formal studies of language mixing has been to propose concrete restrictions. A key finding of the current investigations is that English noun stems are easily and frequently used in AmNo. As discussed above, these stems are treated no differently than a Norwegian stem in the structure. However, in parallel to the discussion of lexical items' flexibility in Section 4.1.7, any stem from any language would, in theory, be available for insertion in the open positions, but harmony and convention will typically guide our choices (Åfarli, 2007; Marantz, 2013; Lohndal, 2014). Thus, when in Norwegian language mode, stems associated with Norwegian will presumably be preferred.

From an exoskeletal perspective, restrictions on language mixing primarily come into play concerning the functional components of a structure. The results of this dissertation show that while English noun stems occur frequently in AmNo, the functional items are by and large Norwegian. This follows from the general feature-matching requirements of the model, and even in a situation of similar functional features, a speaker's language mode will give precedence to the exponents of the main language. Language mixing involving functional components is therefore unexpected. Still, all three articles contained in the current dissertation, and in particular Article 3, address the issue of English functional items occurring in AmNo. This concerns the plural suffix *-s* and the articles *the* and *a,* which are unexpected exponents if the structures contain feature bundles of definiteness, number, and gender. These cases have here been considered reanalyses of the grammatical structures in AmNo, which entails that, after long-lasting contact with English, the functional features may in some cases be weakened, rearranged, or even erased from the structure. Consequently, English functional exponents turn out to be equally well-suited for insertion. Considering the dominance of English both in the society and for the individual speaker, we can then assume that English exponents are easily accessed, even when the speakers are in a Norwegian language mode.

6.2.2. Analyses of grammatical gender

The investigations into mixing between a grammatical gender language like Norwegian and a language where gender is an alien category to nouns, like English, have also led me to draw specific conclusions concerning gender. The main interest of the current dissertation is the formal analysis of gender, more specifically its structural position and how gender agreement is accomplished among different components of the DP structure. The current dissertation argues in favor of a syntactic approach to grammatical gender where gender is assumed to be a feature of the functional structure of the noun phrase. This is discussed presently. A different question concerns the basis for gender assignment. This is not addressed specifically in the three articles contained in the dissertation, but I will nevertheless discuss it briefly in Section 6.2.3. The reason is that the AmNo mixing data shed light on one specific proposed hypothesis about gender assignment in language mixing, namely the hypothesis that gender is "translated" from the corresponding noun in the gender language.

A much-debated question in the literature concerns the locus of gender in a non-lexical analysis. In general, we can identify three potential analyses: that gender is a feature of the root, a feature of the nominalizer, or a feature of the higher functional structure. In the following, I discuss these analyses and some potential challenges posed for them by data from language mixing (see also Kramer, 2016 for a review).

In traditional analyses, gender is considered an inherent feature of an individual lexical item. A main argument in favor of such an analysis is that a given noun typically belongs to one specific gender, thus the value of a noun's gender feature will not change in different structural surroundings. Analyses of gender as part of the root in an exoskeletal framework would nevertheless be problematic both theoretically and empirically. Since roots are here considered to be devoid of all grammatical features, such an approach is theoretically dismissed.[42] Moreover, if the gender feature were in fact a property of the root, considering that a root can become a noun, a verb, or an adjective, depending on which categorizer it is combined with, we would be unable to account for the fact that only nouns have gender. Empirically, the data from language mixing in AmNo suggest that gender cannot be provided by the root. A gender-less English noun mixed into Norwegian is still assigned to a gender category, visible on functional morphemes, which is not easily attributed to the English root.

[42] Notice that some approaches within DM may consider grammatical gender a feature related to the root (see, e.g., Embick, 2015: 42). Such approaches thus assume only partially bare roots (cf. Gallego's, 2014 typology of approaches to roots, as discussed in Section 4.1.6), and differ from the totally bare approach taken by, e.g., Borer (2014) and followed in this dissertation.

And, as the AmNo material shows, English nouns are distributed across the three genders in Norwegian (Haugen, 1953; Hjelde, 1996), so a potential default mechanism is unlikely.

The second approach is to analyze gender as a feature of the category-defining head, the nominalizer (e.g., Alexiadou, 2004; Julien, 2005; Acquaviva, 2009; Kramer, 2014, Anagnostopoulou, 2017). Such an analysis describes the assignment of gender as an essential part of turning a root into a noun, as knowing a noun also means knowing the gender of it. Through investigating the gender system of Amharic, Kramer (2014) proposes an analysis where natural gender is considered an interpretable gender feature and grammatical gender an uninterpretable feature, both introduced by the nominalizing head, *n*. Licensing conditions will then determine which roots can be combined with which *n*, of which there are different "flavors". A related but slightly different analysis is found in Alexiadou (2004). The languages taken under consideration by Alexiadou are Hebrew, Greek, Italian, and Spanish, and the proposal is, briefly summarized, that there are two types of gender: one type which is fixed as an inherent property of the noun stem already in the lexicon, and a second type which varies according to its referent. The crucial distinction is that the latter has a [+human] or [+animate] feature.

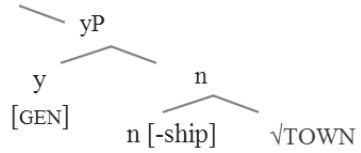
Analyses like these are, nevertheless, not as straightforward for Norwegian, where gender is non-transparent and arbitrary (Rodina & Westergaard, 2013, 2015a, b). Take for instance Kramer's (2014) analysis. It is built on a language that relies heavily on natural gender, where a typical pattern is that one root can have both a masculine and a feminine form, depending on the natural gender of its referent. In cases where the natural gender is unknown, or the referent does not have natural gender, the nominal is masculine by default. In the case of Norwegian, gender assignment is not equally dependent on or affected by natural gender. As an example, we can consider the noun *barn* 'child', which is neuter in Norwegian regardless of the natural gender of the child.

An additional argument against analyzing gender as a property of the nominalizer comes from mixing that also includes derivational suffixes. Examples from AmNo are shown in (41).

(41) a. et township (N) - 'a township'
b. den store building (M/F) - 'the big building'
c. government-en (M) - 'the government'

These examples show English noun stems mixed into AmNo, and interestingly the derivational suffixes are also English. Hence, the item drawn from English in this case is actually the complex of a root and the category-defining head. A potential structural representation of (41a) is provided in (42) below.

(42)



township

In this example, both the root and the nominalizer are drawn from English, as neither realization is common in the Norwegian vocabulary. When these stems enter an otherwise Norwegian structure, they are nevertheless assigned gender, which is then difficult to attribute to the (English) nominalizing head. Instead, the gender feature must be situated somewhere in the functional structure above the stem, represented by yP in (42).

Importantly, what these examples also show is that the stem is an available component for mixing. The subsequent strong hypothesis is that, at least for AmNo, the stem is always the item being mixed.[43] Notice, however, that the weaker hypothesis, that both roots and stems are available for mixing, is also compatible with the exoskeletal framework (see, e.g., González-Vilbazo & López, 2011; Alexiadou et al., 2015; Alexiadou, 2017 for such an analysis).

Regardless of which hypothesis turns out to be correct, mixed derived nouns show that gender should be placed higher in the functional structure, which is the third alternative. Several different analyses are possible depending on the structure one employs and its different projections. Ritter (1993) proposes that gender is, at least for Romance languages, positioned in the number projection. An attempt at decomposing Norwegian plural nouns, as in (43) below, may indicate that gender variation is indeed connected to the plural projection.

[43] Article 2 of the dissertation offers two additional arguments in favor of stems being mixed: 1) In the verbal domain, mixing between Norwegian and Turkish (Türker, 2000) and between Norwegian and Mandarin Chinese (Åfarli & Jin, 2014) show that it is typically the infinitival form that is mixed, not the bare root. 2) Mixed items often come with an already established and specific conceptual content, which presumably arises from the process of categorization. The latter argument is more speculative than the arguments based on infinitival or derivational forms, but is nevertheless worth consideration.

This is especially apparent for the masculine (43a) and feminine (43b), where the plural suffix varies and the definite suffix is alike.

(43)	a.	bil-**a**-**ne** (M)	car-PL-DF	'the cars'
	b.	jakk-**e**-**ne** (F)	jacket-PL-DF	'the jackets'
	c.	hus-**a** (N)	house-PL.DF	'the houses'

Decomposition of Norwegian singular nouns, on the other hand, as in (44), would suggest that gender is positioned in the definiteness projection.

(44)	a.	bil-**en** (M)	car-DF	'the car'
	b.	jakk-**a** (F)	jacket-DF	'the jacket'
	c.	hus-**et** (N)	house-DF	'the house'

Analyzing gender as a feature of the number projection in plurals and as a feature of the definiteness projection in singulars is, however, not very desirable. Alternatively, gender could be analyzed as the head of its own projection, GenP, as proposed in Picallo (1991, 2008) and Nygård and Åfarli (2015). A complication for such an analysis is that, at least in Norwegian, gender is never realized as a separate morpheme, so the projection in question will always have an empty realization. A final alternative is combining the features definiteness, number and gender in one feature bundle in a common functional projection above the stem. The appropriate exponent of such a feature bundle then depends on the interplay between the values of the features. This is the analysis that is employed in the current dissertation. I return to this in Section 6.2.4.

A related analysis is provided by Nygård, Riksem, and Åfarli (2015), under the label "distributed gender". This analysis does not consider gender to be a feature of one specific projection, but rather distributed across the different projections in the nominal structure. An emerging question, which is in fact relevant for all syntactic analyses of grammatical gender, then concerns the assignment of a specific gender to a specific noun; on what basis would the gender feature be established in such cases? An analysis compatible with the exoskeletal approach taken in the current dissertation has been proposed by Picallo (2008) and Nygård and Åfarli (2015). They argue that gender is a functional projection turning conceptual class or entity categorization into a linguistic or grammatical component. Nygård and Åfarli (2015) dub this process "feature construal" (see also Nygård, 2013). In other words, just as the

number feature is the linguistic appropriation of the non-linguistic numeric property of the noun, a gender feature is the linguistic appropriation of the conceptual content of a root or stem. Combining the hypothesis of feature construal with the hypothesis of distributed gender implies that the value of the gender feature is established, or construed, in the lowest instance of the gender feature in the structure based on the conceptual content of the root or stem. The higher instantiations will subsequently receive valuation through a probe–goal relation.

6.2.3. Gender assignment through translation?

An interesting question concerning the mixing of English stems into AmNo nominal structure regards their gender assignment. This issue is not specifically addressed in the current dissertation, but I will briefly discuss one previously proposed (and debated) strategy, namely the idea that the gender assigned to English nouns is "translated" from the analogous gender of the corresponding Norwegian noun (see, e.g., Jake et al., 2002; Liceras, Fuertes, Perales, Pérez-Tattam, & Spradlin, 2008; Parafita Couto et al., 2015 for discussion). I argue that this hypothesis is dubious for the AmNo data for a number of reasons.

A first, and quite trivial, challenge for such a hypothesis is connected to the idea that the speaker resorts to English words when (s)he cannot recall the appropriate Norwegian one, and that this is a typical reason for mixing in the first place. Yet, it seems unlikely that a speaker would be able to translate the gender of a forgotten Norwegian noun onto the English replacement.[44]

A second and somewhat more fundamental challenge is that many of the English nouns occurring in the AmNo material may in fact translate into several different Norwegian versions. Take for instance the following examples, where the words in italics show (some) possible translations of the boldfaced English noun, and their respective genders:

(45) a. **government** → *stat* (M), *regjering* (F), *styre(sett)* (N)
b. **shed** → *skur* (N), *skjul* (N), *bod* (F)
c. **field** → *åker* (M), *eng* (F), *mark* (F), *jorde* (N)

In these cases, it is not possible to determine for certain which Norwegian noun the speaker potentially was translating from, and therefore it is not possible to confirm the translation.

[44] Studies of the "tip of the tongue" phenomenon suggest that speakers may have generated the correct structure, potentially with the correct gender feature, even though they have to resort to a different phonological form (see e.g., Vigliocco, Antonini, & Garrett, 1997). Space prevents further discussion here.

Moreover, to complicate matters further, gender distribution varies greatly among Norwegian dialects, meaning that one noun may actually be assigned different genders in different dialects. As the dialect background for these speakers is difficult to determine, the "correct" gender of the nouns is equally difficult to determine. Hence, it is in general problematic (if not impossible) to determine whether such online strategies are being used.

A third argument disfavoring the translation hypothesis comes from the actual distribution of gender to the English nouns in AmNo. In order to investigate this, I have compared the gender assignment to all singular English nouns occurring in AmNo with gender identified by functional morphology or associated words, and compared these to a potential Norwegian equivalent.[45] The results are presented in (46).

(46)

Overlapping gender?	
Gender corresponding to the Norwegian equivalent	61.2 %
Gender diverging from the Norwegian equivalent	38.8 %

These numbers show that only 61.2% of the English nouns can presumably be considered a translation from Norwegian when it comes to gender assignment. The remaining 38.8% are assigned a gender different from its Norwegian equivalent. A related question then concerns whether these 38.8% result from an overgeneralization to masculine, which is the most frequent gender in Norwegian. In order to see whether this is the case, I have investigated this subpart separately, and the distribution is presented in the table below.

(47)

Diverging gender = overgeneralized to masculine?	
Overgeneralization to masculine	62.8 %
Different patterns	37.2 %

Among these nouns, 62.8% have either a neuter or a feminine Norwegian equivalent but are assigned masculine gender when an English noun is used, thus indicating an overgeneralization. Still, the remaining 37.2% show a divergence going in different directions, for instance from masculine to feminine or neuter. What the data in (46) and (47) indicate is that, at least for a considerable part of the English nouns, their assigned gender is

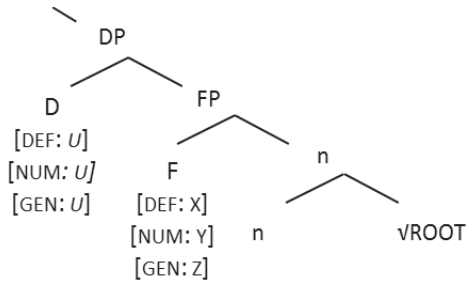
[45] I should be the first to point out that these translations are obviously not absolute, precisely due to the dialectal differences and the multiple alternative translations, as discussed above. However, I have used a probable translation in order to provide indications of any translation strategies. These translations and the subsequent distributions are nevertheless not intended to establish independent results, simply indications.

not based on a translation from the Norwegian equivalent. And, moreover, overgeneralization to the masculine alone cannot explain the cases diverging from their potential Norwegian translations. Thus, gender assignment seems to be an independent process in AmNo, potentially in line with the process of feature construal, as discussed in the previous subsection.

6.2.4. FP

The specific DP model employed for the analyses in the current dissertation is based on, but nevertheless different from, the model proposed by Julien (2005), as discussed in Section 4.3.2. The structure I have employed is presented in (48) below.

(48)



In accordance with Julien's (2005) model, the derivation starts by combining a root and a category-defining head, yielding the stem of the phrase. Notice, however, that the categorizer and the stem is labeled with a little n (not a capital N), to be consistent with most DM approaches. Additional similarities are that the topmost layer is a DP,[46] and that additional functional projections will be merged below this DP in phrases with attributive adjectives or weak quantifiers. The most significant difference from Julien (2005) is that the structure in (48) makes use of a functional projection, FP, whose head is considered to be the terminal of a functional feature bundle consisting of definiteness, number, and gender. In other words, the projection FP combines features which in previous work have been analyzed as the heads of two (Julien, 2005) or three (Nygård & Åfarli, 2015) separate projections. FP may, on the one hand, arguably be considered an abbreviation for ease of exposition in the analyses, and

[46] More specifically, the DP is the topmost obligatory projection. Julien (2005) also proposes additional projections for demonstratives and strong quantifiers. These are not employed in the current analyses.

thereby cover a more complex structure. Still, the FP analysis does also invite a discussion of the functional components of contemporary Norwegian nominal structure, especially when combined with data from language mixing.

The analysis put forth in Julien (2005) provides a separate projection for definiteness and number, and gender is attributed to the nominalizer. The Norwegian functional suffixes are also typically referred to as a plural suffix and a definite suffix (Faarlund et al., 1997). Some illustrative examples were provided in (43) above. In the examples in (43a) and (43b), *-ne* can be considered the definite suffix, whereas *-a* and *-e* are plural suffixes for masculine and feminine, respectively. These suffixes are considered exponents of individual functional features in Julien's (2005) structure. This approach gets somewhat more difficult when encountering neuter, as in (43c), where plurality and definiteness seem to have melded together into one exponent. Moreover, when dealing with singular phrases (44), only a definite suffix is assigned, raising the question of whether the number projection, when valued singular, is always realized as null, or is not there at all. The proposal involving three separate projections, as put forth in Nygård and Åfarli (2015), will encounter the same difficulties concerning the gender projection which is also realized as null.

As already mentioned above, part of the motivation for employing one functional projection with a combined feature bundle in the current dissertation has been its ease of exposition. The main objective for this dissertation is not to investigate the fine-grained details of Norwegian DPs, but rather to analyze how English nouns acquire Norwegian functional properties, for instance a functional suffix. For this purpose, the single projection is sufficient; the FP successfully accounts for the functional properties materializing as a suffix on the English noun stem inserted into a Norwegian nominal structure. Moreover, it also enables the analysis of gender as a property of a functional projection above the item being mixed, i.e., the stem, as discussed above. However, even though the data investigated for the current dissertation are not sufficiently rich to bring forth new evidence concerning the internal functional structure of the Norwegian DP, language mixing may shed light on the issue. For instance, if the nominal domain consists of several separate functional projections, one would expect to see mixing between them. This pattern is uncommon in the data investigated here, but more data is needed to draw any conclusions. Detailed investigations of these cases would then potentially illuminate the basic components of the nominal domain. Such investigations are, nevertheless, beyond the scope of the current dissertation and rather a topic for future research, which I will discuss briefly in the next subsection.

6.2.5. The road ahead

Investigations of AmNo, language mixing, and noun phrases are still far from exhausted by the current dissertation. Future studies should, for instance, consider language mixing within other domains of the sentence and grammar more generally, and compare the synchronic and diachronic patterns of both mixed and unmixed AmNo speech. This, among other things, will provide a more comprehensive picture of the language and the phenomenon of language mixing.

An interesting venue for further investigations is studying the patterns of language mixing in AmNo at the individual level. The present dissertation has been exclusively concerned with the group level. Thus, I have not discussed individual results, which would provide new insights and a more nuanced picture. For instance, throughout the investigations conducted for the present dissertation, a general observation is that some speakers mix the languages more frequently than others. Future studies should address the potential correlation between the frequency of mixing and an individual's overall proficiency in AmNo. In addition, individual studies would contribute more details to the picture of the observed diachronic changes, as discussed on a group level in Article 3. For instance, are the changes limited to a small group of speakers? And what different strategies, if any, can be identified in cases where they have difficulties arriving at the expected Norwegian realizations? A concrete example can be pointed out from the realization of definiteness, namely whether some speakers consistently omit the functional suffix but realize the determiner, or if the variations are accidental. However, the size of CANS (at the time of writing, summer 2017) is a limitation in this respect. The amount of data for the individual speakers is unlikely to provide enough data to provide reliable results. Lohndal and Westergaard (2016) discuss these problems in their study of grammatical gender in AmNo, where they tentatively present and discuss individual results.

Further investigations of the Norwegian DP, with or without language mixing, would also be a relevant subject for future research. One issue of particular interest is the functional projections of the DP. As mentioned above, language mixing may provide evidence to adjudicate the question of whether there should be one, two, or three functional projections. For instance, if there are several projections holding functional features, one would expect that it is possible to mix between them. Following the Subset Principle, the English plural *-s* would be a potential realization of a projection containing a plural feature. Consequently, one would expect cases where the plural suffix is English and the definite suffix is Norwegian. In fact, Haugen (1953: 542) makes note of examples where the English plural *-s* seems to be an

active affix and occurs in combination with a Norwegian definite suffix. In contemporary non-heritage Norwegian, such cases are also occasionally observed. One example is *cardigan-s-ane* 'cardigan-PL-DF.PL.M' as mentioned in footnote 17 above. Interestingly, the Norwegian exponent for definiteness is here the full *–ane*, specified for both plurality and definiteness, and not *-ne* as would be expected from the decomposition in (43) above. In other words, plurality is seemingly realized both by the English and the Norwegian suffix. The question of the functional structure of a Norwegian DP is therefore not settled and a more detailed investigation of these cases may provide new insights in this case.

Investigations of other components in the nominal domain would also be an interesting topic for future research. Adjectives and possessive pronouns, for instance, are related components that also display significant differences between English and Norwegian. As discussed in Section 4.3, adjectives in Norwegian agree with the noun in definiteness, number, and gender, and at least two potential questions may be investigated here. First, can we find examples of English adjectival stems mixed into AmNo (or non-heritage Norwegian), and if so, are these inflected according to the Norwegian system? And secondly, in case of deviances from the Norwegian pattern, do they show influences from English? Possessive pronouns may be particularly interesting, as they may be either pre- or post-nominal in Norwegian. The latter alternative is the most frequent, and also requires realization of the definite suffix on the noun stem. Westergaard and Anderssen (2015) have already investigated possessive pronouns and provide evidence against the hypothesis that contact with English has shifted the preference toward the pre-nominal position. As an extension, it would be interesting to investigate the realization of possessive pronouns in cases of language mixing in order to see if any new pattern may emerge here.

All these proposals are based on the idea that language mixing may provide a unique window into the underlying grammatical structures. The interplay between Norwegian and English exponents in the nominal structure is then arguably a fruitful domain for investigating both the details of nominal structures and the restrictions and mechanisms of language mixing in general.

7 Final remarks

The three articles contained in this dissertation belong both theoretically and empirically to the same area of research; they all address AmNo noun phrases that show mixing between Norwegian and English, and they do so from an exoskeletal perspective of grammar. The overall goal is, as pointed out in Section 1, to describe and to provide formal analyses of this linguistic phenomenon. In keeping with this goal, the three articles address separate subareas: Article 1 is primarily a theoretical discussion of the proposed model and the overall analytical approach to language mixing, whereas Articles 2 and 3 provide synchronic and diachronic studies of the empirical material, respectively. Hence, the three articles provide independent results, as summarized in Section 2, and are at the same time interrelated through the common object of inquiry. In addition, this cover article has aimed to bind the articles together by providing a discussion of the common theoretical background and methodological considerations. Finally, the results of the individual investigations have been united and some overall findings and implications have been pointed out.

In conclusion, I argue that the three articles that constitute the bulk of this dissertation, supported by the cover article, provide novel insight into the object under investigation. Moreover, taken together, they support the proposed hypothesis that language mixing in AmNo is in fact systematic and that an exoskeletal approach to grammar is a well-suited analytical tool for gaining insight into these patterns.

References

Åfarli, Tor A. (2007). Do Verbs have Argument Structure? In E. Reuland, T. Bhattacharya, & G. Spathas (Eds.), *Argument Structure* (pp. 1–16). Amsterdam: John Benjamins.

Åfarli, Tor A. (2015a). A Syntactic Model for the Analysis of Language Mixing Phenomena: American Norwegian and Beyond. In B. R. Page & M. T. Putnam (Eds.), *Moribund Germanic Heritage Languages in North America* (pp. 12–33). Leiden: Brill. doi: 10.1163/9789004290211_003

Åfarli, Tor A. (2015b). Hybrid verb forms in American Norwegian and the analysis of the syntactic relation between the verb and its tense. In J. B. Johannessen & J. C. Salmons (Eds.), *Germanic Heritage Languages in North America* (pp. 161–177). Amsterdam: John Benjamins. doi: 10.1075/silv.18

Åfarli, Tor A. & Jin, Fufen (2014). Syntactic Frames and Single-word Code-switching: A Case Study of Norwegian–Mandarin Chinese. In T. A. Åfarli & B. Mæhlum (Eds.), *The Sociolinguistics of Grammar* (pp. 153–170). Amsterdam: John Benjamins.

Åfarli, Tor A. & Subbarao, Karumuri V. (in press). Models of grammar and the outcomes of long-term language contact: Language mixing in Dakkhini. In Z. Estrada-Fernández, A. Alvarez-Gonzalez, & C. Chamoreau (Eds.), *Diverse scenarios of syntactic complexity: inter and intra typological diversity.* Amsterdam: John Benjamins.

Abney, Steven P. (1987). *The English Noun Phrase in its Sentential Aspect* (Doctoral dissertation). Massachusetts Institute of Technology, Cambridge, MA.

Ackerman, Farrell; Stump, Gregory T., & Webelhuth, Gert (2011). Lexicalism, periphrasis, and implicative morphology. In R. D. Borsley & K. Börjars (Eds.), *Non-transformational syntax: Formal and explicit models of grammar* (pp. 325–358). Malden, MA: Blackwell-Wiley.

Acquaviva, Paolo (2009). Roots and lexicality in Distributed Morphology. *York Papers in Linguistics,* (10)*: York-Essex Morphology Meeting*, 1–21.

Adger, David (2003). *Core Syntax: A Minimalist Approach*. Oxford: Oxford University Press.

Adger, David & Svenonius, Peter (2011). Features in minimalist syntax. In C. Boeckx (Ed.), *The Oxford handbook of linguistic minimalism* (pp. 27–51). Oxford: Oxford University Press.

Alexiadou, Artemis (2001). *Functional Structure in Nominals: Nominalization and ergativity.* Amsterdam: John Benjamins.

Alexiadou, Artemis (2004). Inflection class, gender and DP-internal structure. In G. Müller, L. Gunkel, & G. Zifonun (Eds.), *Explorations in Nominal Inflection* (pp. 21–50). Berlin: Mouton de Gruyter.

Alexiadou, Artemis (2017). Building verbs in language mixing varieties. *Zeitschrift für Sprachwissenschaft,* 36 (1), 165–192. doi: 10.1515/zfs-2017-0008

Alexiadou, Artemis; Anagnostopoulou, Elena, & Schäfer, Florian (2015). *External arguments in transitivity alternations: a layering approach*. Oxford: Oxford University Press.

Alexiadou, Artemis; Haegeman, Liliane, & Stavrou, Melita (2007). *Noun Phrase in the Generative Perspective.* Berlin: Mouton de Gruyter.

Alexiadou, Artemis & Lohndal, Terje (2017). On the division of labor between roots and functional structure. In R. D'alessandro, I. Franco, & Á. J. Gallego (Eds.), *The Verbal Domain* (pp. 85–102). Oxford: Oxford University Press.

Alexiadou, Artemis & Lohndal, Terje (in press). The structural configurations of root categorization. In L. Bauke, A. Blümel, & E. Groat (Eds.), *Labels and Roots* (pp. 203–231). Berlin: Mouton de Gruyter.

Alexiadou, Artemis; Lohndal; Terje, Åfarli, Tor A., & Grimstad, Maren B. (2015). Language Mixing: A Distributed Morphology approach. In T. Bui & D. Özyildiz (Eds.), *Proceedings of NELS 45* (pp. 25–38). CreateSpace Independent Publishing Platform.

Anagnostopoulou, Elena (2017). Accusative case morphology conditioned by gender. Presented at Workshop in honor of David Pesetsky's 60th birthday, MIT, February 11, 2017.

Anagnostopoulou, Elena & Samioti, Yota (2014). Domains within words and their meanings: A case study. In A. Alexiadou, H. Borer, & F. Schäfer (Eds.), *The Syntax of Roots and the Roots of Syntax* (pp. 81–111). Oxford: Oxford University Press.

Anderson, Stephen (1982). Where's morphology? *Linguistic Inquiry, 13*, 571–612.

Anderson, Stephen (1992). *Amorphous Morphology.* Cambridge: Cambridge University Press.

Arad, Maya (2003). Locality constraints on the interpretation of roots: The case of Hebrew denominal verbs. *Natural Language & Linguistic Theory*, *21* (4), 737–778.

Arad, Maya (2005). *Roots and Patterns: Hebrew Morpho-syntax*. Dordrecht: Springer.

Baker, Mark (1985). The Mirror Principle and Morphosyntactic Explanation. *Linguistic Inquiry, 16* (3), 373–416.

Bandi-Rao, Shoba & den Dikken, Marcel (2014). Light Switches: On v as a pivot in codeswitching, and the nature of the ban on word-internal switches. In J. MacSwan (Ed.), *Grammatical Theory and Bilingual Codeswitching* (pp. 161–184). Cambridge, MA: MIT Press.

Bauer, Laurie (2014). Grammaticality, acceptability, possible words and large corpora. *Morphology, 24* (2), 83–103. doi: 10.1007/s11525-014-9234-z

Beavers, John & Sells, Peter (2013). Constructing and supporting a linguistic analysis. In R. J. Podesva & D. Sharma (Eds.), *Research Methods in Linguistics* (pp. 397–421). Cambridge: Cambrigde University Press.

Belazi, Hedi M.; Rubin, Edward J., & Toribio, Almeida J. (1994). Code Switching and X-Bar Theory. *Linguistic Inquiry, 25,* 221–237.

Benmamoun, Elabbas; Montrul, Silvina, & Polinsky, Maria (2013). Heritage languages and their speakers: Opportunities and challenges for linguistics. *Theoretical Linguistics, 39* (3–4), 129–181.

Benson, Erica J. (2001). The neglected early history of codeswitching research in the United States. *Language and Communication, 21* (1), 23–36.

Blondeau, Hélène (2013). Studying language over time. In R. J. Podesva & D. Sharma (Eds.), *Research Methods in Linguistics* (pp. 494–518). Cambridge: Cambrigde University Press.

Boeckx, Cedric (2006). *Linguistic Minimalism: Origins, Concepts, Methods, and Aims*. Oxford: Oxford University Press.

Borer, Hagit (1984). *Parametric syntax: case studies in Semitic and Romance languages*. Dordrecht: Foris.

Borer, Hagit (2003). Exo-skeletal vs. Endo-skeletal Explanations: Syntactic Projections and the Lexicon. In J. More & M. Polinsky (Eds.) *The nature of explanation in linguistic theory* (pp. 31–67). Stanford, CA: CSLI Publication.

Borer, Hagit (2005a). *Structuring Sense I: In Name Only*. Oxford: Oxford University Press.

Borer, Hagit (2005b). *Structuring Sense II: The Normal Course of Events*. Oxford: Oxford University Press.

Borer, Hagit (2013). *Structuring Sense III: Taking Form*. Oxford: Oxford University Press.

Borer, Hagit (2014). The Category of Roots. In A. Alexiadou, H. Borer, & F. Schäfer (Eds.), *The Syntax of Roots and the Roots of Syntax.* Oxford: Oxford University Press.

Borer, Hagit (2017). The Generative Word. In J. McGilvray (Ed.) *The Cambridge Companion to Chomsky* (2nd ed., pp. 110–133). Cambridge: Cambridge University Press.

Brøseth, Heidi (2007). *A neo-constructional approach to computer-oriented talk* (Doctoral dissertation). Norwegian University of Science and Technology, Trondheim.

Cantone, Katja F. & MacSwan, Jeff (2009). Adjectives and word order: A focus on Italian–German codeswitching. In L. Isurin, D. Winford, & K. de Bot (Eds.) *Multidisciplinary Approaches to Code Switching* (pp. 243–278). Amsterdam: John Benjamins. doi: 10.1075/sibil.41

Cantone, Katja. F. & Müller, Natascha (2008). Un nase or una nase? What gender marking within switched DPs reveals about the architecture of the bilingual language faculty. *Lingua, 118* (6), 810–826. doi:10.1016/j.lingua.2007.05.007

Carter, Diana; Deuchar, Margaret; Davies, Peredur, & Parafita Couto, María C. (2011). A Systematic Comparison of Factors Affecting the Choice of Matrix Language in Three Bilingual Communities. *Journal of Language Contact, 4* (2), 153–183. doi: 10.1163/187740911x592808

Chomsky, Noam (1957). *Syntactic Structures.* The Hague: Mouton.

Chomsky, Noam (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: The MIT Press.

Chomsky, Noam (1970). Remarks on Nominalization. In P. S. Rosenbaum & R. A. Jacobs (Eds.), *Readings in English transformational grammar* (pp. 184–221). Waltham, MA: Ginn.

Chomsky, Noam (1986). *Knowledge of Language: Its Nature, Origin and Use*, New York: Praeger.

Chomsky, Noam (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, Noam (2000). Minimalist inquiries: the framework. In R. Martin, D. Michaels, & J. Uriagereka (Eds.), *Step by Step: Essays on minimalist syntax in honor of Howard Lasnik* (pp. 89–155). Cambridge, MA: MIT Press.

Chomsky, Noam (2001). Derivation by phase. In M. Kenstowicz (Ed.), *Ken Hale: a life in language* (pp. 1–52). Cambridge, MA: MIT Press.

Chomsky, Noam (2005). Three Factors in Language Design. *Linguistic Inquiry, 36* (1), 1–22.

Chomsky, Noam (2008). On phases. In R. Freidin, C. P. Otero, & M. L. Zubizarreta (Eds.), *Foundational Issues in Linguistic Theory: Essays in Honor of Jean-Roger Vergnaud* (pp. 133–166). Cambridge, MA: MIT Press.

Chomsky, Noam (2013). Problems of projection. *Lingua,* 130, 33–49.

Cinque, Guglielmo (1994). On the evidence for partial N-movement in the Romance DP. In G. Cinque, J. Koster, J.-Y. Pollock, L. Rizzi, & R. Zanuttini (Eds.), *Paths toward Universal Grammar: Studies in Honor of Richard S. Kayne* (pp. 85–110). Washington DC: Georgetown University Press.

Cinque, Guglielmo (1999). *Adverbs and functional heads: A cross-linguistic perspective*. Oxford: Oxford University Press.

Clark, Eve & Clark, Herbert (1979). When nouns surface as verbs. *Language,* 55 (4), 767–811.

de Belder, Marijke (2013). Collective mass affixes: When derivation restricts functional structure. *Lingua 126*, 32–50.

Delsing, Lars-Olof (1993). *The Internal Structure of Noun Phrases in the Scandinavian Languages. A Comparative Study* (Doctoral dissertation). University of Lund.

Di Sciullo, Anne-Marie; Muysken, Pieter, & Singh, Rajendra (1986). Government and Code-Mixing. *Journal of Linguistics, 22*, 1–24.

Embick, David (2015). *The Morpheme: A Theoretical Introduction*. Berlin: De Gruyter

Embick, David & Marantz, Alec (2008). Architecture and Blocking. *Linguistic Inquiry, 39* (1), 1–53.

Embick, David & Noyer, Rolf (2001). Movement Operations after Syntax. *Linguistic Inquiry, 32* (4), 555–595.

Embick, David & Noyer, Rolf (2007). Distributed Morphology and the syntax morphology interface. In G. Ramchand & C. Reiss (Eds.), *The Oxford Handbook of Linguistic Interfaces* (pp. 289–324). Oxford: Oxford University Press.

Enger, Hans-Olav (2004). On the relation between gender and declension – A diachronic perspective from Norwegian. *Studies in Language*, *28*, 51–82.

Enger, Hans-Olav (2009). The role of core and non-core semantic rules in gender assignment. *Lingua, 119,* 1281–1299.

Faarlund, Jan Terje (2009). On the history of definiteness marking in Scandinavian. *Journal of Linguistics, 45* (3), 617–639.

Faarlund, Jan T., Lie, Svein, & Vannebo, Kjell I. (1997). *Norsk referansegrammatikk* [Norwegian Reference Grammar]. Oslo: Universitetsforlaget.

Flaten, Nils (1900). Notes on the American-Norwegian with vocabulary. *Dialect Notes Vol. 2*, 1900–1904, 115–126.

Flom, George T. (1900). English elements in Norse dialects of Utica, Wisconsin. *Dialect Notes Vol. 2*, 1900–1904, 257–268.

Flom, George T. (1903). The Gender of English Loan-Nouns in Norse Dialects in America; A Contriution to the Study of the Development of Grammatical Gender. *The Journal of English and Germanic Philology, 5* (1), 1–31.

Flom, George T. (1926). English Loanwords in American Norwegian as Spoken in the Koshkonong Settlement, Wisconsin. *American Speech, 1* (10), 541–558

Frengstad, Lars (2016). *Å verbe: en eksoskeletal analyse av norske creative verbkonstruksjoner* [To verb: an exoskeletal analysis of Norwegian creative verb constructions] (Master's thesis). Norwegian University of Science and Technology, Trondheim.

Fuller, Janet M. & Lehnert, Heike (2000). Noun phrase structure in German–English codeswitching: Variation on gender assignment and article use. *International Journal of Bilingualism, 4* (3), 399–420.

Gallego, Ángel J. (2014). Roots and phases. In A. Alexiadou, H. Borer, & F. Schäfer (Eds.) *The Syntax of Roots and the Roots of Syntax.* (pp. 192–207). Oxford: Oxford University Press.

Gardner-Chloros, Penelope (2009). *Code-switching.* Cambridge: Cambridge University Press.

Goldberg, Adele (1995). *Constructions: A Construction Grammar Approach to Argument Structure.* Chicago: University of Chicago Press.

Goldberg, Adele (2006). *Constructions at Work.* Oxford: Oxford University Press.

Goldrick, Matthew; Putnam, Michael, & Schwarz, Laura (2016a). Coactivation in bilingual grammars: A computational account 1049 of code mixing. *Bilingualism: Language and Cognition, 19*, 857–876. doi: 10.1017/S1366728915000802

Goldrick, Matthew; Putnam, Michael, & Schwarz, Laura (2016b). The future of code mixing research: Integrating psycholinguistics and formal grammatical theories. *Bilingualism: Language and Cognition, 19,* 903–906. doi: 10.1017/S1366728916000390

González-Vilbazo, Kay; Bartlett, Laura; Downey, Sarah; Ebert, Shane; Heil, Jeanne; Hoot, Bradley; Koronkiewicz, Bryan, & Ramos, Sergio (2013). Methodological Considerations in Code-Switching Research. *Studies in Hispanic and Lusophone Linguistics, 6* (1), 119–138. doi: 10.1515/shll–2013-1143

Gónzalez-Vilbazo, Kay & López, Luis (2011). Some properties of light verbs in code-switching. *Lingua*, *121*, 832–850.

Gónzalez-Vilbazo, Kay & López, Luis (2012). Little v and parametric variation. *Natural Language and Linguistic Theory*, *30*, 33–77.

Gries, Stefan Th. & Newman, John (2013). Creating and using corpora. In R. J. Podesva & D. Sharma (Eds.) *Research Methods in Linguistics* (pp. 257–287). Cambridge: Cambrigde University Press.

Grimstad, Maren B. (2017). The codeswitching/borrowing debate: Evidence from English-origin verbs in American Norwegian. *Lingue e Linguaggio*, *XVI* (1)*,* 3–34. doi: 10.1418/86999

Grimstad, Maren B.; Lohndal, Terje, & Åfarli, Tor A. (2014). Language mixing and exoskeletal theory: A case study of word-internal mixing in American Norwegian. *Nordlyd, 41* (2), 213–237.

Grimstad, Maren B.; Riksem, Brita R.; Lohndal, Terje, & Åfarli, Tor A. (in press). Lexicalist vs. exoskeletal approaches to language mixing. *The Linguistic Review.*

Grosjean, François (1998) [2004]. Studying bilinguals: Methodological and conceptual issues. *Bilingualism*: *Language and Cognition, 1,* 131–149. [And in T. K. Bhatia & W. C. Ritchie (Eds.). *The Handbook of Bilingualism* (pp. 32–63), 2004. Oxford: Blackwell Publishing].

Grosjean, François (2001). The Bilingual's Language Modes. In J. L. Nicol (Ed.), *One Mind, Two Languages: Bilingual Language Processing* (pp. 1–22). Oxford: Blackwell.

Grosjean, François (2013). Bilingual and Monolingual Language Modes. In C. A. Chapelle (Ed.), *Encyclopedia of Applied Linguistics* (pp. 1–9). Malden: Blackwell.

Haegeman, Liliane & Guéron, Jacqueline (1999). *English Grammar: a generative perspective.* Oxford: Blackwell.

Hale, Kenneth & Keyser, S. Jay (1993). On argument structure and the lexical expression of syntactic relations. In K. Hale & S. J. Keyser (Eds.) *The View from Building 20: Essays in linguistics in honor of Sylvian Bloomberger* (pp. 53–110). Cambrigde, MA: MIT Press.

Halle, Morris (1997). Distributed morphology: Impoverishment and fission. *MIT Working Papers in Linguistics, 30: PF: Papers at the Interface*, 425–449.

Halle, Morris & Marantz, Alec (1993). Distributed Morphology and the Pieces of Inflection. In K. Hale & S. J. Keyser (Eds.), *The View from Building 20: Essays in linguistics in honor of Sylvian Bloomberger* (pp. 111–176). Cambridge, MA: MIT Press.

Halle, Morris & Marantz, Alec (1994). Some Key Features of Distributed Morphology. *MIT Working Papers in Linguistics, 21*, 275–288.

Hankamer, Jorge & Mikkelsen, Line H. (2002). A morphological analysis of definite nouns in Danish. *Journal of Germanic Linguistics, 14* (2), 137–175.

Hankamer, Jorge & Mikkelsen, Line H. (2005). When movement must be blocked: A reply to Embick and Noyer. *Linguistic Inquiry, 36* (1), 85–125.

Harley, Heidi (2005). How Do Verbs Get Their Names? Denominal verbs, Manner Incorporation, and the Ontology of Verb Roots in English. In N. Erteschik-Shir & T. Rapoport (Eds.) *The Syntax of Aspect: Deriving Thematic and Aspectual Interpretation* (pp. 42–64). Oxford: Oxford University Press.

Harley, Heidi (2009). The morphology of nominalizations and the syntax of vP. In M. Rathert & A. Giannadikou (Eds.) *Quantification, Definiteness and Nominalization* (pp. 320–342). Oxford: Oxford University Press.

Harley, Heidi (2014). On the Identity of Roots. *Theoretical Linguistics, 40,* 225–276

Harley, Heidi & Noyer, Rolf (1999). Distributed Morphology. *Glot International, 4,* 3–9.

Harley, Heidi & Noyer, Rolf (2000). Formal versus Encyclopedic Properties of Vocabulary: Evidence from Nominalisations. In B. Peeters (Ed.) *The Lexicon-encyclopedia interface* (pp. 349–374). Amsterdam: Elsevier.

Haugen, Einar (1953). *The Norwegian Language in America*: *A Study in Bilingual Behavior* (Vol. I & II). Philadelphia: University of Philadelphia Press

Haugen, Einar (1967). *The Norwegians in America: A Students' Guide to Localized History*. New York: Teachers College Press

Haugen, Einar (1992). A Language Survey That Failed: Seip and Selmer. *American Speech,, 67* (3), 330–336.

Hellan, Lars (1986). The Headedness of NPs in Norwegian. In P. Muysken & H. van Riemsdijk (Eds.) *Features and Projections* (pp. 89–122). Dordrecht: Floris Publications. doi: 10.1515/9783110871661-005

Herring, Jon R.; Deuchar, Margaret; Parafita Couto, María C., & Moro Quintanilla, Mónica (2010). 'I saw the *madre*': evaluating predictions about codeswitched determiner-noun sequences using Spanish–English and Welsh–English data. *International Journal of Bilingual Education and Bilingualism, 13* (5), 553–573.

Hjelde, Arnstein (1992). *Trøndsk talemål i Amerika* [*Trønder* dialect in America]. Trondheim: Tapir.

Hjelde, Arnstein (1996). The gender of English nouns in American Norwegian. In P. S. Ureland & I. Clarkson (Eds.), *Language Contact across the Atlantic* (pp. 297–312). Tübingen: Max Niemeyer Verlag.

Jake, Janice L.; Myers-Scotton, Carol, & Gross, Steven (2002). Making a minimalist approach to codeswitching work: Adding the Matrix Language. *Bilingualism: Language and Cognition, 5* (1), 69–91.

Jake, Janice L.; Myers-Scotton, Carol, & Gross, Steven (2005). A response to MacSwan (2005): Keeping the Matrix Language. *Bilingualism: Language and Cognition,* 8 (3), 271–276.

Johannessen, Janne B. (2015a). 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). NEALT Proceedings Series 23.

Johannessen, Janne B. (2015b). Attrition in an American Norwegian heritage speaker. In J. B. Johannessen & J. C. Salmons (Eds.), *Germanic Heritage Languages in North America* (pp. 46–71). Amsterdam: John Benjamins. doi: 10.1075/silv.18

Johannessen, Janne B. & Laake, Signe (2017). Norwegian in the American Midwest: A Common Dialect? *Journal of Language Contact, 10,* 5–21. doi: 10.1163/19552629-01001002

Johannessen, Janne B. & Larsson, Ida (2015). Complexity Matters: On Gender Agreement in Heritage Scandinavian. *Frontiers in Psychology,* 6:1842. doi: 10.3389/fpsyg.2015.01842

Johannessen, Janne B. & Salmons, Joseph (2012). Innledning [Introduction]. *Norsk Lingvistisk Tidsskrift,* 30 (2), 139–148.

Johannessen, Janne B. & Salmons, Joseph (2015). The study of Germanic heritage languages in the Americas. In J. B. Johannessen & J. C. Salmons (Eds.), G*ermanic Heritage Languages in North America* (pp. 1–17). Amsterdam: John Benjamins. doi: 10.1075/silv.18

Julien, Marit (2005). *Nominal Phrases from a Scandinavian Perspective.* Amsterdam: John Benjamins.

Kåsen, Andre; Olsen, Eirik; Rødvand, Linn I. S., & Tengesdal, Eirik (2016). Transkripsjons- og translittereringsveiledning for Norsk i Amerika [Instructions for transcriptions and transliteration in *Norsk i Amerika*]. Retrieved from: http://www.tekstlab.uio.no/norskiamerika/Transkripsjons-translittereringsveiledning-norskiamerika.pdf [accessed July 28, 2017]

Kamwangamalu, Nkonko M. (1997). Language contact, code-switching, and I-languages: Evidence from Africa. *South African Journal of Linguistics, 15,* 45–51

Kastner, Itamar (2017). Templatic morphology as an emergent property: Roots and functional heads in Hebrew, ms. Humboldt-Universität zu Berlin, available at http://ling.auf.net/lingbuzz/003103

Khayitova, Sofiya (2016). *V2 i amerikanorsk – ufullstendig innlæring eller språkforvitring*? [V2 in American Norwegian – incomplete acquisition or attrition?] (Master's thesis). University of Oslo.

Kramer, Ruth (2014) Gender in Amharic: a morphosyntactic approach to natural and grammatical gender. *Language Sciences, 43,* 102–115.

Kramer, Ruth (2016). The location of gender features in the syntax. *Language and Linguistics Compass, 10* (11), 661–677. doi: 10.1111/lnc3.12226

Kupisch, Tanja & Rothman, Jason (2016). Terminology matters! Why difference is not incompleteness and how early child bilinguals are heritage speakers. *International Journal of Bilingualism.* doi: http://dx.doi.org/10.1177%2F1367006916654355

Labov, William (1972). *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.

Lardiere, Donna (2000). Mapping Features to Forms in Second Language Acquisition. In J. Archibald (Ed.), *Second Language Acquisition and Linguistic Theory* (pp. 102–129). Malden, MA: Blackwell.

Larsson, Ida & Johannessen, Janne B. (2015). Incomplete Acquisition and Verb Placement in Heritage Scandinavian. In R. S. Page & M. T. Putnam (Eds.), *Moribund Germanic Heritage Languages in North America* (pp. 153–189). Leiden: Brill. doi: 10.1163/9789004290211_008

Levinson, Lisa (2007). *The Roots of Verbs* (Doctoral dissertation). New York University.

Liceras, Juana M.; Fuertes, Raquel F.; Perales, Susana; Pérez-Tattam, Rocío, & Spradlin, Kenton T. (2008) Gender and gender agreement in bilingual native and non-native grammars: A view from child and adult functional–lexical mixings. *Lingua, 118*, 827–851.

Lipski, John M. (1985). *Linguistic Aspects of Spanish–English Language Switching.* Tempe, AZ: Center for Latin American Studies, Arizona State University.

Lohndal, Terje. (2012). *Without Specifiers: Phrase Structure and Argument Structure* (Doctoral dissertation). University of Maryland.

Lohndal, Terje (2013). Generative grammar and language mixing. *Theoretical Linguistics, 39,* 215–224.

Lohndal, Terje (2014). *Phrase structure and argument structure: A case-study of the syntax semantics interface*. Oxford: Oxford University Press.

Lohndal, Terje & Westergaard, Marit (2016). Grammatical Gender in American Norwegian Heritage Language: Stability or attrition? *Frontiers in Psychology,* 7:344. doi: 10.3389/fpsyg.2016.00344

Longobardi, Giuseppe (2001). How comparative is semantics? A unified parametric theory of bare nouns and proper names. *Natural Language Semantics, 9,* 335–369.

Lovoll, Odd S. (1999). *The Promise of America: A History of the Norwegian-American People* (Revised ed.). Minneapolis/London: University of Minnesota Press.

Lundskær-Nielsen, Tom & Holmes, Philip (2010). *Danish: a comprehensive grammar* (2nd ed.). London: Routledge.

Lødrup, Helge (2011). Hvor mange genus er det i Oslo-dialekten? [How many genders are there in the Oslo dialect?]. *Maal og Minne, 2,* 120–136.

MacSwan, Jeff (1999). *A minimalist approach to intra-sentential code-switching*. New York: Garland.

MacSwan, Jeff (2000). The architecture of the bilingual faculty: evidence from intrasentential code switching. *Bilingualism, 3,* 37–54.

MacSwan, Jeff (2005a). Codeswitching and generative grammar: A critique of the MLF model and some remarks on "modified minimalism". *Bilingualism: Language and Cognition, 8,* 1–22.

MacSwan, Jeff (2005b). Remarks on Jake, Myers-Scotton and Gross's response: There is no "matrix language". *Bilingualism*: *Language and Cognition, 8* (2), 277–284.

MacSwan, Jeff (2009). Generative approaches to code-switching. In B. E. Bullock & A. J. Toribio (Eds.) *The Cambridge Handbook of Linguistic Code-switching* (pp. 309–335). Cambridge: Cambridge University Press.

MacSwan, Jeff (2014). Programs and Proposals in Codeswitching Research: Unconstraining Theories of Bilingual Language Mixing. In J. MacSwan (Ed.), *Grammatical Theory and Bilingual Codeswitching* (pp. 1–33). Cambridge, MA: MIT Press

MacSwan, Jeff & Colina, Sonia (2014). Some Consequences of Language Design. Codeswitching and the PF interface. In J. MacSwan (Ed.), *Grammatical Theory and Bilingual Codeswitching* (pp. 185–210). Cambridge, MA: MIT Press.

Mahootian, Shahrzad (1993). *A null theory of code-switching* (Doctoral dissertation). Northwestern University, Evanston, IL.

Marantz, Alec. (1996). "Cat" as a phrasal idiom. Unpublished ms., MIT

Marantz, Alec (1997). No Escape from Syntax: Don't Try Morphological Analysis in the Privacy of Your Own Lexicon. *University of Pennsylvania Working Papers in Linguistics, 4* (1), 201–225.

Marantz, Alec (2013). Verbal argument structure: Events and participants. *Lingua, 130*, 152–168.

Montrul, Silvina (2008). *Incomplete Acquisition in Bilingualism. Re-examining the Age Factor*. Amsterdam: John Benjamins.

Montrul, Silvina (2012). Is the heritage language like a second language? In L. Roberts, C. Lindqvist, C. Bardel, & N. Abrahamsson (Eds.) *EUROSLA Yearbook, 12* (pp. 1–29). Amsterdam: John Benjamins. doi: 10.1075/eurosla.12.03mon

Montrul, Silvina (2016). *The Acquisition of Heritage Languages*. Cambridge: Cambridge University Press

Moro Quintanilla, Mónica (2014). The Semantic Interpretation and Syntactic Distribution of Determiner Phrases in Spanish–English Codeswitching. In J.MacSwan (Ed.), *Grammatical Theory and Bilingual Codeswitching* (pp. 213–226). Cambridge, MA: MIT Press.

Muysken, Pieter (2000). *Bilingual Speech.* Cambridge: Cambridge University Press

Muysken, Pieter (2015). Déjà VooDoo or New Trails Ahead? Re-evaluating the Mixing Typology Model. In R. Torres Cacoullos, N. Dion, & A. Lapierre (Eds.), *Linguistic Variation: Confronting Fact and Theory* (pp. 242–261). New York: Routledge.

Myers-Scotton, Carol (1993). *Duelling Languages: Grammatical Structure in Code Switching*. Oxford: Oxford University Press.

Myers-Scotton, Carol (2002). *Contact Linguistics: Bilingual Encounters and Grammatical Outcomes*. Oxford: Oxford University Press.

Norsk i Amerika [Norwegian in America] (2017) Retrieved from http://www.tekstlab.uio.no/norskiamerika/english/field-work.html [accessed June 29, 2017]

Nygård, Mari (2013). *Discourse ellipsis in spontaneously spoken Norwegian* (Doctoral dissertation). Norwegian University of Science and Technology, Trondheim.

Nygård, Mari & Åfarli, Tor A. (2015). On the structure of gender assignment. *Indian Linguistics, 76* (1–2), 67–76.

Nygård, Mari; Riksem, Brita R., & Åfarli, Tor A. (2015). What language mixing can tell us about the syntax of gender. Presented at Linguistic Complexity in the Individual and Society (LCIS Conference), Trondheim, October 14–15, 2015.

Parafita Couto, María C.; Munarriz, Amaia; Epelde, Irantzu; Deuchar, Margaret, & Oyharçabal, Beñat (2015). Gender conflict resolution in Basque–Spanish mixed DPs. *Bilingualism: Language and Cognition, 18* (2), 304–323.

Pascual y Cabo, Diego & Rothman, Jason (2012). The (Il)logical Problem of Heritage Speaker Bilingualism and Incomplete Acquisition. *Applied Linguistics, 33* (4), 1–7.

Picallo, M. Carme (1991). Nominals and Nominalizations in Catalan. *Probus, 3*, 279–316.

Picallo, M. Carme (2008). Gender and Number in Romance. *Lingue e Linguaggio, VII*, 47–66.

Pierantozzi, Cristina (2012). Agreement within early mixed DP: What mixed agreement can tell us about the bilingual language faculty. In C. Gabriel & K. Braunmüller (Eds.), *Multilingual Individuals and Multilingual Societies* (pp. 137–152). Amsterdam: John Benjamins.

Polinsky, Maria (2006). Incomplete acquisition: American Russian. *Journal of Slavic Linguistics, 14,* 191–262.

Polinsky, Maria (2011). Reanalysis in adult heritage language: A case for attrition. *Studies in Second Language Acquisition, 33,* 305–328.

Polinsky, Maria (2016). Structure vs. use in heritage language. *Linguistics Vanguard, 2* (1). doi: https://doi.org/10.1515/lingvan-2015-0036

Polinsky, Maria & Kagan, Olga (2007). Heritage languages: In the 'wild' and in the classroom. *Language and Linguistics Compass, 1* (5), 368–395.

Poplack, Shana (1980). "Sometimes I'll start a sentence in Spanish Y TERMINO EN ESPAÑOL": Toward a typology of code-switching. *Linguistics*, *18* (7/8), 581–618. doi: 10.1515/ling.1980.18.7-8.581

Poplack, Shana (2004). Code-Switching. In U. Ammon, N. Dittmar, K. J. Mattheier, & P. Trudgill (Eds.), *Sociolinguistics: an international handbook of the science of language and society: Vol. 2* (pp. 589–596). Berlin: Mouton de Gruyter.

Poplack, Shana & Dion, Nathalie (2012). Myths and facts about loanword development. *Language Variation and Change, 24* (3), 279–315.

Poplack, Shana; Sankoff, David, & Miller, Christopher (1988). The social correlates and linguistic processes of lexical borrowing and assimilation. *Linguistics, 26*, 47–104.

Putnam, Michael T. & Sánchez, Liliana. (2013). What's so incomplete about incomplete acquisition? *Linguistic Approaches to Bilingualism, 3* (4), 478–508. doi:10.1075/lab.3.4.04put

Pylkkänen, Liina (2008). *Introducing Arguments.* Cambridge, MA: MIT Press.

Ramchand, Gillian (2008). *Verb Meaning and the Lexicon: A First Phase Syntax*. Cambridge: Cambridge University Press.

Rappaport Hovav, Malka & Levin, Beth (1998). Building verb meanings. In M. Butt & W. Geuder (Eds.), *The Projection of Arguments* (pp. 97–134). Stanford: CSLI

Riksem, Brita R. (2017). Language Mixing and Diachronic Change: American Norwegian Noun Phrases Then and Now. *Languages,* 2 (2). doi: 10.3390/languages2020003

Riksem, Brita R. (in press). Language Mixing in American Norwegian Noun Phrases. *Journal of Language Contact, 12* (1).

Riksem, Brita R.; Grimstad, Maren B.; Lohndal, Terje, & Åfarli, Tor A. (in press). Language mixing within verbs and nouns in American Norwegian. *Journal of Comparative Germanic Syntax.*

Ritter, Elizabeth (1993). Where's Gender? *Linguistic Inquiry, 24,* 795–803.

Rizzi, Luigi & Cinque, Guglielmo (2016). Functional categories and syntactic theory. *Annual Review of Linguistics 2*, 139–163. doi: 10.1146/annurev-linguistics-011415-040827

Roberts, Ian (2007). *Diachronic Syntax.* Oxford: Oxford University Press.

Rodina, Yulia & Westergaard, Marit (2013). The acquisition of gender and declension class in a non-transparent system: monolinguals and bilinguals. *Studia Linguistica, 67* (1), 47–67. doi:10.1111/stul.12012

Rodina, Yulia & Westergaard, Marit (2015a). Grammatical Gender in Norwegian: Language Acquisition and Language Change. *Journal of Germanic Linguistics, 27* (2), 145–187. doi:10.1017/S1470542714000245

Rodina, Yulia & Westergaard, Marit (2015b). Grammatical gender in bilingual Norwegian–Russian acquisition: The role of input and transparency. *Bilingualism, 20* (1), 197–214. doi: 10.1017/S1366728915000668

Rothman, Jason (2007). Heritage speaker competence differences, language change, and input type: Inflected infinitives in Heritage Brazilian Portuguese. *International Journal of Bilingualism, 11* (4), 359–389.

Rothman, Jason (2009). Understanding the nature and outcomes of early bilingualism: Romance languages as heritage languages. *International Journal of Bilingualism, 13* (2), 155–163.

Rothman, Jason & Treffers-Daller, Jeanine (2014). A prolegomenon to the construct of the native speaker: heritage speaker bilinguals are natives too! *Applied Linguistics, 35* (1), 93–98.

Rødvand, Linn I. S. (2017). *Empirical investigations of grammatical gender in American Heritage Norwegian* (Master's thesis). University of Oslo.

Sankoff, David & Poplack, Shana (1981). A Formal Grammar for Code-Switching. *International Journal of Human Communication*, 14, 3–46.

Schütze, Carson T. (2011). Linguistic evidence and grammatical theory. *Wiley Interdisciplinary Rewiew: Cognitive Science,* 2 (2), 206–221. doi: 10.1002/wcs.102

Scontras Gregory; Fuchs, Zuzanna, & Polinsky, Maria (2015). Heritage language and linguistic theory. *Frontiers in Psychology 6*:1545. doi: 10.3389/fpsyg.2015.01545

Sorace, Antonella (2011). Pinning down the concept of "interface" in bilingualism. *Linguistic approaches to bilingualism*, *1* (1), 1-33.

Stowell, Tim (1981). *Origins of Phrase Structure* (Doctoral dissertation). Massachusetts Institute of Technology, Cambridge, MA

Svenonius, Peter (2012). Spanning, ms. University of Tromsø, available at ling.auf.net/lingBuzz/001501.

Svenonius, Peter (2016) Spans and Words. In D. Siddiqi & H. Harley (Eds.) *Morphological Metatheory* (pp. 201–222). Amsterdam: John Benjamins.

Szabolcsi, Anna (1983). The Possessor that ran away from Home. *The Linguistic Review, 3* (1), 89–102. doi: 10.1515/tlir.1983.3.1.89

Travis, Lisa (1984). *Parameters and Effects of Word Order Variation* (Doctoral dissertation). Massachusetts Institute of Technology, Cambridge, MA

Treffers-Daller, Jeanine (2005). Evidence for insertional codemixing: Mixed compounds and French nominal groups in Brussels Dutch. *International Journal of Bilingualism*, *9*, 477–508.

Trosterud, Trond (2001). Genus i norsk er regelstyrt [Gender in Norwegian is rule bound]. *Norsk lingvistisk tidsskrift, 19*, 29–58.

Türker, Emel (2000). *Turkish–Norwegian codeswitching: evidence from intermediate and second generation Turkish immigrants in Norway* (Doctoral dissertation). University of Oslo.

van Hout, Angeliek (1996). *Event Semantics and Verb Frame Alternations* (Doctoral dissertation), Tilburg University.

Vangsnes, Øystein A. (1999). The identification of functional architecture (Doctoral dissertation). University of Bergen.

Venås, Kjell (1993). On the choice between two written standards in Norway. In E. H. Jahr (Ed.), *Language Conflict and Language Planning* (pp. 263–278). Berlin: Mouton de Gruyter.

Vigliocco, Gabriella; Antonini, Tiziana, & Garrett, Merrill F. (1997). Grammatical gender is on the tip of Italian tongues. *Psychological science*, *8* (4), 314–317.

Vikør, Lars S. (1995). *The Nordic Languages: Their Status and Interrelations* (Nordic Language Secretariat Publication 14). Oslo: Novus Press.

Westergaard, Marit & Anderssen, Merete (2015). Word order variations in Norwegian possessive constructions. In J. B. Johannessen & J. C. Salmons (Eds.), G*ermanic Heritage Languages in North America* (pp. 21–45). Amsterdam: John Benjamins. doi: 10.1075/silv.18

Williams, Edwin (1981). On the notions *Lexically related* and *Head of a Word. Linguistic Inquiry,* 12 (2), 245–274.

Woolford, Ellen (1983). Bilingual Code-Switching and Syntactic Theory. *Linguistic Inquiry, 14,* 520–536.

Yager, Lisa; Hellmold, Nora; Joo, Hyoun-A; Putnam, Michael T.; Rossi, Eleonora; Stafford, Catherine, & Salmons, Joseph (2015). New Structural Patterns in Moribund Grammar: Case Marking in Heritage German. *Frontiers in Psychology*, *6*. doi: 10.3389/fpsyg.2015.01716

Zuidema, Willem & de Boer, Bart (2013). Modeling in the language sciences. In R. J. Podesva & D. Sharma (Eds.) *Research Methods in Linguistics* (pp. 422–439). Cambridge: Cambrigde University Press.

Article 1

Lexicalist vs. exoskeletal approaches to language mixing

Maren Berg Grimstad[1], Brita Ramsevik Riksem[1], Terje Lohndal[1,2] and Tor A. Åfarli[1]
Norwegian University of Science and Technology[1], UiT The Arctic University of Norway[2]

Abstract:
This article presents empirical evidence that disfavors using highly lexicalist minimalist models, such as the one presented in Chomsky (1995), when analyzing language mixing. The data analyzed consist of Spanish–English mixed noun phrases discussed in Moro (2014) as well as Norwegian–English mixed noun phrases and verbs taken from the *Corpus of American Norwegian Speech*. Whereas the lexicalist model in Chomsky (1995) only can explain a subset of the mixing patterns attested in both authentic Spanish–English mixed noun phrases and the American Norwegian corpus, we show that an alternative exoskeletal model can account for all of them. Such a model would entail that rather than assuming lexical items with inherent, functional features that determine the derivation, syntactic structures are generated independently from the lexical items that come to realize them.

Key words:
Agreement, American Norwegian, Exoskeletal, Language mixing, Lexicalism

1 Introduction[1]

In this article, we argue that a lexicalist analysis relying on features being an inherent property of lexical items, as in mainstream analyses within the Minimalist Program, is ill-prepared to explain the phenomenon of language mixing, i.e., intrasentential mixing where linguistic strings contain elements from both a language A and a language B. The main reason is that the feature matching/agreement typically adopted by a lexicalist feature-driven syntax naturally requires matching between elements of the same language, viz. the probe and the goal need the same features. Contrary to that requirement, we will show that items drawn from different lexicons often do not have the matching/agreeing features that are required for convergence given such an analysis, yet they frequently co-occur in language mixing. We argue that lexicalist-type feature matching theories generally predict that language mixing will be extremely restricted, contrary to fact.

The focus of this article is critical since we will concentrate on the shortcomings of lexicalist feature-driven syntax as a tool for the analysis of language mixing. This relates to

[1] We are grateful to audiences at a range of conferences around the world for feedback on various incarnations of ideas contained in this paper. We would also like to thank the two anonymous reviewers for constructive feedback that we believe made the paper better.

the debate between MacSwan (2000, 2005) and Jake, Myers-Scotton, and Gross (2002, 2005) concerning the utility of a minimalist approach to language mixing. However, the current analysis will improve on both approaches in that it develops an exoskeletal analysis which provides a formal and more descriptively adequate generative analysis of the data. Given the programmatic nature of the Minimalist Program, we consider our exoskeletal analysis a variety of minimalism, albeit a non-lexicalist variety.

We will argue that an exoskeletal approach to language mixing can account for the data that we claim are problematic for the lexicalist approach, and further that it correctly predicts that language mixing is ubiquitous in language. The latter is not surprising, given the overwhelming evidence that both grammars are active simultaneously in the mind of bilinguals (see Kroll & Grollan 2014 for an overview). Particular exoskeletal analyses have been defended in detail in our other work (see, e.g., Grimstad, Lohndal, & Åfarli 2014, Alexiadou, Lohndal, Åfarli, & Grimstad 2015). In this article, we will present a general outline of how an exoskeletal analysis works in order to demonstrate that it provides a better alternative to the analysis of language mixing.[2]

The article is organized as follows. We start out in Section 2 by discussing different types of theories of language mixing. In Section 3, we discuss the main characteristics of the standard version of the Minimalist Program and in particular its possible relevance to mixing phenomena. Section 4 provides a detailed criticism of one particular minimalist lexicalist analysis of mixing phenomena that has been proposed, namely Moro (2014), which seeks to explain Spanish–English mixing within the DP. In Section 5, we investigate a possible minimalist lexicalist approach to mixing phenomena within the DP as found in American Norwegian. Section 6 provides an exoskeletal analysis which is argued to be superior both descriptively and theoretically. In Section 7, we consider whether the lexicalist approach works better for mixing in the verbal domain, concluding that it does not, and we suggest an exoskeletal analysis of that as well. Section 8 concludes the article.

[2] We will not discuss how to capture the fact that certain patterns of mixing are more common than others. For an approach that incorporates a probabilistic distribution of attested forms, see Goldrick, Putnam, and Schwarz (2016).

2 Theories of language mixing

There are essentially two types of language mixing theories: those that posit special machinery to handle mixing data and those that do not.[3] The latter are the so-called null theories or constraint-free theories, see, e.g., Mahootian (1993), MacSwan (1999, 2014), González-Vilbazo and López (2011, 2012), Pierantozzi (2012), Bandi-Rao and den Dikken (2014), Grimstad et al. (2014), Åfarli (2015a), Merchant (2015), and Alexiadou (2017).

According to Mahootian (1993: 3), a null theory of what we label language mixing asserts that mixing is not constrained by any special mechanisms or principles specific to mixing, and that "exactly the same principles which apply to monolingual speech apply to codeswitching". That a theory of language mixing should be a null theory is important simply because the internalized grammars that we postulate should be able to account for all sorts of natural language outcomes, including language mixing. If we are forced to postulate special mechanisms to account for language mixing outcomes, that would mean that neither the internalized grammar that we assume nor the special mechanisms that we adopt are on the right track. This is simply a question of theoretical parsimony. Muysken (2000: 3) states that

> [t]he challenge is to account for the patterns found in terms of general properties of grammar. Notice that only in this way can the phenomena of code-mixing help refine our perspective on general grammatical theory. If there were a special and separate theory of code-mixing, it might well be less relevant to general theoretical concerns.

MacSwan (2014: 2–3) claims that whereas many language mixing theorists have considered the attainment of a null theory or constraint-free theory of language mixing to be the ideal, in practice, theories and analyses still have resorted (explicitly or implicitly) to special mechanisms for language mixing, i.e. constraint-based mechanisms in MacSwan's terms. This has often been based on postulating constraints on where language mixing would occur (see Pfaff 1979, Poplack 1980, and Sankoff and Poplack 1981 for important early work), constraints that were unique to mixing as such. See MacSwan (2014: 2 ff.) for extensive discussion of how particular analyses in the history of language mixing theory have fared in this respect. Thus, according to MacSwan, there is an unfulfilled quest in language mixing analyses/theories for constraint-free or null theory solutions, something which he claims is

[3] In this article, we will use the broader term "language mixing" instead of codeswitching. This is mainly to set aside the issue of how to distinguish between codeswitching and nonce-borrowing; see Grimstad et al. (2014) and Grimstad (2017) for discussion.

fulfilled by the standard lexicalist version of the Minimalist Program. This version ensures that "Nothing constrains [code-switching] apart from the requirements of the mixed grammars" (MacSwan 1999, 2014: 18).[4]

We agree that it is important that the analysis of language mixing is based on a null theory. However, we are not convinced that a minimalist lexicalist approach or other lexicalist feature-driven approaches are ideally suited to account for language mixing phenomena, a claim that we will try to substantiate in the remainder of this article. In what follows, we will introduce the standard lexicalist version of the Minimalist Program before turning to how and to what extent both that approach and the exoskeletal approach can explain certain patterns of language mixing.

3 The Minimalist Program

The standard version of the Minimalist Program, or just minimalism, adopts a lexicalist feature-driven model of grammar (Chomsky 1995; see a textbook version in Adger 2003).[5] This is the version of lexicalism that we will discuss in the present paper, even though lexicalism comes in many different guises (see Ackerman, Stump, & Webelhuth 2011). For reasons of space, we limit our focus to this particular version and set aside how other versions of lexicalism potentially could deal with the data in the present paper.

Within the minimalist lexicalist approach, phrase structures are generated or projected based on formal features of lexical items. Thus, the features of these lexical items determine in part the syntactic structure. A mechanism called *Select* in Chomsky (1995) provides a selection of items from the lexicon. These items constitute a Numeration or a Lexical Array (the difference need not concern us here). The computational system then generates a structure based on the numeration/lexical array. The resulting structure is further altered by way of agreement and movement, which in turn is partly driven by feature matching/checking/valuation, that is, the need to value unvalued grammatical features that are driving the derivation. Put differently, functional features are typically unvalued and need to be valued by valued lexical features. To give one example, consider subject–verb agreement in English. The assumption is that the structure looks like in (1), where *U* denotes an unvalued feature.

[4] For MacSwan, there is an important distinction between codeswitching and nonce borrowing. As mentioned, we use "language mixing" as a more theoretically neutral and descriptive label.

[5] It is important to note that minimalism is a program, hence it is hard to accurately characterize what minimalism as such adopts. Nevertheless, we believe that it is fair to say that most practitioners adopt a lexicalist feature-driven view of derivations.

(1) [TP T[NUM:*U*, PERS:*U*, CASE:NOM] [vP [DP she [NUM:SG, PERS:3, CASE:*U*]] [v [VP …]]]]

Movement and agreement ensure that the correct grammatical representation is arrived at, where strikethrough denotes a copy/trace.

(2) [TP [DP she [NUM:SG, PERS:3, CASE:NOM]] T[NUM:SG, PERS:3, CASE:NOM] [vP ~~[DP she[NUM:SG, PERS:3, CASE:*U*]]~~ [v [VP …]]]]

Crucially, feature valuation takes place through Agree, which is an abstract agreement operation that connects a probe to a goal. The Agree mechanism ensures that the same feature values occur in two different places. In Chomsky (1995), a specifier–head relationship, as in (2), had to be established in order to trigger Agree (see also Koopman 2006 for a later defense of the same idea). In later versions, Agree was argued to take place long-distance (Chomsky 2000), so that movement had to be captured through additional movement-triggering features, such as a 'generalized EPP' feature (Chomsky 2001). Crucially, no feature can be sent off to the interfaces without being valued. An unvalued feature causes a crash at the interface.

There are several approaches to language mixing within the Minimalist Program. MacSwan (1999, 2000, 2005, 2009, 2014), Chan (2008), González-Vilbazo and López (2011, 2012), Shim (2013), Bandi-Rao and den Dikken (2014) all pursue different versions, for example. In the present paper, we will focus on the approach most closely associated with MacSwan since this is the version that is most clearly associated with lexicalism.

MacSwan (1999, 2000, 2005, 2009, 2014) relies on the technical approach in Chomsky (1995) in developing his minimalist lexicalist approach to language mixing. Given the assumption that it needs to be a null theory, MacSwan argues that "[…] lexical items may be drawn from the lexicon to introduce features into the lexical array, which must then be valued […] in just the same way as monolingual features must be valued, with no special mechanisms permitted" (MacSwan 2009: 326). Within a lexicalist version of minimalism, differences between languages are attributed to differences regarding lexical and functional items (cf. Borer 1984). As MacSwan (2005: 2) puts it, "[i]n the MP, there are two central components of the syntax: C_{HL}, a computational system for human language, presumed to be invariant across languages, and a lexicon, to which the idiosyncratic differences observed across languages are attributed." Furthermore, "[p]arameters [are] restricted to the lexicon rather than operating on syntactic rules" (MacSwan 2005: 2). This is an important part of MacSwan's approach, because it generates predictions about which patterns that can be mixed

and which that cannot. In the next section, we will consider an example of this which will be used to illustrate MacSwan's approach.

4 A minimalist lexicalist analysis of DPs in Spanish–English mixing

In recent years, there has been a lot of research into gender marking in bilingual grammars, as seen, among others, in Liceras, Fuertes, Perales, Pérez-Tattam, and Spradlin (2008), Cantone and Müller (2008), Parafita Couto, Munarriz, Epelde, Deuchar, and Oyharçabal (2015); Valdés Kroff (2016), Valdés Kroff, Dussias, Gerfen, Perrotti, and Bajo (2016), and Johnson Fowler (2017). We will scrutinize one particular paper because it explicitly adopts a minimalist lexicalist approach, namely Moro (2014). This paper investigates Spanish–English mixing in DPs in a linguistic variety spoken in Gibraltar from the point of view of MacSwan's approach. The article contains very little information about the status of Spanish and English in the linguistic community that the data are drawn from, or even about the immediate linguistic context of the nominal strings that she considers (a point of some importance, as we will note below), so we have at the outset to take her data at face value.

Moro (2014) considers two possible types of mixing between D and N: one where D is Spanish and N is English, as in (3), and another where D is English and N is Spanish, as in (4). According to her, only the first pattern is well-formed in the mixing variety that she considers.

(3) a. el **employer**
'the employer'

b. la **washing machine**
'the washing machine'

(4) a. ***the** casa
'the house'

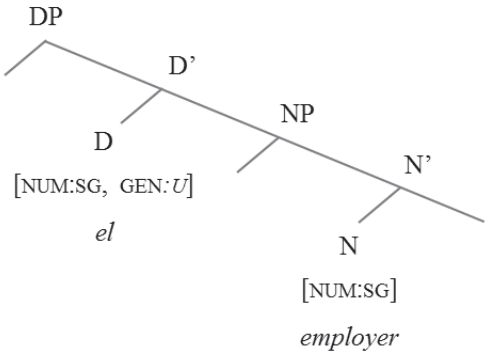
b. ***the** vecina
'the neighbor'

Spanish is a grammatical gender language whereas English is not, and Moro implicitly assumes that the *-er* in (3a) does not indicate natural gender. The gender difference will play a crucial role in the analysis, as we will see momentarily.

Now, consider first Moro's analysis of data of the type in (3), using (3a) as our example. D exhibits gender and number, while the English noun lacks the gender feature. Moro (2014:

223) analyzes this in the following minimalist lexicalist manner, exploiting valued and unvalued features: "I assume […] that the number and gender features of the determiner enter the derivation unvalued and have to be valued via Agree with the corresponding N." Although Moro does not provide structures, we assume that the structure of (3a) is as shown in (5), where only relevant features are included.

(5)



As can be seen, the NUM category on D is valued as SG by the corresponding category on N. The GEN category on D, however, lacks a corresponding valued category elsewhere in the structure and thus appears to remain unvalued. Note that despite this, the representation in (5) is the representation of a well-formed mixing pattern, according to Moro. She suggests that "[a]ccordingly, the unvalued features number and gender in the Spanish determiner can be valued via Agree with the English noun because the former bears the full set of these features (number and gender)" (Moro 2014: 223).

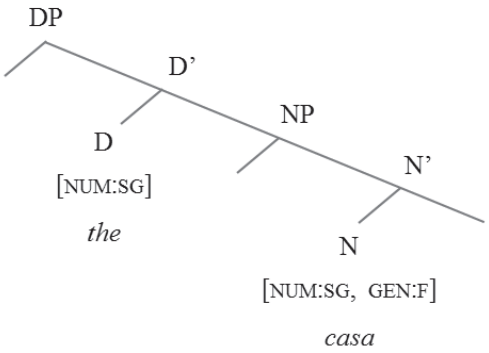
This is just stipulated, and Moro does not explain exactly how the unvalued GEN feature in D can be valued by a non-existent GEN feature on the English noun, given that the determiner "bears the full set of these features (number and gender)." In our view, there would in fact not be any problems for the analysis if D failed to bear the so-called full set of features, as Moro suggests. Rather, a problem arises precicely when the English noun does not bear the inherently valued feature required for valuation of the corresponding feature on D, as in this case. Therefore, it seems to us that Moro's solution invokes an 'impossible' theoretical rule or principle, namely a principle that amounts to a claim that an unvalued feature may be valued by a non-existent feature, i.e., that [GEN:*U*] in "*el* [NUM:*U*, GEN:*U*]" can be valued by [NUM:SG] in the noun *employer*.

Now, consider the mixing illustrated in (4), where the determiner is English and the noun is Spanish. We use (4a) as our example, repeated here for convenience.

(4) a. ***the** casa
'the house'

Notice, crucially, that the combination of the English determiner *the* and a Spanish noun is not well-formed according to Moro. Moro's analysis will presumably be something like the one shown in (6).

(6)



Here, N bears NUM and GEN features (since the noun is Spanish), whereas the English D does not bear a GEN feature. Given standard assumptions about feature checking, this example should be well-formed; there are no unvalued features left that could cause a crash at the interfaces.

However, this is an unwanted result as far as Moro is concerned, since this precise mixing pattern is assumed to be ungrammatical in her data. Her solution is seen in the following quotation: "On the contrary, the derivation crashes in the case of the English determiner and the Spanish noun because the feature set of the English determiner is incomplete (it lacks the gender feature)" (Moro 2014: 223). This, to us, appears to be another stipulation without any empirical or theoretical justification. First of all, we cannot see any justification for assuming something like a "complete set of features for D" that holds cross-linguistically. Moreover, since English *the* does not contain/express gender in the first place, it should not be problematic that it does not contain GEN. In fact, a more natural assumption would be that it is problematic for *the* to contain a gender feature. *Prima facie*, one should

think that it would be possible for, e.g., [NUM:*U*] in *the* to be valued by [NUM:SG] in *casa* irrespective of the other valued or unvalued features involved. In general, an analysis that makes valuation of a particular feature dependent on the presence of features of a completely different type is in need of strong independent motivation. Unless such motivation is forthcoming, such an analysis should be discarded.

Moreover, as Liceras et al. (2008) make clear, Moro's (2014) (cited as 2001 in Liceras et al. 2008) empirical claim is factually wrong. Contrary to Moro's claim, instances like (4) are in fact attested in spontaneous production (see the reviews in Liceras, Spradlin, & Fuertes 2005, Liceras et al. 2008, and Pierantozzi 2012). For reasons of space, we will not delve into that debate here, but simply assume that (4) contains well-formed mixing data that have to be accounted for, contrary to Moro's claim. Liceras et al. (2005, 2008) furthermore argue that the Spanish determiner is *preferred*. They propose to account for this preference by suggesting a Grammatical Features Spell-out Hypothesis (GFSH), which claims that functional categories containing highly 'grammaticized' features will be chosen. Since Spanish determiners contain more features than English determiners, the speaker will choose the former. Note that the GFSH is a hypothesis about production preferences guided by a grammatical mechanism on the PF side.

In summary, we have shown that Moro's (2014) minimalist lexicalist analysis of DP-internal language mixing is empirically inadequate. We conclude that a lexicalist feature-based analysis is the wrong tool for analyzing language mixing in the nominal domain, the reason being that such analyses require more feature matching than is actually found in mixing. In the next section, we will look at data from American Norwegian which will lend further support to this conclusion.

5 A possible minimalist lexicalist analysis of mixing in American Norwegian DPs

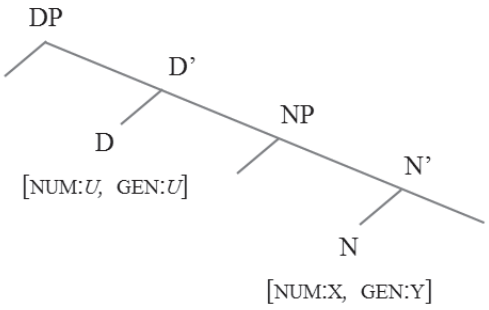
In this section, we will consider language mixing in American Norwegian DPs as a way of solidifying the conclusions reached in Section 4. We will first briefly give a description of American Norwegian before we attempt to envision how a minimalist lexicalist model of the sort adopted by MacSwan and Moro could possibly handle mixing of English forms into American Norwegian DPs.

American Norwegian is a heritage variety of Norwegian spoken in North America (mainly in the US) by immigrants who came from Norway roughly from the 1850s until the

1920s, as well as their Norwegian-speaking descendants. In other words, American Norwegian is a minority language existing in the midst of a language community heavily dominated by English. A common factor for the speakers in question is that American Norwegian is their L1 and in many cases their only language up until school age. However, through schooling and regular interaction in the community, English has come to be their dominant language. As a consequence of their language situation, these speakers often produce linguistic outcomes showing a mixture of the two languages. In the following, we will turn to analyses of such mixed linguistic outcomes in the nominal domain. The corpus we exploit for data is the *Corpus of American Norwegian Speech* (CANS) (Johannessen 2015), developed by the Text Laboratory at the University of Oslo. At the moment, this corpus comprises recordings from 50 individual speakers.

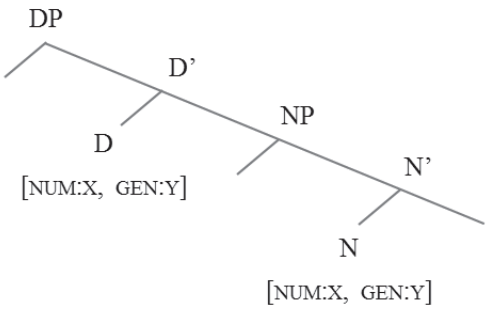
Given that American Norwegian is a heritage variety of Norwegian, let us first take a look at how the mechanism of feature valuation proceeds in a Norwegian DP without mixing. Like Spanish, Norwegian is a grammatical gender language, and we use a simplified DP structure which only contains a D-projection and an N-projection. As we are, for the moment, concerned with a minimalist lexicalist type analysis, we make the lexicalist assumption that the agreement features NUM and GEN are inherent in N, and that DP internal agreement comes about as a result of the agreement features in N valuing the corresponding unvalued features in D. This is illustrated in (7), where [NUM:X, GEN:Y] in N denotes the particular inherently fixed agreement features in question, and where [NUM:*U*, GEN:*U*] denotes the corresponding agreement features in D.

(7)



After valuation of the unvalued features in D by a probe–goal relation, the resulting agreement structure will be as shown in (8), where N and D have identical feature values.

(8)



Consider now a concerete example:

(9) a. dette hus-et
this.SG.N house-SG.DF.N
'this house'

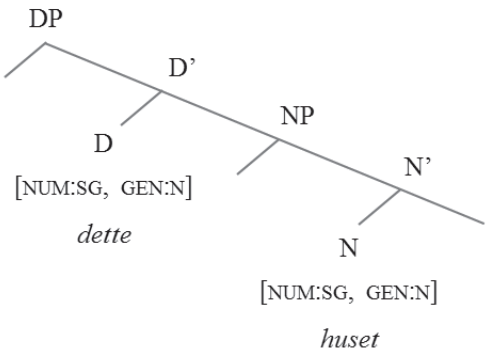
b. *denne hus-et
this.SG.M/F house-SG.DF.N

c. *dette hus-a
this.SG.N house-PL.DF.N

(9a) shows the correct agreement pattern inside the DP, with both N and D marked as SG and N. (9b) is ungrammatical due to an agreement mismatch, the N being marked as SG.N, whereas D is marked as SG.M/F. In other words, there is a gender mismatch between N and D that cannot be generated given the Agree mechanism. (9b) therefore fails to be generated, as desired. The same goes for (9c), except in this case there is a number mismatch, N being marked as PL while D is marked as SG. (9a) is shown in (10), which is the structure after valuation of the unvalued features in D.[6]

[6] Notice that (9a)/(10) exhibit double definiteness, i.e. definiteness realized both by a determiner and by a functional suffix on the noun, which is characteristic for Norwegian DPs, see Julien (2003, 2005) for discussion. Still, the definiteness feature is not shown in the representation (10) (nor in our subsequent representations of the Norwegian DP) for expository purposes, since what we concentrate on here is the logic of Moro's (2014) analysis of DP internal mixing, where definiteness is left out of consideration. In Section 6, double definiteness will play a crucial role in motivating our exoskeletal structure for the DP.

(10)



Let us next consider feature valuation in American Norwegian DPs in which English elements are mixed into the otherwise Norwegian string. In American Norwegian, it is common for an English noun to be embedded under a Norwegian determiner, e.g., an indefinite article. This is exemplified in (11). The information in parenthesis behind each American Norwegian example is a reference to the speaker in the CANS corpus who uttered that specific phrase, and the mixed English noun is boldfaced.[7] Note that gender is not fixed across speakers and is not in general identical to their Norwegian translational equivalents.[8]

(11) a. en **blanket** (rushford_MN_01gm)
a.M blanket
'a blanket'

b. ei **nurse** (coon_valley_WI_02gm)
a.F nurse
'a nurse'

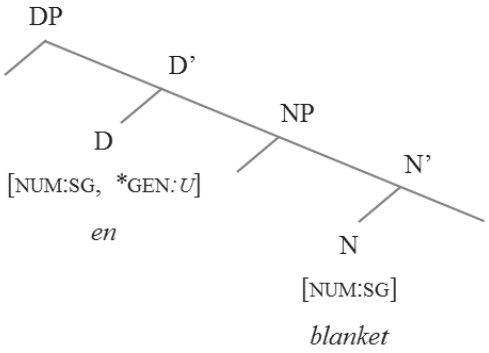
c. et **crew** (westby_WI_03gk)
a.N crew
'a crew'

Adopting a standard minimalist lexicalist analysis, D has unvalued GEN and NUM features that must be valued by the corresponding fixed values on N, as explained above for standard Norwegian. However, there is a problem with this analysis given mixing cases like those in (11). Since English nouns do not have a gender feature, the GEN feature of D remains

[7] We have not indicated what the Norwegian counterparts of the English words would be, but they are very different from the English ones.
[8] A few examples of this are: *choiren*.M (coon_valley_WI_07gk) – *koret*.N (Norwegian); *ferryen*.M (harmony_MN_04gm) – *ferga*.F (Norwegian); *et*.N *title* (stillwater_MN_01gm) – *ein*.M *tittel* (Norwegian).

unvalued and the derivation will crash, contrary to the fact that such structures are common and therefore should converge.[9] These examples are parallel to the Spanish cases in (3), where Moro (2014) stipulated a solution where the "full set" of features in D would enable valuation from the English noun. As discussed in Section 4, we do not see how an unvalued GEN feature in D can be valued by a non-existing GEN feature in N, and therefore reject such an analysis. The ill-formed structural representation of (11a) is given in (12), the ill-formedness being indicated by an "*" on the offending feature category.

(12)



Now, consider other comparable examples where D is a demonstrative or determiner, illustrated in (13) and (14), respectively.[10]

(13) a. denne **cheese** (blair_WI_04gk)
this.M/F cheese
'this cheese'

b. denne **heritage tour-**en (flom_MN_01gm**)**
this.M heritage tour-SG.DF.M
'this heritage tour'

c. dette **computer business** (harmony_MN_01gk)
this.N computer business
'this computer business'

[9] An alternative approach could be to argue that the speaker has internalized two lexical items for each English noun: one with gender and one without gender. We assume that this analysis is implausible, in part because some speakers of American Norwegian vary in their gender assignment (Lohndal & Westergaard 2016).

[10] From the perspective of (European) Norwegian, we would, in data like (13) and (14), expect (near) obligatory double definiteness, that is, we would expect the English nouns to have a post-nominal Norwegian definite suffix as well. This is what we find in earlier varieties of American Norwegian, but in contemporary varieties of American Norwegian, we find double definiteness only in approximately half of the relevant cases. The remaining half lacks either the suffix or the determiner, which probably is an effect of attrition. Space prevents us from discussing this further, but see Riksem (2017).

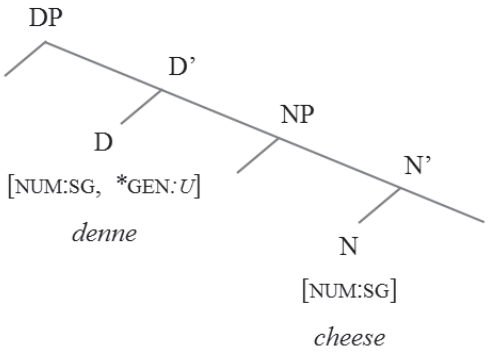
(14) a. alt det gamle **stuff** (chicago_IL_01gk)
all the.N old stuff
'all the old stuff'

b. denne digre **chopper**-en (blair_WI_01gm)
this.M huge chopper-SG.DF.M
'the huge chopper'

c. den samme **lodg**-en (vancouver_WA_03uk)
the.M same lodge-SG.DF.M
'the same lodge'

Again, the problem is that the GEN feature of D cannot be valued because the English noun does not bear a gender feature, cf. the ill-formed structural representation of (13a) in (15).

(15)



Now, consider American Norwegian DPs where the Norwegian–English mixing pattern is switched, so to speak. We will focus on patterns where there is a mix consisting of an English definite article and a Norwegian noun, as seen in the examples in (16).

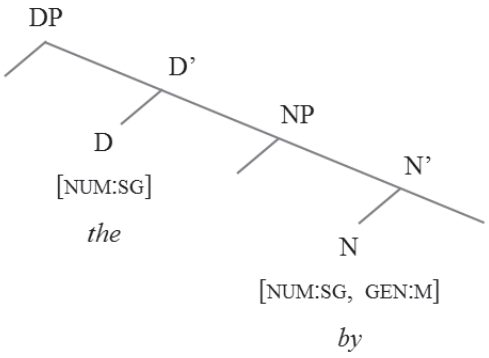
(16) a. **the** by (chicago_IL_01gk)
the city.SG
'the city'

b. **the** gård (vancouver_WA_01gm)
the farm.SG
'the farm'

c. **the** penger (albert_lea_MN_01gk)
the money.PL
'the money'

In these examples, there is a Norwegian noun with gender and number features, but the English definite article *the* probably bears only a number feature and in any case lacks a gender feature. The structural representation of (16a) is shown in (17).

(17)



As can be seen in (17), there are no features that remain unvalued in this representation. As such, this representation should be deemed well-formed, at least as far as feature valuation is concerned. Note that these strings are parallel to the Spanish–English cases in (4), which Moro (2014) judged to be ungrammatical. The rationale for this conclusion was that the "incomplete" set of features in D caused an inability of feature valuation from N. Again, as discussed in Section 4, we do not see how the lack of a GEN feature in D should prevent NUM in N from valuating NUM in D. (17) illustrates that after such a valuation, there are no unvalued features left in the structure that could make the derivation crash. As discussed above, Liceras et al. (2008) argue against Moro on empirical grounds, which also aligns with the evidence found in American Norwegian.

As seen in this subsection, attempting to analyze the American Norwegian data under a lexicalist approach is problematic. On the one hand, if we were to accept Moro's analysis, mixed phrases containing Norwegian D and English N would be acceptable, whereas phrases with English D and Norwegian N would be unacceptable. Rejecting Moro's analysis based on the discussion in Section 4 and employing standard minimalist mechanisms of valuation, on the other hand, would apparently reverse the picture; the latter pattern would be acceptable and the former unacceptable. The crucial fact is nevertheless that both mixing patterns are attested in our corpus. In the next section, we will propose an alternative analysis.

6 An exoskeletal analysis of language mixing in DPs

We argue that an exoskeletal approach to grammar provides a more adequate analysis of language mixing. The following outlines such an analysis and demonstrates how the previously discussed data may be analyzed.

Exoskeletal approaches to grammar, also known as generative, neo-constructivist approaches, have been developed by several scholars, e.g., van Hout (1996), Marantz (1997, 2013), Borer (2005a, b, 2013), Åfarli (2007), Ramchand (2008), Lohndal (2012, 2014), and Alexiadou, Anagnostopoulou, and Schäfer (2015). The details of the proposed models and analyses vary, but crucially, they share a common core, namely that the structures are generated independently from the lexical items that will come to realize them. Hence, lexical items do not have inherent, functional features that determine the derivation of the structures, but functional features are instead defined by the syntactic structure. Although exoskeletal models primarily have been applied to monolingual data, the approach has proven to be quite successful in analyzing language mixing, see, e.g., González-Vilbazo and López (2011, 2012), Pierantozzi (2012), Bandi-Rao and den Dikken (2014), Grimstad et al. (2014), Åfarli (2015a), Alexiadou, Lohndal, et al. (2015), Merchant (2015), Alexiadou (2017), and Riksem (in press).

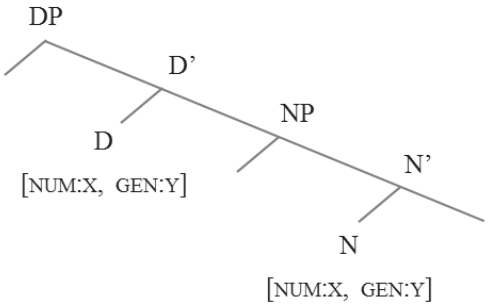
Furthermore, the model employed in this article builds on certain aspects of Distributed Morphology (DM) (see, e.g., Harley & Noyer 1999, 2000; Marantz 1997; Embick 2000; Alexiadou 2001; Embick & Noyer 2007). The core and common ingredient is the assumption that syntactic structures consist of features that define the morphosyntactic properties. In addition, what is typically known as the lexicon is distributed in DM across three separate lists: 1) syntactic terminals, 2) vocabulary items, and 3) encyclopedic information. These are accessed at different points thoughout the derivation. The structure then distinguishes between two types of terminals: functional features or feature bundles, and designated slots for root/stem insertion.[11] At Spell-Out, morpho-phonological exponents, known in DM as vocabulary items, are inserted and give the structure its phonological realization. This process is radically different for the two types of syntactic terminals. Following the Subset Principle (Halle 1997), functional exponents are required to match all or a subset of the functional features specified in the structure to be inserted. If two exponents are eligible for insertion, the one matching the most features will win. This ensures that a given feature or feature bundle in

[11] We will not go into the discussion of roots (their nature, when they are inserted, etc.) in this article. Note, however, that the NP is a simplification and most likely contains more structure – for example a nominalizing phrase, nP, with a bare root in its complement space, as linguists working within Distributed Morphology would typically assume.

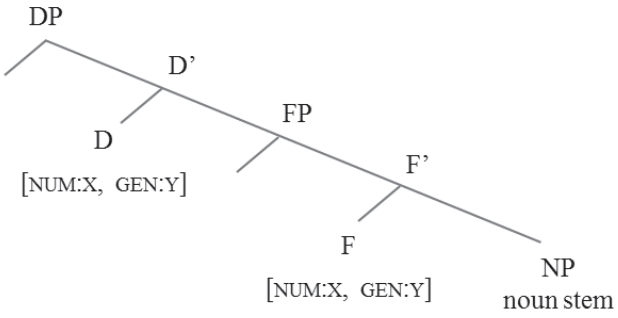
the structure is spelled out by the most appropriate exponent available. Substantial exponents, i.e., roots or stems, are instead inserted into designated slots in the structure without such feature matching requirements, except that they need to match the relevant category feature.

With this model, we can now provide an analysis of the mixed American Norwegian DPs. Instead of (8), repeated below, where the inflectional properties are a property of the noun itself, we assume (18), where the inflectional properties are generated under a functional projection called F, and the noun stem is generated in the complement domain of that F.

(8)



(18)



Let us briefly review some evidence in favor of the FP projection. A core piece of evidence comes from the existence of double definiteness in Norwegian and American Norwegian. Consider (19):

(19) den gamle mann-en
the-DF.SG.M old man-DF.SG.M
'the old man'

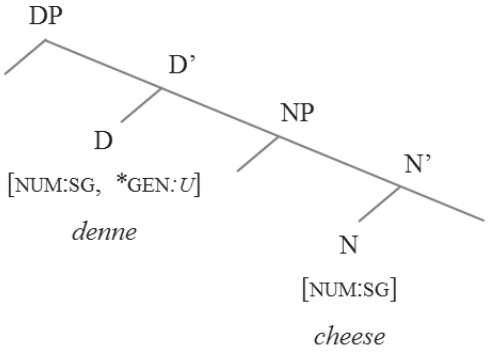
In (19), definiteness is encoded both pre- and postnominally. Julien (2005), building on an extensive review of previous research into the nominal phrase in Norwegian, argues that there is a functional projection for each of the two definiteness features. The prenominal one is situated in D, whereas she suggests that the postnominal one serves as the head of a definiteness projection in the lower domain of the DP. The adjective, situated as the specifier of its own projection (αP in Julien 2005), then agrees with the features of the definiteness and D heads. Julien also assumes that there is a separate NumP. In (18), FP is a different label for the lower definiteness projection, which also encompasses number. We have collapsed both definiteness and number onto one head, both because we have not been able to find evidence in American Norwegian for distinguishing them into two projections, and because the choice between one or two projections is not crucial for present purposes. Furthermore, we assume that FP is always present, but its feature content differs across varities. In English, FP only has a NUM feature, whereas the Norwegian and American Norwegian FP both have NUM, GEN and DEF features. Just like subject–verb agreement at the sentential level differs between English and Norwegian, with Norwegian not exhibiting features for subject–verb agreement, so does the presence of the functional features in the nominal domain.

Given (18), noun stems from any language will acquire the inflectional properties of the language that specifies the syntactic frame.[12] Thus, the prediction is that mixing is fairly free, as is actually the case in American Norwegian. Let us consider (13a) *denne cheese,* "this cheese". As the structure in (15) shows, repeated below, the lexicalist approach fails to predict this pattern, as the feature matching process will leave the GEN feature of D unvalued. The exoskeletal representation in (20) fares better. The functional features are here determined by the syntactic structure, which in this case is Norwegian. Given the Subset Principle, inserted functional exponents must provide the best match to the functional features in the structure. In this case, the Norwegian exponent will provide the best match, as the structure holds a GEN feature. The complement position of F, on the other hand, is available for insertion of an

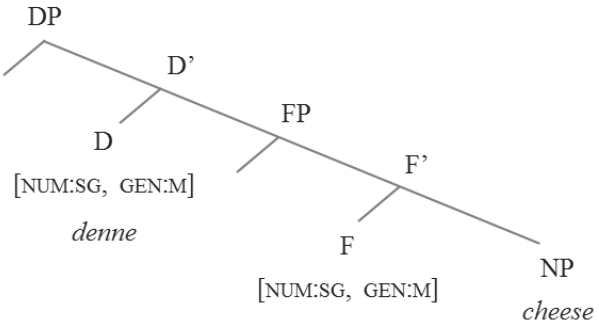
[12] Please note that although we will speak of an English or a Norwegian structure in the remainder of this article, this is merely informal: We do not assume any "language features" whereby syntactic features are annotated for language. Rather, we assume that the "language mode" of the speaker will determine what kind of features are selected as the basis for the abstract syntactic structure. The idea is that the speaker is attempting to speak a certain language, e.g., American Norwegian, and that this manifests itself by virtue of the overall structure of the sentence mimicking or resembling this language. Depending on the features, different structures and thereby potentially different morphosyntactic realizations will be produced. We set aside the precise implementation of this important issue for future research.

English noun stem.[13] Mixing patterns like these can thus be characterized as English stems being inserted into Norwegian structure, and are in fact the most frequent in the corpus.

(15)



(20)



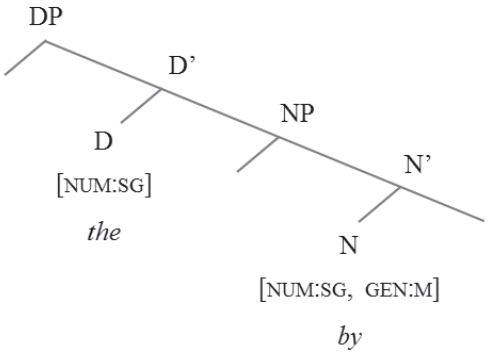
As mentioned in footnote 10, earlier American Norwegian (and also European Norwegian) would typically have double definiteness in examples corresponding to (15). As shown in (13b) and (14b, c), that is also very often the case in contemporary American Norwegian; although as noted, there is vacillation among contemporary speakers (see Riksem 2017 for an analysis of instances where the postnominal definite suffix is missing, arguing in favor of features being either rearranged or erased from the structure). However, it is important to point out that the presence of double definiteness in American Norwegian mixing examples like (13b) and (14b, c) provides further empirical support for an exoskeletal analysis. The

[13] Note that this model would work just as well for the data in this paper if the smallest lexical building blocks for the syntax were uncategorized roots merged with a categorizer. For ease of exposition, we will assume that the mixed unit is a categorized stem, but see Riksem, Grimstad, Lohndal, and Åfarli (in press) for a discussion of whether these mixed items below word-level in American Norwegian are roots or stems.

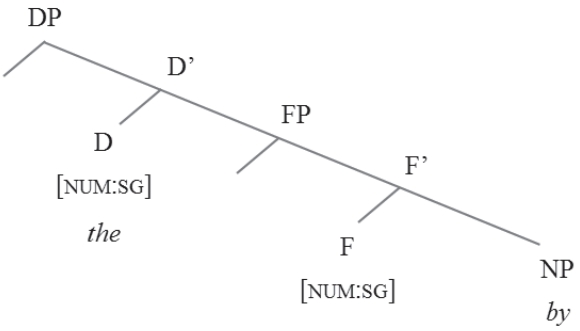
reason is that the definiteness suffix bears number and gender features which must be Norwegian, thus providing evidence for the existence of the F head as part of the exoskeletal frame. See Riksem (in press) for further data and analysis.

Considering the second pattern of determiner–noun mixing in American Norwegian, i.e., where the determiner is English and the noun Norwegian, this, too, can be analyzed with an exoskeletal model. The lexicalist version of (16a) *the by* is repeated in (17), followed by an exoskeletal representation in (21).

(17)



(21)



As the structures show, cases of English D plus Norwegian N are successfully analyzed in both frameworks, despite Moro's claim that such combinations are ungrammatical. The crucial difference is, however, that whereas the exoskeletal model is capable of accounting for both mixing patterns attested in the corpus, the lexicalist analysis will fail in cases like (15).

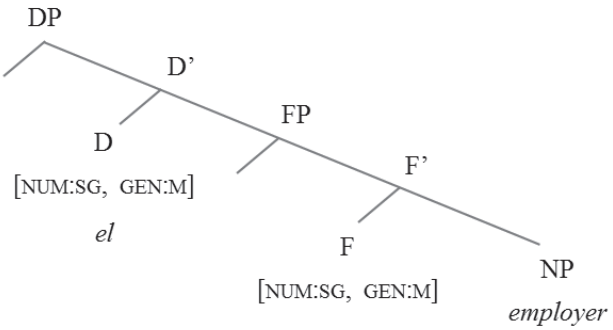
The process of generating the structure in (21) is not considered to be quite the same as the one for the structure in (20). Whereas the main language in (20) is Norwegian, hence the

GEN feature in D and F, the main language for the DP in (21) is assumed to be English, meaning neither D nor F have a GEN feature. Evidence in favor of this is the absence of double definiteness, i.e., the speaker says *by* 'city' and not *byen* 'the city', which together with the English determiner suggests that the determiner causes a change into English for the rest of the noun phrase. In other words, what we assume we have here is a Norwegian stem inserted into an English structure, the opposite of what we have seen so far.[14] This English DP can further be inserted into the DP slot of a larger Norwegian utterance, which is the case for this particular example:

(22) Jeg husker ikke **the** by der vi stoppet. (chicago_IL_01gk)
I remember not the city there we stopped
'I don't remember the city where we stopped.'

As for the Spanish–English mixing data, the mixing pattern that is well-formed according to Moro (2014) corresponds to the American Norwegian structure (20), see (23):

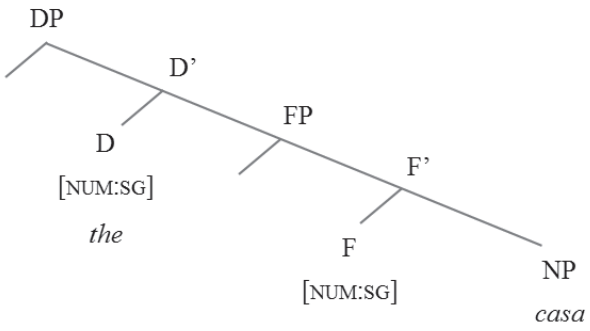
(23)



In addition, the pattern that Moro (2014) claims is ill-formed is also, like the American Norwegian (21), predicted to be grammatical given an exoskeletal analysis, cf. (24).

[14] Another possible solution would be to assume that the structure below D is Norwegian, so that the feature bundle in F is unchanged, whereas the one in D contains only the NUM feature. Thus, the feature bundle in D allows insertion of the English determiner. We have a few examples of structures with English determiner, English or Norwegian noun stem, and Norwegian definiteness suffix, i.e., double definiteness, such as *the roaden*, "the road-DF" and *the andre dagen*, "the other day-DF". However, we are talking about only a couple of examples found in the corpus, so we will leave this issue aside for now (see Riksem 2017). The key takeaway is that both of these solutions are compatible with an exoskeletal analysis, and both possibilities may also co-exist in the contemporary American Norwegian speech community.

(24)



As mentioned in Section 4, Moro (2014) provides little information concerning the status of Spanish and English in the linguistic community she draws her data from. There is also virtually nothing about the informants, the collection of data or even the immediate linguistic context of the nominal strings in question. This lack of information is problematic in several ways. From an exoskeletal perspective, in a linguistic community where Spanish is the main language spoken, it is expected that DPs like *el employer* will be more common that those like *the casa*. This is simply because if you are speaking Spanish, the structures will be Spanish as well, whereas you would need an English DP for the Subset Principle to license an English determiner like *the*.

Moreover, social factors such as prestige are likely to impact the notion of something being acceptable versus unacceptable, which is a crucial point of discussion when employing acceptability judgments in language mixing. The alleged unacceptability of phrases like *the casa* may thus actually stem from sociolinguistic norms in the language community. In order to give a thorough analysis of language mixing patterns, one must consider both which language constitutes the main one as well as other factors that may influence the judgements of mixed phrases. See Liceras et al. (2008) for further discussion of preferences in the linguistic production of speakers who mix English and Spanish.

7 An extension: Lexicalism does not fare better in the verbal domain

So far we have seen that mixing patterns in the nominal domain provide evidence against a minimalist lexicalist analysis and that an exoskeletal analysis is descriptively more adequate. In this section, we will briefly show that mixed verb forms in American Norwegian provide

additional evidence against a minimalist lexicalist analysis, and, subsequently, that the way the mixed verbs pattern is predicted by an exoskeletal analysis.

An example of mixed verb forms is illustrated in (25).

(25) vi bare satt der og **watch-**a da (sunburg_MN_03gm)
we just sat there and watch-PAST then
'We just sat there watching then.'

Here, the verb stem, *watch*, is clearly English, while the tense inflection is clearly Norwegian, *-a* being a past tense suffix belonging to the main class of Norwegian weak verbs. (26) provides additional data.

(26) a. **teach-**er (rushford_MN_01gm)
teach-PRES

b. **play-**de (coon_valley_WI_03gm)
play-PAST

The main pattern we find in CANS when lone English verbs are mixed into otherwise Norwegian utterances is that the mixed English verb stems appear with the appropriate Norwegian inflection, as shown in (25) and (26). This is in accordance with what we already saw for nouns in, e.g., (13b) and (14b, c), where the English nouns for the most part occur in exactly the position their Norwegian counterparts would in the noun phrases, with appropriate inflections.

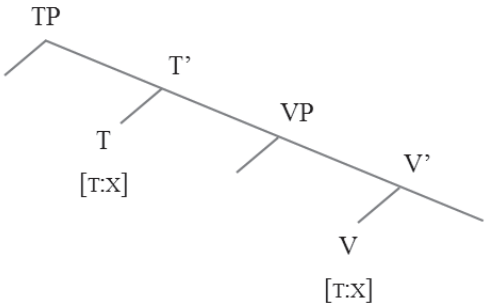
Since the verbal stem and the tense affix belong to different languages, it is not self-evident that standard analyses of the T–V relation can explain these data. Chomsky (1995) provides a classic formulation of the theoretical possibilities when analyzing the T–V relation, which again amounts to an exoskeletal vs. lexicalist analysis:

> The main verb typically "picks up" the features T and Agr […], adjoining to an inflectional element to form [V I]. There are two ways to interpret the process, for a lexical element *a*. One is to take *a* to be a bare, uninflected form; PF rules are then designed to interpret the abstract complex [*a* I] as a single inflected phonological word. The other approach is to take *a* to have inflectional features in the lexicon as an intrinsic property (in the spirit of lexicalist phonology); these features are then checked against the inflectional element I in the complex [a I]. (Chomsky 1995: 195)

Chomsky embraces the second solution, assuming that the inflected form of the verb already is created in the lexicon and subsequently inserted into the syntax with an inherent feature bundle, i.e., fully tensed. This amounts to a full-blooded lexicalist analysis of the T–V relation, and is also the one MacSwan makes use of.[15]

Just as we saw for the DPs, the mechanism of feature checking or valuation plays a crucial role within the lexicalist analysis of the T–V relation. In order to prohibit arbitrary insertion of tensed verbs, a given tensed form that is inserted into syntax must be checked against a corresponding feature in T to ensure that it occurs in a structurally correct position.[16] (27) shows a relevant structure ("x" denotes a particular tense feature value).

(27)



However, there are at least three problems with this lexicalist analysis of the T–V relation.

The first problem is how to account for the overwhelming occurrence of Norwegian tense suffixes on English verbal stems if the verb has "inflectional features in the lexicon as an intrinsic property". As argued in Åfarli (2015b: 168–169), this assumption would make us expect that an English verb stem should have English tense inflection, and it remains a mystery that the inflection instead is Norwegian.

[15] Note that MacSwan himself actually escapes the whole issue of word-internal mixing by claiming that they are not proper examples of codeswitching (or language mixing) at all. Instead, they are what he calls "nonce borrowings" (after Poplack, Sankoff, & Miller 1988), and they come about by being copied from one mental lexicon to the other, thus receiving the appropriate feature bundle. This solution is not falsifiable (unless neural imaging one day were to show that that is not how the brain does it), but for an extensive discussion of this and other positions, see Grimstad et al. (2014) and Grimstad (2017).

Furthermore, if one assumes with Distributed Morphology and other theories that syntax operates "all the way down", i.e., word-internally, then words are constructed in the syntax in the sense that the stem (or root) is inserted in one position, the inflectional morpheme in another, and syntactic or post-syntactic operations ensure that they combine or appear adjacent to one another. From this perspective, the American Norwegian data can be easily accommodated, as we will show below.

[16] We will not go into technicalities concerning checking vs. valuation here. For the purpose of this section, we just notice that somehow it must be ensured that the lexically given tense of the verb itself will be identical to the corresponding tense specified in T.

The second problem concerns the feature checking between T and V. An English verb will, given lexicalist assumptions, bear not only a tense feature, but also agreement features. The most standard assumptions is that those features of the verb must be checked against corresponding features in T. However, if an English verb is employed in a Norwegian structure, T will only contain a tense feature. Thus, the putative agreement features of the verb cannot be checked. This is parallel, *mutatis mutandis*, to the failure of feature checking in the nominal domain that we discussed earlier.

The third problem has to do with the richness of mixing varieties. As was the case for the DPs, the main pattern for mixed verbs seen in (25) and (26) is not the only type found in the corpus. We have found 292 lone English verbs occurring in otherwise Norwegian utterances in the corpus, and out of those, 210 had Norwegian inflectional suffixes whereas 82 did not, instead seemingly appearing with the appropriate English morphology.[17] Examples are given in (28), where # marks a pause.

(28) a. så e # I kunne ikke # e **sing** # (north_battleford_SK_02gk)
so e # I could not # e sing #
'so I couldn't sing'

b. og han **sends** # han er # (vancouver_WA_01gm)
and he sends # he is #
'and he sends'

c. e v- vi # **translated** "Synnøve Solbakken" (gary_MN_01gm)
e w- we # translated "Synnøve Solbakken"
'we translated (the novel) Synnøve Solbakken'

Since these appear to be English verbs with English inflection, (28b) even displaying the appropriate 3.PERS.SG. suffix *-s*, we can assume that the entire TP structure is English.[18] Importantly, since both the earlier patterns and the patterns in (28) are attested in the corpus, we would like an analysis that is flexible enough to handle both.

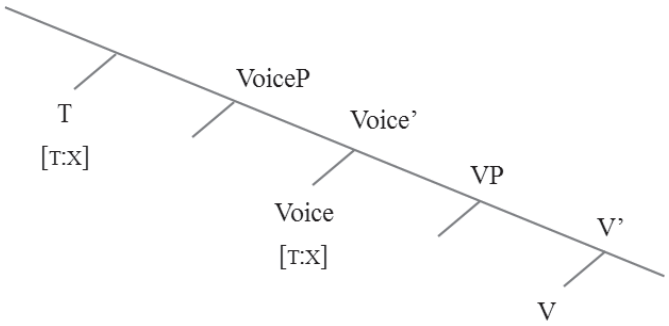
We will now sketch an exoskeletal analysis of mixed verb forms and show that it can handle the problems reviewed above and make the right predictions for American Norwegian. As in our exoskeletal analysis of nouns, we will assume an exoskeletal analysis of verbs

[17] Note that even though the written Norwegian standards, *Bokmål* and *Nynorsk*, both have tense suffixes for all weak verb forms, many of those are omitted in several Norwegian dialects. We have not checked this further (and as these are heritage speakers, placing them dialect-wise is often not possible); but we can assume that several of these 82 verb forms without Norwegian tense suffixes in fact are correct, dialectal forms of Norwegian, reducing this group even further. See Grimstad (2017) for further discussion.
[18] This hypothesis is further elaborated on and supported in Grimstad (2017).

where the inflectional properties of the verb are not an intrinsic property of the verb itself, contrary to the lexicalist analysis, but of a functional head whose complement contains the verb stem. Thus, the language of the structure and its inflectional elements may belong to a language different from the stem, as is the case for the examples in (25) and (26).

Let's start with a structure where the verb phrase of a simple clause is c-commanded by T and the verb phrase itself consists of a Voice-projection (Kratzer 1996) with a VP in its complement domain, as in (29).[19]

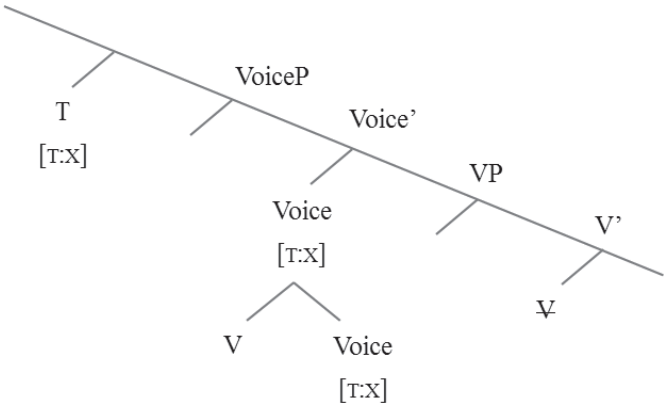
(29)



Focusing on the T–Voice–V relation for Norwegian verb phrases, we assume, as mentioned, that the tense feature is generated under T and that Voice contains a corresponding tense feature that is checked against the feature in T. V acquires tense by obligatory movement to Voice, giving us the structure sketched in (30).

[19] Note that the VoiceP between VP and TP in fact mirrors the FP between NP and DP, and that just like we saw for NP, VP is a simplification which most likely contains more structure. See Harley (1995), Alexiadou, Anagnostopoulou, et al. (2006, 2015), Folli & Harley (2007), Pylkkänen (2008), Ramchand (2008), and others.

(30)



Note that V, being a stem, can be inserted into the verb phrase from any language, like N can be inserted from any language into a given noun phrase structure.

The main exoskeletal point of this analysis is that the generation of the tense inflection is divorced from the generation of the verb stem, and that these two elements are syntactically integrated during the derivation. Thus, the first and second problems for a lexicalist analysis, as noted earlier, are solved. A Norwegian tense affix is expected on the English verb, since the Voice projection, containing the lower tense feature, is part of the Norwegian structure. Furthermore, feature checking between T and the verb, now situated in Voice, is unproblematic because both T and Voice belong to the same abstract syntactic frame and thus contain corresponding features, whereas the verb stem alone may belong to another language.

The exoskeletal analysis can also deal adequately with the other attested verb pattern, illustrated in (28), where the mixed English verb has English inflection although the rest of the clause may be Norwegian. Such a pattern may occur if the main exoskeletal structure in fact is English, with Norwegian phrases inserted in argument and adjunct positions, which is a possible mixing pattern given this analysis.

Summing up this brief section on mixing in the verbal domain, our hypothesis is that just like in the nominal domain, the exoskeletal analysis can account for all the attested mixing patterns in the corpus whereas the lexicalist one is problematic in several respects, as noted. This clearly favors the exoskeletal approach.[20]

[20] Additional support for the exoskeletal analysis comes from argument structure patterns in the verbal domain. These patterns are also unexpected given a lexicalist analysis, see Åfarli (2015a: 18–20) for discussion.

8 Conclusion

According to MacSwan (2014), generative theories prior to the Minimalist Program did not provide sufficient theoretical tools for implementing an adequate language mixing or codeswitching analysis; prior attempts were simply doomed to fail, seeing as the formal model they built on was wrong to begin with. MacSwan contrasts this approach with his own, the minimalist lexicalist one, which he finds more satisfying:

> Within the [Minimalist Program], structures are built from a stock of lexical items, with lexical insertion […] taking place at the outset. This important development permits [codeswitching] researchers to probe the structural consequences of particular lexical items from specific languages […]. (MacSwan 2014: 18)

By scrutinizing one paper explicitly adopting such a minimalist lexicalist approach, Moro (2014), we contest the claim that syntactic structure is dependent on the language of the lexical items involved. In fact, we show that Moro's analysis of Spanish–English mixing between a determiner and a noun fails to predict the observed patterns: Under standard minimalist feature valuation mechanisms the pattern Moro argues is well-formed, i.e. a Spanish D and English N, would crash the derivation, whereas the alleged ill-formed pattern, an English D and Spanish N, actually should converge.

Adopting an exoskeletal approach to grammar instead, a model that entails the independent generation of syntactic structures, we successfully analyze both patterns. This is further supported by data showing Norwegian–English mixing between a determiner and a noun in the heritage language American Norwegian, as well as mixing in the verbal domain in the same language. For instance, Sections 5–6 illustrate that an English noun inserted into American Norwegian has no power to influence the overall syntactic structure of the phrase, not even on its own inflection; and likewise, Section 7 shows that the same is true for English verbs inserted into otherwise Norwegian phrases in American Norwegian.

Thus, it remains to be seen what empirical shortcomings exoskeletal models might be subject to cross-linguistically, but as we have tried to argue in this article, they seem to remain our most viable options for analyzing language mixing data at present.

References

Ackerman, F., G. Stump, & G. Webelhuth. 2011. Lexicalism, periphrasis, and implicative morphology. In R. D. Borsley & K. Börjars (eds.), *Non-transformational syntax: Formal and explicit models of grammar*, pp. 325–358. Malden, MA: Blackwell-Wiley.

Adger, D. 2003. *Core Syntax*. Oxford: Oxford University Press.

Åfarli, T. A. 2007. Do Verbs have Argument Structure? In E. Reuland, T. Bhattacharya & G. Spathas (Eds.), *Argument Structure*, pp. 1–16. Amsterdam: John Benjamins.

Åfarli, T. A. 2015a. A Syntactic Model for the Analysis of Language Mixing Phenomena: American Norwegian and Beyond. In B. R. Page & M. T. Putnam (Eds.), *Moribund Germanic Heritage Languages in North America*, pp. 12–33. Leiden: Brill.

Åfarli, T. A. 2015b. Hybrid verb forms in American Norwegian and the analysis of the syntactic relation between the verb and its tense. In J. B. Johannessen & J. C. Salmons (Eds.), *Germanic Heritage Languages in North America: Acquisition, attrition and change,* pp. 161-177. Amsterdam: John Benjamins.

Alexiadou, A. 2001. *Functional structure in nominals: nominalisation and ergativity*. Amsterdam: John Benjamins.

Alexiadou, A. 2017. Building verbs in language mixing varieties. *Zeitschrift für Sprachwissenschaft*, 36, 165–192.

Alexiadou, Artemis, E. Anagnostopoulou, & F. Schäfer. 2006. The properties of anticausatives crosslinguistically. In M. Frascarelli (Ed.), *Phases of Interpretation*, pp. 187–212. Berlin: Mouton de Gruyter.

Alexiadou, A., E. Anagnostopoulou, & F. Schäfer. 2015a. *External arguments in transitivity alternations: a layering approach*. Oxford: Oxford University Press.

Alexiadou, A., T. Lohndal, T. A. Åfarli & M. B. Grimstad. 2015b. Language Mixing: A Distributed Morphology approach. In T. Bui & D. Özyildiz (Eds.), *Proceedings of NELS 45*, pp. 25–38. Create Space.

Bandi-Rao, S. & M. Den Dikken. 2014. Light Switches: On v as a pivot in codeswitching, and the nature of the ban on word-internal switches. In J. MacSwan (Ed.), *Grammatical Theory and Bilingual Codeswitching*, pp. 161–184. Cambridge, MA: MIT Press.

Borer, H. 1984. *Parametric Syntax*. Dordrecht: Foris.

Borer, H. 2005a. *Structuring Sense I: In Name Only*. Oxford: Oxford University Press.

Borer, H. 2005b. *Structuring Sense II: The Normal Course of Events*. Oxford: Oxford University Press.

Borer, H. 2013. *Structuring Sense III: Taking Form*. Oxford: Oxford University Press.

Cantone, K. F. & N. Müller. 2008. Un nase or una nase? What gender marking within switched DPs reveals about the architecture of the bilingual language faculty. *Lingua,* 118, 810–826.

Chan, B. H-S. 2008. Code switching, word order and the lexical/functional category distinction. *Lingua*, 118, 777–809.

Chomsky, N. 1995. *The Minimalist Program.* Cambridge, MA: MIT Press.

Chomsky, N. 2000. Minimalist inquiries: the framework. In R.Martin, D. Michaels & J. Uriagereka (Eds.), *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik,* pp. 89–155. Cambridge, MA: MIT Press.

Chomsky, N. 2001. Derivation by Phase. In M. Kenstowicz (Ed.), *Kenneth Hale: A Life in Language*, pp. 1–52. Cambridge, MA: MIT Press.

Embick, D. 2000. Features, syntax, and categories in the Latin perfect. *Linguistic Inquiry,* 31*,* 185–230.

Embick, D. & R. Noyer. 2007. Distributed Morphology and the syntax morphology interface. In G. Ramchand & C. Reiss (Eds.), *The Oxford Handbook of Linguistic Interfaces*, pp. 289–324. Oxford: Oxford University Press.

Folli, R., & H. Harley. 2007. Causation, Obligation, and Argument Structure: On the Nature of Little v. *Linguistic Inquiry,* 3*,* 197–238.

Goldrick, M., M. Putnam, & L. Schwarz. 2016. Coactivation in bilingual grammars: A Gradient Symbolic Computation account of code mixing. *Bilingualism: Language and Cognition*, 19, 857–876.

Gónzalez-Vilbazo, K. & L. López. 2011. Some properties of light verbs in code-switching. *Lingua*, 121, 832–850.

Gónzalez-Vilbazo, K. & L. López. 2012. Little *v* and parametric variation. *Natural Language and Linguistic Theory*, 30, 33–77.

Grimstad, M. B. 2017. The CS/borrowing debate: Evidence from English-origin verbs in American Norwegian. *Lingue e Linguaggio* XVI: 3–34.

Grimstad, M. B., T. Lohndal, & T. A. Åfarli. 2014. Language mixing and exoskeletal theory: A case study of word-internal mixing in American Norwegian. *Nordlyd*, 41*,* 213–237.

Halle, M. 1997. Distributed morphology: Impoverishment and fission. In B. Bruening, Y. Kang & M. McGinnis (Eds.), *MITWPL 30: Papers at the Interface*, pp. 425–449. Cambridge: MITWPL.

Harley, H. 1995. Subjects, Events, and Licensing. Doctoral dissertation, MIT.

Harley, H. & R. Noyer. 1999. Distributed Morphology. *Glot International,* 4, 3–9.

Harley, H. & R. Noyer. 2000. Licensing in the non-lexicalist lexicon. In B. Peeters (Ed.), *The Lexicon/Encyclopedia Interface*, pp. 349–374. Amsterdam: Elsevier Press.

Jake, J. L., C. Myers-Scotton, & S. Gross. 2002. Making a minimalist approach to codeswitching work: Adding the Matrix Language. *Bilingualism: Language and Cognition,* 5*,* 69–91.

Jake, J. L., C. Myers-Scotton, & S. Gross. 2005. A response to MacSwan (2005): keeping the matrix language. *Bilingualism: Language and Cognition,* 8*,* 271–276.

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, May 11–13, 2015, Vilnius, Lithuania.* NEALT Proceedings Series 23.

Johnson-Fowler, C. 2017. Cross-language interaction during language production in L1 German-L2 Italian bilinguals. Doctoral dissertation, Penn State University.

Julien, M. 2003. Double definiteness in Scandinavian. *Nordlyd,* 31*,* 230–244.

Julien, M. 2005. *Nominal Phrases from a Scandinavian Perspective.* Amsterdam: John Benjamins.

Koopman, H. 2006. Agreement configurations: In defense of "Spec head". In C. Boeckx (Ed.), *Agreement Systems*, pp. 159–199. Amsterdam: John Benjamins.

Kratzer, A. 1996. Severing the External Argument from the Verb. In J. Rooryck, & L. Zaring. (Eds.), *Phrase Structure and the Lexicon*, pp. 109–137. Dordrecht: Kluwer.

Kroll, J. F. & T. Gollan. 2014. Speech planning in two languages: What bilinguals tell us about language production. In M. Goldrick, V. Ferreira, & M. Miozzo (Eds.), The Oxford Handbook of Language Production, pp. 165–181. Oxford: Oxford University Press.

Liceras, J. M., R. F. Fuertes, S. Perales, R. Pérez-Tattam, & K. T. Spradlin. 2008. Gender and gender agreement in bilingual native and non-native grammars: A view from child and adult functional-lexical mixings. *Lingua,* 118*,* 827–851.

Liceras, J. M., K. T. Spradlin, & R. F. Fuertes. 2005. Bilingual early functional-lexical mixing and the activiation of formal features. *International Journal of Bilingualism*, 9*,* 227–252.

Lohndal, T. 2012. Without Specifiers: Phrase Structure and Argument Structure. Doctoral dissertation, University of Maryland.

Lohndal, T. 2014. *Phrase structure and argument structure: A case-study of the syntax semantics interface*. Oxford: Oxford University Press.

Lohndal, T. & M. Westergaard. 2016. Grammatical Gender in American Norwegian Heritage Language: Stability or Attrition? *Frontiers in Psychology*. doi: http://dx.doi.org/10.3389/fpsyg.2016.00344

MacSwan, J. 1999. *A minimalist approach to intrasentential code switching.* New York: Garland.

MacSwan, J. 2000. The architecture of the bilingual faculty: evidence from intrasentential code switching. *Bilingualism,* 3*,* 37–54.

MacSwan, J. 2005. Codeswitching and generative grammar: A critique of the MLF model and some remarks on "modified minimalism". *Bilingualism: Language and Cognition,* 8*,* 1–22.

MacSwan, J. 2009. Generative approaches to codeswitching. In B. E. Bullock & A. J. Toribio (Eds.), *Cambridge Handbook of Linguistic Codeswitching,* pp. 309–335. Cambridge: Cambridge University Press.

MacSwan, J. 2014. Programs and proposals in codeswitching research: Unconstraining theories of bilingual language mixing. In J. MacSwan (Ed.), *Grammatical theory and bilingual codeswitching*, pp. 1–33. Cambridge, MA: MIT Press.

Mahootian, S. 1993. A null theory of code switching. Doctoral dissertation, Northwestern University.

Marantz, A. 1997. No Escape from Syntax: Don't Try Morphological Analysis in the Privacy of Your Own Lexicon. In A. Dimitriadis, L. Siegel, C Surek-Clark, & A. Williams (Eds.), *Proceedings of the 21st Penn Linguistics Colloquium*, pp. 201–225. University of Pennsylvania, UPenn Working Papers in Linguistics.

Marantz, A. 2013. Verbal argument structure: Events and participants. *Lingua,* 130*,* 152–168.

Merchant, J. 2015. On ineffable predicates: Bilingual Greek-English code-switching under ellipsis. *Lingua*, 166, 199–213.

Moro, M. Q. 2014. The semantic interpretation and syntactuic distribution of determiner phrases in Spanish–English codeswitching. In J. MacSwan (Ed.), *Grammatical theory and bilingual codeswitching*, pp. 213–226. Cambridge, MA: MIT Press.

Muysken, P. 2000. *Bilingual Speech. A Typology of Code-Mixing*. Cambridge: Cambridge University Press.

Parafita Couto, M., A. Munarriz, I. Epelde, M. Deuchar, & B. Oyharçabal. 2015. Gender conflict resolution in Spanish–Basque mixed DPs. *Bilingualism: Language and Cognition*, 18, 304–323.

Pfaff, C. W. 1979. Constraints on Language Mixing: Intrasentential Code-Switching and Borrowing in Spanish/English. *Language*, 55, 291–318.

Pierantozzi, C. 2012. Agreement within early mixed DP: What mixed agreement can tell us about the bilingual language faculty. In C. Gabriel & K. Braunmüller (Eds.), *Multilingual Individuals and Multilingual Societies*, pp. 137–152. Amsterdam: John Benjamins.

Poplack, S. 1980. "Sometimes I'll start a sentence in Spanish Y TERMINO EN ESPAÑOL": Toward a typology of code-switching. *Linguistics* 18, 581–618.

Poplack, S., D. Sankoff, & C. Miller. 1988. The social correlates and linguistic processes of lexical borrowing and assimilation. *Linguistics,* 26*,* 47–104.

Pylkkänen, L. 2008. *Introducing Arguments*. Cambridge, MA: MIT Press.

Ramchand, G. 2008. *Verb Meaning and the Lexicon: A First Phase Syntax*. Cambridge: Cambridge University Press.

Riksem, B. R. In press. Language Mixing in American Norwegian Noun Phrases. *Journal of Language Contact 12* (1).

Riksem, B. R. 2017. Language Mixing and Diachronic Change: American Norwegian Noun Phrases Then and Now. *Languages* 2, doi: 10.3390/languages2020003

Riksem, B. R., M. B. Grimstad, T. Lohndal, & T. A. Åfarli. In press. Language mixing within verbs and nouns in American Norwegian. *Journal of Comparative Germanic Syntax*.

Sankoff, D. & S. Poplack. 1981. A Formal Grammar for Code-Switching. *International Journal of Human Communication*, 14, 3–46.

Shim, J. Y. 2013. Deriving Word Order in Code-Switching: Feature Inheritance and Light Verbs. Doctoral dissertation, The City University of New York.

Valdés Kroff, J. R. 2016. Mixed NPs in Spanish–English bilingual speech: Using a corpus-based approach to inform models of sentence processing. In R. E. Guzzardo Tamargo, C. Mazak & M. C. Parafita Couto (Eds.), *Spanish–English codeswitching in the Caribbean and the U.S.*, 281–300. Amsterdam: John Benjamins.

Valdés Kroff, J. R., P. E. Dussias, C. Gerfen, L. Perrotti, & T. Bajo. 2016. Experience with code-switching modulates the use of grammatical gender during sentence processing. *Linguistic Approaches to Bilingualism*, 7, 163–198.

van Hout, A. 1996. Event Semantics and Verb Frame Alternations. Doctoral dissertation [TILDIL Dissertation Series], Tilburg University.

Article 2

Language Mixing in American Norwegian Noun Phrases

Brita Ramsevik Riksem

Norwegian University of Science and Technology

Abstract:
This article investigates the morphosyntax of American Norwegian noun phrases that show mixing between Norwegian and English and proposes a formal analysis of these. The data show a distinct pattern characterized by English content items occurring together with Norwegian functional material such as determiners and suffixes. In the article, it will be argued that an exoskeletal approach to grammar is ideally suited to capture this empirical pattern. This framework crucially separates the realization of functional and non-functional terminals in an abstract, syntactic structure. Insertion of functional exponents is restricted by feature matching, whereas insertion into non-functional terminals is radically less restrictive. English exponents for noun stems are thus easily inserted into open positions in the structure, whereas functional exponents are typically drawn from Norwegian, as these are better matches to feature bundles comprising definiteness, number, and gender. In addition to the typical mixing pattern, the article also addresses an unexpected empirical phenomenon, the occurrence of the English plural *-s*, and proposes a possible analysis for this using the exoskeletal framework. The formal analysis of American Norwegian noun phrases also exemplifies how an exoskeletal approach complies with the ideal of a null theory of language mixing.

Keywords:
American Norwegian, exoskeletal analysis, language mixing, morphosyntax, noun phrase

1 Introduction[1]

This article has two main goals. The first goal is to provide a detailed analysis of American Norwegian noun phrases that show language mixing of Norwegian and English. The second goal is to demonstrate that an exoskeletal model is ideally suited to capture the empirical patterns.

Apart from the domain of second language acquisition, most previous work, and in fact much current work, on formal grammar has, at least implicitly, been focused on the ideal speaker/listener (as described in Chomsky, 1965). This strategy has certainly enabled scholars to achieve insight into the language faculty and has carved the theoretical foundation of generative syntax (Lohndal, 2013). Currently however, research is moving beyond this idealization and languages and linguistic phenomena often labelled as peripheral, such as

[1] I am grateful to two anonymous reviewers for helpful comments and suggestions. Moreover, I would like to thank Tor Anders Åfarli, Terje Lohndal, and Maren Berg Grimstad for valuable input and discussion.

language mixing, are given increasing attention. Accounting for such phenomena and patterns will without a doubt give a more realistic and nuanced picture of what the speaker actually encounters and will "help refine our perspective on general grammatical theory" (Muysken, 2000: 3).

This article provides a contribution by investigating the morphosyntax of noun phrases that show mixing between Norwegian and English in the heritage language American Norwegian (henceforth AmNo). This is a variety of Norwegian spoken by Norwegian settlers who came to the US a century or more ago, as well as their descendants, and even today some people still speak this variety. Thus, AmNo exists as a minority language in a society where the dominating language is English and the contact between Norwegian and English has resulted in outcomes often showing a mixture of the two. In this article, I show that mixed noun phrases follow a clear and predictable pattern, and I argue that an exoskeletal approach to grammar provides a good analytical tool for analyzing them.

The organization of the article is as follows. In Section 2, the empirical focus of the article, AmNo, is introduced, as well as the corpus that will be investigated. The theoretical background is presented in Section 3 before a formal model for the AmNo noun phrase is proposed in Section 4. Section 5 applies the theoretical framework and the model to the empirical data showing how an exoskeletal approach can provide a formal analysis of language mixing in the AmNo nominal domain. In addition to analyzing the typical mixing patterns, the more unexpected occurrence of the English plural suffix *-s* is discussed in Section 6. Section 7 concludes the article.

2 The heritage language American Norwegian

As already introduced, the empirical interest of this article is the heritage language AmNo. The term *heritage language* describes a language situated in the midst of a more dominant language community and can be defined in the following way: "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" (Rothman, 2009: 156). As in a monolingual situation, acquisition of a heritage language is based on naturalistic input, but the result may be qualitatively different due to degraded input conditions and influence from the majority language, and additionally the lack of formal education (see, e.g., Montrul, 2008; Rothman, 2009; Benmamoun, Montrul, & Polinsky, 2013). Speakers of such varieties are accordingly referred to as *heritage speakers*,

and they are considered native speakers of the heritage language as it is "acquired from naturalistic exposure, in early childhood and in an authentic social context/speech community" (Rothman & Treffers-Daller, 2014: 95). Nevertheless, their competence in adulthood often differs from speakers growing up in a society where this language is dominant. In the literature, this divergence has been analyzed as incomplete acquisition (see, e.g., Polinsky, 2006; Montrul, 2008) or attrition (see, e.g., Polinsky, 2011; Pascual y Cabo & Rothman, 2012).[2] For the purposes of the present article, the differences between these two concepts need not concern us.

As mentioned in the introduction, AmNo is the variety spoken by Norwegian immigrants who settled in the US in the period roughly from the mid-1800s until the 1920s, and also their descendants. During this period of time, a total of over 800,000 Norwegians immigrated to the US. Many of them settled in the Midwest area, where large Norwegian communities were established. The Norwegian language was actively used in these communities, e.g., in churches and in newspapers, and importantly also as the home language. In fact, due to these Norwegian settlements, few of the original immigrants became bilingual, but learned only as much English as was needed to get by (Haugen, 1953). However, the necessity of knowing English gradually expanded and English was established as the language spoken in commercial activities, in larger social groups, and in schools, whereas Norwegian was limited to domestic use and to the local community. With time, also this usage of Norwegian decreased in favor of English.[3] Today, AmNo is a moribund language only spoken by a few elderly individuals.

AmNo has been documented and studied in several publications (e.g., Haugen, 1953; Hjelde, 1992, 1996; Grimstad, Lohndal, & Åfarli, 2014; Eide & Hjelde, 2015; Westergaard & Anderssen, 2015; Åfarli, 2015a, b; Johannessen, 2015a, b; Larsson & Johannessen, 2015; Lohndal & Westergaard, 2016). As many of these publications demonstrate, the growing contact with English has left its mark on the variety, exemplified by, for instance, the occurrence of English items and signs of attrition. In this article, I will focus on the co-occurrence of English and Norwegian items in AmNo noun phrases, a phenomenon referred to as language mixing.

[2] Although their competence may resemble that of adult L2 learners, heritage speakers typically outperform L2 learners of the language (Pascual y Cabo & Rothman, 2012; Benmamoun et al., 2013)

[3] To illustrate, approximately 650,000 Americans declared Norwegian as the language of their childhood home in 1940, whereas 81,000 reported to use Norwegian at home in 1990 (Haugen, 1953; Hjelde, 2000; see also Johannessen & Salmons, 2015).

2.1 The Corpus of American Norwegian Speech

AmNo data have been collected in several rounds. Einar Haugen carried out extensive fieldwork in the 1930s and 1940s, presented in his seminal *The Norwegian Language in America* (1953), and more data was collected in the 1980s by Arnstein Hjelde (Hjelde, 1992). The most recently collected corpus is the *Corpus of American Norwegian Speech*, henceforth CANS (Johannessen, 2015b), and this is the dataset under investigation in this article. This corpus is created at the Text Laboratory at the University of Oslo, and it currently comprises recordings of 50 individual speakers.[4] The majority of these speakers are US-born (two are Canadian), and even though their first language (L1) is AmNo, their dominant language at the present time is English. Due to the bilingual nature of the speakers, the corpus is a rich source of language mixing between Norwegian and English. The speakers are also primarily elderly people, ranging from 70 to 100 years old, and even though the frequency of speaking AmNo varies extensively, from speaking it daily to not having practiced it for many years, many speak it more or less fluently (Johannessen & Salmons, 2012).

CANS is made available as a searchable database online. The material is transcribed at two levels: a broad phonological transcription and an orthographic transcription. The latter is standardized to the Norwegian written standard *Bokmål*.[5] In addition, the corpus also offers sound and video files, allowing the researcher to listen to the actual recording of each speaker. Unfortunately, the corpus does enable searching directly for single English items. Instead, one can apply the search option "x" which returns all items not found in the Norwegian (*Bokmål*) dictionary, among them a considerable amount of English items. Such a search will, however, also provide various dialect words which need to be separated out. When searching the corpus for the current study, the tag "x" was used and subsequently, relevant noun phrases were manually selected. In addition to Norwegian items and/or non-nouns, proper nouns, fixed expressions, and immediate repetitions were eliminated. To ensure that the remaining data are in fact cases of language mixing, both levels of transcriptions, as well as the context and sound files were considered, as these are all factors that may contribute to settling the issue.

This search leaves 1265 English nouns in the corpus.[6] Seventy-five of these occur without an immediate context, thus making it impossible to tell whether or not they are cases of language mixing. One hundred fifty-six nouns are found in an English context, either as

[4] The collections started in 2010 and the corpus is still growing.
[5] Written Norwegian has two standardizations, *Bokmål* and *Nynorsk*. *Bokmål* is used by the majority of Norwegians. See Venås (1993) and Vikør (1995) for more on the Norwegian language situation.
[6] Longer sequences of English are already excluded in the basic search opportunities in the corpus.

part of an English phrase or larger English piece of structure. Among these, all except one[7] have English inflection, suggesting that the speaker is not engaged in mixing, but has in fact switched to speaking English in these cases. The largest share of English nouns, 1034 in number, occur inside an otherwise Norwegian context, and these are thus the most interesting cases for the question of language mixing and the current article.

In the next step, these 1034 English nouns appearing in a Norwegian context have been sorted according to the functional items or affixes they appear together with. Thirteen of them show a combination of both English and Norwegian functional material. A total of 93 cases occur with only English inflection, more specifically the English plural *-s,* in the sample. These are discussed in Section 6. The majority of the nouns in the Norwegian context, 730 cases, are utterly embedded into a Norwegian structure, showing no English inflection. True enough, about half of them do not show any inflectional morphemes at all, but they are also found in a context where Norwegian nouns typically do not have inflectional morphemes. The other half, however, occur with Norwegian functional material such as determiners and functional suffixes. These constitute the typical mixing pattern of AmNo noun phrases and the main interest in the current article.[8]

In the following presentation of data, orthographic transcriptions are employed. However, in cases of discrepancy between the two layers of transcriptions, the phonological version and the corresponding recording are preferred. To get an idea of the type of data investigated in the current article, some examples of mixed AmNo noun phrases are presented in (1)–(3) below. The English items are highlighted in bold, and the information in parenthesis identifies the speaker in CANS. In the transcriptions, pauses are marked with "#" and hesitation with "e". Notice also that a detailed glossing is only provided for the items that are relevant for this article, namely the mixed noun phrases.[9]

(1) a. det var # var ei # **nurse** fra e ## E3[10] (coon_valley_WI_02gm)
it was was a.INDF.SG.F nurse from E3
'it was a nurse from E3'

[7] *The roaden* (rushford_MN_01gm) has both an English definite article and a Norwegian definite suffix.
[8] Notice that 198 of the English nouns lack functional material even though it would be expected in their Norwegian context. Such lack of functional material is also interesting, but beyond the scope of the present article. See, e.g., Riksem (2017) instead.
[9] The following annotations are used in the glosses throughout the article: DEF: definiteness, DF: definite, INDF: indefinite, NUM: number, PL: plural, SG: singular, GEN: gender M: masculine, F: feminine, N: neuter, COMP: comparative.
[10] "E3" represents the name of a place which has been anonymized.

b. nå må du ha en **permit** (westby_WI_06gm)
now must you have a.INDF.SG.M permit
og en **licence**
and a.INDF.SG.M licence
'now you need a permit and a licence'

c. og så er det et **township** (flom_MN_01gm)
and then is it a.INDF.SG.N township
'and it's a township'

(2) a. han skal # lære å leie denne kalven på (coon_valley_WI_06gm)
he shall learn to lead this calf on
fair-a
fair-DF.SG.F
'he shall learn to lead this calf at the fair'

b. de vil ikke lage noe med (blair_WI_07gm)
they will not make anything with
dette gaml-e **stuff-**et
this.DF.SG.N old-DF.SG.N stuff-DF.SG.N
'they won't make anything with this old stuff'

c. den **resort**-en som vi hadde (stillwater_MN_01gm)
the.DF.SG.M resort-DF.SG.M that we had
'the resort that we had'

d. for mange folk og # stor-e (glasgow_MT_01gm)
too many people and big-INDF.PL.M
truck-er
truck-INDF.PL.M
'too many people and big trucks'

(3) a. å celebrate **birthday**-en hennes (coon_valley_WI_06gm)
to celebrate birthday-DF.SG.M hers
'to celebrate her birthday'

b. vi e satt på **deck**-en hans (westby_WI_01gm)
we sat on deck-DF.SG.M his
'we sat on his deck'

The overall pattern in (1)–(3) is English nouns incorporated into otherwise Norwegian structures.[11] In (1), English nouns appear with a Norwegian indefinite article, and notice already here that the articles reveal that these nouns have been assigned to different gender

[11] This pattern is applicable also for other domains of the sentence. See, e.g., Grimstad et al. (2014) for discussion of the verbal domain.

categories, despite the fact that English nouns do not have gender. (2) shows cases where English nouns receive a Norwegian suffix in the same order as for Norwegian nouns. (2a) and (2b) have a definite, singular suffix, and also in these cases we see that the suffixes differ depending on the gender[12]. (2c) has an indefinite, plural suffix. Notice also that (2b) and (2c) show more complex DP structures involving a demonstrative and a determiner respectively. These are also Norwegian in structure, even though the noun is English. The data in (3) further show that the mixed nominal phrases have a Norwegian word order with a post-nominal possessive, unlike English.

The main objective of the present article is to propose a formal analysis of these and similar mixed noun phrases, which will involve mixing that occurs between words, such as between the determiner and the noun, and also word-internal mixing, such as the cases where an English noun stem occurs with a Norwegian functional suffix. Notice, however, that the question of possessives, as in (3), is not discussed in this article (see Westergaard & Anderssen, 2015 instead).

3 Theoretical background

This section introduces the theoretical foundation of the article, which is an exoskeletal approach to grammar. Before turning to that part, I will briefly discuss the phenomenon of language mixing and two previously suggested ways of approaching it. Both analyses have their advantages and disadvantages, and I will suggest an exoskeletal model as an alternative analysis, capturing the essential insights of both previous analyses.

3.1 Analyzing language mixing

Following Lohndal (2013: fn. 2), I take language mixing to "describe a situation where a speaker produces linguistic outcomes constituted by a mixture of elements from two or more languages". The term is related to a range of other terms describing similar phenomena, for instance the commonly used terms "code-switching" and "borrowing". In the literature, there is a discussion of what identifies and distinguishes these phenomena, concerning the length of the utterance, phonological integration, and frequency, as well as whether the token is borrowed only for the moment or has become a more or less established loan word (see, e.g., Grimstad et al. 2014, and references therein for discussion). A relevant question is,

[12] In the literature, there is a discussion concerning the functional suffix and whether it is a marker for gender or for declension class. I assume it to involve a gender feature.

nevertheless, whether it is necessary to adopt such distinctions. In fact, some suggest that code-switching and borrowing could be considered nuances of the same phenomenon, and that borrowed items are in fact established code-switches (e.g., Myers-Scotton, 1993, 2002; Stammers & Deuchar, 2012). For the purpose of this article, I do not formally separate between the two, and I will therefore employ the more general term "language mixing", basically describing phenomena where items which can be linked to different languages appear together in the same sentence, sometimes even in the same word.

In general, there are two main positions on how to analyze language mixing. The first claims that mixing is a process that requires additional theoretical primitives, whereas the second position argues that mixing should be accounted for by the same principles used to account for monolingual data (a null theory). Each camp is represented by prominent frameworks, which I will briefly introduce in the following paragraphs.

In her research on language mixing, Myers-Scotton (1993, 2002) finds a persistent asymmetry between the languages involved: One language enjoys the status as the main language or Matrix Language (ML), and is responsible for both word order and providing the inflectional or functional morphemes of the utterance, whereas the other language(s), the Embedded Language(s) (EL), may contribute content items. In other words, while the ML sets the frame for an utterance, including functional morphemes, the EL is primarily a source of content items. This has been formulated as the *Matrix Language Frame Model* (MLF model) which predicts that in the case of language mixing, the surface morpheme order, as well as all functional morphemes, will be determined by the ML. Only content items may occasionally be drawn from the EL.

The MLF model is empirically convincing. In the AmNo data in (1)–(3), for instance, Norwegian establishes its role as the ML, providing both word order and the relevant determiners and inflectional affixes. English, on the other hand, functions as a source for content items. The same pattern is found in mixing between other language pairs (see, e.g., Myers-Scotton, 1993, 2002; Kamwangamalu, 1997).

A closer look at the MLF model, however, reveals theoretical weaknesses. The main objection to the MLF model is the fact that it is a model designed specifically to deal with language mixing (MacSwan, 2000, 2005; Åfarli, Grimstad, & Subbarao, 2013). In other words, even though the framework is argued to be valid also for monolingual speech, it provides no independent evidence for this claim, reducing its potential of being a general model of grammar. Another complaint concerns the specific reference the model makes to the separate languages involved and the mixing situation itself. MacSwan (2014) argues that

specific languages cannot be included as primitives in the analysis as the grammar is formally blind to such distinctions. Instead, mixing is a generalization of the output. The crucial question in this discussion revolves around bilingual competence and whether or not this involves an additional component allowing the speaker to mix the languages that (s)he masters. Among many language mixing researchers, there is a general consensus that one should not develop specialized mechanisms and constraints for mixing (e.g., Mahootian, 1993; Belazi, Rubin, & Toribo, 1994; MacSwan, 1999, 2000, 2005, 2014). The basic argument is that we only have one language faculty and this is responsible for all production, monolingual as well as bilingual.

The second position on how to account for language mixing thus aims to develop a theory and a model that can account for both mixed and un-mixed language production by using the same principles. This is known as a null theory (Mahootian, 1993) or constraint free approach to language mixing. MacSwan (1999, 2000, 2005, 2014) is a strong advocate for such an approach and proposes an analysis of mixing based on a lexicalist approach within the Minimalist Program. This approach has in turn been criticized for not accommodating the observed asymmetry between the languages involved in language mixing in a convincing way (Jake, Myers-Scotton, & Gross, 2002, 2005). I will not go into the details of the framework proposed by MacSwan, as this is not employed in the current article (see Grimstad, Riksem, Lohndal, & Åfarli, in press, for a review). The quest for a null theory of language mixing, on the other hand, remains, and the task is to come up with a model able to unify the empirical asymmetry observed by Myers-Scotton with the ideal of a null theory. In this article, I argue that an exoskeletal approach allows precisely that: merging the essential insights from both the MLF model and the null theory account.

3.2 An exoskeletal approach to grammar

The model I propose and employ in this article falls within exoskeletal approaches to grammar, which is best described as a family of approaches, also known as generative, neo-constructivist approaches. Such approaches have been developed in different ways by several scholars, e.g., van Hout (1996), Marantz (1997, 2013), Borer (2005a, b, 2013), Åfarli (2007), Ramchand (2008), Lohndal (2012, 2014), and Alexiadou, Anagnostopoulou, and Schäfer (2015). Like any ordinary family, they have somewhat different assumptions and motivations, but all varieties share the same theoretical core, namely the assumption that syntactic structures are to some degree generated independently from the lexical items that realize

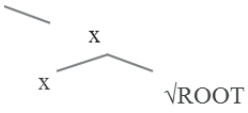
them. Consequently, the model proposed and employed in the current article will differ in its details from many of the above cited references, but the core assumption remains.

The particular model proposed here is what can be called a late-insertion exoskeletal model, inspired primarily by the work of Borer, Åfarli, Lohndal, and Marantz, as cited above. In such a model, the syntactic structure is assumed to form a skeleton (or template or frame; the name is not crucial) which determines the morphosyntactic information of the phrase. Lexical items, on the other hand, are inserted late into designated positions in the structure. Importantly, this approach and related models are motivated based on monolingual data, and assumptions are not designed specially to deal with language mixing. Moreover, the model also implements certain important insights from Distributed Morphology (DM) (see, e.g., Harley & Noyer, 1999; Alexiadou, 2001; Embick & Noyer, 2007). In DM, the content of the lexicon is distributed across three separate lists: one for syntactic terminals, one for vocabulary items, and one for encyclopedic information. These are accessed at different points throughout the derivation, limiting the information that is available at a given point. This mirrors the core assumptions of exoskeletal approaches: In the first stage of the derivation, an abstract, syntactic skeleton is generated, whereas lexical items are accessed at a later, second stage.[13]

The abstract syntactic skeleton that is generated in stage one is assumed to have two types of terminals: roots and functional features or feature bundles. In the literature, there is an ongoing discussion on the nature of roots (see, e.g., Harley, 2014 and other articles in the same special issue). Following Arad (2005), I assume that roots are atomic elements of the syntactic structure, devoid of all grammatical features as well as underspecified for semantics and phonology. This entails that roots have core semantics, but are not specified for word class, which instead will arise from roots being structurally combined with a category defining head, or categorizer (Marantz, 1997; Arad, 2005; Pylkkänen, 2008; Embick & Marantz, 2008). Consequently, one root may surface as different categories depending on the syntactic context. For instance, the verb *braid* and the noun *braid* are considered exponents of the same root √BRAID, but combined with different category defining heads. Structurally, this is displayed in (4) where the categorizer x and the root combine and form the stem x (see Alexiadou & Lohndal (in press) for a discussion of the structural configuration).

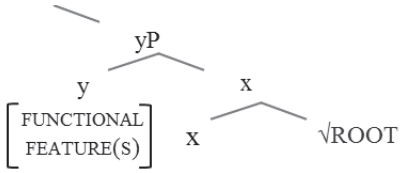
[13] The third stage, encountering encyclopedic information, will not be addressed in this article.

(4)



The second type of syntactic terminal comprises functional features or feature bundles, holding the relevant syntactic features. This is illustrated in (5), where (4) is expanded with a functional projection, yP.

(5)



The content of the functional feature bundle will vary according to the overall phrase it is included in, as well as according to language. For instance, the features of a noun phrase will differ from those of a verb phrase, and the features of noun phrases will vary across languages. A general assumption is that the language faculty makes available a full set of features from which a particular grammar can be made by activating a selection (Adger, 2003). A particular language is thus characterized by the combination of features that are activated and how they are combined in bundles. A Norwegian noun phrase, for instance, typically involves the features definiteness, number and gender, which helps us separate it from languages not involving the same feature composition. Basically, this means that structures and features themselves are not language specific, but a given selection may be. Referring to something as a "Norwegian structure" thus only means that the structure holds features and feature bundles typically associated with Norwegian.

The next stage in the derivation is Spell-Out, the process of inserting vocabulary items, or phonological exponents, into the syntactic terminals. Inserting exponents into terminals housing functional features or feature bundles is a process regulated by the *Subset Principle:*[14]

[14] Terminals holding functional features or feature bundles are known as *morphemes* in the DM literature.

> The phonological exponent of a Vocabulary item is inserted into a morpheme in the terminal string if the item matches all or a subset of the grammatical features specified in the terminal morpheme. Insertion does not take place if the Vocabulary item contains features not present in the morpheme. Where several Vocabulary items meet the conditions for insertion, the item matching the greatest number of features specified in the terminal morpheme must be chosen. (Halle, 1997: 428)

By this principle, insertion of functional material is competitive. In the vocabulary, the language user will have a wide repertoire of functional exponents, each paired with a set of conditions for insertion. In the insertion process, the functional exponent matching the greatest number of features specified in the structure must be chosen. To illustrate, let us consider the Norwegian indefinite articles, where the speaker has three possible exponents:[15]

(6)

[INDF, SG, M]	↔	**en**
[INDF, SG, F]	↔	**ei**
[INDF, SG, N]	↔	**et**

The composition of the feature bundle in the structure will determine the insertion of an exponent. The exponents in (6) share the features indefinite and singular, but when the structure is specified for masculine gender, *en* is inserted, whereas *ei* and *et* are inserted in feminine or neuter cases respectively (see also Harley & Noyer, 1999 for discussion and additional examples). Notice that being the *best* match does not necessarily mean having a *complete* match of the features in a bundle, and in addition that insertion is blocked if the exponent is specified for features not present in the structure.

In contrast to Spell-Out of functional terminals, Spell-Out of the stem, i.e., the root and the categorizer, is radically less restrictive. According to Arad (2005), and scholars following her, a root alone is unavailable for Spell-Out, meaning that a root only can be realized by an exponent in combination with the categorizer. Since the resulting stems do not involve complex functional feature bundles,[16] the possible realizations are more numerous; insertion is not limited by feature matching requirements. I thus assume that these positions, as well as adjuncts and specifiers, constitute what I will call "open slots" in the structure. Into these positions, content items from any language are easily inserted. The term "open" must,

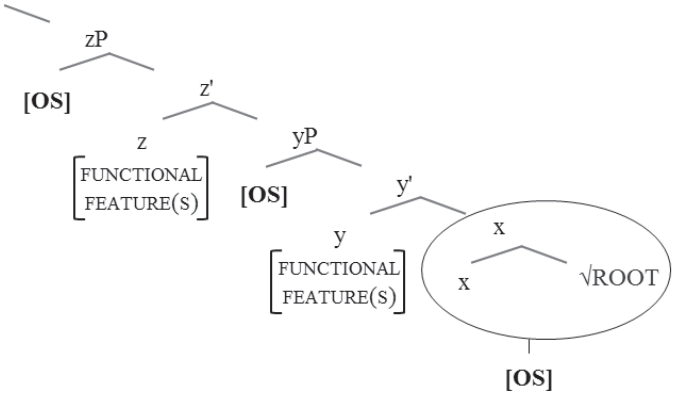
[15] These are the standardized exponents of *Bokmål*. *Nynorsk* and many dialects, in Norwegian as well as in AmNo, will provide different alternatives.

[16] Some will argue that a category is also a grammatical feature (e.g., Adger, 2003), but for the present article, this has no decisive consequences, as the major categories, among them nouns, are known in most languages.

however, be used with some reservations, since these positions may involve certain restrictions too (see, e.g., Åfarli & Subbarao, in press).

In practice, this means that the language user may insert elements from any available lexicon into the open slots, whereas Spell-Out of functional features is strictly regulated by the Subset Principle. The representation of a larger piece of structure is given in (7), where functional heads contain functional features or feature bundles, whereas specifiers, adjuncts, and the stem position constitute open slots [OS]. In this article, I focus on the (nominal) stem position and the associated functional projections.

(7)



In Section 5, I will discuss how the distinction between the two types of syntactic terminals and the process of inserting vocabulary items into them can account for the typical mixing patterns in AmNo noun phrases. Before turning to that discussion, I will suggest a syntactic structure for AmNo noun phrases.

4 The structure of American Norwegian noun phrases

In this section, I formulate an exoskeletal model for the nominal domain that can be used to analyze AmNo mixed noun phrases. This builds on previous research on Norwegian noun phrases, primarily Julien (2005), as well as what the AmNo data can tell us.

Before going into details of the model, notice that there are already a number of studies of language mixing within the nominal domain in the literature, including Jake, Myers-Scotton, and Gross (2002); Herring, Deuchar, Parafita Couto, and Moro (2010); Moro (2014); and Parafita Couto, Munarriz, Epelde, Deuchard, and Oyharçabal (2015). These studies discuss mixing between many different language pairs, although none involving Norwegian,

as well as different problems that may arise in mixing the different languages within a noun phrase. However, these papers adopt a rather different theoretical framework than the present article, and for reasons of space, it is not possible to compare frameworks here.

Returning to the case of AmNo, the word order, as well as most of the functional and lexical items, is recognized as Norwegian. This establishes Norwegian as the main language, providing the syntactic skeleton and the relevant functional features. Julien (2005) has conducted a thorough investigation of the Norwegian noun phrase, and concludes that its maximal expansion includes prenominal determiners, weak quantifiers and/or adjectives, possessive pronouns that can be either pre- or post-nominal, and finally post-nominal PPs. The noun itself is inflected for definiteness, number, and gender, realized as a functional suffix in all cases except indefinite singulars. The potential of Norwegian noun phrases is illustrated in a simplified version in (8), showing a case where the possessive surfaces post-nominally.

(8) [Determiner [weak quantifier [adjective [noun [possessive [PP]]]]]]

The model I propose and employ is anchored in Julien's work, with additional developments of the framework. I will not go into possessives or PPs in this article, and therefore I focus on somewhat less complex phrases such as (9).

(9)	den	gaml-e	maskin-a[17]	(fargo_ND_01gm)
	the.DF.SG.F	old-DF.SG.F	machine-DF.SG.F	
	'the old machine'			

The model can be said to consist of three layers: i) the root and the nominalizer, ii) the functional features, and iii) the higher structure. I will discuss these separately in the following paragraphs.

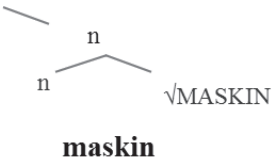
4.1 The root and the nominalizer

At the bottom of the structure, a root combines with a category defining head, as in (4) above, and more specifically in this case the root combines with a nominalizer, n. This structural combination forms the nominal stem, in (10), which is then spelled out by the exponent

[17] The standardized transcription of this particular utterance is "den gamle maskinen" with the masculine definite suffix *-en*. The phonological transcriptions, however, disclose the feminine definite suffix *-a*.

maskin. Note that in this and in subsequent examples I have included the exponent in the presentation as well as the abstract categories. The exponent is written in bold.

(10)



In the case of language mixing in AmNo noun phrases, there are two options concerning the units of mixing, i.e., the items drawn from a different language than of the structural frame. Either the stem (root + categorizer) is being mixed, or the root itself. Space prevents a full discussion of this important issue, but in what follows I briefly present two arguments in favor of stems being mixed.

The first argument is taken from the verbal domain. In general, mixing in the verbal domain follows the same pattern as mixing in the nominal domain. For instance in AmNo English items occur with Norwegian inflectional suffixes. However, studies by Türker (2000) and Åfarli and Jin (2014) show that when Norwegian verbal content items are mixed into Turkish and Mandarin Chinese structures respectively, it is not a bare form that is used, but the Norwegian infinitival forms, including the suffix *-e*. This suggests that also the categorizer is drawn from Norwegian in these cases, and that the mixed element is already categorized as a verb.

A second related argument builds on the conceptual meaning of the mixed pieces. Consider the noun *chair*, meaning the chairperson at an event. This noun is commonly mixed into Norwegian, and when it is, it also brings with it that particular, conceptual meaning. Recall that roots are considered to be without any grammatical features or phonological information, and that they also lack any semantic interpretation except some fundamental core (Arad, 2005). It then appears unlikely that the root *chair* still carries such a specific, conceptual content, and instead, I assume that this is something that arises in the combination with a categorizer (see also Grimstad et al. 2014).

4.2 The functional features

Norwegian nouns are inflected for definiteness, number, and gender. Definiteness and number constitute individual functional projections in Julien's (2005) analysis, whereas gender is analyzed as a property of the nominal stem, introduced by the nominalizer. However, mixing data also shed light on the question of gender. I will consider two broad alternatives for analyzing gender in order to motivate the next layer in the DP structure (Riksem, 2015).

The first alternative is to analyze gender as an inherent property of the root itself or of the nominal stem. Analyzing gender as a property of the root is theoretically dismissed in an exoskeletal framework, where the root is devoid of grammatical features. Additionally, considering that a given root can become a noun, a verb, or an adjective depending on the categorizing head it is combined with, it is unexpected that only nouns have gender, if this was in fact a property of the root. Another possibility is therefore to analyze gender as a feature of the categorizer n (see, e.g., Alexiadou, 2004, 2011; Kramer, 2014). This makes gender a property of the nominal stem, and the assignment is described as an essential part of turning a root into a noun. In her analysis, Kramer (2014) proposes that n has a gender feature with different values depending on whether it is natural gender or grammatical gender, and that licensing conditions will determine which root can be combined with which n. However, Kramer's analysis is based on a language that relies heavily on natural gender, Amharic, but as the Norwegian gender system is basically grammatical and arbitrarily assigned, it is not clear how easily the analysis would transfer to the present data.

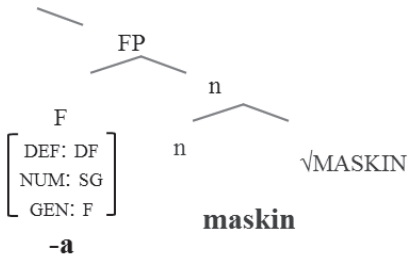
More importantly, the mixing data from AmNo provide empirical counter-arguments to such analyses. Despite their usual lack of gender, English noun stems mixed into Norwegian structures, are assigned a gender category, expressed on functional suffixes and associated words. The data further show a distribution of English stems across the three genders in Norwegian, which opposes a possible default assignment. How would this be possible if the assignment of gender were to depend on the English stem? Instead, it follows that gender in AmNo cannot be a feature of neither the root nor the categorizer, but must be found somewhere in the functional structure of the noun phrase.

The second alternative to analyzing gender is then precisely that gender must be located somewhere above the n complex. Possible analyses are, for instance, that gender is a head in its own functional projection (Picallo, 1991, 2008; Nygård & Åfarli, 2015) or a feature of another functional projection (Ritter, 1993).[18]

[18] A related question which then arises concerns how a noun is assigned a specific gender if not being a quality of the individual root or stem. This is discussed under 5.1.

Data presented here, however, does not clearly reveal the fine-grained structure and the exact locus of gender. Neither are such details relevant for the purpose of this article. Norwegian functional suffixes may be internally complex, but for the current analysis it is sufficient to note that the English stems occur with Norwegian functional suffixes and associated words signalizing the features definiteness, number, and gender as a complex. Hence, I propose a common functional projection, named FP (for "functional projection"), housing a bundle of all three features. The next layer of the structure will then be like the one in (11).

(11)

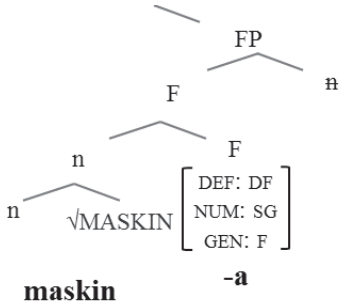


Each feature in the feature bundle in F has a set of possible values in Norwegian: Definiteness can be definite or indefinite, number can be singular or plural, and gender can be masculine, feminine, or neuter. The interplay between these three features sets the requirements for Spell-Out, and the most suitable exponent is inserted, cf. the Subset Principle. A necessary assumption in such analyses is that the functional exponents of these terminals express what features can be found in that specific head. In the case of (11), knowing that *-a* is the Norwegian exponent for definite, singular, and feminine, is the base for assuming the presence of these specific feature values in the structure. In Spell-Out of such a feature bundle, Norwegian exponents offer a more precise match than the English alternative, being specified for all three features in the bundle. This is why functional exponents are generally picked from the same language as the one providing the structural skeleton, regardless of the stem.

The exponent of F in Norwegian is furthermore suffixed to the noun stem. This suggests that the stem moves to F, a movement that is considered obligatory in Norwegian noun phrases (Julien, 2005). Without going into an elaborate discussion, I assume that this is triggered by another feature of F, which is also in line with Julien's framework, and I also

assume that this is a case of head movement (Travis, 1984; Roberts, 2010), where n moves to F and combines with the functional suffix. The next stage of the derivation is then displayed in (12).

(12)



As shown in (12), the internal structure of F is now more complex. In subsequent structures, however, I will simplify this presentation by placing both exponents, for the stem and for F, together under F.

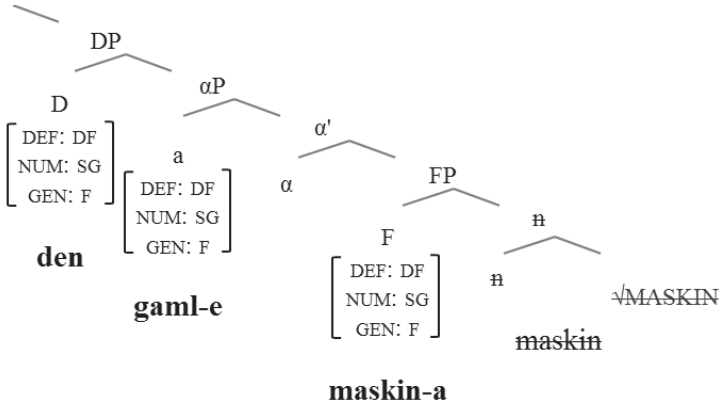
4.3 The higher structure

Following Julien (2005), I assume that the higher structure contains the projections αP, CardP and DP. αP and CardP are both optional projections present in cases with adjectives or weak quantifiers (such as *mange* 'many' or various numerals). More specifically, Julien argues that adjectives and weak quantifiers constitute separate phrases and that these are generated in the specifier position of αP and CardP respectively. On top of the structure is the DP layer.[19] The head of this projection, as well as possible heads in spec-αP and spec-CardP are generated with a bundle of unvalued features corresponding to those found in F, and these are valued through a probe–goal relation (Chomsky, 2000; Adger, 2003). In order to make the phrase referential, the Norwegian DP projection must also be made visible by overt material in either D or spec-DP. This requirement is met by moving the FP complex to spec-DP, or by inserting a separate determiner or demonstrative in D (Julien, 2005).

[19] In addition, Julien (2005) argues that there are additional projections for strong quantifiers and demonstratives. As such subdivisions are not crucial for this article I will simply analyze both determiners and demonstratives as exponents of D.

The complete structure of the noun phrase in (9) is presented in (13) below. In this structure, the features in D and spec-αP[20] have already been valued and the appropriate exponents are inserted.

(13)



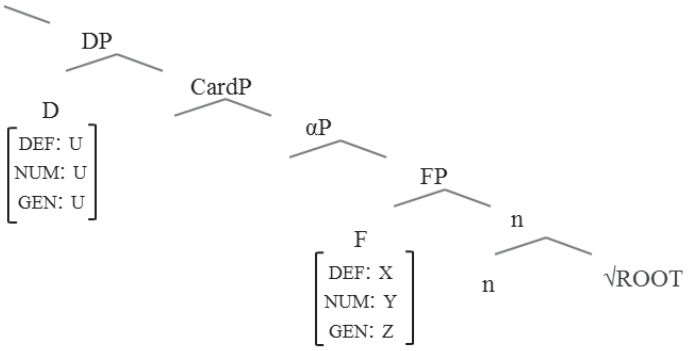
The structure in (13) also shows the phenomenon double definiteness, i.e., the co-occurrence of a definite determiner and a definite suffix. This term is typically reserved for phrases just like (13), involving a modifying adjective or a weak quantifier. These categories will be generated in the specifier position of αP or CardP, which will prevent FP from moving past it, and this requires that some other overt material is inserted into D or spec-D in order to make the phrase referential. The result is that definiteness is spelled out both in F and in D (Julien, 2005). As will be obvious as we go into the analyses, also noun phrases without such an adjective or weak quantifier may have a similar output, with the double marking of definiteness.

5 Analysis

In this section, I will employ the exoskeletal model suggested in the previous section to analyze a selection of mixed AmNo noun phrases in CANS. The structural starting point for all of these phrases is displayed in (14).

[20] Adjectives are discussed in some more detail in 5.4.

(14)



For each unique phrase, the features in F will be specified for the relevant values. The features in D are generated as unvalued and get their valuation through a probe–goal relation with F. When the phrase also contains an adjective or a weak quantifier, similar unvalued feature bundles will be generated in the specifier position of αP or CardP respectively. In the second stage, the relevant phonological exponents are inserted into the syntactic terminals. The Subset Principle will regulate the exponents of the functional feature bundles, which in the AmNo cases will be chosen from the Norwegian pool of functional exponents, due to the feature matching requirements. The stem, on the other hand, is generated without such features and thus constitute an open slot in the structure. As long as it has nominal specifications, content items from any language may easily be inserted into this position. In the following subsections, I discuss different typical mixing patterns in the AmNo material.

5.1 English stem with Norwegian indefinite article

The first type of data I will address is a frequent and typical mixing pattern in AmNo noun phrases, namely indefinite singulars, where an English noun stem occurs with a Norwegian article. A handful of examples is given in (15).

(15)	a.	en	**permit**	(westby_WI_06gm)
		a.INDF.SG.M	permit	
		'a permit'		
	b.	en	**licence**	(westby_WI_06gm)
		a.INDF.SG.M	licence	
		'a licence'		

c.	en		**chainsaw**	(blair_WI_07gm)
	a.INDF.SG.M		chainsaw	
	'a chainsaw'			
d.	en		**screen**	(westby_WI_06gm)
	a.INDF.SG.M		screen	
	'a screen'			
e.	en		**apartment**	(chicago_IL_01gk)
	a.INDF.SG.M		apartment	
	'an apartment'			
f.	ei	#	**nurse**	(coon_valley_WI_02gm)
	a.INDF.SG.F		nurse	
	'a nurse'			
g.	et		**crew**	(westby_WI_03gk)
	a.INDF.SG.N		crew	
	'a crew'			
h.	et		**township**	(flom_MN_01gm)
	a.INDF.SG.N		township	
	'a township'			
i.	et		**title**	(stillwater_MN_01gm)
	a.INDF.SG.N		title	
	'a title'			

A first thing to notice in (15) is that these mixed phrases occur with three different indefinite articles, *en* (M), *ei* (F), and *et* (N), which tells us that they are assigned to the three different gender categories in Norwegian. The most frequently-occurring gender of Norwegian is masculine (Lohndal & Westergaard, 2016). This is apparent also among the English stems in CANS investigated in the current article; most English stems are assigned masculine gender. However, a notable number of feminine and neuter indefinite articles are also found with English noun stems. This is a clear indication that gender assignment in the mixed phrases cannot be a mere default mechanism.

These data then support the analysis discussed in Section 4, that gender is part of the structural representation of the noun phrase. If gender were truly an inherent quality of the stem, the observed pattern would be difficult to explain, as English stems could not be expected to provide a gender feature, covering three different values, when this is an alien category to English nouns. One possible analysis of this suggests that when a speaker mixes languages as in (15), the speaker has established two entries for each noun in the lexicon: one without gender, which is the English version, and one which has been assigned a gender

feature, forming the Norwegian version. This, however, emerges as an uneconomical analysis, having two entries for an item with exactly the same meaning. An analysis where gender is structurally assigned is thus more convincing.

An immediate question is then how gender in these cases is established. In this question, I follow Nygård and Åfarli (2015), considering it a process of *feature construal*, where conceptual properties of the stem will materialize as grammatical properties in the structure. This is also discussed a bit further in Section 6. Other possible strategies for assigning gender may be that it is based on the phonology of the stem and ease of pronunciation, or on analogy with the corresponding stem from the other language. While the former is plausible, the latter does not coincide with the AmNo data. Some counterexamples are:

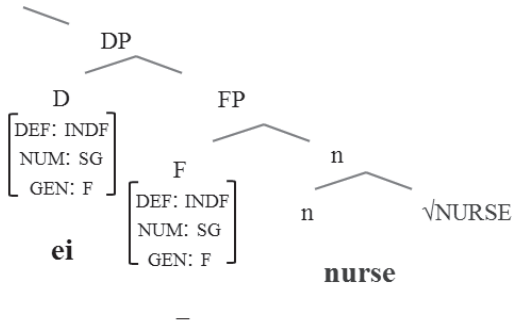
(16)

	Norwegian	**American Norwegian**	**English**
a.	kor-et.N	choir-en.M	the choir
b.	ferg-a.F	ferry-en.M	the ferry
c.	en.M tittel	et.N title	a title

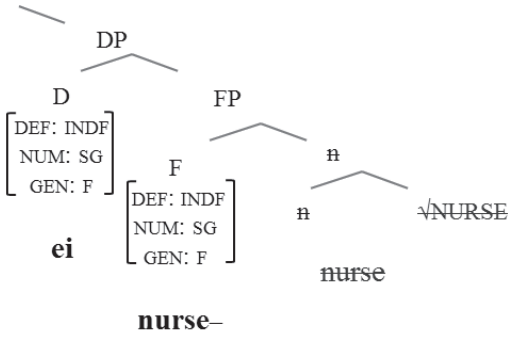
These discrepancies between the gender assigned mixed stems in AmNo and the gender of the corresponding Norwegian stem indicates that gender assignment in language mixing is not simply a translation from the one language to the other.

Let us now consider the phrase in (15f) *ei nurse.* The structural analysis of this phrase after the valuation of D is shown in (17) where (17a) shows the structure prior to movement, and (17b) shows the structure after the stem has moved to F. Again, the exponents are included and boldfaced.

(17) a.



b.



The English stem exponent *nurse* is here incorporated into a Norwegian structure and appears with a Norwegian indefinite article. The stem, in the complement position of F, constitutes an open slot in the structure, allowing the insertion of the English exponent. The higher levels of the phrase, on the other hand, contain features typically associated with Norwegian. FP is the first of the structural projections above n, and in this particular example the feature bundle holds the features indefinite, singular, and feminine, which strictly restricts the insertion of an exponent. The Norwegian pool of functional exponents will in (17) and similar cases provide a better match, and thus be inserted, since they correspond to all three features in the bundle, unlike the English alternatives.

Indefinite singulars in Norwegian do not have an overt exponent of F, and F is thus not overtly realized. This leaves the indefinite article as the main indicator of the specific feature values in the phrase, since the features in D have been valued by the features in F. Knowing that the Norwegian indefinite article *ei* is indefinite, singular, and feminine, is thus the base for assuming the presence of such features in F and subsequently in D. In (17b) the stem has completed the obligatory movement to F, something that will be more apparent in cases with an overt exponent in F. The indefinite article *ei* is inserted into D.[21]

The rest of the data in (15) will have a similar structure as (17) only varying according to gender. Neither will have an overt exponent in F, but the difference will manifest itself in the realization of the indefinite article in D.

[21] Julien (2005) argues that the indefinite article originates in a WQP in spec-CardP and then moves to the DP domain to provide it with overt material. For the purpose of this article, I simplify this analysis by inserting the indefinite article into D.

5.2 English stem with Norwegian suffix

(18) shows another typical mixing pattern in AmNo noun phrases: an English stem receiving a Norwegian functional suffix. The data in (18) are definite singulars, and by looking at the different suffixes assigned, *-en* (M), *-a* (F), and *-et* (N), we also see variation according to gender.

(18) a. **road**-en (webster_SD_02gm)
road-DF.SG.M
'the road'

b. **choir**-en (coon_valley_WI_07gk)
choir-DF.SG.M
'the choir'

c. **ferry**-en (harmony_MN_04gm)
ferry-DF.SG.M
'the ferry'

d. **fair**-a (coon_valley_WI_06gm)
fair-DF.SG.F
'the fair'

e. **bluff**-a[22] (westby_WI_01gm)
bluff-DF.SG.F
'the bluff'

f. **stuff**-et (blair_WI_07gm)
stuff-DF.SG.N
'the stuff'

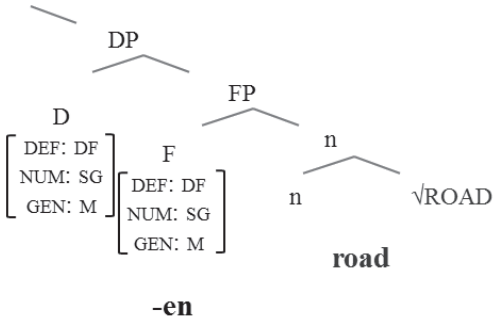
g. **cover-**et (coon_valley_WI_04gm)
cover-DF.SG.N
'the cover'

h. **shed-**et (westby_WI_06gm)
shed-DF.SG.N
'the shed'

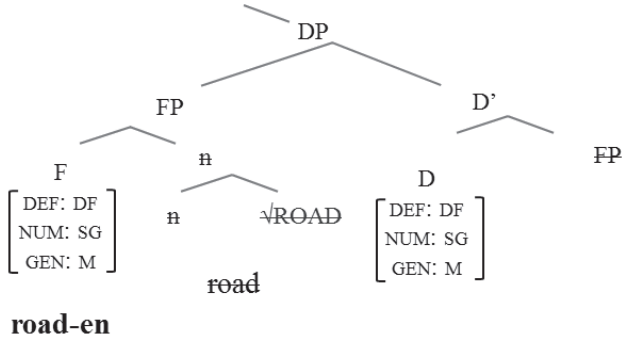
Taking a closer look at the example in (18a) *roaden*, this gives us a structure like in (19) after the valuation of the functional features in D. Again (19a) shows the structure prior to any movement, whereas (19b) shows the post-movement structure.

[22] This noun is also found once with masculine gender: *en bluff* (westby_WI_06gm).

(19) a.



b.



The English exponent *road* is first inserted into the stem position n in the structure as shown in (19a). The functional head F is specified for definite, singular, masculine, which is spelled out by the Norwegian suffix *-en*, due to this being the most suitable exponent given the Subset Principle. (19b) displays the movement within the phrase. First the stem moves to F to receive the suffix, yielding the complex form *roaden.* Thereafter, in order to supply the DP domain with overt material, FP moves to spec-DP (as discussed in Julien, 2005, although with different labels).

So far, only noun phrases in singular have been addressed. When it comes to plural phrases, both the indefinite and the definite ones will have the functional exponent realized as a suffix. Some examples are given in (20).

(20) a. **truck-**er (glasgow_MT_01gm)
truck-INDF.PL.M
'trucks'

b. **farmer**-a[23] (wanamingo_MN_04gk)
farmer-INDF.PL.M
'farmers'

c. **sportsgam[e]**-an[24] (westby_WI_02gm)
sportsgame-DF.PL.M
'the sports games'

d. tobakk**shed**-a (coon_valley_WI_06gm)
tobacco_shed-DF.PL.N
'the tobacco sheds'

These examples follow the same pattern as the previous examples of English stem exponents being incorporated into the open slots of a structure holding Norwegian functional feature bundles and consequently being assigned Norwegian functional material. Norwegian plural suffixes also vary according to gender. In (20), three out of four examples are masculine, though realized with different functional exponents requiring an additional comment. A feature bundle in F consisting of indefinite, plural, and masculine is typically realized by the suffix *-er*, as is the case in (20a). This is true for the written standard *Bokmål,* in which the utterance is transcribed, and for many dialects. In *Bokmål* and in these dialects, this is also the same exponent as for the feminine counterpart, making the exact determination of the gender in (20a) somewhat more challenging. Due to the default status of masculine gender, I assume that this example is masculine. Several other Norwegian dialects as well as the second written standard, *Nynorsk,* however, differentiate more distinctly between genders in the plural suffixes by using *-ar* or *-ane* for indefinite, masculine and definite, masculine respectively, in contrast to *-er* and *-ene* for the feminine cases. The examples in (20b) and (20c) can thus more reliably be analyzed as masculine, as the suffixes we see here are such variations of masculine suffixes.

The compound in (20d) *tobakksheda* comprises a Norwegian item, *tobakk* 'tobacco', and an English one, *shed.* This example is definite and neuter, which is typically realized by the suffix *-a* in Norwegian. Notice that indefinite neuter plurals are difficult to attest in the corpus. These do not have an overt suffix in Norwegian, making it hard to separate them from potential bare forms. Because of this, I will not discuss these in the current article.

The bottom line is nevertheless that also plural phrases follow the expected pattern. The only difference is the composition of the feature bundle in the structure, which will consequently require the insertion of a different exponent.

[23] This specific suffix appears in the phonological transcription.
[24] This specific suffix appears in the phonological transcription.

5.3 English stem with a Norwegian determiner or demonstrative

In cases where the FP complex does not move to the DP domain, as it does in (19b), the phrase needs some other overt material in D or spec-DP in order to be referential. This is typically done by inserting a determiner or a demonstrative in D, which is what the data in (21) show. The outcome is then that definiteness is expressed both by the determiner or the demonstrative and by the functional suffix, the phenomenon referred to as double definiteness.

(21) a. den **track**-en (westby_WI_02gm)
that.DF.SG.M track-DF.SG.M
'that track'

b. den **resort**-en (stillwater_MN_01gm)
that.DF.SG.M resort-DF.SG.M
'that resort'

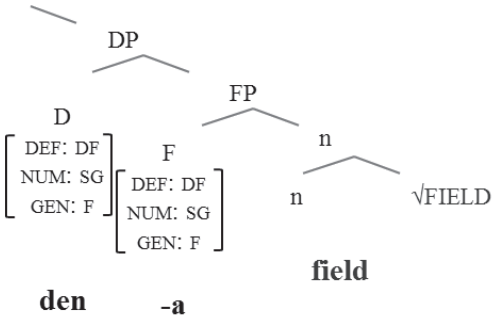
c. den e **field**-a[25] (coon_valley_WI_02gm)
that.DF.SG.F field-DF.SG.F
'that field'

d. det **pastur[e]**-et[26] (coon_valley_WI_03gm)
that.DF.SG.N pasture-DF.SG.N
'that pasture'

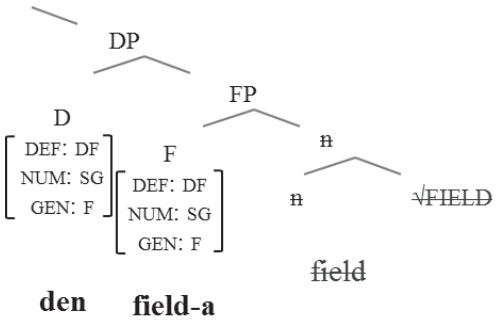
The examples in (21) are all singular, but with different genders. (21d) shows the most obvious gender attributes, visible on both the demonstrative *det* and the suffix *-et.* In (21a–c), on the other hand, all three examples have the exponent *den*, which is the common demonstrative for masculine and feminine. The gender distinction in these cases is rather detected on the suffix, parallel to the examples in (18) above. The structure of (21c) *den fielda* is displayed in (22).

[25] This noun is also found once with neuter gender: *field-et* (rushford_MN_01gm), and recall that "e" only is the transcription of hesitation.
[26] The context of this phrase suggests another possible analysis, namely that *det* is a formal subject, and that *pasturet* is an elaboration of this, but the context does not provide sufficient information in order to settle this question. However, constructions like the one suggested in (21d) are very frequent in Norwegian, making it a highly plausible analysis.

(22) a.



b.



(22a) shows the structure after the features in D have been valued, but before any movement has taken place. In the same way as in earlier examples, the English stem exponent *field* is inserted into the open slot of the structure. The two feature bundles in the structure are realized by exponents matching the most features, which, due to the features given, is a Norwegian functional exponent specified for all of the three relevant features. (22b) shows that the stem moves to F where the functional exponent is suffixed to it, but instead of moving further, as in (19b), the stem remains under F and D is spelled out by a separate demonstrative.

5.4 Noun phrases with an attributive adjective

More complex noun phrases, like those involving an adjective, also follow the expected pattern from the proposed exoskeletal model. Again, an English stem exponent is incorporated into a Norwegian structure, but this does not affect the remaining parts of the structure. Now, however, there will be an additional position, spec-αP, for the adjectival stem.

These positions may be internally complex, but at least, I assume that also adjectival stems are formed by the structural combination of a root and a categorizer.

In a Norwegian structure, the adjectival position is generated with a bundle of unvalued features corresponding to those in F. These are valued throughout the derivation, and realized by corresponding functional exponents. Before I move on to the data, a short introduction to the Norwegian adjectival inflection is in order. This category is split into a strong and a weak inflection. The weak inflection only has one form *-e* and is realized whenever the phrase is definite. The strong inflection, on the other hand, is realized in indefinite phrases, and is sensitive to gender and number. This gives us the following repertoire of exponents:

(23) Strong adjectival inflection

[INDF, SG, M/F]	↔	**-**
[INDF, SG, N]	↔	**-t**
[INDF, PL]	↔	**-e**

Weak adjectival inflection

[DEF]	↔	**-e**

As the data in (24) show, this pattern is maintained even in the mixed phrases.

(24) a. en rik **farmer** (flom_MN_02gm)
a.INDF.SG.M rich.INDF.SG.M farmer
'a rich farmer'

b. en stor **hook** (harmony_MN_02gk)
a.INDF.SG.M big.INDF.SG.M hook
'a big hook'

c. ei pussig **his- s- # story** (flom_MN_01gm)
a.INDF.SG.F odd.INDF.SG.F history
'an odd history'

d. et gammel-t **brewery** (flom_MN_01gm)
a.INDF.SG.N old-INDF.SG.N brewery
'an old brewery'

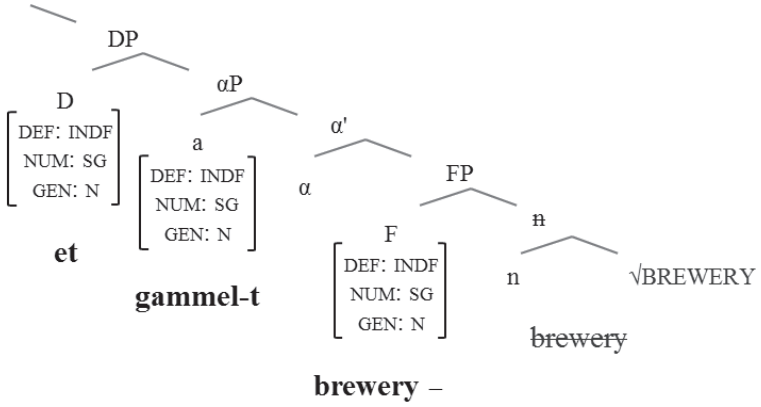
e. denne digr-e **chopper**-en[27] (blair_WI_01gm)
this.DF.SG.M huge-DF.SG.M chopper-DF.SG.M
'this huge chopper'

[27] The complete phrase is *denne digre traktoren og chopperen* 'this huge tractor and chopper', but since the conjunction does not matter for the analysis, the former part is not included.

f. den best-e aure**creek**-en[28] (coon_valley_WI_06gm)
the.DF.SG.M best-DF.SG.M trout_creek-DF.SG.M
'the best trout creek'

The structure of (24d) *et gammelt brewery* is presented in (25), and following Julien (2005), the adjective is generated in the specifier position of the functional phrase αP. Notice that a fully detailed structure for this position is not presented here, as this is not the main objective of this article. The important fact is that the adjective, or rather an accompanying functional projection, is generated with a bundle of unvalued features which are valued by the corresponding features in FP. Hence, the composition of features in F has consequences also for the inflection of the adjective and ensures agreement across the noun phrase.

(25)

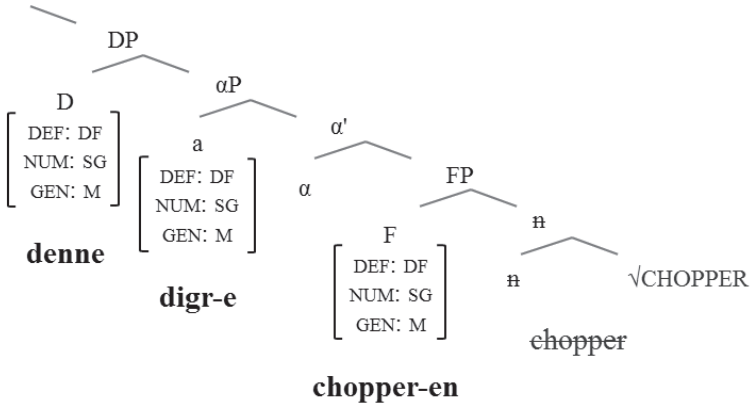


In the case of (25), the adjective has the suffix *-t* (strong inflection) which is the exponent for indefinite, singular, neuter in this position. Apart from the presence of the adjective, the analysis in (25) is parallel to the one in (17) above. The English exponent *brewery* is first inserted into the open slot and then moved to F. F does not have an overt exponent in indefinite, singular phrases, and D is spelled out by the appropriate indefinite article.

The structure in (26) shows another example involving an adjective, namely (24e) *denne digre chopperen*. In this case, the phrase is definite, which triggers the weak adjectival inflection in Norwegian, realized by the exponent *-e*.

[28] This is another compound consisting of a Norwegian item, *aure* 'trout', and an English one, *creek*, meaning that this is a place to go trout fishing.

(26)



In the lower part, (26) is parallel to (19). Again we see an English stem exponent being inserted into the open slot in a structure, and Norwegian exponents are inserted into the functional feature bundles as they are the best matches for the features in question. The stem is moved to F and is connected to the suffix. This is the only possible movement in (26), since the adjective in spec-αP prevents FP from moving to spec-DP. D is then spelled out by a separate demonstrative, *denne.*

Mixing in the more complex structures, as those involving adjectives, adds extra evidence to the assumption that the noun stem, or its exponent, does not affect the feature composition or realization of the higher functional structure. To emphasize the pattern of English stem exponents being incorporated into otherwise Norwegian structures in AmNo, notice that English items may appear as adjectives in AmNo, as the data in (27) show. Also in these cases the inflectional morphology is provided by Norwegian.

(27) a. det er ikke noen **small**-e farm-er (westby_WI_06gm)
it is not any small-INDF.PL.M farm-INDF.PL.M
noe mer
any more
'there isn't any small farms anymore'

b. det ble **easy**-ere om ei stund (webster_SD_02gm)
it became easy-COMP in a while
'it got easier after a while'

5.5 Interim summary and conclusion

In this section, the exoskeletal model developed in Section 4 has been tested against a variety of AmNo noun phrases. The model incorporates some core assumptions in an exoskeletal approach, namely i) that abstract, syntactic structures are generated independently from the items that realize them, ii) that this structure has two types of terminals, functional features or feature bundles and open slots consisting of a root together with a categorizer, and finally iii) that insertion into the functional terminals is regulated by the Subset Principle, whereas insertion into the open slots is less restricted, allowing a wide range of realizations. Importantly, the model does not employ any mechanisms especially for language mixing, which comply with the ideal of a null theory.

The asymmetry in the observed data is predicted by the model: Content items from any available language are easily inserted into open slots in the structure. Functional items, on the other hand, are typically drawn from the language of the structure, as these are tailor-made to spell out specific functional features or feature bundles. These predictions are borne out in the typical mixing pattern of AmNo noun phrases. However, in the next section I will consider exceptional AmNo data that require further discussion.

6 The plural *-s*

There is one particular English functional exponent that is found in the AmNo material, and that is the English plural *-s*. Some examples are given in (28).

(28) a. det er for mange # **lawyers** (sunburg_MN_03gm)
it is too many # lawyers
'there are too many lawyers'

b. og så andre andre **tools** (sunburg_MN_03gm)
and then other other tools
'and other tools'

c. jeg må nå bake sikkert en fem seks **pies** (coon_valley_WI_07gk)
I must now bake surely a five six pies
'I surely have to bake five to six pies'

d. de var store **cookies** vet du (wanamingo_MN_04gk)
they were big cookies you know
'they were big cookies, you know'

e. jeg tegner # **cartoons** (sunburg_MN_03gm)
I draw # cartoons
'I draw cartoons'

f. Norge har fem **dialects** (portland_ND_01gm)
Norway has five dialects
'Norway has five dialects'

As can be seen from these data, the expressions where the *-s* appear are otherwise Norwegian, which gives reason to assume that the structures and feature bundles that are generated are those typically associated with a Norwegian structure. On the basis of the discussion in the previous sections, the pattern in (28) is unexpected; Norwegian exponents should be a better match for the features in the structure since they could match all features, whereas the English plural *-s* only matches with plurality. In this section, I will discuss this problem, and finally propose an analysis for such phrases using the exoskeletal framework.

The occurrence of the English plural *-s* mixed into a non-English structure is not limited to Norwegian–English mixing, nor is it a new phenomenon. Myers-Scotton (1993) finds this pattern in her data with mixing between English and Bantu languages, and she describes it as a production error; the plural morpheme must have become part of the same lexical entry as its head. Moreover, her MLF model does not exclude appearance of functional morphemes from the EL as long as the ML version of it also is present, which is supported by data showing a double plural marking in such cases. In the AmNo context, the occurrence of the plural *-s* is previously observed by Haugen (1953)[29] and by Hjelde (1992). As discussed in 2.1, ninety-three cases where the English plural *-s* is used as the sole functional exponent in plural phrases within otherwise Norwegian contexts are found in CANS. This strongly suggests that we no longer can regard it a production error.

A closer look at the phrases in (28) reveals that they have two things in common: they are indefinite and they involve English stems.[30] This provides a first step towards an analysis.

[29] Haugen also observes a difference among his informants in their usage of the plural *-s:* What he calls *pre-bilingual borrowers* are speakers who have learnt English in adulthood and thus are not recognized as truly bilingual. In this group, the *-s* is taken to be part of the nominal stem, and the Norwegian suffix is added as well. However, in the second group, the *childhood bilinguals,* the plural *-s* is occasionally used as the sole exponent of plurality. As the informants in CANS are all 2nd or later generations of AmNo speakers, who are born in the US or Canada and have learnt English in school, I consider them part of the second group.

[30] Actually, there are a few examples which are definite and a few that involve Norwegian stems in CANS. I have found 5 cases of the former, some examples being *disse pill-s-ane* 'these pills' (westby_WI_03gk) which has both an English and a Norwegian functional suffix, and *disse homesteader-s* 'these homesteaders' (stillwater_MN_01gm), which has only the English suffix. Of the latter, Norwegian stems with an English plural suffix, I have found 4 cases. A couple of examples are *spiseplass-es* 'dining areas' (coon_valley_WI_01gk), and

A question of interest in this context concerns the nature of the definiteness feature: what does it really mean to be indefinite? One hypothesis, assumed by Julien (2005), is that there is no indefinite feature, just the absence of the definite one. She supports this among other things by referring to the Norwegian adjectival inflection:

> The fact that adjectives in indefinite DPs inflect like predicative adjectives could be taken to mean that the realisation of the adjectival agreement never makes reference to indefiniteness features. Alternatively, […] it might be that indefiniteness simply means absence of any definiteness feature, so that adjectival phrases contained in indefinite DPs have the same feature makeup as predicative adjectival phrases, which never have a definiteness feature since they are not inserted in a nominal environment. (Julien, 2005: 46)

If this is the case, it would mean that the feature bundles of F in indefinite phrases have one feature less for the exponent to match, which leaves us with gender as the main difference between the English and the Norwegian functional projection, and thereby also the main factor in the analysis of the English plural *-s* in the AmNo material.

I propose an analysis where the plural *-s* is considered a gender neutral alternative. In the following, I will put forward arguments to support this. Starting with the broad picture, gender differences are not particularity prominent in Norwegian plural phrases. In fact, neither adjectives nor weak quantifiers accompanying such indefinite phrases differentiate between genders in their plural form. This is true also for the AmNo data in (28) where *mange* 'many', *andre* 'other', *fem seks* 'five six', and *store* 'big' would all be the same regardless of the gender feature in F. This is typical for many gender systems, as described by Dahl (2000: 582): "Singular and plural gender systems often differ in the number of distinctions made: in structuralist terms, some distinctions may be 'neutralized' in the plural". Such a neutralization of gender is thus true for the prenominal adjectives and weak quantifiers in Norwegian noun phrases.

Still, gender is normally of crucial importance for the realization of the functional suffix in plural Norwegian noun phrases, which is exactly where the *-s* is inserted. In indefinite, plural contexts, Norwegian offers three possible exponents, against English's single exponent, displayed in (29).

---

innvandrer-s 'immigrants' (flom_MN_01gm). These data are in fact very interesting, but beyond the scope of the present article.

(29) Norwegian exponents[31]

[PL, M]	↔	**-ar**
[PL, F]	↔	**-er**
[PL, N]	↔	**-**

English exponents

[PL]	↔	**-s**

One possible analysis of the phrases in (28) could then be that a neutralization of gender, also in the case of the functional suffix, is in progress in AmNo. Without the gender feature present, Norwegian and English exponents would have the same criteria for insertion, making the English plural *-s* just as suitable as a Norwegian alternative. This, however, would imply that the *-s* is an equally adequate alternative also in cases with a Norwegian stem, which is not what the typical pattern in the data show. The vast majority of nouns stems receiving the plural *-s* in AmNo are realized by an English exponent, which should not be disregarded from the analysis.

A second possible analysis thus takes us into a discussion of the establishment of gender in the functional feature bundle. Although gender is not considered a property of the noun stem, Picallo (2008) and Nygård and Åfarli (2015) propose that the conceptual "class" or "entity category" associated with the stem nevertheless plays a facilitating role in the establishment of the value of the gender feature. Following Nygård and Åfarli (2015), I assume that the Norwegian gender category is ultimately anchored in biological gender. But once gender is established as a grammatical category in a language, any noun must be adapted into the system. Considering the nouns not denoting a biological gender, the establishment of grammatical gender turns out to be more or less arbitrary and something that must be learned. Nevertheless, once the connection is established between a stem and a specific grammatical gender, it tends to stick. In the case of Norwegian, I assume that this connection between a stem and the gender it is typically associated with is well established and dependable, and that this link is maintained through input and production.

Heritage speakers of AmNo, however, experience fundamentally different conditions in terms of input and production, which has consequences for the establishment of gender. In his studies of the *trønder* dialect, Hjelde (1992) finds that gender on Norwegian nouns in AmNo is relatively stable. Newer studies of gender on Norwegian nouns in AmNo are not entirely consistent in their conclusions. Johannessen and Larsson (2015) investigate 34 speakers in

[31] These exponents are typical for the written standard *Nynorsk*, as well as for several dialects, and they show the most obvious gender difference. The main argument will remain even if one used other dialects or *Bokmål*.

CANS and find an overall stability in gender. However, they also report a tendency to overgeneralize to the masculine. Lohndal and Westergaard (2016), on the other hand, conducts a cross-cutting investigating of all 50 speakers in CANS and find a more crucial change, or even erosion, in the gender system. The contrast between the two studies is striking, and might be, as pointed out by Lohndal and Westergaard (2016) themselves, at least partially explained by the definition of gender employed, where Johannessen and Larsson (2015) includes the definite suffix as a gender marker, which Lohndal and Westergaard (2016) do not.

Concerning the English nouns, these are essentially not associated with a specific gender, or even the gender category. Moreover, the input to rely on is scarce, making the gender assignment a process that happens quite spontaneously in the mixing cases, which may result in ambiguity, especially in plural cases where no hints are provided by accompanying adjectives or weak quantifiers either. This opens the door to the English functional exponent as a proper substitute. A first reason for this is precisely its appearance as a gender neutral alternative where the speaker does not have to make a decision. And considering the bilingual nature of these speakers, it is not surprising that the English exponent is easily accessible. Additionally, inserting a Norwegian exponent, which *is* specified for a particular gender feature, could in fact mean violating the Subset Principle if that gender feature does not have a counterpart in the structure. Instead, the English plural *-s* turn out to be the most appropriate exponent under the Subset Principle, matching a subset of the relevant features.

In a way then, one can say that gender is being neutralized also in the case of the functional suffix, and that this is a consequence of the problems related to establishing a gender value based on an English noun stem. A further speculation may be that as the gender system changes, the *-s* would become increasingly attractive as a gender neutral alternative also for the Norwegian nouns, which we could see the emergence of in the examples mentioned in footnote 30. For the time being, however, this analysis captures the two present characteristics of phrases with the plural *-s*: The lack of the indefiniteness feature explains why the *-s* appears mostly in indefinite phrases, and the uncertainty connected to the gender assignment when involving an English stem, accounts for the fact that this almost exclusively occurs on precisely English nouns.

7 Summary and conclusion

This article started out with two main goals, namely to provide an analysis of AmNo noun phrases that show mixing between Norwegian and English, and to demonstrate that the empirical insights can be captured in an explanatory way by an exoskeletal approach to grammar.

The contact between the heritage language AmNo and the majority language English in the US has left its mark on AmNo as, among other things, many English items being mixed into it. In this article, I have approached noun phrases showing such a mix and provided a formal analysis for these. The typical pattern is characterized by English content items occurring together with Norwegian functional material such as determiners and suffixes. This pattern is accounted for by an exoskeletal model which separates principally between the realization of functional and non-functional terminals in an abstract syntactic structure. English stem exponents are freely inserted into open slots in the structure, whereas functional exponents are inserted by a principle of feature matching. In the case of AmNo, this accounts for the fact that functional exponents are drawn from Norwegian; they simply provide the best match to the relevant feature bundles. A discussion of a more unexpected pattern, the occurrence of the English functional exponent *-s* shows how also this phenomenon may be analyzed in an exoskeletal model.

The analyses of the empirical material in question in this article corroborate the exoskeletal approach to grammar. This framework is primarily motivated by a series of studies of monolingual data (cf. the references in 4.1). However, any model needs to be constantly tested against new data, and the ability to reject falsification tells us how strong that model is. This article shows how the proposed model is able to explain different patterns even when encountering more peripheral data, such as language mixing. This establishes the model as a null theory of language mixing. At the same time, the notion of the structural frame and the realization of functional exponents regulated by the Subset Principle, offers an explanation to the asymmetry that is observed between languages involved in language mixing. As a result, the exoskeletal model is able to combine the insights from different theoretical frameworks in analyzing language mixing.

References

Åfarli, Tor A. 2007. Do Verbs have Argument Structure? In Eric Reuland, Tanmoy Bhattacharya, & Giorgos Spathas (Eds.), *Argument Structure (Linguistics Today 108)*, 1–16. Amsterdam: John Benjamins.

Åfarli, Tor A. 2015a. A Syntactic Model for the Analysis of Language Mixing Phenomena: American Norwegian and Beyond. In B. Richard Page & Michael T. Putnam (Eds.), *Moribund Germanic Heritage Languages in North America (Empirical Approaches to Linguistic Theory 8),* 12–33. Leiden: Brill.

Åfarli, Tor A. 2015b. Hybrid verb forms in American Norwegian and the analysis of the syntactic relation between the verb and its tense. In Janne B. Johannessen & Joseph C. Salmons (Eds.), *Germanic Heritage Languages in North America (Studies in Language Variation 18),* 161–177. Amsterdam: John Benjamins. doi: 10.1075/silv.18

Åfarli, Tor A., Maren B. Grimstad, & Karumuri V. Subbarao. 2013. Dakkhini and the Problem of a Matrix Language Frame in Sustained Syntactic Contact. Paper presented at International Conference on Language Contact in India, February 6–8, 2013, Deccan College Postgraduate and Research Institute, Pune.

Åfarli, Tor A. & Fufen Jin. 2014. Syntactic Frames and Single-word Code-switching: A Case Study of Norwegian–Mandarin Chinese. In Tor A. Åfarli & Brit Mæhlum (Eds.), *The Sociolinguistics of Grammar (Studies in Language Comparison 154)*, 153–170. Amsterdam: John Benjamins.

Åfarli, Tor A. & Karumuri V. Subbarao. In press. Models of grammar and the outcomes of long-term language contact: Language mixing in Dakkhini. In Zarina Estrada-Fernández, Albert Alvarez-Gonzalez, & Claudine Chamoreau (Eds.), *Diverse scenarios of syntactic complexity: inter and intra typological diversity.* Amsterdam: John Benjamins.

Adger, David. 2003. *Core Syntax: A Minimalist Approach*. Oxford: Oxford University Press.

Alexiadou, Artemis. 2001. *Functional Structure in Nominals: Nominalization and ergativity (Linguistics Today 42).* Amsterdam: John Benjamins.

Alexiadou, Artemis. 2004. Inflection class, gender and DP-internal structure. In Gereon Müller, Lutz Gunkel, & Gisela Zifonun (Eds.), *Explorations in Nominal Inflection (Interface Explorations 10)*, 21–50. Berlin: Mouton de Gruyter.

Alexiadou, Artemis. 2011. Remarks on the morpho-syntax of code-switching. Proceedings of the 9th international conference on Greek linguistics, October 29–31, 2009, University of Chicago, 44–55.

Alexiadou, Artemis, Elena Anagnostopoulou, & Florian Schäfer. 2015. *External arguments in transitivity alternations: a layering approach*. Oxford: Oxford University Press.

Alexiadou, Artemis & Terje Lohndal. In press. The structural configurations of root categorization. In Leah Bauke, Andreas Blümel, & Erich Groat (Eds.), *Roots and Label (Studies in Generative Grammar 128)*, 203–231. Berlin: Mouton de Gruyter.

Arad, Maya. 2005. *Roots and Patterns: Hebrew Morpho-syntax (Studies in Natural Language and Linguistic Theory 63).* Dordrecht: Springer.

Belazi, Hedi M., Edward J. Rubin, & Almeida J. Toribio. 1994. Code Switching and X-Bar Theory. *Linguistic Inquiry* 25: 221–237.

Benmamoun, Elabbas, Silvina Montrul, & Maria Polinsky. 2013. Heritage languages and their speakers: Opportunities and challenges for linguistics. *Theoretical Linguistics* 39 (3–4): 129–181.

Borer, Hagit. 2005a. *Structuring Sense I: In Name Only*. Oxford: Oxford University Press.

Borer, Hagit. 2005b. *Structuring Sense II: The Normal Course of Events*. Oxford: Oxford University Press.

Borer, Hagit. 2013. *Structuring Sense III: Taking Form*. Oxford: Oxford University Press.

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: The MIT Press.

Chomsky, Noam. 2000. Minimalist inquiries: the framework. In Roger Martin, David Michaels, & Juan Uriagereka (Eds.), *Step by Step: Essays on minimalist syntax in honor of Howard Lasnik,* 89–155. Cambridge, MA: MIT Press.

Dahl, Östen. 2000. Elementary gender distinctions. In Barbara Unterbeck, Matti Rissanen, Terttu Nevalainen, & Mirja Saari (Eds.), *Gender in grammar and cognition II: Manifestations of gender*, 577–593. Berlin: Mouton de Gruyter.

Eide, Kristin M. & Arnstein Hjelde. 2015. Borrowing modal elements into American Norwegian. In Janne B. Johannessen & Joseph C. Salmons (Eds.), *Germanic Heritage Languages in North America (Studies in Language Variation 18)*, 256–280. Amsterdam: John Benjamins. doi: 10.1075/silv.18

Embick, David & Alec Marantz. 2008. Architecture and Blocking. *Linguistic Inquiry* 39 (1): 1–53.

Embick, David & Rolf Noyer. 2007. Distributed Morphology and the syntax morphology interface. In Gillian Ramchand & Charles Reiss (Eds.), *The Oxford Handbook of Linguistic Interfaces*, 289–324. Oxford: Oxford University Press.

Grimstad, Maren B., Terje Lohndal, & Tor A. Åfarli. 2014. Language mixing and exoskeletal theory: A case study of word-internal mixing in American Norwegian. *Nordlyd* 41 (2): 213–237.

Grimstad, Maren B., Brita R. Riksem, Terje Lohndal, & Tor A. Åfarli. In press. Lexicalist vs. exoskeletal approaches to language mixing. *The Linguistic Review.*

Halle, Morris. 1997. Distributed morphology: Impoverishment and fission. *MIT Working Papers in Linguistics 30: PF: Papers at the Interface*, 425–449.

Harley, Heidi. 2014. On the Identity of Roots. *Theoretical Linguistics* 40: 225–276

Harley, Heidi & Rolf Noyer. 1999. Distributed Morphology. *Glot International* 4: 3–9.

Haugen, Einar. 1953. *The Norwegian Language in America*: *A Study in Bilingual Behavior* (Vol. I & II). Philadelphia: University of Philadelphia Press.

Herring, Jon R., Margaret Deuchar, M. Carmen Parafita Couto, & Mónica Moro Quintanilla. 2010. 'I saw the *madre*': evaluating predictions about codeswitched determiner–noun sequences using Spanish–English and Welsh–English data. *International Journal of Bilingual Education and Bilingualism* 13 (5): 553–573.

Hjelde, Arnstein. 1992. *Trøndsk talemål i Amerika*. Trondheim: Tapir.

Hjelde, Arnstein. 1996. The gender of English nouns in American Norwegian. In P. Sture Ureland & Iain Clarkson (Eds.), *Language Contact across the Atlantic (Linguistische Arbeiten 359),* 297–312. Tübingen: Max Niemeyer Verlag.

Hjelde, Arnstein. 2000. Norsk i Amerika. Downloadable at http://nabo.nb.no/trip?_b=EMITEKSTandurn=%22URN:NBN:no-nb_emidata_1187%22 (accessed July 13, 2015)

van Hout, Angeliek. 1996. Event Semantics of Verb Frame Alternations. PhD dissertation, Tilburg University.

Jake, Janice L., Carol Myers-Scotton, & Steven Gross. 2002. Making a minimalist approach to codeswitching work: Adding the Matrix Language. *Bilingualism: Language and Cognition* 5 (1): 69–91.

Jake, Janice L., Carol Myers-Scotton, & Steven Gross. 2005. A response to MacSwan (2005): Keeping the Matrix Language. *Bilingualism: Language and Cognition* 8 (3): 271–276.

Johannessen, Janne B. 2015a. Attrition in an American Norwegian heritage speaker. In Janne B. Johannessen & Joseph C. Salmons (Eds.), *Germanic Heritage Languages in North America, (Studies in Language Variation 18),* 46–71. Amsterdam: John Benjamins. doi: 10.1075/silv.18

Johannessen, Janne B. 2015b. The Corpus of American Norwegian Speech (CANS). In Béata Megyesi (ed.): Proceedings of the 20th Nordic Conference of Computational Linguistics, NODALIDA 2015, May 11–13, 2015, Vilnius, Lithuania. NEALT Proceedings Series 23.

Johannessen, Janne B. & Ida Larsson. 2015. Complexity Matters: On Gender Agreement in Heritage Scandinavian. *Frontiers in Psychology* 6:1842. doi: 10.3389/fpsyg.2015.01842

Johannessen, Janne B. & Joseph Salmons. 2012. Innledning. *Norsk Lingvistisk Tidsskrift* 30 (2): 139–148.

Johannessen, Janne B. & Joseph Salmons. 2015. The study of Germanic heritage languages in the Americas. In Janne B. Johannessen & Joseph C. Salmons (Eds.), G*ermanic Heritage Languages in North America, (Studies in Language Variation 18),* 1–17. Amsterdam: John Benjamins. doi: 10.1075/silv.18

Julien, Marit. 2005. *Nominal Phrases from a Scandinavian Perspective (Linguistics Today 87).* Amsterdam: John Benjamins.

Kamwangamalu, Nkonko M. 1997. Language contact, code-switching, and I-languages: Evidence from Africa. *South African Journal of Linguistics* 15: 45–51

Kramer, Ruth. 2014. Gender in Amharic: a morphosyntactic approach to natural and grammatical gender. *Language Sciences* 43: 102–115.

Larsson, Ida & Janne B. Johannessen. 2015. Incomplete Acquisition and Verb Placement in Heritage Scandinavian. In Richard S. Page & Michael T. Putnam (Eds.), *Moribund Germanic Heritage Languages in North America (Empirical Approaches to Linguistic Theory 8)*, 153–189. Leiden: Brill.

Lohndal, Terje. 2012. Without Specifiers: Phrase Structure and Argument Structure. PhD dissertation, University of Maryland.

Lohndal, Terje. 2013. Generative grammar and language mixing. *Theoretical Linguistics* 39: 215–224.

Lohndal, Terje. 2014. *Phrase structure and argument structure: A case-study of the syntax semantics interface (Oxford Studies in Theoretical Linguistics 49)*. Oxford: Oxford University Press.

Lohndal, Terje & Marit Westergaard. 2016. Grammatical Gender in American Norwegian Heritage Language: Stability or attrition? *Frontiers in Psychology* 7: 344. doi: 10.3389/fpsyg.2016.00344

MacSwan, Jeff. 1999. *A minimalist approach to intra-sentential code-switching*. New York: Garland.

MacSwan, Jeff. 2000. The architecture of the bilingual faculty: evidence from intrasentential code switching. *Bilingualism* 3: 37–54.

MacSwan, Jeff. 2005. Codeswitching and generative grammar: A critique of the MLF model and some remarks on "modified minimalism". *Bilingualism: Language and Cognition* 8: 1–22.

MacSwan, Jeff. 2014. Programs and Proposals in Codeswitching Research: Unconstraining Theories of Bilingual Language Mixing. In Jeff MacSwan (ed.), *Grammatical Theory and Bilingual Codeswitching,* 1–33. Cambridge, MA: MIT Press

Mahootian, Shahrzad. 1993. A null theory of code-switching. PhD dissertation, Northwestern University.

Marantz, Alec. 1997. No Escape from Syntax: Don't Try Morphological Analysis in the Privacy of Your Own Lexicon. *University of Pennsylvania Working Papers in Linguistics* 4(1): 201–225.

Marantz, Alec. 2013. Verbal argument structure: Events and participants. *Lingua* 130: 152–168.

Montrul, Silvina. 2008. *Incomplete Acquisition in Bilingualism. Re-examining the Age Factor (Studies in Bilingualism 39)*. Amsterdam: John Benjamins.

Moro Quintanilla, Mónica. 2014. The Semantic Interpretation and Syntactic Distribution of Determiner Phrases in Spanish-English Codeswitching. In Jeff MacSwan (ed.), *Grammatical Theory and Bilingual Codeswitching,* 213–226.Cambridge, MA: MIT Press.

Muysken, Pieter. 2000. *Bilingual Speech: A Typology of Code-Mixing.* Cambridge: Cambridge University Press

Myers-Scotton, Carol. 1993. *Dueling Languages: Grammatical Structure in Code Switching*. Oxford: Oxford University Press.

Myers-Scotton, Carol. 2002. *Contact Linguistics: Bilingual Encounters and Grammatical Outcomes*. Oxford: Oxford University Press.

Nygård, Mari & Tor A. Åfarli. 2015. On the structure of gender assignment. *Indian Linguistics* 76 (1–2): 67–76.

Parafita Couto, M. Carmen, Amaia Munarriz, Irantzu Epelde, Margaret Deuchar, & Beñat Oyharçabal. 2015. Gender conflict resolution in Basque–Spanish mixed DPs. *Bilingualism: Language and Cognition* 18(2): 304–323.

Pascual y Cabo, Diego & Jason Rothman. 2012. The (Il)logical Problem of Heritage Speaker Bilingualism and Incomplete Acquisition. *Applied Linguistics* 33 (4): 1–7.

Picallo. M. Carme. 2008. Gender and Number in Romance. *Lingue e Linguaggio* VII: 47–66.

Polinsky, Maria. 2006. Incomplete acquisition: American Russian. *Journal of Slavic Linguistics* 14: 191–262

Polinsky, Maria. 2011. Reanalysis in adult heritage language: A case for attrition. *Studies in Second Language Acquisition* 33: 305–328

Pylkkänen, Liina. 2008. *Introducing Arguments (Linguistic Inquiry Monographs 49)*. Cambridge, MA: MIT Press.

Ramchand, Gillian. 2008. *Verb Meaning and the Lexicon: A First Phase Syntax (Cambridge Studies in Linguistics 116)*. Cambridge: Cambridge University Press.

Riksem, Brita R. 2015. Gender (in American Norwegian). Paper presented at Distributed Morphology and Language Mixing (DAAD project Stuttgart–Trondheim), March 23, 2015, University of Stuttgart.

Riksem, Brita R. 2017. Language Mixing and Diachronic Change: American Norwegian Noun Phrases Then and Now. *Languages* 2 (2). doi:10.3390/languages2020003

Ritter, Elizabeth. 1993. Where's Gender? *Linguistic Inquiry* 24: 795–803.

Rothman, Jason. 2009. Understanding the nature and outcomes of early bilingualism: Romance languages as heritage languages. *International Journal of Bilingualism* 13 (2): 155–163.

Rothman, Jason & Jeanine Treffers-Daller. 2014. A prolegomenon to the construct of the native speaker: heritage speaker bilinguals are natives too! *Applied Linguistics* 35 (1): 93–98.

Roberts, Ian. 2010. *Agreement and Head Movement: Clitics, Incorporation, and Defective Goals (Linguistic Inquiry Monographs 59).* Cambridge, MA: MIT Press.

Stammers, Jonathan R. & Margaret Deuchar. 2012. Testing the nonce borrowing hypothesis: Counter-evidence from English-origin verbs in Welsh. *Bilingualism: Language and Cognition* 15 (3): 630–643.

Travis, Lisa. 1984. Parameters and Effects of Word Order Variation. PhD Dissertation, MIT.

Türker, Emel. 2000. Turkish–Norwegian codeswitching: evidence from intermediate and second generation Turkish immigrants in Norway. PhD dissertation, University of Oslo.

Venås, Kjell. 1993. On the choice between two written standards in Norway. In Ernst H. Jahr (ed.), *Language Conflict and Language Planning*, 263–278. Berlin: Mouton de Gruyter.

Vikør, Lars S. 1995. *The Nordic Languages: Their Status and Interrelations* (Nordic Language Secretariat Publication 14). Oslo: Novus Press.

Westergaard, Marit & Merete Anderssen. 2015. Word order variations in Norwegian possessive constructions. In Janne B. Johannessen & Joseph C. Salmons (Eds.), G*ermanic Heritage Languages in North America (Studies in Language Variation 18),* 21–45. Amsterdam: John Benjamins. doi: 10.1075/silv.18

Article 3

languages

MDPI

Article

Language Mixing and Diachronic Change: American Norwegian Noun Phrases Then and Now

Brita Ramsevik Riksem

Department of Language and Literature, Faculty of Humanities, Norwegian University of Science and Technology, 7491 Trondheim, Norway; brita.riksem@ntnu.no; Tel.: +47-7359-6987

Academic Editors: Ji Young Shim, Tabea Ihsane and Maria del Carmen Parafita Couto
Received: 20 December 2016; Accepted: 3 April 2017; Published: 20 April 2017

Abstract: This article investigates the diachronic development of language mixing within noun phrases in the heritage language American Norwegian. By comparing data collected in the 1930s and 1940s with recently collected data, I present and discuss patterns showing systematic changes, specifically concerning the categories number and definiteness. Moreover, I propose two potential analyses of these patterns based on an exoskeletal approach to grammar. This theoretical framework crucially separates the abstract syntactic structure from its phonological exponents, and the analyses that are discussed consider both the structure and the exponents as the origins of the change.

Keywords: American Norwegian; diachronic change; exoskeletal approach to grammar; language mixing; noun phrase

1. Introduction

Language mixing, in the form of utterances consisting of both English and Norwegian items, is a typical attribute of the heritage language American Norwegian (AmNo). This variety of Norwegian finds its origin in the language of the many immigrants who settled in North America in the century prior to 1920, and it is still spoken today by some of their descendants. Previous studies have shown that mixing of English and Norwegian typically involves English content items occurring together with Norwegian functional material [1–5]. This article pursues the question of whether or not these mixing patterns are persistent over time, and it presents data showing that systematic, diachronic changes can be found. Furthermore, it explores changes in the underlying grammar that can potentially explain the observed patterns.

The focus of this article is AmNo noun phrases showing a mix of English and Norwegian items. Comparisons of newly collected data with data from the 1930s and 1940s show overall stability in the main patterns of mixing. Still, some systematic changes are found. Examples of language mixing within AmNo noun phrases are given in (1), where the data in (1a–c) show the typical mixing pattern, i.e., English content items with Norwegian functional material, and (1d–e) are examples that I will argue are the results of diachronic change: omission of functional morphology and the use of English functional items. Notice that the accompanying references show which corpus the utterance is drawn from, either Einar Haugen's collections from the 1930s and 1940s [6] or the recently established Corpus of American Norwegian Speech (CANS) [7], as well as the associated page number, in the case of Haugen [6], or informant code, in the case of CANS [7]. The two corpora will be introduced and discussed in more detail in Section 3. Moreover, all English items throughout this article are boldfaced, and I use English spelling in all examples even though Haugen [6] uses a more phonetic spelling.

The underscore in (1d) and subsequent examples, indicates the position of missing functional material, from the point of view of Norwegian.[1]

1. a. det andre **crew**-et [6] (p. 571)
the.N other crew-DF.SG.N
'the other crew'.

b. eg fekk arbeid på **railroad**-en [6] (p. 590)
I got work at railroad-DF.SG.M
'I got a job at the railroad'.

c. ikke mange **party**-er [6] (p.587)
not many party-INDF.PL.M/F
'not many parties'.

d. den stor-e **building_** [7]; chicago_IL_01gk
the.M/F big-DF building

e. mange **lawyer-s** [7]; sunburg_MN_03gm
many lawyer-PL
'many lawyers'.

The outline of the article is as follows. Section 2 introduces AmNo, both in terms of its historical background and the material that is available, and in Section 3, the two corpora under investigation are presented, as well as some methodological concerns. This somewhat lengthy introduction to the empirical material is intended to give the reader some insight into the environment and conditions surrounding AmNo during its lifespan, as well as to establish the comparability of the two corpora under investigation. Section 4 presents the theoretical backdrop of the article and provides an analysis of the typical mixing patterns. Data showing diachronic change are presented in Section 5, and possible analyses of these changes are proposed and discussed in Section 6. Section 7 concludes the article.

2. The Heritage Language American Norwegian

AmNo is a Norwegian variety that emerged in communities of Norwegian immigrants who settled in North America (mainly the U.S.) roughly from the mid-1800s until the 1920s, and is still spoken by some of their descendants. This section provides an overview of some main events in the period of Norwegian immigration and the immigrants' new lives in America, as well as an introduction to the available AmNo data. For a more comprehensive discussion of the AmNo language and society see [6–10] and references therein.

2.1. Historical Background

The first Norwegian immigrants to America left Norway in 1825, and in the years between approximately 1850 and 1920, this escalated into a mass migration. According to Haugen [6] (pp. 28–29), as many as 810,000 Norwegians immigrated to the U.S. in the period from 1836–1930, a number nearly equal to the entire population of Norway in 1800. Upon arriving in the U.S., many Norwegian immigrants settled in the Midwest (in particular Wisconsin, Illinois, and Minnesota), gradually forming large Norwegian settlements, where important institutions such as churches, hospitals, retirement homes, and newspapers were quickly established [10].

---

[1] The following annotations are used in the glosses: DEF: Definiteness, DF: Definite, INDF: Indefinite, NUM: Number, PL: Plural, SG: Singular, GEN: Gender, M: Masculine, F: Feminine, N: Neuter. I have only provided a detailed glossary for the relevant noun phrases.

The conditions for, or necessity of, speaking English changed over the years following the first wave of immigration. The very first immigrants were forced to learn the language of the new country, English, in order to settle and live there. However, as the Norwegian settlements grew, this necessity diminished, and one could basically find everything one needed within the Norwegian-speaking community. Engaging in work, politics or social life outside the Norwegian settlement, on the other hand, required knowledge of English, and the children went to English-language schools. Subsequently, AmNo gradually turned into a language primarily used in the home and the church, the spheres most shielded from the English-speaking environment.

From the 1920s onward, the climate surrounding the Norwegian language in America changed. Immigration slowed down, Norwegian newspapers ceased publication, and major social and religious institutions switched to English as their main language. Moreover, the language became an obstacle for children, who typically entered school as AmNo monolinguals and faced teaching conducted in English. These issues, reinforced by a stigma against speaking Norwegian or speaking English with a foreign accent, led many parents to choose not to pass the language on to the next generation [6,11]. This severe decline notwithstanding, AmNo is still spoken in some areas, especially in the rural areas of the Midwest.

2.2. Heritage Languages

The Norwegian-speaking communities in the U.S. were always islands within the larger community in which English was, and is, the dominant language. Such immigrant language communities, situated in the midst of a larger, dominating language community, are recognized as heritage languages (HL), and their speakers as heritage speakers (HS). Definitions of these terms is given by Rothman [12]:

> A language qualifies as a *heritage language* if it is a language spoken at home or otherwise readily available for young children, and crucially this language is not a dominant language of the larger (national) society. [. . .] From a purely linguistic point of view, we assume that an individual qualifies as a heritage speaker, if and only if he or she has some command of the heritage language acquired naturalistically.
>
> Rothman [12] (p. 156)

As can be understood from this definition, a heritage language is acquired in childhood through naturalistic input, and HS are therefore considered native speakers of the HL [13]. However, at some point, typically when starting school, the speaker is introduced to the dominant language of the community, which in most cases eventually becomes the HS' own dominant language. This makes HS an interesting group of language users. On the one hand, the heritage language is their native language, but compared to monolinguals of the non-heritage variety of the language in question, they often do not reach the same level of competence. On the other hand, their competence might resemble that of L2 learners of the language, but HS will typically outperform L2 learners in many areas (see [14,15]). This tension has been discussed in several works and attributed to incomplete acquisition [16,17], or attritio [14,18]. Others have suggested that the grammar of HL should not be considered incomplete or impaired, just different, e.g., [19]. I return to these questions in Section 6.

In the case of AmNo specifically, the speakers in question are native speakers of AmNo, who acquired English as an L2. For the majority, English has also been their dominant language throughout most of their life. When they speak AmNo, it is clear that it is a variety of Norwegian; the majority of both lexical and functional items are Norwegian. However, English items occur frequently. Much work has been done documenting and researching AmNo, most of which focuses on the Norwegian properties of the language (see [9,10] and references therein). Language mixing in AmNo has also been investigated [1–5], and this is the phenomenon under investigation in the current article as well. The novelty of the current article, however, lies in a detailed investigation of aspects of the nominal

domain, providing the first systematic diachronic study of language mixing in AmNo. In the next subsection, I present the available data.

2.3. Data

The following timeline in Figure 1 gives a rough overview of the available AmNo material.

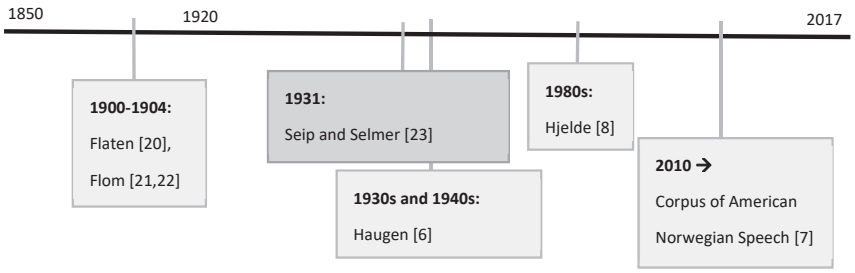


Figure 1. Overview of American Norwegian (AmNo) data.

Already around the turn of the 20th century, AmNo had gained the attention of researchers, when Nils Flaten and George Flom both published articles about the language variant [20–22]. Neither Flaten nor Flom collected large corpora, but in their articles, they included lists of English words occurring in AmNo. In 1931, Didrik A. Seip and Ernst W. Selmer interviewed and recorded several AmNo speakers, but unfortunately, this material was neither used much nor maintained very well. Many of the recordings were unfortunately broken or lost, and the quality of what remains is quite poor (remaining recordings are available in [23]). In the 1930s and 1940s, Einar Haugen carried out extensive fieldwork, which is presented and discussed in his two-volume work *The Norwegian Language in America* [6]. Arnstein Hjelde collected new data in the 1980s, and he was especially interested in a specific Norwegian dialect, *trøndersk* [8]. The most recent data collection effort started in 2010, under the auspices of the Norwegian in America (NorAmDiaSyn) project, and is still ongoing at the time of this writing. These data have been made available in the online CANS [7] created at the Text Laboratory at the University of Oslo, Norway. The diachronic comparisons in this article are based primarily on the material collected by Haugen [6] and the material in CANS [7], as these are the most extensive corpora and include a variety of dialects. These two corpora are introduced in the next section.

3. Introducing the Corpora and the Method

3.1. Haugen (1953)

Einar Haugen collected data from 1936 to 1948 [6]. At this time, the usage of Norwegian was already declining, and many cornerstone institutions in the Norwegian settlements, e.g., newspapers, social networks, and churches, were debating, or in fact carrying out, a switch to English as their main language. Nevertheless, Haugen describes communities where Norwegian was still spoken, churches occasionally had services in Norwegian, and the Norwegian newspaper *Decorah-Posten* was still circulated [6] (pp. 605–617). Although there was considerable variation among the communities, it would be fair to say that, in general, there was still a vital environment for the Norwegian language at the time of Haugen's data gathering.

Haugen's material consists of data from 260 informants, mainly from Wisconsin, collected through questionnaires, field notes and recordings. The first volume of his work is primarily a discussion of the AmNo society, whereas the second deals specifically with the linguistic data. The most relevant parts for the current article are the chapter discussing the grammar of English loanwords (i.e., what I refer to as mixed items), the selected vocabulary of English loans, which comprise 10% of the 3000 items he registered, and the appendix presenting the communities and informants studied.

Although some of Haugen's recordings are available online [24], I rely on his written materials and his own discussion of them, as the recordings are not transcribed and not of the best sound quality.

3.2. The Corpus of American Norwegian Speech

The most recent corpus available at present, CANS [7], captures the language as spoken nearly 100 years after the decline of immigration. Speakers were recruited through advertisements specifically seeking Norwegian-speaking persons whose ancestors had emigrated from Norway prior to 1920 and who had learnt the language at home from family members. Most informants recruited for this collection came from remote locations in the Midwest, where the Norwegian culture is still evident in cafes, shops, folk music, and handcrafting [9,25]. Usage of the language, on the other hand, varies. Some informants reported that they speak AmNo on a daily basis, whereas others might not have spoken AmNo since their parents passed away several years prior. All informants were, however, relatively fluent in AmNo [9]. Due to the challenges they faced, for instance at school, many have refrained from passing the language on to the next generation, meaning that these speakers may represent the last generation of AmNo speakers.

CANS is available online, and recordings of 50 individuals have so far been transcribed and published [7].[2] The corpus has two levels of transcription, one broad phonological transcription and one standardized transcription (*Bokmål*[3]), and sound and video files are provided. Individual items in the corpus are tagged with a variety of different grammatical categories, making it searchable. However, English items are not tagged in an equally detailed manner, and I have thus conducted certain specific searches to find these. The tag "x" provides all items not found in the Norwegian dictionary, which includes the English items, and through a process of manually sorting these items, 1265 English nouns remain.[4] These were subsequently sorted according to context. Seventy-five English nouns occur without any context at all and 156 nouns appear in a smaller (e.g., a phrase) or larger English context. Most interesting for the purpose of the present article are the 1034 English items that are found in an otherwise Norwegian context. The following discussions will be based on this sample.

3.3. Some Methodological Considerations

The two corpora presented in the subsections above clearly capture AmNo at two different stages of its development. Some methodological remarks should be made regarding the composition of the informant groups, and the form of the CANS corpus.

First, when discussing HS, first generation immigrants are typically not included as these speakers have acquired the language in circumstances with more exposure, and with no pressure from a dominant language [15]. On the other hand, one can argue that first generation immigrants should be included as they too are speakers of a minority language in their current society, and that their language may show contact-induced differences similar to other HS (see, e.g., [3,14]). In this article, I do not present arguments supporting either side of this issue. However, I include first generation immigrants in the group of HS from the 1930s and 1940s, as Haugen does not separate these speakers from the others in his material. His description and discussion of AmNo grammar are thus based on a heterogeneous group of AmNo speakers. Still, Haugen provides a complete list of his informants and to which generation they belong. This list reveals that the majority are in fact second or later generation speakers of AmNo, thus unquestionably HS. I therefore assume that Haugen's overall

---

[2] The collection of data is still ongoing as of the time of this writing (early 2017), and the corpus will be expanded in the future.
[3] Bokmål is one of the two written standards of Norwegian. See [26,27] for discussion of the Norwegian language situation.
[4] Proper nouns, fixed expressions, and repetitions within the same immediate utterance have been excluded from the count, and for words with a potential lexical overlap between English and Norwegian, I have used the sound files to determine whether they have an English-like or a Norwegian-like pronunciation, and sorted them accordingly.

findings represent a typical heritage speaker, and I rely on Haugen's evaluations and commentaries as an authentic description of AmNo at the time.

A second concern, especially relevant when doing diachronic comparisons of the two corpora, is the fact that there is no established family link between the two groups of speakers. As far as we know, the speakers in CANS [7] are neither the same speakers, nor the children of the speakers in Haugen [6], meaning that we lack information about their input and competence throughout life. Nevertheless, based on the CANS speakers' ages and locations, we can assume that the group of speakers discussed by Haugen represent the parents/grandparents from whom the CANS speakers learnt AmNo, and thus the grammar discussed by Haugen represent the input that the CANS speakers received. In other words, even if the two corpora are not directly connected in terms of family relations, a comparison of the two will still show the general development of AmNo over these years.

A brief comment should also be offered regarding the new corpus and the data drawn from it. As an online corpus, CANS is not fixed in the same way as other corpora and may be subject to additions, updates, and improvements. In practice, this means that the details of the corpus may change over time. The data presented and discussed in this article were drawn from the corpus in April 2016, and I have not considered any later updates. In addition, random searches have demonstrated that a few English items are incorrectly not tagged "x". In order to make the data employed in this article as clear as possible, these data are *not* included in the numerical description above, but I will occasionally use them as examples of specific phenomena. A footnote is provided in these specific cases.

In Section 5 and onwards, I compare the data collected by Haugen and in CANS and present changes in the patterns of language mixing within noun phrases. Before going into this material, I will briefly introduce the theoretical background for the article in the next section, as well as a description and illustrative examples of what can be considered the *typical* or *main* pattern of language mixing in AmNo noun phrases. This will serve as the foundation for investigating potential changes.

4. Theoretical Background

4.1. Language Mixing

In this article, I employ the term "language mixing" to describe the phenomenon under investigation, namely the occurrence of English items in AmNo.[5] This type of mixing is what Myusken [28] (p. 3) refers to as insertion, i.e., the "insertion of material [...] from one language into a structure from the other language", and occurs quite frequently in AmNo noun phrases, forming a recognizable pattern where English nouns appear with Norwegian determiners and suffixes in a Norwegian word order [1,4]. Examples of this are presented in (1a–c), repeated here as (2) for convenience.

2. a. det andre **crew**-et [6] (p. 571)
the.N other crew-DF.SG.N
'the other crew'.

b. eg fekk arbeid på **railroad**-en [6] (p. 590)
I got work at railroad-DF.SG.M
'I got a job at the railroad'.

c. ikke mange **party**-er [6] (p.587)
not many party-INDF.PL.M/F
'not many parties'.

[5] The terms "code-switching" and "borrowing" are also frequently used to describe this phenomenon. See [1] for discussion of these terms and how they relate to each other.

From a formal perspective, there are two main ways of approaching and analyzing language mixing: to posit special constraints to account for mixing data [29,30], or to assume that mixing is constrained by the same principles as monolingual speech [31–37]. In the literature, the latter approach is referred to as a Null Theory [31] or constraint-free approach to language mixing [33].

I assume that Null Theory should be the null hypothesis. However, key empirical insights from the other model appear too essential to be overlooked. Myers-Scotton [29] observes that one of the languages involved is more prominent in cases of language mixing. This is referred to as the Matrix Language (ML), and it provides both word order and functional morphemes in the mixed utterances. The other language(s), the Embedded Language(s) (EL), can only contribute content items. I argue that an exoskeletal model, which I will employ in this article, can account for these asymmetric contributions and at the same time be a Null Theory of language mixing. Although I acknowledge this empirical asymmetry, an essential distinction is that I nevertheless do not adopt Myers-Scotton's notion of ML and EL as theoretical primitives. Instead, I use the terms "main" and "secondary" language quite informally as descriptive or observational terms.

4.2. Exoskeletal Approaches to Grammar

To analyze these data, I employ a late-insertion exoskeletal model. The term "exoskeletal" unites a family of grammatical analyses [38–48]. These works may differ in terms of how they account for details in the syntactic structure and its derivation, but the shared, fundamental core is the assumption that abstract syntactic structures are generated independently of the lexical items that will realize them. These approaches are all motivated by monolingual data, meaning that they are not specially designed to handle language mixing, but do nevertheless prove to be good analytical tools for bilingual grammars. The specific model employed in the current article relates mainly to the works by Borer [31–43], Åfarli [44], Lohndal [46,47], and Marantz [38,39]. Additionally, the current approach also incorporates insights from Distributed Morphology (DM), e.g., [49–51], especially concerning the process of late insertion, which I will discuss below. In DM, the lexicon is split into three separate lists: syntactic terminals, vocabulary items, and encyclopedic information. The encyclopedia holds "world-knowledge", which is not relevant for the grammar, and thus is not discussed in this article. The first and second lists, however, are important in the late-insertion exoskeletal model and how language mixing is analyzed.

The first list holds abstract syntactic components, which are used to build structures, forming a syntactic frame or template for the sentence. There are two different types of terminals in this list: roots and functional features or feature bundles. The properties of roots and how they are structured in the syntax is a much-debated question (see, e.g., [52]), however not one that I will delve into here as it is not crucial for the purpose of the current discussions and analyses. Importantly, roots are considered devoid of any grammatical features. Roots also therefore lack a lexical category, which is instead syntactically assigned. Following Marantz [39], Arad [53], Pylkkänen [54], and Embick and Marantz [55], I assume that the category is assigned by combining the root with a category-defining head, constituting a complex I will informally refer to as the stem. Moreover, I assume that roots have some core yet underspecified phonological and semantic properties (cf. [53]).

Functional features are the second type of syntactic terminals in this list, and they are considered properties of the abstract syntactic structure. Moreover, features in this context are restricted to formal morphological features, and these may be bundled in different projections.

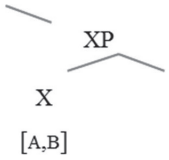
Phonological content is provided in the process of Spell-Out, or in DM, Vocabulary Insertion. In this process, vocabulary items, or phonological exponents, from the second list are accessed and inserted. For the two types of syntactic terminals, this process is radically different. Following Arad [53], I assume that a root alone is unavailable for Spell-Out and can only be phonologically realized in combination with a category-defining head. In other words, the stem is spelled out as one unit. This position in the structure emerges as relatively open, with few restrictions for insertion, meaning that content items from any language and of any kind may in principle be inserted.

Spell-Out of functional features or feature bundles, on the other hand, is a more restricted process, regulated by the Subset Principle:[6]

> The phonological exponent of a Vocabulary item is inserted into a morpheme in the terminal string if the item matches all or a subset of the grammatical features specified in the terminal morpheme. Insertion does not take place if the Vocabulary item contains features not present in the morpheme. Where several Vocabulary items meet the conditions for insertion, the item matching the greatest number of features specified in the terminal morpheme must be chosen.
>
> Halle [56]

In other words, insertion of functional exponents is a competitive process, where the exponent that best matches the features specified in the syntactic terminal wins and is inserted. However, the exponent cannot be specified for any features that are not represented in the structure. The structure in (3) serves as a simplified illustration of this process.

3.



Here, the phonological exponent of X should be the best possible match to the feature bundle [A,B]. In case a complete match is available, this will rule out any alternative exponents specified for only [A] or [B]. Furthermore, an exponent with the specifications [A,B,C] would not be allowed for insertion at this terminal, as the feature [C] is not part of the syntactic structure.[7]

These different restrictions on Spell-Out of functional and substantial material will capture the empirical asymmetry in language mixing: content items from any language can be inserted in the stem position, whereas the most appropriate functional exponents typically are provided by the language of the syntactic frame. Hence, content items from any language are predicted to acquire the functional properties of the language specifying the syntactic frame. Notice, however, that this does not mean that structures bear language tags in our grammars. Instead, structures are composed of functional features, and a specific language is recognized by the features that are active in the language and how they are combined [59]. In other words, when describing something as a Norwegian structure, I mean a structure composed of features in a combination that it is typically associated with Norwegian.

In the next subsections, I will introduce the structure of the AmNo noun phrase (DP) and the typical mixing patterns seen in AmNo to demonstrate how a late-insertion exoskeletal model offers an insightful analysis of these data.

4.3. The Structure of (American) Norwegian Noun Phrases

Norwegian is the main language in AmNo and provides the structural frames. In this section, I will therefore introduce and discuss the Norwegian DP structure and thereafter employ this framework in an analysis of mixed AmNo noun phrases.

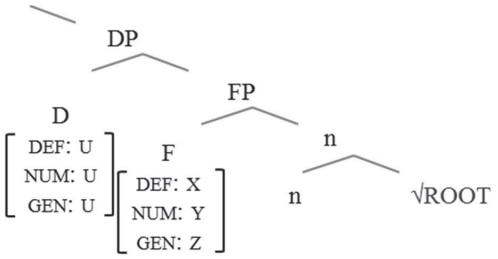
---

[6] Terminals holding functional features or feature bundles are referred to as morphemes in the DM literature.

[7] The mechanisms presented here imply Underspecification, which plays an important role in DM. The basic assumption is that vocabulary items are underspecified for syntactico-semantic features. Hence, one vocabulary item can spell out several syntactic positions, but in cases where multiple exponents compete for the same position, the more specified one is inserted. As pointed out by an anonymous reviewer, other studies have shown that bilinguals simultaneously activate elements from both languages, and a model has been proposed in which multiple elements may be present simultaneously in a position in the linguistic structure, referred to as co-activation or blends. See [57,58] for discussion of such an analysis.

Norwegian DPs, like Scandinavian DPs in general, can be quite complex, and they have been thoroughly studied in various works [60–62]. The obligatory components of the Norwegian noun phrase are the stem (i.e., the root together with its categorizer), one (or more) functional projections above the stem, and finally a D layer.[8] Norwegian nouns are inflected for three functional categories: definiteness, number and gender, which will be recognizable through affixes and associated words in the noun phrase. The basic structure employed in this article is presented in (4) (see [4] for a more elaborate discussion of the different projections in this model, or [62] for an in-depth study of Norwegian DPs in general).

4.



At the bottom of this structure is the stem, which is composed of a root and a category-defining head, in this case a nominalizer. Following from the discussion in the previous subsection, I assume that the root needs to be merged with such a categorizer in order to be spelled out.

Immediately above the stem, we find a functional projection (F) holding a bundle of the features gender, number and definiteness. In the literature, there are various alternatives as to how these are structured, for example with two [62] or three separate projections [63]. For the purpose of the analyses in the current article, however, such a detailed structure is not necessary. Moreover, the AmNo data exploited here do not provide new insight into the division of the functional features in the structure, so number and definiteness are combined into one projection. The most debatable issue in (4) is presumably gender, as a notable part of the literature argues that gender is a property of the nominalizer, thus part of the nominal stem [62,64,65]. Language mixing gives reason to argue that gender is positioned higher in the structure. Consider, for instance, English derived stems like *settlement*, *township*, and *building*, which are attested in the AmNo material [6–8]. Assuming that the derivational suffixes are realizations of the nominalizer, n, these data show that stems are available for mixing. Thus, if gender were considered a property of the stem, we would not expect the pattern where English stems are mixed into AmNo and assigned to different gender categories.[9]

I thus assume that the stem is generally the item being drawn from the secondary language in language mixing, and that gender is positioned in the higher functional structure of the Norwegian DP together with number and definiteness. The interplay between the functional features in F will determine which functional exponent is most appropriate for insertion. Furthermore, in a Norwegian DP the stem complex obligatorily moves to F, possibly due to some nominal feature, meaning that the exponent of F will materialize as a suffix on the noun stem [62].

On top of the noun phrase is a D projection, holding a feature bundle of the corresponding unvalued functional features. These get their valuation through a probe-goal relation (Agree) between

---

[8] The Norwegian DP may also include weak quantifiers, adjectives, pre- or post-nominal possessors and post-nominal prepositional phrases. A discussion of these is beyond the scope of the current article. See [62] for details.

[9] An alternative analysis could be that in case of language mixing the speaker has established two separate entries for nouns in their list of vocabulary items, one without gender (the English version) and one with gender (the Norwegian version). Due to the uneconomical status of this analysis, I will not pursue it.

D and F.[10] Noun phrases constituting arguments in Norwegian typically require an overt realization of the DP domain. This is accomplished either by moving FP to Spec-DP, or by inserting a separate determiner or demonstrative in D (see [62] for discussion).[11] The latter alternative results in the phenomenon known as double definiteness, i.e., the co-occurrence of definiteness in the determiner and in the suffix. In phrases involving an adjective or a weak quantifier, double definiteness is obligatory as FP is prevented from moving to the DP domain by intervening projections [62].

4.4. Typical Mixing Patterns in AmNo Noun Phrases and How to Analyze Them

Since Norwegian is the main language in AmNo, we can expect to find mixed noun phrases with a Norwegian structure and Norwegian functional exponents into which English stems are incorporated. This is, in fact, the pattern described by both Flaten [20] and Haugen [6]:

> Some words are, indeed, used without any appreciable difference in pronunciation, but more generally the root, or stem, is taken and Norse inflections are added as required by the rules of the language.
>
> Flaten [20] (p. 115)

> A single form is usually imported and is then given whatever endings the language requires to make it feel like a proper word and to express the categories which this particular language requires its words to express.
>
> Haugen [6] (p. 440)

Moreover, this is also the main pattern of mixing in the most recently collected material [4]. As discussed above, this pattern is predicted by the exoskeletal model: AmNo speakers produce structures with functional features typical for Norwegian, the main language. The Subset Principle requires these to be spelled out by the most appropriate exponents, namely the Norwegian functional exponents. The stem, on the other hand, is drawn from English, and acquires Norwegian functional properties by being inserted into such a structure.

Some examples are shown in (5), where English stems occur with a Norwegian indefinite article (5a), or with a Norwegian functional suffix and in a Norwegian word order (5b–d). Note that even though the noun phrase alone is shown here, these DPs are all part of larger Norwegian utterances.

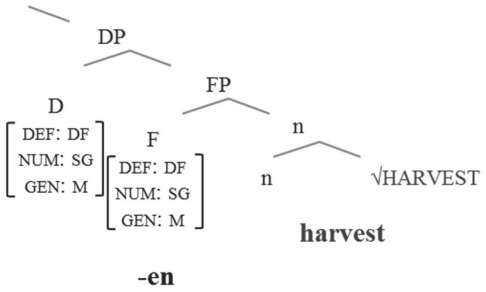
5.	a.	et	rent	**towel**	[6] (p. 601)
		a.INDF.SG.N	clean.INDF.SG.N	towel	
		'a clean towel'.			
	b.	**harvest**-en			[6] (p. 579)
		harvest-DF.SG.M			
		'the havest'.			
	c.	**field**-a			[6] (p. 575)
		field-DF.SG.F			
		'the field'.			
	d.	**trunk**-en	min		[6] (p. 603)
		trunk-DF.SG.M	my		
		'my trunk'.			

[10] In cases involving a weak quantifier or an adjective, these will be generated in separate projections between D and F and will also have unvalued corresponding features. See [62] for discussion.

[11] Note that Julien [62] proposes a separate projection for demonstratives. However, for convenience, I analyze both determiners and demonstratives as exponents of D (see also [66]).

The exoskeletal model serves as a good analytical tool for these cases of mixing, and as an example, the structure of (5b) *harvesten* 'the harvest' prior to movement of the stem complex is presented in (6).

6.



The structure generated in (6) is a typical Norwegian structure, where feature bundles composed of definiteness, number, and gender are present in F and D, and in (6) the features of D have already been valued by Agree with F. The structure also shows the inserted phonological exponents (boldfaced). The mechanisms are as follows: the English lexical item *harvest* has been inserted into the stem position, which is possible since this position does not have strict requirements for insertion. The functional feature bundle in F, on the other hand, is spelled out by a Norwegian exponent offering a complete match with the relevant features: definite, singular, masculine. In the next step (not shown here), I assume that the stem obligatorily moves to F, yielding the complex form *harvesten*, and that this complex subsequently moves to Spec-DP in order to fulfil the interpretability requirements of the DP domain in Norwegian (see [62]).

This brief overview of the analysis of the typical mixing pattern in AmNo noun phrases serves two purposes. First, it demonstrates that the late-insertion exoskeletal model is a good analytical tool for analyzing this type of language mixing. See also [1,4] for a more in-depth discussion and analysis of the typical mixing pattern in AmNo. Second, this discussion is relevant as the basis for investigating potential diachronic changes in language mixing, which is the topic of the following sections.

5. Diachronic Change

In this section, I compare data from Haugen [6] and CANS [7] and show that diachronic changes can be found in the mixing patterns. Due to the limits of this article, I will not discuss the DP exhaustively, but focus on how gender, number, and definiteness are realized by suffixes on the noun stem or on determiners or demonstratives in D. The data are discussed separately: Haugen in Section 5.1, and CANS in Section 5.2. In the former subsection, I also include a brief introduction to how gender, number, and definiteness are typically realized in a Norwegian structure.[12] Please recall that when referring to specific examples, data from Haugen [6] are accompanied by the page number where the examples can be found, and data from CANS [7] by the informant code.

---

[12] Notice that this article discusses the data on a population level, considering the two corpora as two different stages in the development of AmNo. There are without a doubt individual differences in both groups, and studying individuals would possibly yield additional insights. However, discussing changes on a population level, as in the present article, will provide a general overview of potential changes and their development, which is beneficial to a study on the individual level in the future.

5.1. Haugen (1953)

5.1.1. Gender

Gender in Norwegian is non-transparent. This means that one cannot tell the gender of a noun from the phonological or semantic properties of the noun itself. Instead, gender is revealed by affixes and associated words.[13] Previous studies have documented and mapped the gender distribution of nouns in both non-heritage Norwegian [69] and in AmNo [6,70–73]. Without going into the details of these studies, they all establish that masculine is the predominant gender of Norwegian nouns, accounting for 50% or more of nouns, whereas feminine and neuter each cover a smaller percentage, which may wary across different dialects.

In mixed AmNo phrases, English nouns are also assigned to one of the three genders in Norwegian, despite the fact that English nouns do not have gender.[14] Table 1 shows the distribution in Haugen's material.

Table 1. Gender distribution among English nouns in Haugen (1953).

Haugen [6]	
M	71.6%
F	1.6%
N	8.2%
Alternating gender	18.6%

M: Masculine; F: Feminine; N: Neuter.

Haugen bases these numbers on a sample of 317 noun stems in his material. All three genders are used, and similar to the distribution of the native Norwegian vocabulary, masculine is the most frequent gender. In this selection, 59 nouns, or 18.6% of the total, vacillated between genders, which is not surprising considering that many nouns are assigned different genders in different Norwegian dialects (see, e.g., [75]).

5.1.2. Number

Plurality is typically expressed as a functional suffix both in Norwegian and in English, and the Norwegian plural suffix additionally varies according to gender. In Haugen's [6] material, English nouns in plural phrases typically occur with a Norwegian suffix. In fact, Haugen states that a loanword "almost universally [was] given the most common plural ending of the gender to which it had been assigned" [6] (p. 450). Some examples are provided in (7).

---

[13] Notice that there is discussion in the literature concerning whether the definite suffix in Norwegian is a marker for gender or rather for declension class [67,68]. I assume that the suffix expresses gender, and will analyze it accordingly.

[14] The interest of this article is the distribution across the different genders, and not the process of how an individual noun is assigned a specific gender. This presumably relies on a number of factors not addressed in the present article, such as phonology, conceptual content, convention, and it can vary among different varieties of Norwegian. See [61,74] for an approach that is compatible with the late-insertion exoskeletal model.

7. a. **piec[e]**-ar [6] (p. 450)
piece-INDF.PL.M

b. **creek**-ar [6] (p. 450)
creek-INDF.PL.M

c. **bluff**-er [6] (p. 563)
bluff-INDF.PL.F

d. **field**-er [6] (p. 757)
field-INDF.PL.F

e. **team**- [6] (p. 450)
team-INDF.PL.N

f. **store**- [6] (p. 598)
store-INDF.PL.N

However, one English inflectional form is attested in Haugen's material, and that is the plural suffix *-s*. In accounting for the usage of this suffix, Haugen splits the speakers into two groups: pre-bilingual borrowers and childhood bilinguals. Pre-bilingual borrowers are those who acquired English in adulthood, and are not considered "true" bilinguals. Haugen suggests that these speakers were not aware of the plural value of *-s*, consequently producing cases where the *-s* is present both in singular and plural, e.g., in *cookies* (used in both SG and PL), and with Norwegian suffixes in addition, e.g., *car-s-ar* 'car-PL-INDF.PL.M' and *bean-s-en* 'bean-PL-DF.SG.M' [6] (pp. 450–451). Haugen concludes that these speakers took the *-s* to be part of the noun stem.

The second group, the childhood bilinguals, occasionally uses the *-s* in its correct plural function and as a replacement for a Norwegian alternative. This is, according to Haugen, limited to indefinite cases, and foreshadowing the diachronic development, Haugen comments that the usage "naturally increased as time went on" [6] (p. 451).

5.1.3. Definiteness

As discussed above, definiteness in Norwegian is expressed both in F and in D. Due to the stem complex obligatorily moving to F, the exponent of F in a definite phrase materializes as a functional suffix on the noun stem, commonly called the definite article. The realization of D, on the other hand, can be fulfilled either by FP moving further to Spec-DP or by spelling out D with a separate determiner or demonstrative.

Concerning definiteness, Haugen gives two clear restrictions for AmNo: "Whether words were singular or plural [...] they had to add the N[orwegian] *definite article* under appropriate circumstances" [6] (p. 451) and "E[nglish] *the* would not be acceptable" [6] (p. 451). In other words, in definite phrases, realization of Norwegian functional exponents is obligatory. Some examples are given in (8).

8. a. **railroad**-en [6] (p. 590)
railroad-DF.SG.M

b. **field**-a [6] (p. 575)
field-DF.SG.F

c. det **crew**-et [6] (p. 571)
the.N crew-DF.SG.N

In the next subsection, I consider the more recently collected data and show how some of the patterns and restrictions discussed by Haugen have changed.

5.2. Corpus of American Norwegian Speech

This section provides data from CANS [7] showing patterns of language mixing deviating from the ones attested in Haugen [6]. The basis of the discussion is the 1034 English nouns occurring in a Norwegian context in CANS, see Section 3.2. above. Thus, due to the relatively limited amount of data in the corpus, the following presentation serves primarily to describe a trend of diachronic change.

5.2.1. Gender

When accounting for gender distribution in the most recent AmNo material, I have considered all singular forms where gender is revealed by the indefinite article or the definite suffix. Although plural forms are also sensitive to gender, these are excluded due to the syncretism of plural masculine and feminine in many Norwegian dialects and in the written standard, Bokmål. What remains is a sample of 292 nouns. Their distribution is presented in Table 2.[15]

Table 2. Gender distribution among English nouns in Corpus of American Norwegian Speech (CANS).

	CANS [7]
M	66.1%
F	6.5%
N	6.2%
Alternating gender	21.2%

Similar to Haugen's findings presented in Table 1, the category of alternating gender in Table 2 includes the nouns that vacillate between genders. This group covers roughly one fifth of the nouns, whereas 66.1% of the nouns are masculine, 6.5% feminine and 6.2% neuter.[16]

Comparing these numbers with earlier material, the distribution of gender appears to be relatively stable; see Table 3.[17] Generally, around 70% of the nouns are masculine, whereas feminine and neuter each are assigned to less than 15% of the nouns. The group of nouns with alternating genders in the most recent material is quite large, which may indicate some uncertainty in the gender system (see [73] for discussion). However, as there is no clear developmental pattern or obvious diachronic change, the question of gender will not be discussed further in this article.

Table 3. Development of gender distribution among English nouns in American Norwegian (AmNo).

	Flom [68]	**Haugen [6]**	**Hjelde [69]**	**CANS [7]**
M	71%	71.6%	70.7%	66.1%
F	5%	1.6%	10.5%	6.5%
N	16%	8.2%	15.7%	6.5%
Alternating gender	8%	18.6%	3.1%	21.2%

[15] This distribution concerns English nouns mixed into AmNo. For gender distribution among Norwegian nouns in CANS, see [72,73].

[16] The numbers are based on tokens in the selection. Counting types instead would provide a slightly, but not radically, different picture with 79% masculine, 7.6% feminine, 7.6% neuter and 5.7% alternating gender.

[17] Notice that Hjelde's [71] numbers for feminine and neuter are slightly higher than in the other distributions, which may be due to the fact that Hjelde isolated one specific dialect in his study.

5.2.2. Number

As discussed above, Haugen [6] separated his informants into two groups: pre-bilingual borrowers and childhood bilinguals. All of the speakers represented in CANS were born in the US or Canada and belong to the latter group, and similar to Haugen's findings for this group, both Norwegian plural suffixes and the English plural suffix *-s* are used by the CANS speakers. Some examples are given in (9), where (9a,b) show phrases with the Norwegian suffixes, and (9c–f) show cases with the English suffix.

9.	a.	ti ten	**kid**-er kid-INDF.PL.M/F	[7]; portland_ND_01gm
	b.	**boss**-er boss-INDF.PL.M/F		[7]; coon_valley_WI_06gm
	c.	mange many	**lawyer-s** lawyer-PL	[7]; sunburg_MN_03gm
	d.	fem five	**dialect-s** dialect-PL	[7]; portland_ND_01gm
	e.	andre other	**tool-s** tool-PL	[7]; sunburg_MN_03gm
	f.	alle slags all kinds of	**pill-s** pill-PL	[7]; westby_WI_02gm

In light of Haugen's [6] study, the fact that both Norwegian and English plural suffixes are attested in CANS is not surprising. What is interesting as a possible sign of diachronic change, however, is the distribution of these two realizations. Haugen does not provide any quantitative measures of the distribution, but since loanwords "almost universally" were given Norwegian plural suffixes, we must assume that the English plural suffix was used in a clear minority of cases. In CANS, on the other hand, this picture is reversed. Out of all 175 plural phrases involving an English noun, 103 are realized with the English plural *-s*, compared to 37 cases with the Norwegian suffix. The remaining 35 phrases are realized without any plural suffix, which I will return to below. Among the cases with the plural *-s*, the vast majority are indefinite phrases, as in the examples in (9c–f) above. However, in CANS [7] the *-s* occasionally occurs in definite phrases. Some examples are given in (10).

10.	a.	alle disse all these	**minute-s** minute-PL	[7]; stillwater_MN_01gm
	b.	disse these	**lutefisk dinner-s** lutefisk dinner-PL	[7]; westby_WI_03gk
	c.	de samme the same	**gene-s**[18] gene-PL	[7]; flom_MN_02gm

[18] This phrase is not included in the count described in Section 3.2, due to being part of a repetition.

The pattern in (10) is not found in Haugen [6], and the attestations in the new corpus are not frequent (10 attested examples). A common property is, nevertheless, that in addition to the plural suffix *-s*, they have a Norwegian exponent of definiteness in the higher projection, D.[19]

A third pattern, not mentioned by Haugen, is plural phrases without any suffix at all. This pattern is found primarily in indefinite cases, as shown in (11).

11.	a.	fem seks five six	**hour_** hour	[7]; chicago_IL_01gk
	b.	flere more	**store_** store	[7]; westby_WI_03gk
	c.	mange many	**memorial_** memorial	[7]; webster_SD_01gm

It is, however, challenging to provide a reliable estimate of the prevalence of such cases, since they are often impossible to confirm as plurals. The examples in (11) are given away by their quantifiers.[20]

5.2.3. Definiteness

Concerning definiteness, the majority of the relevant cases in CANS [7] behave the same way as described by Haugen [6] (and expected in a Norwegian structure): they receive the definite suffix as expected, and the determiner or demonstrative is present in relevant cases [4]. Nevertheless, two patterns of change are found.

The first pattern of change is omission of the functional suffix. Among definite singulars, 98 phrases occur without the definite suffix. Some examples follow in (12).

---

[19] One phrase may, based on its context, be considered an instance where the *-s* occurs alone in a definite phrase: *hun har tickets* 'she has the tickets' ([7]; coon_valley_WI_02gm), but it is the sole example of its kind.

[20] Two possible definite cases are also attested: *disse garter snake_* 'these garter snake' ([7]; sunburg_MN_03gm) and *disse deer_* 'these deer' ([7]; stillwater_MN_01gm). However, since the latter is a possible English realization of plural *deer*, and the former is produced after hesitation, this pattern is very limited.

12.	a.	den that.M/F	**school_** school	[7]; gary_MN_01gm
	b.	den that.M/F	**birdhouse_** birdhouse	[7]; coon_valley_WI_12gm
	c.	denne this.M/F	**cheese_**[21] cheese	[7]; blair_WI_04gk
	d.	den the.M/F	store **building_** big building	[7]; chicago_IL_01gk
	e.	det the.N	gamle **stuff_** old stuff	[7]; chicago_IL_01gk
	f.	det the.N	norske **settlement_** Norwegian settlement	[7]; albert_lea_MN_01gk
	g.	det the.N	første **trip_** first trip	[7]; westby_WI_06gm
	h.	**nephew_** nephew	min my.M/F	[7]; portland_ND_02gk
	i.	**family_** family	min my.M/F	[7]; portland_ND_01gm

(12a–g) show cases where the characteristic double definiteness in Norwegian is expected, but definiteness is only expressed by the determiner or demonstrative.[22] The latter two examples, (12h,i), show phrases with a post-nominal suffix where the definite suffix is expected, but omitted.[23] Notice that this pattern is not exclusive to the mixed phrases, as there are examples of the definite suffix being omitted in "all-Norwegian" phrases also, as in (13).[24] This might indicate that the change is not directly connected to the mixed items, but is rather a more general process.

---

[21] This phrase is not included in the count presented in Section 3.2, as the item *cheese* is not tagged "x", even though *cheese* is not a common Norwegian vocabulary item.

[22] Notice that the adjectives in question also show that the phrase is definite, as they have the weak inflection *-e*, which is typical for definite cases.

[23] Family terms are often used without the definite suffix, e.g., *far min* 'my father' and *mor mi* 'my mother', but arguably this is not equally common with *nevø* 'nephew', or with *familie* 'family' as in (12h,i).

[24] Norwegian noun phrases can be realized with only a determiner or demonstrative and without the functional suffix, primarily in a formal or written-like style. As these informants are not formally educated in Norwegian, I consider it possible, but not very likely, that they are using this style of speech.

13. a. denne skole_ [7]; harmony_MN_01gk
this.M/F school

b. den sommer_ [7]; coon_valley_WI_12gm
that.M/F summer

c. dette land [7]; chicago_IL_01gk
this.N country

d. dette brød_ [7]; blair_WI_07gm
this.N bread

e. disse nabolag_ [7]; chicago_IL_01gk
these neighborhood

The second change is the usage of the English determiner *the*, which was described as unacceptable by Haugen. Although not frequent, 20 cases are attested in CANS where the determiner *the* occurs together with a Norwegian noun or in an otherwise Norwegian structure. Some examples are given in (14), and all these examples are part of a larger Norwegian utterance.

14. a. **the** by [7]; chicago_IL_01gk
the city

b. **the** ungdom [7]; harmony_MN_01gk
the youth

c. **the** gamle kirke [7]; chicago_IL_01gk
the old church

d. **the** penger [7]; albert_lea_MN_01gk
the money

A common property of the phrases in (14) is that they additionally lack the definite suffix, which would have been expected in a corresponding Norwegian noun phrase. However, in about half of the 20 phrases, the English determiner *the* co-occurs with such a Norwegian definite suffix, as in (15).

15. a. **the** gård-en [7]; gary_MN_01gm
the farm-DF.SG.M

b. **the** rest-en [7]; vancouver_WA_03uk
the rest-DF.SG.M

c. **the** andre dag-en [7]; vancouver_WA_01gm
the other day-DF.SG.M

d. **the** samme tid-a [7]; albert_lea_MN_01gk
the same time-DF.SG.F

As CANS enables the researcher to listen to the recordings of these AmNo speakers, it is worth mentioning that the phonology of the determiner varies, and a possible objection could concern the similarities between *the* and the Norwegian neuter determiner *det*. These two may in fact sound quite similar, especially if *the* is pronounced with an alveolar stop instead of a dental fricative. However, two arguments support the analysis of these as English determiners. First, a prenominal determiner would, in most cases, be an alien element in a Norwegian structure without the definite suffix as

in (14) (see also footnote 24). Secondly, the phrases where a prenominal determinative *is* expected in Norwegian, e.g., in the phrases requiring double definiteness, are primarily masculine, as in (15a–c). This means the appropriate Norwegian determiner would be *den*, which is not equally similar to the English *the*.

5.2.4. The Indefinite Article

In addition to the comparisons of Haugen [6] and CANS [7] so far, a brief comment on the indefinite article needs to be added. Haugen [6] does not mention or discuss any generalizations or irregularities concerning the indefinite article. Thus, we must assume that its usage follows an expected Norwegian pattern in Haugen's material. In the new corpus, however, the domain of the indefinite article also seems to be subject to diachronic change, and in parallel to the discussion of definite phrases above, these changes materialize as either omitting the article (60 attested cases) or, in a few cases, using the English *a* (I have found 8 such cases). Some examples are provided in (16), with the relevant context included.

16.	a.	så then	du fikk you got	_ **candybar** a candybar	[7]; webster_SD_01gm
	b.	han hadde #[25] he had	_ stor a big	**steam engine** steam engine	[7]; rushford_MN_01gm
	c.	det er it is	_ **bluebird** a bluebird	som sitter ute that sits outside	[7]; coon_valley_WI_01gk
	d.	**a** a	stort hus big house		[7]; albert_lea_MN_01gk
	e.	**a** a	spiker nail		[7]; flom_MN_02gm

These patterns are less frequent than in the definite phrases, but they still do occur.

5.3. Interim Summary of the Findings

In this section, I have compared Haugen [6] to CANS [7] and presented systematic changes between the corpora concerning the categories gender, number, and definiteness in mixed noun phrases. Concerning gender, the main interest in this article is the distribution across the three genders of Norwegian, where no remarkable change was found when comparing today's AmNo to Haugen [6] and other previous AmNo collections. Hence, I decided to focus on number and definiteness in the rest of the article.

In the case of number, Haugen describes two patterns concerning the realization of plurality. In most cases, the appropriate Norwegian suffix is added, but among a subgroup of his informants, the childhood bilinguals, the English suffix *-s* is attested. In CANS both these patterns are attested, and the English plural *-s* is used in most cases. In addition, a new pattern is attested, namely the omission of a functional suffix in plural phrases.

Concerning definiteness, two patterns that are unattested in Haugen's material are found in CANS: definite functional suffixes are omitted in several phrases, and the English determiner *the* is sometimes used instead of a Norwegian alternative. Interestingly, omission of functional material and

[25] # marks a brief pause.

usage of English determiners are also attested in indefinite phrases in CANS, which is something that is not discussed by Haugen.

Summing up, the data show two main patterns of change in the AmNo noun phrases:

1. Omission of functional suffixes, both in plural and/or definite cases
2. Usage of English functional exponents

In the next section, I will continue the discussion of the patterns that diverge from the typical pattern of mixing in AmNo, and explore how the observed changes can potentially be explained by changes in the underlying grammar. Furthermore, I briefly address some limitations when it comes to investigating diachronic changes in a language like AmNo.

6. Analysis and Discussion

From the perspective of the exoskeletal model, two different scenarios can explain changes like the ones presented in Section 5.2: on the one hand, we could assume that the structure is intact, but the exponents have changed. On the other hand, we could assume that the observed change is a result of the structure itself changing. Both scenarios would disrupt the process of insertion, facilitating realizations diverging from the expected patterns. In this section, I explore these two alternatives separately.

6.1. Change in the Exponent

In the first scenario, we assume that the abstract syntactic frame is intact, and the observed change is caused by the functional exponents and/or their conditions for insertion. Support for this alternative is found in the Missing Surface Inflection Hypothesis (MSIH) [76,77]. This hypothesis was proposed based on evidence from second language acquisition and claims that the absence of overt morphology does not necessarily mean the absence of functional categories in the syntax. Instead, the lack of overtly realized functional exponents may be due to the learner not having established the complete set of exponents or by a failure to meet matching conditions between the exponent and the structure [76]. For the AmNo speakers, we can in a parallel manner assume that the structure is generated as expected, but that their repertoire of phonological exponents and corresponding versatility concerning insertion may be reduced, creating obstacles in the spell-out process.

A key word in the MSIH is avoidance, as the learner is taken to prefer a missing form over a faulty inflection [76]. In other words, when the speaker is in doubt, she will, consciously or not, avoid inserting any exponent in order to prevent mismatches.[26] Considering the AmNo data discussed above, such a strategy of avoidance could explain the cases where the speaker omits functional suffixes. Take for instance the examples in (17), where the Norwegian definite suffix is omitted, but the phrase is still accompanied by a Norwegian determiner or demonstrative.

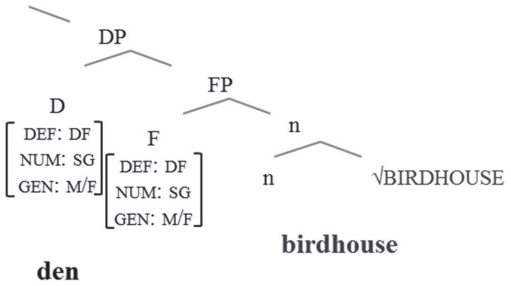
17.	a.	den that.M/F	**birdhouse_** birdhouse	[7]; coon_valley_WI_12gm
	b.	den the.M/F	store **building_** big building	[7]; chicago_IL_01gk
	c.	det the.N	første **trip_** first trip	[7]; vancouver_WA_01gm

These data may serve as evidence of the presence of an underlying structure even if the overt morphology is lacking, as argued by the MSIH. The argument follows from the assumption that the

[26] According to Gass and Selinker [78], avoidance is a typical phenomenon in L2 acquisition.

features in D are valued through a probe-goal relation with the features in F: the determiners in (17) vary according to gender, (17a,b) being either masculine or feminine, and (17c) neuter. As the valuation of the gender feature in D requires a corresponding gender feature in F, the gender feature must be specified in F, presumably together with number and definiteness, even if this feature bundle is not realized by a phonological exponent. A possible structure for (17a) is shown in (18). As in similar cases discussed above, the stem complex will move to F obligatorily, but because the functional exponent is avoided in this position, the stem will surface without a functional suffix.

18.



Nevertheless, the hypothesis does not necessarily entail that the speaker does not know the appropriate exponent at all. Such an approach would imply that the speaker never uses the functional suffix, which is easily tested by checking all relevant noun phrases produced by the speaker in question. A random check of the speakers who produced the examples in (17) shows that this implication is strongly questionable. These speakers do produce the definite suffixes in other similar phrases, suggesting that they do have this exponent in their list of vocabulary items. The realization, however, is variable, both in mixed and unmixed phrases, indicating that they are experiencing difficulties with the connection between the exponent and the features in the structure.

Furthermore, usage of the English plural *-s* can also be considered an effect of a similar avoidance strategy: the speaker avoids a potential mismatch, for instance with the gender feature in the Norwegian structure, by using an exponent from their dominant language. This is possible since the inserted exponent, given the Subset Principle, does not have to match *all* features in the feature bundle; matching with a subset is sufficient. In comparison, English does not have an alternative exponent to replace the definite suffix, leaving omission as the only available avoidance strategy. However, the English determiner *the* could be a replacement in cases when the speaker is unsure about which Norwegian determiner to insert.

The MSIH is therefore one potential approach to analyzing changing or diverging linguistic patterns, and incorporating it into the exoskeletal model provides an analysis like the one in (18). Reduced exposure to and practice in AmNo emerge as probable factors that could cause a reduced repertoire of functional exponents and increased uncertainty in how to the use them. However, a concern is that the MSIH lacks clear predictions as to where and how the missing inflections will take place, as well as clear restrictions in the model, and quite problematically, anything could potentially be explained as avoidance. In the next section, I will discuss the possibility that these diachronic changes are caused by certain changes in the syntactic structure.

6.2. Change in the Structure

The second scenario that could explain the observed changes, seen from an exoskeletal perspective, is that the structures themselves may be changing. This has been suggested in studies of other HL, e.g., heritage Russian [18,79], heritage Spanish [80], and heritage German [81], all of which conclude that the heritage language in question seems to have fundamentally different structures than its native

counterpart. Polinsky [18] suggests that the changes she finds between heritage and non-heritage speakers of Russian, and between children and adult HS, are the result of a structural reanalysis of the heritage grammar. She further contends that this is a process taking place over the lifespan of the HS in the absence of consistent input. In a similar vein, Putnam and Sánchez [19] argue in favor of a reanalysis of heritage grammars. In their analysis, the levels of activation for comprehension and production purposes play a crucial role in the development and maintenance of a heritage grammar; difficulties for HS can be due to reduced activation and availability of functional features, complicating the exercise of mapping them in the ways expected in monolingual variants of the language. The result may be a progressive reassembly of the features.

In the case of AmNo and the patterns of diachronic change in language mixing, the data suggest that such a structural reanalysis of grammar could be going on. In the exoskeletal model, this could take the form of features or feature bundles either being rearranged or erased from the structure, which would in turn have consequences for insertion of functional exponents. In many cases, a rearrangement of the structural outfit of the DP would mean that certain Norwegian functional exponents would not fit anymore. Given the Subset Principle, phonological exponents holding features not specified in the structure cannot be inserted, and supposing that the structural outfit of the noun phrase is changed, a Norwegian exponent could turn out to be "too specific", i.e., specified for features not present in the structure and thus blocked from insertion. In fact, changes in the structural composition of the noun phrase would instead allow, or even give preference to, insertion of English exponents.

As an example, consider the usage of the English plural *-s*. A couple of examples are given in (19).

19.	a.	mange	**lawyer-s**	[7]; sunburg_MN_03gm
		many	lawyer-PL	
	b.	fem	**dialect-s**	[7]; portland_ND_01gm
		five	dialect-PL	

Norwegian functional suffixes are typically also specified for gender in the plural, whereas associated words such as adjectives, quantifiers and determiners are not. The use of the English plural inflection could thus be seen as an indication that the representation of gender is diminished for the functional suffixes. If so, the Norwegian exponents for the suffix would be blocked from insertion due to holding a gender feature not specified in the structure, and the English exponent would be the preferred alternative (see [4] for discussion of the plural *-s* in AmNo).

Importantly, the development of reanalyzed structural patterns in a heritage language is described as a gradual process, potentially one where the dominating language gradually takes the place of the original structure, which is a typical trajectory in the development of a minority language, e.g., [19,79]. Lower exposure to lexical items in the heritage language means lower levels of activation of certain functional features. This, combined with an increased exposure to the dominant language, makes the features of the heritage language vulnerable for replacement [19]. In the case of AmNo, English has a dominating role both for the individual speakers and in the community at large. Hence, the possibility of English structures taking over for AmNo structures is not an unlikely scenario. This is also supported by the way in which some changes take form. For instance, the omission of Norwegian definite suffixes complies with an English structure where such suffixes do not exist, and the usage of English functional material suggests that the feature bundles in the structure are designed in such a way that these are considered the most appropriate exponents, following the Subset Principle.

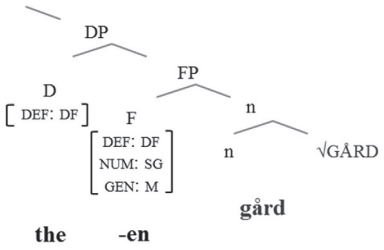
The gradual nature of the change is especially striking in the definite phrases with the English determiner *the*, where some patterns appear to be in an intermediate stage. As discussed in Section 5.2.3., half of the attested phrases occurred with both the English determiner and the Norwegian definite suffix, whereas the second half followed a typical English pattern realizing only the determiner. In the former group of these examples, one can argue that English influence is ongoing, allowing the insertion of an English determiner, but not yet complete, as the Norwegian functional suffix indicates

an underlying typical Norwegian feature bundle, spelled out by a Norwegian functional exponent. In the latter group, however, the influence of English is more pronounced as these examples follow a typical English DP pattern. In fact, as I will argue below, these examples may be described as English structures with Norwegian stems incorporated into them. Some examples are given in (20), where (20a,b) represent the intermediate stage and (20c,d) the potential full English influence in AmNo nominal structures:[27]

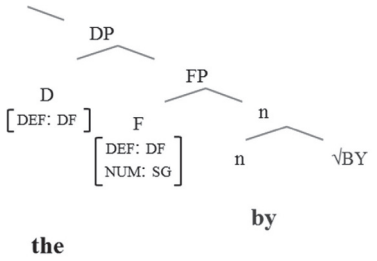
20.	a.	**the**	gård-en	[7]; gary_MN_01gm
		the	farm-DF.SG.M	
	b.	**the**	rest-en	[7]; vancouver_WA_03uk
		the	rest-DF.SG.M	
	c.	**the**	by	[7]; chicago_IL_01gk
		the	city	
	d.	**the**	ungdom	[7]; harmony_MN_01gk
		the	youth	

Possible structures for the two stages of mixing in (20) are shown in (21) and (22). The former shows the intermediate stage, represented by (20a), and the latter shows a case where a Norwegian noun is inserted into an English structure, as may be the case in (20c). Notice also that (21) and (22) show structures prior to movement, and that the stem complex will move to F.

21.



22.



[27] An alternative approach suggests that these are cases where the Norwegian determiner has been relexified by the English determiner *the*. Even though the process of relexification may be a considerable factor in language development and change, I argue that this is not plausible in these specific cases as a (Norwegian) determiner would not typically be expected in cases like (20).

In both (21) and (22), the feature bundle in D is reduced compared to its Norwegian counterpart, allowing the insertion of the English determiner. The main difference is found in F, where gender is presumably a key component. If a gender feature were present in the underlying structure, Norwegian functional exponents would be preferred over the English alternatives. However, if the structure has been reanalyzed and the gender feature is weakened, then the Norwegian exponent would be blocked, since inserting exponents specified for features other than those present in the structure would constitute a violation of the Subset Principle. Hence, I assume that the gender feature is preserved in F in the intermediate cases like (21), whereas in cases like (22), the displacement of the Norwegian structure by English has progressed further, eliminating the gender feature. The structure in (22) may now be considered an English structure where a Norwegian noun stem is inserted.

Parallel to the discussion in Section 6.1. above, positing changes in the syntactic structure is one possible approach to analyzing diachronic changes in AmNo. Considering its language environment over the past century, combined with the change going in a more English direction (e.g., without definite suffixes and with English functional exponents), it appears promising to analyze the changes as a structural reanalysis due to influence from English. Nevertheless, the two scenarios for change are not necessarily mutually exclusive. On the contrary, they may be two parallel trajectories to language change, and the observed change may be the result of a combination of the two.

Notice, however, that my discussion of changes in AmNo is based on a relatively limited sample of mixed noun phrases, and future expansions of CANS will bring new data, potentially corroborating the patterns discussed in this article. Studies of individual speakers and of diachronic changes in other domains of the grammar would also provide a clearer picture of the development and the potential impact of English in the structural reanalysis of AmNo.

6.3. The Nature of the Change

This article is primarily concerned with the explanation of the observed diachronic changes in AmNo as possible effects of changes in its grammar. However, a related question concerns the historical and sociolinguistic conditioning of these changes. As this is not the main focus of the current article, I will not go into an elaborate discussion of this question, but there are some crucial limitations to be addressed when investigating diachronic changes in a language like AmNo.

Cross-linguistic influence from the dominating language, English, has already been introduced and discussed in Section 6.2. In addition, changes in heritage grammars are often considered to be the result of incomplete acquisition or attrition [14,16–18,80]. Incomplete acquisition suggests that the HS, due to being introduced to the dominant language, experience a delay or break in the acquisition of the heritage language, hindering them from developing it in the same way as monolingual speakers of that variety [16,17]. Attrition, on the other hand, refers to a weakening or loss of linguistic competence that the speaker once mastered [14,18].

In the case of AmNo and its development over the past decades, there are some factors preventing us from determining which of these scenarios best reflect the linguistic situation. As already discussed in Section 3, speakers described in both Haugen's [6] material and in CANS [7] are descendants of immigrants who came to North America prior to 1920, and the corpora thus enable a comparison over a span of decades. The lack of (established) relationship between the speakers in the two corpora, also discussed in Section 3, is nevertheless a limiting factor. In order to study an effect of incomplete acquisition carefully, one needs data about the input of the learner, and a study of attrition requires data from the early production of the speaker, neither of which are available from the two corpora under consideration here. Since there is no established relationship between the speakers in the two corpora, we cannot study the younger speakers' input to evaluate their acquisition. Moreover, as the speakers in CANS were already adults and elderly people at the time of recording, we are unable to determine whether their grammars have been stable throughout their lives or if they have lost linguistic skills due to attrition.

Also, in order to properly investigate cross-linguistic influence, more data would be required, documenting for instance the speakers' competence in their dominant language. This is not provided by any of the corpora. Nevertheless, as discussed above, the dominance of English both in the individual speakers and in the larger language community suggests that the speakers of AmNo would be experiencing an influence from English to some degree.

Comparing the two corpora thus means studying HL in a retrospective fashion. On the one hand, this enables a study of different stages in its development, but at the same time potential conclusions are limited due to the lack of a relation or direct link between them. As already mentioned in Section 3, however, based on the speakers' ages and places of origin, we can argue that the speakers from the 1930s and 1940s represent the type of input that the CANS speakers received, and thus establish an indirect link between them. In addition, as the speakers in Haugen [6] in the vast majority of cases used loanwords with the appropriate Norwegian inflection, the diverging patterns attested in CANS can in fact be considered a diachronic development in AmNo. Nevertheless, as the developmental trajectory cannot be traced for the individual speakers, these diachronic changes are best described as tendencies of change in the language community.

7. Conclusions

This article has investigated the heritage language AmNo and whether its patterns of language mixing are persistent over time. A comparison of mixed AmNo noun phrases from the 1930s and 1940s [6] and the present [7] shows that the overall pattern of language mixing is stable, but some systematic diachronic changes are attested. The purpose of this article has thus been twofold: first, to describe the changes, focusing on changes in the exponence of number and definiteness, and second, to explore potential changes in the underlying grammar which could explain the observed changes.

The diachronic changes in the categories number and definiteness can be summed up as two main patterns: omission of functional exponents and usage of English functional exponents such as the plural suffix *-s* and the determiner *the*. These patterns are studied based on an exoskeletal approach to grammar where the main component is a separation of the abstract, syntactic structure and the phonological exponents realizing it. The article then discusses two possible scenarios for how to account for the observed changes: they could be due to changes in the phonological exponents, i.e., their conditions for insertion into the syntactic structure, or they could be due to a change in the underlying syntactic structure itself. Both alternatives would disrupt the process of insertion, and they are not necessarily mutually exclusive. However, based on the observed patterns of change in the data, I argue that a structural reanalysis of AmNo grammar is occurring. Moreover, the article also discusses why one should be careful when drawing conclusions concerning diachronic changes in AmNo. Although the two corpora under consideration provide valuable insights into AmNo at two different stages in its development, they are nevertheless not directly connected in terms of family relations between speakers, and the nature of the changes is therefore not easily determined. Future studies of individual competences, however, will presumably provide more knowledge of changes in the underlying grammars, and how they can explain the observed patterns.

In a broader context, the present article shows that the patterns of language mixing are stable over time, although not completely resistant to change. The observed changes in AmNo support this analysis, suggesting that (heritage) grammars may change under conditions of reduced input and activation. This takes place as a gradual reanalysis of the structures under the influence of the dominant language, and may be reinforced by a diminishing repertoire of functional exponents. Moreover, to reach insights into the nature of change in heritage grammars, input, competence, and gradual change should be sufficiently documented and taken into consideration.

Acknowledgments: I would like to thank three anonymous reviewers for very helpful comments and suggestions. I am also grateful to Tor Anders Åfarli (Norwegian University of Science and Technology, Norway) and Terje Lohndal (Norwegian University of Science and Technology & UiT The Artic University of Norway, Norway) for valuable discussions as well as detailed and constructive comments.

Conflicts of Interest: The author declares no conflict of interest.

Abbreviations

Glossary of Linguistic Codes Used in the Glosses

DEF	Definiteness
DF	Definite
F	Feminine
GEN	Gender
INDF	Indefinite
M	Masculine
N	Neuter
NUM	Number
PL	Plural
SG	Singular

References

1. Grimstad, M.B.; Lohndal, T.; Åfarli, T.A. Language mixing and exoskeletal theory: A case study of word-internal mixing in American Norwegian. *Nordlyd* **2014**, *41*, 213–237. [CrossRef]
2. Alexiadou, A.; Lohndal, T.; Åfarli, T.A.; Grimstad, M.B. Language Mixing: A Distributed Morphology approach. In *NELS 45, Proceedings of the Forty-Fifth Annual Meeting of the North East Linguistic Society, MIT, Cambridge, MA, USA, 31 October–2 November 2014*; Bui, Ö., Ed.; GLSA: Amherst, MA, USA, 2015; pp. 25–38.
3. Åfarli, T.A. Hybrid verb forms in American Norwegian and the analysis of the syntactic relation between the verb and its tense. In *German Heritage Languages in North America*; Johannessen, J.B., Salmons, J.C., Eds.; John Benjamins Publishing Company: Amsterdam, The Netherlands; Philadelphia, PA, USA, 2015; pp. 161–177.
4. Riksem, B.R. Language Mixing in American Norwegian Noun Phrases. *J. Lang. Contact* **2017**, in press.
5. Riksem, B.R.; Grimstad, M.B.; Lohndal, T.; Åfarli, T.A. Language mixing within verbs and nouns in American Norwegian. *J. Comp. Ger. Linguist.* **2017**, in press.
6. Haugen, E. *The Norwegian Language in America: A Study in Bilingual Behavior*; University of Philadelphia Press: Philadelphia, PA, USA, 1953; Volumes I–II.
7. Johannessen, J.B. The Corpus of American Norwegian Speech (CANS). In *NEALT Proceedings Series Vol. 23, Proceedings of the 20th Nordic Conference of Computational Linguistics (NoDaLiDa 2015), Institute of the Lithuanian Language, Vilnius, Lithuania, 11–13 May 2015*; Megyesi, B., Ed.; ACL Anthology: Stockholm, Sweden, 2015; pp. 297–300.
8. Hjelde, A. *Trøndsk talemål i Amerika*; Tapir: Trondheim, Norway, 1992.
9. Johannessen, J.B.; Salmons, J. Innledning. *Norsk Lingvistisk Tidsskrift* **2012**, *30*, 139–148.
10. Johannessen, J.B.; Salmons, J. The study of Germanic heritage languages in the Americas. In *German Heritage Languages in North America*; Johannessen, J.B., Salmons, J.C., Eds.; John Benjamins Publishing Company: Amsterdam, The Netherlands; Philadelphia, PA, USA, 2015; pp. 1–17.
11. Lovoll, O.S. *Norwegian Newsarticles in America*; Minnesota Historical Society Press: St. Paul, MN, USA, 2010.
12. Rothman, J. Understanding the nature and outcomes of early bilingualism: Romance languages as heritage languages. *Int. J. Biling.* **2009**, *13*, 155–163. [CrossRef]
13. Rothman, J.; Treffers-Daller, J. A prolegomenon to the construct of the native speaker: Heritage speaker bilinguals are natives too! *Appl. Ling.* **2014**, *35*, 93–98. [CrossRef]
14. Pascual y Cabo, D.; Rothman, J. The (Il)logical problem of heritage speaker bilingualism and incomplete acquisition. *Appl. Ling.* **2012**, *33*, 1–7.
15. Benmamoun, E.; Montrul, S.; Polinsky, M. Heritage languages and their speakers: Opportunities and challenges for linguistics. *Theor. Ling.* **2013**, *39*, 129–181. [CrossRef]
16. Montrul, S. *Incomplete Acquisition in Bilingualism. Re-examining the Age Factor*; John Benjamins Publishing Company: Amsterdam, The Netherlands; Philadelphia, PA, USA, 2008.
17. Polinsky, M. Incomplete acquisition: American Russian. *J. Slav. Ling.* **2006**, *14*, 191–262.

18. Polinsky, M. Reanalysis in adult heritage language: A case for attrition. *Stud. Second Lang. Acquis.* **2011**, *33*, 305–328. [CrossRef]
19. Putnam, M.T.; Sánchez, L. What's so incomplete about incomplete acquisition? A prolegomenon to modeling heritage language grammars. *Ling. Approaches Biling.* **2013**, *3*, 478–508. [CrossRef]
20. Flaten, N. Notes on the American-Norwegian with vocabulary. *Dialect Notes* **1900–1904**, *2*, 115–126.
21. Flom, G.T. English elements in Norse dialects of Utica, Wisconsin. *Dialect Notes* **1900–1904**, *2*, 257–268.
22. Flom, G.T. English loanwords in American Norwegian as spoken in the Koshkonong settlement, Wisconsin. *Am. Speech* **1926**, *1*, 541–558. [CrossRef]
23. Seip, D.A.; Selmer, E.W. Seip og Selmers voksrull-opptak av norsk-amerikanske informanter. Available online: http://www.tekstlab.uio.no/nota/NorDiaSyn/dialekt_seip_og_selmer.html (accessed on 6 April 2017), audio files with Norwegian dialects.
24. Haugen, E. Einar Haugens opptak av norskamerikanere i 1935–1948. Available online: http://tekstlab.uio.no/norskiamerika/opptak/haugen.html (accessed on 6 April 2017), audio files with Norwegian dialects.
25. Hjelde, A. "Folkan mine, dæm bære snakke norsk"—norsk i Wisconsin frå 1940-talet og fram til i dag. *Norsk Lingvistisk Tidsskrift* **2012**, *30*, 183–203.
26. Venås, K. On the choice between two written standards in Norway. In *Language Conflict and Language Planning*; Jahr, E.H., Ed.; Mouton de Gruyter: Berlin, Germany, 1993; pp. 263–278.
27. Vikør, L.S. *The Nordic Languages: Their Status and Interrelations*; Nordic Language Secretariat Publication No. 14; Novus Press: Oslo, Norway, 1995.
28. Muysken, P. *Bilingual Speech*; Cambridge University Press: Cambridge, UK, 2000.
29. Myers-Scotton, C. *Duelling Languages: Grammatical Structure in Code Switching*; Oxford University Press: Oxford, UK, 1993.
30. Myers-Scotton, C. *Contact Linguistics: Bilingual Encounters and Grammatical Outcomes*; Oxford University Press: Oxford, UK, 2002.
31. Mahootian, S. A Null Theory of Code-Switching. Ph.D. Dissertation, Northwestern University, Evanston, IL, USA, 1993.
32. Belazi, H.M.; Rubin, E.J.; Toribio, A.J. Code switching and X-bar theory. *Ling. Inq.* **1994**, *25*, 221–237.
33. MacSwan, J. *A Minimalist Approach to Intrasentential Code Switching*; Garland Press: New York, NY, USA, 1999.
34. MacSwan, J. The architecture of the bilingual faculty: Evidence from intrasentential code switching. *Bilingualism* **2000**, *3*, 37–54. [CrossRef]
35. MacSwan, J. Codeswitching and generative grammar: A critique of the MLF model and some remarks on "modified minimalism". *Biling. Lang. Cognit.* **2005**, *8*, 1–22. [CrossRef]
36. MacSwan, J. Generative approaches to code-switching. In *The Cambridge Handbook of Linguistic Code-switching*; Bullock, B.E., Toribio, A.J., Eds.; Cambridge University Press: Cambridge, UK, 2009; pp. 309–335.
37. MacSwan, J. Programs and proposals in codeswitching research: Unconstraining theories of bilingual language mixing. In *Grammatical Theory and Bilingual Codeswitching*; MacSwan, J., Ed.; MIT Press: Cambridge, MA, USA, 2014; pp. 1–33.
38. van Hout, A. Event Semantics of Verb Frame Alternations. Ph.D. Dissertation, Tilburg University, Tilburg, The Netherlands, 1996.
39. Marantz, A. No escape from syntax: Do not try morphological analysis in the privacy of your own lexicon. In *University of Pennsylvania Working Papers in Linguistics, Proceedings of the 21st Annual Penn Linguistics Colloquium, University of Pennsylvania, Philadelphia, PA, USA, 22–23 February 1997*; Dimitriadis, S., Surek-Clark, W., Eds.; Penn Linguistics Club: Philadelphia, PA, USA, 1997; pp. 201–225.
40. Marantz, A. Verbal argument structure: Events and participants. *Lingua* **2013**, *130*, 152–168. [CrossRef]
41. Borer, H. *Structuring Sense I: In Name Only*; Oxford University Press: Oxford, UK, 2005.
42. Borer, H. *Structuring Sense II: The Normal Course of Events*; Oxford University Press: Oxford, UK, 2005.
43. Borer, H. *Structuring Sense III: Taking Form*; Oxford University Press: Oxford, UK, 2013.
44. Åfarli, T.A. Do verbs have argument structure? In *Argument Structure*; Reuland, E., Bhattacharya, T., Spathas, G., Eds.; John Benjamins Publishing Company: Amsterdam, The Netherlands; Philadelphia, PA, USA, 2007; pp. 1–16.
45. Ramchand, G. *Verb Meaning and the Lexicon: A First Phase Syntax*; Cambridge University Press: Cambridge, UK, 2008.

46. Lohndal, T. Without Specifiers: Phrase Structure and Argument Structure. Ph.D. Dissertation, University of Maryland, College Park, MD, USA, 2012.
47. Lohndal, T. *Phrase Structure and Argument Structure: A Case-Study of the Syntax Semantics Interface*; Oxford University Press: Oxford, UK, 2014.
48. Alexiadou, A.; Anagnostopoulou, E.; Schäfer, F. *External Arguments in Transitivity Alternations: A Layering Approach*; Oxford University Press: Oxford, UK, 2015.
49. Harley, H.; Noyer, R. Distributed Morphology. *Glot Int.* **1999**, *4*, 3–9.
50. Alexiadou, A. *Functional Structure in Nominals: Nominalization and Ergativity*; John Benjamins Publishing Company: Amsterdam, The Netherlands; Philadelphia, PA, USA, 2001.
51. Embick, D.; Noyer, R. Distributed morphology and the syntax morphology interface. In *The Oxford Handbook of Linguistic Interfaces*; Ramchand, G., Reiss, C., Eds.; Oxford University Press: Oxford, UK, 2007; pp. 289–324.
52. Harley, H. On the identity of roots. *Theor. Ling.* **2014**, *40*, 225–276. [CrossRef]
53. Arad, M. *Roots and Patterns: Hebrew Morpho-Syntax*; Springer: Dordrecht, The Netherlands, 2005.
54. Pylkkänen, L. *Introducing Argument*; MIT Press: Cambridge, MA, USA, 2008.
55. Embick, D.; Marantz, A. Architecture and blocking. *Ling. Inq.* **2008**, *39*, 1–53. [CrossRef]
56. Halle, M. Distributed morphology: Impoverishment and fission. In *MIT Working Papers in Linguistics 30. PF: Articles at the Interface*; Bruening, B., Kang, Y., McGinnis, M., Eds.; MIT: Cambridge, MA, USA, 1997; pp. 425–449.
57. Goldrick, M.; Putnam, M.; Schwarz, L. Coactivation in bilingual grammars: A computational account of code mixing. *Biling. Lang. Cognit.* **2016**, *19*, 857–876. [CrossRef]
58. Goldrick, M.; Putnam, M.; Schwarz, L. The future of code mixing research: Integrating psycholinguistics and formal grammatical theories. *Biling. Lang. Cognit.* **2016**, *19*, 903–906. [CrossRef]
59. Embick, D. *The Morpheme: A Theoretical Introduction*; Mouton de Gruyter: Berlin, Germany; Boston, MA, USA, 2015.
60. Delsing, L.-O. The Internal Structure of Noun Phrases in the Scandinavian Languages. A Comparative Study. Ph.D. Dissertation, University of Lund, Lund, Sweden, 1993.
61. Vangsnes, Ø.A. The Identification of Functional Architecture. Ph.D. Dissertation, University of Bergen, Bergen, Norway, 1999.
62. Julien, M. *Nominal Phrases from a Scandinavian Perspective*; John Benjamins Publishing Company: Amsterdam, The Netherlands; Philadelphia, PA, USA, 2005.
63. Nygård, M.; Åfarli, T.A. On the structure of gender assignment. *Indian Ling.* **2015**, *76*, 67–76.
64. Kramer, R. Gender in Amharic: A morphosyntactic approach to natural and grammatical gender. *Lang. Sci.* **2014**, *43*, 102–115. [CrossRef]
65. Alexiadou, A. Inflection class, gender and DP-internal structure. In *Explorations in Nominal Inflection*; Müller, G., Gunkel, L., Zifonun, G., Eds.; Mouton de Gruyter: Berlin, Germany, 2004; pp. 21–50.
66. Myklebust, A. "Hva er de derre greiene der?": En syntaktisk analyse av komplekse demonstrativ i muntlig norsk. Master's Thesis, Norwegian University of Science and Technology, Trondheim, Norway, 2012.
67. Enger, H.-O. On the relation between gender and declension—A diachronic perspective from Norwegian. *Stud. Lang.* **2004**, *28*, 51–82. [CrossRef]
68. Lødrup, H. Hvor mange genus er det i Oslo-dialekten? *Maal og Minne* **2011**, *2*, 120–136.
69. Trosterud, T. Genus i norsk er regelstyrt. *Norsk lingvistisk tidsskrift* **2001**, *19*, 29–58.
70. Flom, G.T. The gender of English loan-nouns in Norse dialects in America; A contribution to the study of the development of grammatical gender. *J. Eng. Germ. Philol.* **1903**, *5*, 1–31.
71. Hjelde, A. The gender of English nouns in American Norwegian. In *Language Contact across the Atlantic*; Ureland, P.S., Clarkson, I., Eds.; Max Niemeyer Verlag: Tübingen, Germany, 1996; pp. 297–312.
72. Johannessen, J.B.; Larsson, I. Complexity matters: On gender agreement in heritage Scandinavian. *Front. Psychol.* **2015**, *6*, 1842. [CrossRef] [PubMed]
73. Lohndal, T.; Westergaard, M. Grammatical gender in American Norwegian heritage language: Stability or attrition? *Front. Psychol.* **2016**, *7*. [CrossRef] [PubMed]
74. Picallo, M.C. Gender and number in Romance. *Lingue e Linguaggio* **2008**, *VII*, 47–66.
75. Enger, H.-O. The role of core and non-core semantic rules in gender assignment. *Lingua* **2009**, *119*, 1281–1299. [CrossRef]
76. Lardiere, D. Mapping features to forms in second language acquisition. In *Second Language Acquisition and Linguistic Theory*; Archibald, J., Ed.; Blackwell: Malden, MA, USA; Oxford, UK, 2000; pp. 102–129.

77. Prévost, P.; White, L. Missing surface inflection or impairment in second language acquisition? Evidence from tense and agreement. *Second Lang. Res.* **2000**, *16*, 103–133. [CrossRef]
78. Gass, S.M.; Selinker, L. *Second Language Acquisition. An Introductory Course*, 3rd ed.; Routledge: New York, NY, USA, 2008.
79. Polinsky, M. Structure vs. use in heritage language. *Ling. Vanguard* **2016**, *2*. [CrossRef]
80. Scontras, G.; Fuchs, Z.; Polinsky, M. Heritage language and linguistic theory. *Front. Psychol.* **2015**, *6*. [CrossRef] [PubMed]
81. Yager, L.; Hellmold, N.; Joo, H.-A.; Putnam, M.T.; Rossi, E.; Stafford, C.; Salmons, J. New structural patterns in moribund grammar: Case marking in heritage German. *Front. Psychol.* **2015**, *6*, 1716. [CrossRef] [PubMed]