CASE AND GENDER LOSS IN GERMANIC, ROMANCE, AND BALKAN SPRACHBUND LANGUAGES

by

MOFAREH D. ALHAZMI

A DISSERTATION

Presented to the Department of Linguistics
and the Division of Graduate Studies of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

December 2022

DISSERTATION APPROVAL PAGE

Student: Mofareh D. Alhazmi

Title: Case and Gender Loss in Germanic, Romance, and Balkan Sprachbund Languages

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Linguistics by:

Cynthia Vakareliyska	Chairperson
Vsevolod Kapatsinski	Core Member
Tyler Kendall	Core Member
Martha Bayless	Institutional Representative

and

Krista Chronister	Vice Provost for Graduate Studies

Original approval signatures are on file with the University of Oregon Division of Graduate Studies.

Degree awarded December 2022.

DISSERTATION ABSTRACT

Mofareh D. Alhazmi
Doctor of Philosophy
Department of Linguistics
December 2022
Title: Case and Gender Loss in Germanic, Romance, and Balkan Sprachbund Languages

My dissertation investigates the loss of morphological case and grammatical gender in the Germanic, Romance, and Balkan Sprachbund languages. Crucial language-internal and language-external motivations are considered. To illustrate the changes of morphological cases, the languages are divided into historical stages. Every change in nominal inflection between stages is attributed to either sound change or analogical change; these choices are justified through consideration of historical sound changes and the motivations behind analogical processes. The changes are also discussed in terms of their effects on number syncretism, case and gender mergers, order of case loss, and the relationship between gender and declension.

These motivations can be classified as language-internal or language-external. Phonological, morphosyntactic, and semantic factors are among the former. Different types of sound change can neutralize inflection differences, but two closely related types, prosodic change, and vowel reduction have been suggested as key causes in case and gender loss in IE languages. A usual direction of change in morphological case loss includes variation between two or more cases in one or more functions, followed by functional narrowing and occasionally a complete functional merger of the case markings. Similarly, there can be differences between a case and an analytic construction, which can lead to the former being replaced by the latter in some or all functions.-External motivations for case and gender loss include the kinds of contact conditions that cause or accelerate simplification in internal developments. Essential contact situation is the establishment of a sprachbund, or linguistic region, which usually entails structural convergence among surrounding languages during a long period of profound contact. Interactions among number, case, and gender are analyzed using original quantitative measures of number syncretism on nouns and gender syncretism on agreement targets. Overall, the results of my study support the general hypothesis that the loss of case and gender categories can be

explained by the neutralization of distinctions in these categories as a direct result of sound change and by the profiling of a more relevant category through analogical processes.

This dissertation includes previously unpublished material.

ACKNOWLEDGMENTS

I would like to express my deepest appreciation to my advisor Professor Cynthia Vakareliyska for her constant support and guidance throughout the writing process of my dissertation. Her assistance makes this dissertation possible. I could not have undertaken this journey without her assistance.

I also would like to express my gratitude to my committee members: Professor Vsevolod Kapatsinski for providing invaluable suggestions that helped guide this entire dissertation; Professor Tyler Kendall for his helpful advice and contribution on ways to strengthen my research, and Professor Martha Bayless for her insightful comments and suggestions.

I would like to thank my late parents. They always supported me when they were alive; unfortunately, they cannot see me graduate. I also would like to express my thanks to my wife, and my children for their support and patience; they put up with me and missed a lot of events. Special thanks to my brothers for their help and encouragement.

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

CHAPTER I
INTRODUCTION

1.1. The Problem

Varying degrees of morphological case loss are present in just about all[1] of the descendants of Proto-Indo-European (PIE). No modern Indo-European (IE) language preserves all eight of the original PIE cases. The inflectional morphology of a noun in IE languages shows a continuous shift over time in several families. Throughout the languages' reconstructable history, most IE families have undergone morphological changes from synthetic to analytic. The paradigm of singular, dual, and plural numbers has been reduced to singular and plural, with the dual number being eliminated or syncretized with the plural. Some IE languages and language families have experienced a much greater loss of case inflections than others. The Germanic language family has generally experienced a relatively high degree of case loss; a few of these languages maintain up to four morphological cases, but most standard varieties have lost morphological case on nouns entirely. Many of these have also undergone an accompanying reduction in the number of grammatical genders. Swedish, Danish, and Dutch experienced a merger of masculine and feminine into a common gender. The Romance language family has also experienced significant morphological case loss on nouns, in most languages a total loss of case inflections, as well as changes to the gender system. The Western Romance languages lost the neuter, merging with the masculine. Finally, the Balkan Sprachbund languages have undergone certain common developments, often differing from those in related languages outside the Sprachbund. Some of these developments involved the loss of morphological case distinctions, but most Balkan Sprachbund languages still retain some case marking on nouns.

1.2. Scope

To my knowledge, my study is the first comprehensive comparison of case and gender loss and simplification in 18 IE languages in a single study. Earlier efforts have been confined to a single branch or a few languages. A family of closely related languages with vastly different outcomes concerning case and gender is the Germanic. This family is traditionally divided into three branches: North, West, and East; all descend from Proto-Germanic (PGmc). For PGmc

[1] Sanskrit maintained all eight cases normally reconstructed for PIE.

languages, my investigation is limited to a collection of North and West Germanic languages. The North Germanic languages all descend from Old Norse (ON). ON and its descendants can be divided into Western and Eastern branches. The oldest well-attested varieties in the Western branch of ON include Old Icelandic (OIc), which is often simply referred to as ON, and Old Norwegian (ONw). The oldest attested varieties in the Eastern branch are Old Swedish (OSw) and Old Danish (ODan), referred to together as Old Scandinavian (OSc). The West Germanic languages include Old High German (OHG), Old Saxon (OS), Old Dutch (ODu), also known as Old Low Franconian, Old English (OE), and their descendants.

In contrast to the Germanic languages, most of the Romance languages have developed very similarly in terms of case and gender, with the notable exception of Romanian. These languages are the descendants of Vulgar Latin (VL), the language spoken by the common people of the Roman Empire and many of its former territories, while Classical Latin (CL) was the literary form based on an earlier stage of the language. The major modern Romance languages are Portuguese, Spanish, French, Italian, and Romanian. Romance varieties can be categorized in several ways. Western Romance refers to Portuguese, Spanish, French, and related varieties, as well as varieties of Italian. This term excludes Romanian and related Romance languages varieties (Aromanian, etc.) which is a member of the Balkan Sprachbund, a group of geographically close languages that vary in their genetic relationships but have developed certain common grammatical features. As commonly defined, the Balkan Sprachbund consists of modern[2] Bulgarian, Macedonian, Romanian, Greek, and Albanian. Balkan Romani, Balkan Turkish, and Bosnian-Croatian-Montenegrin-Serbian (BCMS), or at least its southeastern dialects, which are sometimes also considered Balkan Sprachbund languages, have not undergone significant case loss.

[2] The term *modern* is used to refer to the most recent major division in a language's development. This period generally spans around 500 years and continues until the present. For some language such Slavic languages, Romanian and Albanian *modern* starts between 17th and 19th centuries. The term *contemporary* refers to a language as it is used today; this term is usually used in combination with *standard* because it is difficult to accurately represent the various ongoing developments in contemporary non-standard varieties. However, non-standard varieties are discussed where they help to clarify developments in the standard or otherwise provide insight that cannot be gained from the standard alone.

1.3. Purpose of the Study

My study aims to draw conclusions behind case and gender loss in the Indo-European subfamilies under investigation. In order to achieve this, numerous motivations have been proposed to account for the loss of case and gender inflections both within specific languages and across many different languages. These motivations can be grouped into language-internal and language-external factors. The former includes phonological, morphosyntactic, and semantic factors. Various types of sound change can neutralize the distinctions between inflections, but two closely related types, prosodic change, and vowel reduction, have been proposed as major factors in case and gender loss in IE languages. Other internal motivations involve the interaction among morphology, syntax, and semantics. Markedness, relevance, analogy, and syncretism are all factors in paradigm organization, so they also relate to case and gender loss, as does the merging of declension classes. Another proposed motivation is the grammaticalization of demonstratives into definite articles; this process appears to be associated with case loss in some languages but to help preserve case distinctions in others. A common path of development in the loss of a morphological case involves variation between two or more cases in one or more functions, followed by a functional narrowing of one case at the expense of another, and sometimes a full functional merger of the cases. Similarly, variation can occur between a case and an analytic construction, often leading to the replacement of the former by the latter in some or all functions. Semantic and syntactic motivations are discussed for these developments, such as overlapping functions and changes in frequency, but they may instead be consequences of changes in form that result from phonological or analogical processes.

External motivations for case and gender loss concern the types of contact situations that give rise to or accelerate internal processes discussed above, particularly those relating to morphosyntax and semantics, but also sound change. One important contact situation is dialect contact, which can lead to simplification through semicommunication or koineization; more generally, it allows innovations in one dialect to spread to others, even if they lack the underlying motivation for the change. Another is the formation of a sprachbund, or linguistic area, which generally involves structural convergence among neighboring languages over an extended period of intense contact; structural convergence can also occur outside of sprachbunds. Proposed internal and external motivations are evaluated and extended in chapter III, along with proposals for how these factors interact.

1.4. Research Question

The main study considers a wide breadth of motivations, but the focus of the investigation is phonological and morphosyntactic factors. The research question is the extent to which these two factors can account for the degree to which case and gender categories are lost in a selection of Germanic, Romance, and Balkan Sprachbund languages.

1.5. Definition of Terms

1.5.1. Morphological Case

My study is primarily concerned with the morphological case. Vakareliyska (1994:7) defines case as a semantic property of verbs that is expressed morphologically in Indo-European (IE) languages as a marking on noun phrases (NPs) governed by the verbs. According to this approach, semantic case is not a syntactic relationship between NPs, but a set of physical or psychological relationships that are perceived among the referents of the NPs (ibid., 8). The distinction between case and case marking treats case as a cognitive phenomenon that exists whether or not it is morphologically expressed. By this definition, a case has been lost when its formal expression has completely merged with another case. This does not necessarily mean that the function expressed by that form has disappeared from the language. In most instances, the functions of a lost case are either taken on by one or more other cases or expressed by analytic means such as word order, adpositions, or clitics. In the languages under investigation, some case distinctions on pronouns are always retained even when all case distinctions have been lost on nouns. My investigation focuses on case marking on nouns and other elements of the NP.

1.5.2. Grammatical Gender

A concept closely connected to morphological case marking is that of grammatical gender. Grammatical gender is a system of noun classes assigned based on common features of their constituents and reflected through the agreement of syntactically associated words (see, for example, Hockett 1958:231, Corbett 1991:1, 4). The features that determine the grammatical gender of a noun can be semantic, morphological, or phonological (Corbett 1991:3). The role of grammatical gender in a language often goes hand-in-hand with case because of portmanteau gender/case inflections; hence the two features are often correlated in the deterioration or strengthening of either of their roles. This is particularly true when morphological criteria are

prominent in gender assignment, as in most IE languages: when the markers distinguishing different declensional paradigms are lost, gender is no longer overt on the lexical level. The same loss of markers on agreement targets such as adjectives and articles results in the loss of gender agreement itself, and without an agreement, there is no grammatical gender (ibid., 315).

1.5.3. Declensional Paradigm

A declensional paradigm is a structured collection of inflectionally connected forms that comprise syntactically relevant and/or syntactically driven morphology. Forms themselves can be organized into a series of cells based on a specific root they to which they belong. The information of the cells can be thought of as a set of features that include, but are not limited to, case, number, and gender (see Blake 2004:203, Joseph 2009:49).

1.5.4. Profiling/Relevance

Profiling a category makes the markers of a category more salient, particularly those associated with its marked values; in syntagmatic profiling, they become longer, more visible, or spread to the root, as in umlaut; in paradigmatic profiling, more allomorphs arise to mark the relevant category. In both types, marked plurals are more salient in opposition to unmarked singular Kürschner and Nübling (2011:362). Both types apply to the profiling of number marking in the Germanic, Romance, and Balkan Sprachbund languages. Inherited syncretism, sound change, and analogical processes all contribute to the neutralization of distinctions within these morphological categories in specific paradigms. However, these categories vary in semantic relevance to the nouns and agreement targets on which they occur and, therefore, in the degree to which these neutralizations are tolerated. Thus, analogical change generally increases the salience of the more semantically relevant categories at the expense of the less relevant categories (see Bybee 1985).

1.5.5. Markedness

Markedness concerns specific measures that include a feature or a category that is unmarked in regard to another. The use of a sign in a particular category makes this category marked in apposition to an unmarked one. The marked features can be observable paradigmatically or

syntagmatically, as with marking plural in many IE languages (see Gvozdanović 1989:59, Tiersma 1982:832). My study is limited to this notion of markedness.

1.5.6. Functional Narrowing

Functional narrowing is another conceivable effect of case variation that causes the loss of one or more case markings. The idea of functional narrowing is that a category loses its functions in favor of another category. Barðdal (2008:166) highlights its associations with the more general notion of productivity. She attributes the effects of productivity to the type frequency, semantic coherence, and the interaction between the two. Meiser (1992:195) refers to this evolution as a shift of extensions in reference to the changing semantic extension, or set of functions, of both cases involved. This phenomenon is referred to as functional narrowing in my study.

1.5.7. Formal and Functional Syncretism

There are various ways to measure the functionality of an inflectional system. One of the most basic is the degree of syncretism, and this will be the primary measure of the interaction among categories and their functionality in my study. Except where noted, my study adopts Baerman et al.'s (2005:2-4) definition of *syncretism* as the failure to mark a syntactically relevant distinction under specific morphological conditions. Different types of syncretism can be distinguished based on the features that are neutralized, i.e., case, number, and gender. Meiser (1992) provides an informative overview of several important concepts related to syncretism and analogical changes to inflectional paradigms. He uses the term *syncretism* to refer to the functional merging of two (or more) categories, i.e., a complete neutralization of morphological distinctions between them. This means that the functional distinction between the morphemes is lost. Several other authors who use the term in this way are cited in my study, e.g., Luraghi (1987).

Wackernagel (1920:302-304) proposed three motivations for functional mergers: paradigmatic inconsistencies, sound change, and (partial) semantic overlap. Meiser argues that sound change alone rarely leads to a functional merger; more often, sound change creates inconsistencies in the marking of categories among different morphological classes, and these spread as forms in other paradigms are leveled by analogy. In addition, he argues that a functional merger must accompany this formal merger for two categories to be considered merged and that functional mergers can arise through semantic overlap, with an intermediate step

of some degree of free variation, or when one category takes over functions that are losing productivity for another category as the latter is reduced to relic functions and sometimes, subsequently lost. When a category is lost, it is usually compensated for by content words, function words, or word order (1992:191-196). Sometimes this compensation is optional, but these analytic means must exist before the category is lost, i.e., as a competing means of expression or reinforcement to the category (see Plank 1979:619). The morphosyntactic consequences of functional mergers are also considered in my study. When categories merge, only the function of the marked category is generally compensated for. For example, when the Dat and Loc merged in the development of Greek and Germanic languages, prepositions compensated for the Loc function, but the inherited Dat functions remained without compensation.

1.5.8. Refunctionalization/Exaptation

Various outcomes can apply to the morphological markers of two categories that merge. All of the forms of one category can survive in the functions of both, a mix of forms can survive in different paradigms, and more rarely, forms of both can continue to be used in free variation. If only the forms of one category are used for the combined functions, the forms of the other category are sometimes refunctionalized. Refunctionalization (or exaptation) is when a form that is about to disappear or used with other meanings is repurposed to express a new function (see Lass 1997:316-324, Pato 2018:23).

1.5.9. Principles of Selection

If a mix of forms survives, the choice of particular forms in different morphological contexts can be attributed to several tendencies or principles of selection: the iconic principle, system adequacy, and distinctive strength. These principles evolved from natural morphology, designating the connections within paradigms. These principles are a response to inconsistencies in declension classes caused by accidental changes as well as functional mergers. The iconic principle is the preference for more salient forms in more marked categories to mark a category semantically (Meiser 1992:198-208). The principle of system adequacy describes a preference for consistent paradigm organization, e.g., consistently marking a category with suffixes, not a mix of suffixes, prefixes, and/or stem-inflection. Finally, the principle of distinctive strength is

the increased chance of survival for an overstable marker, i.e., a phonetically distinctive marker in one morphological class that tends to spread to other classes. (Wurzel 1984:87, 139, 209). When a system is disrupted by sound change, any remaining forms are refunctionalized to involve developments such as profiling more relevant categories as well as using principles of selection. Otherwise, the forms fall into disuse, depending on the particular conditions.

1.5.10. Grammaticalization

Heine and Reh (1984:15) define grammaticalization as a development whereby language components lose semantic content, pragmatic value, syntactic flexibility, and phonetic content. (Aristar 1997:385) states that when form and meaning are not suited, a common reanalysis includes the extension of functions from one morpheme to another. Campbell (2013:281-283) provides examples of grammaticalization: phonological reduction, such as vowel reduction since the reduced forms no longer represent their lexical meaning. The use of prepositions instead of case marking is also a common process. Other mechanisms discussed in my dissertation, such as functional narrowing, refunctionlization, and the use of principles of selection, can be seen as a grammaticalization process (see, for example, Börjars & Vincent 2011, Kiparsky 2012, Traugott 2011, among others for discussion).

1.6. Procedures

1.6.1. Sound Change and Analogical Processes

It has been necessary to set out the case systems in the different chronological periods of each of the languages under study in detail in chapter II. The grammatical information is gathered from relevant sources for each language. Every change in nominal inflection between stages of the languages is attributed to sound change or analogical change. This fine-grained investigation adds contributions to the historical linguistic literature by exploring each nominal ending in every stage in the course of the historical development of the languages under investigation. The relevant sound changes in each language are discussed, and the motivations behind the analogical changes are accounted for to the extent possible. These motivations are traced primarily in the interaction of number, case, and gender, which are generally marked using a single set of inflections.

1.6.2. Number Syncretism, Gender Syncretism, and Gender Criteria

Quantitative analysis in chapter IV of my study measures the interaction among morphological categories involved in analogical change and how it relates to case and gender loss. The methodology involves the comparison of successive stages in the history of a selection of Germanic, Romance, and Balkan Sprachbund languages. For this purpose, I have designed original quantitative measures of number syncretism in noun paradigms and gender syncretism in agreement target paradigms. Two types of number syncretism are investigated: number syncretism due to sound change and number syncretism due to analogical processes. Syncretism between a pair of numbers is measured in terms of how many case forms in one number are the same as at least one case form in the other number; thus, it captures the interactions between case and number in developments to nominal inflection.

Gender syncretism on agreement targets is measured more narrowly in terms of how many genders have the same form within each case and number. The agreement targets examined are limited to demonstratives, articles, and adjectives. Gender types on nominal inflections are also investigated to show the degree of association between gender and declension using criteria for gender types proposed by Kürschner and Nübling (2011).

The general hypothesis presented here is that the loss of case and gender categories in the Germanic, Romance, and Balkan Sprachbund languages is the result not only of sound change that directly neutralizes morphological distinctions involving the lost categories but also of analogical processes in response to the neutralization of more relevant categories, since the leveling of case and gender forms tends to increase the relevance of other categories. That is, if leveling does not occur, case and gender systems are reorganized in one or more other specific ways after the functionality of their original system has been disrupted by sound change. This hypothesis extends Kürschner and Nübling's (2011) proposal, which focused on gender and declension, and their findings, which were limited to a selection of Germanic languages. More specifically, three types of leveling occur in disrupted systems, each triggered by a particular type of syncretism or other morphological patterns. All three types profile a more relevant category.

When a case system has been sufficiently disrupted, case markers that have not been lost to sound change or leveled as part of number profiling are repurposed for more syntactically transparent functions such as phrasal affixes/clitics. Oblique markers in the singular appear to be

particularly prone to this. With respect to these processes, my hypothesis is that decreases in number syncretism due to analogy are correlated with earlier or concurrent increases in number syncretism due to sound change and that more extreme changes to number syncretism are associated with more advanced case loss. The results of the number syncretism analysis generally support these correlations, but notable exceptions have occurred in the development of some languages.

The remaining processes concern grammatical gender. Gender markers can be affected by these processes once the gender system reaches a certain level of opacity. While gender is a lexical category for nouns, it is a morphological category for agreement targets (determiners, adjectives, and pronouns). If gender cannot be reliably predicted based on the form or meaning of a noun, speakers will not be able to remember the gender of a noun unless they are exposed to enough contexts in which the gender is clearly marked on an agreement target within the noun phrase (see De Vos & Vogelaer 2011). The relationship between gender and noun declension in a language can range from total association, where there is a one-to-one correspondence, to total dissociation, where gender is conditioned by more transparent features as opposed to declension class. Despite their differences, these two relationships both involve simplification, particularly a reduction in the number of declension classes. Inbetween are more conservative systems with a partial association or partial dissociation of gender and declension, as in the early stages of all the languages under investigation (see Kürschner & Nübling 2011:377-381). Based on Kürschner and Nübling's hypotheses with respect to a selection of Germanic languages they investigated: the first hypothesis is that a reduction in gender categories accompanies the total dissociation of gender and declension but not their partial dissociation or total association. The second is that case loss is correlated with a reduction in the number of declension classes and, by extension, a shift away from the more conservative gender-declension relationships. A third hypothesis is that the loss of all gender distinctions in the plural is correlated with a dissociation, partial or total, of gender and declension. A final hypothesis is that more conservative gender-declension relationships are associated with lower gender syncretism. The results of the gender analysis in my dissertation refute most of these hypotheses; only the third hypothesis is supported.

My study shows that if a gender distinction cannot be maintained, markers are predicted to undergo one of several processes. One is leveling of gender/declension class distinctions that

profile case category. The second is the leveling of gender/declension that profiles number category. Several changes besides leveling also occur. If the morphological distinctions for a particular category become too reduced by sound change or by leveling that profiles another category, any remaining markers for that category are semanticized, repurposed in other ways, or fall into disuse, depending on the particular conditions. The semanticization of gender categories occurs when all gender distinctions within the NP have become too opaque for first language learners to acquire accurately, i.e., when distinctions are only reliable on pronouns. Semanticization usually involves a shift from lexical/grammatical gender to referential/natural gender. When gender agreement within the NP has been sufficiently disrupted, as described above, a few of the gender markers on agreement targets may not have been lost to sound change or leveled as part of case or number profiling. These are repurposed for discourse functions or fall into disuse if no suitable function can be found.

1.6.3. Studies on Functional Narrowing

Chapter V investigates the assignment of case marking by verbs in three Old English (OE) and Early Middle English (EME) texts: Beowulf, Layamon's Brut, and The Owl and the Nightingale (O&N). The EME texts are representative of the period when significant morphological case ambiguity and case loss first became observable. A quantitative analysis as well as statistical tests, such as the chi-square test and the Fisher exact test, are used to demonstrate how English lost its morphological case over time. This study also compares Vakareliyska's (1990) study of Wernicke's aphasics, who tend to use the core morphological cases instead of the peripheral cases. These two studies provide insight into how morphosyntactic phenomena such as functional narrowing relate to case loss, a function in which case variation can be frequently observed.

1.7. Organization of the Study

The dissertation is organized as follows. In chapter II, the developments in the case systems of a representative selection of Germanic, Romance, and Balkan Sprachbund languages are considered. Subsections for each group of closely related language stages contain tables and discussion of the changes to nominal inflection, with a focus on how sound changes and analogical processes affected number syncretism, as well as the relationship between gender and

declension. Gender syncretism is also considered for agreement targets. Potential motivations for the loss of case and gender inflections are considered in chapter III. Various factors that have been proposed in the literature are evaluated and extended beyond the languages originally studied. Specific examples of some of the morphosyntactic motivations for case loss are discussed in chapter III. Chapter IV quantitatively analyzes how certain phonological and morphosyntactic factors relate to case and gender loss. The first part of the chapter analyzes the effects of sound change and analogical processes on number syncretism and the associations of the resulting developments in number syncretism with case loss. The second part measures how the relationship between gender and noun declension is associated with the number of genders, case loss, and gender syncretism in the core cases of NP-internal agreement targets. Chapter V discusses two studies on functional narrowing and variation of case use. Concluding remarks are presented in chapter VI. The appendices include grammatical descriptions of EME texts investigated, a list of abbreviations used in the dissertation, a transliteration table of old Cyrillic, and information on OE and EME texts investigated.

CHAPTER II.
CASE/GENDER DEVELOPMENTS IN GERMANIC, ROMANCE, AND BALKAN SPRACHBUND LANGUAGES

Language change has been classified into two categories based on formal assessments by the (19th-century) Neo-Grammarians. One is phonological change, which is considered to be regular and can be defined formally as a process of rules. Non-phonological change, on the other hand, such as analogy and reanalysis, is said to be random, unexpected, and definable only by tendencies, not rules. This chapter traces the case systems of relevant Germanic, Romance, and Balkan Sprachbund languages from their proto-languages and/or earliest well-attested stages to the modern period. Case marking on nouns is the focus, but distinctions on pronouns are noted in footnotes for comparison. The chapter identifies perceptible stages in the case loss process, while suggesting relative chronology. In addition to case loss, gender loss is also tracked.

General background information is provided for each stage, followed by a discussion of the sound changes and analogical processes that have had an impact on number, case, and gender distinctions, with a focus on number syncretism in nominal inflection. Gender syncretism on agreement targets is also considered to show how it correlates with gender Types examined in chapter IV. Mergers of declension classes are also discussed, because the development of paradigms over time shows a tendency to reduce the number of paradigms. Tables showing case mergers, i.e., which cases adopted the functions of lost cases, are included where relevant. When a language has lost more than one case in its development, the order of case loss is also examined. This investigation makes a new contribution to the literature by demonstrating in more depth how case and gender loss occur over stages in the evolution of the languages.

The methodology is described in section 2.1. Since all of these languages are Indo-European (IE), an overview of nominal inflection in Proto-Indo-European (PIE) is provided first, in section 2.2. Germanic languages are discussed in section 2.3, Romance languages in section 2.4, and Balkan Sprachbund languages in section 2.5.

2.1. Methodology

Grammatical information is drawn from available sources for every stage of the languages. The sources are chosen based on original scholarship, particularly sources that are authoritative and

recognized in the language field. I list the sources used for each language in the footnotes in each language section. When nominal tables are not available in the sources in a specific stage, I reconstruct the nominal paradigms by applying the sound changes and analogical processes described by the sources. The oldest known stage was used as the starting point for most languages since the investigation is concentrated on the historical case and gender loss. As a result, the chapter did not include some of the proto-languages for showing sound changes and analogical processes, including Proto-Indo-European (PIE) and Proto-Germanic (PGmc). Two were nevertheless added: Proto-Albanian (because no stage prior to Contemporary Standard Albanian (CSA) is known) and Late Proto-Slavic (LPS) (for which Old Church Slavonic (OCS) records were utilized).

Tables are presented showing the changes in noun declension between each pair of successive stages in the individual languages investigated. Tables were not provided for pairs in which the following stage experienced total case loss. The first step was determining which declension classes to include. The sources for the paradigms in a given language often differ in the numbers of classes they include; some minor classes are included in one source but not another for the same language. Therefore, the following criteria were used. Infrequently used declension classes were combined with more frequently used classes or other minor classes when they generally had the same pattern of endings. Otherwise, minor classes only instantiated by a handful nouns in both the earlier and later stage were usually excluded.

The process of demonstrating the nominal changes determines which declension classes corresponded within each pair. This was often clear, but sometimes multiple classes merged into one or were reorganized in other ways. If the merger of two classes involved the adoption of a significant number of forms from both classes, the changes achieving this merger were shown in the respective columns for both. However, if the members of a class went over to another in a later stage while retaining few if any of the reflexes of their original forms, the original class was not considered to have contributed to the later stage. Where relevant to the discussion, these classes that lost productivity are noted in the column of the table for the class that absorbed their members with the phrase “(also some X-stems)”, but their forms in the earlier stage are not included anywhere in the table.

For each declension class in the later stage of a pair, the endings of each case/number form were then compared to the corresponding class or classes in the earlier stage. All changes

that could be fully explained by sound change were marked ">." The remaining changes were marked "»" and analyzed as involving an analogical process of some kind. When no change occurred, one or more forms were listed without either type of arrow. Two or more alternative forms that underwent the same change (or no change) were separated by a comma. If two or more alternative forms underwent different changes, or one underwent a change and the other remained the same, they were separated by a semicolon. However, if one or more analogical changes applied optionally, forms that remained the same as an expected regular development are marked "X > X" to make it clear that they did not participate in the other change(s) listed. If a sound change was expected to apply to a form but did not, the form is marked with "*" and analyzed as involving morphological blocking.

2.2. Proto-Indo-European

PIE was spoken around 4000 BCE. It is reconstructed based on the documented languages of the IE family (Ringe 2006:3). The formation of PIE nouns had the following: (prefix +) root (+ suffix) (+ suffix) + ending. A noun must have a root and an ending (zero morpheme is considered an ending) (Kapović 2017:62). PIE is the reconstructed hypothetical ancestor of all of the language branches considered in this chapter, as well as a number of other language branches. Generally, PIE is reconstructed as distinguishing eight cases morphologically: Nominative (Nom), Vocative (Voc), Accusative (Acc), Genitive (Gen), Dative (Dat), Locative (Loc), Instrumental (Instr), and Ablative (Abl); three genders: masculine, feminine, and neuter; and three numbers: singular, dual, and plural. The Allative is also reconstructed as a distinctive case, based on its use as such in Old Hittite, along with fossilized forms used as adverbs in Greek (see Mallory & Adams 2006:56, Ringe 2006:23-24, Sihler 1995:244-246).

Table 1. Proto-Indo-European Noun Declension

	Consonant-stems (athematic)			
	Masc./Fem.		Neut.	
	Sg.	Pl.	Sg.	Pl.
Nom	-s, -Ø	-es	-Ø	-$\underset{\cdot}{h_2}$
Voc	-Ø	-es	-Ø	-$\underset{\cdot}{h_2}$
Acc	-$\underset{\circ}{m}$	-$\underset{\circ}{m}$s	-Ø	-$\underset{\cdot}{h_2}$
Instr	-b^hi, -mi, -(e)h_1	-$b^h$is, -mīs	-b^hi, -mi, -(e)h_1	-$b^h$is, -mīs
Dat	-ey	-b^hos, -mos	-ey	-b^hos, -mos
Gen	-s, -os, -es	-om (?)	-s, -os, -es	-om (?)
Abl	-s, -os, -es	-b^hos, -mos	-s, -os, -es	-b^hos, -mos
Loc	-i, -Ø	-su	-i, -Ø	-su

	o-stems			
	Masc.		Neut.	
	Sg.	Pl.	Sg.	Pl.
Nom	-o-s	-ōs	-o-m	-e-h_2
Voc	-e-Ø	-ōs	-o-m	-e-h_2
Acc	-o-m	-o-ms	-o-m	-e-h_2
Instr	-o-h_1, e-h_1	-$b^h$is, -mīs	-o-h_1, e-h_1	-$b^h$is, -mīs
Dat	-ōy	-o-b^hos, -o-mos	-ōy	-o-b^hos, -o-mos
Gen	-ī, -osyo	-ōm	-ī, -osyo	-ōm
Abl	-ōt, -āt	-o-b^hos, -o-mos	-ōt, -āt	-o-b^hos, -o-mos
Loc	-o-y, -e-y	(-o-su?)	-o-y, -e-y	(-o-su?)

	eh_2-stems	
	Fem.	
	Sg.	Pl.
Nom	-eh_2-Ø	-eh_2-es
Voc	-eh_2-Ø	-eh_2-es
Acc	-eh_2-m	-eh_2-ms
Instr	-eh_2-b^hi, -eh_2, -h_1?	-eh_2-$b^h$is, -eh_2-mīs
Dat	-eh_2-ey	-eh_2-b^hos, -eh_2-mos
Gen	-eh_2-es, -eh_2-os	-eh_2-om
Abl	-eh_2-es, -eh_2-os	-eh_2-b^hos, -eh_2-mos
Loc	-eh_2-i	-eh_2-su

Table 1. Continued

	i-stems			
	Masc.		Neut.	
	Sg.	Pl.	Sg.	Pl.
Nom	-i-s	-ey-es	-i-Ø	-i- h_2
Voc	-ey-Ø	-ey-es	-i-Ø	-i- h_2
Acc	-i-m	-i-ms	-i-Ø	-i- h_2
Instr	-i-b^hi, -i-h_1	-i-$b^h$is, -i-mīs	-i-b^hi, -i-h_1	-i-$b^h$is, -i-mīs
Dat	-ey-ey	-i-b^hos, -i-mos	-ey-ey	-i-b^hos, -i-mos
Gen	-oy-s	-y-om	-oy-s	-y-om
Abl	-oy-s	-i-b^hos, -i-mos	-oy-s	-i-b^hos, -i-mos
Loc	-ēy-Ø	-i-su	-ēy-Ø	-i-su

	u-stems			
	Masc.		Neut.	
	Sg.	Pl.	Sg.	Pl.
Nom	-u-s	-ew-es	-u-Ø	-u-ms
Voc	-ew-Ø	-ew-es	-u-Ø	-u-ms
Acc	-u-m	-u-ms	-u-Ø	-u-ms
Instr	-u-b^hi, -u-h_1	-u-$b^h$is, -u-mīs	-u-b^hi, -u-h_1	-u-$b^h$is, -u-mīs
Dat	-ew-ey	-u-bhos, u-mos	-ew-ey	-u-bhos, u-mos
Gen	-ow-s	-w-om	-ow-s	-w-om
Abl	-ow-s	-u-bhos, u-mos	-ow-s	-u-bhos, u-mos
Loc	-ēw-Ø	-u-su	-ēw-Ø	-u-su

(Sihler 1995:248)

PIE nouns were divided into two major classes: athematic and thematic, which were marked with consonant stems versus vowel stems, respectively. For athematic nouns, the consonant stems shown in Table 1 above, case endings were attached directly to the consonant-final root. For thematic nouns, a theme vowel came between the root and the case endings (see, for example, Ringe 2006:41, Vakareliyska & Gyllin in press:27-28). The use of different theme vowels led to the formation of the various stem classes reflected in IE languages (shown in Table 1 above), including *o*-stems, *i*-stems, and *u*-stems. These three thematic classes contained masculine and neuter nouns, while feminine nouns were derived from each class with the laryngeal suffix *-h_2, forming the eh_2-stems, ih_2-stems, and uh_2-stems, respectively. The eh_2-stems remained a productive class, becoming *ā*-stems with the loss of laryngeals and compensatory lengthening. The ih_2-stems and uh_2-stems tended to converge with their corresponding masculine paradigms over time (Sihler 1995:245, 249). Although *i*-stems and *u*-stems (and feminine classes derived from them) had a theme vowel, the endings were actually closer to those of the athematic, consonant-stem class (see Vakareliyska & Gyllin in press:27-28, among others, for discussion).

Certain case syncretisms can be observed across the IE daughter languages and can therefore be reconstructed for PIE. It can be assumed that the IE Nom and Voc were always syncretic in the plural; and the Abl was syncretic with the Dat in the plural of all declension classes, and with the Gen in the singular except for *o*-stems. Only three distinctive case forms can be reconstructed for the dual: Nom/Acc/Voc, Gen/Loc, and Dat/Abl/Instr. For neuter nouns, the Nom, Voc, and Acc were always syncretic.

The reconstruction of the functions of the morphological case marking for PIE can be obtained by the semantic analysis of the documented languages of the IE family. The morphological cases had the following main functions. The Nom was primarily used to mark the subject of finite verbs and the complement of 'be' and other linking verbs. The Voc was used for direct address. The Acc was the marker for the direct object (DO). It also indicated motion toward (goal), at least in IE daughters without a distinctive Allative, which had this function if and when it existed. The Gen marked the complement of a noun, including adnominal possession. It most likely also had a partitive function, e.g., when used as the complement of verbs and adjectives. The Dat marked the indirect object (IO), in a broad sense. The Loc indicated location (in space and time). The Instr marked the instrument used in an action. The Abl indicated motion from (source), separation, and standard of comparison. Specific verbs took complements in a case other than the Acc, presumably when another case was a closer fit for the semantic features of the verb. If prepositions existed as such, they likely took different cases depending on their meaning as well (see Clackson 2007:216-224, Kurzová 1993:78-81, Ringe 2006:23-24).

2.3. Changes in Germanic Case Systems

In their development from Proto-Germanic (PGmc), all the Germanic languages experienced some degree of case loss and increased syncretism in the cases that remained. Some Germanic languages have also undergone a reduction or loss of grammatical gender.

2.3.1. Proto-Germanic

PGmc was likely spoken from 500 BCE (Ringe 2006:67). The oldest attested Germanic languages were about 200 CE, the texts of these languages were runic inscriptions (Lass 1994:12) Generally, PGmc is reconstructed as distinguishing six cases morphologically: Nom,

Voc, Acc, Gen, Dat, Instr; the three IE genders: masculine, feminine, neuter; and two numbers: singular and plural.[3] The Voc was usually syncretic with the Nom, and evidence for it is only found in Gothic. PIE Loc and Abl had merged with the Dat in PGmc. The PIE dual was generally lost, except for first and second-person pronouns, which maintain a three-way number distinction.

Table 2. Proto-Germanic Noun Declension[4]

	a-stems				ō-stems		ī/jō-stems	
	Masc.		Neut.		Fem.		Fem.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-az	-ǭz	-ą	-ō	-ō	-ǭz	-ī	-(i)jǭz
Voc	-Ø	-ǭz	-ą	-ō	-ō	-ǭz	-ī	-(i)jǭz
Acc	-ą	-anz	-ą	-ō	-ǫ	-ōz	-(i)jǫ	-(i)jōz
Gen	-as	-ǭ	-as	-ō	-ōz	-ǭ	-(i)jōz	-(i)jǭ
Dat	-ai	-amaz	-ai	-amaz	-ǭi (?)	-ōmaz	-(i)jǭi (?)	-(i)jōmaz
Instr	-ō	-amiz	-ō	-amiz	-ō	-ōmiz	-(i)jō	-(i)jōmiz

	i-stems				u-stems			
	Masc./Fem.		Neut.		Masc./Fem.		Neut.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-iz	-īz	-i	???	-uz	-iwiz	-u	???
Voc	-i?	-īz	-i	???	-u? -au?	-iwiz	-u	???
Acc	-į	-inz	-i	???	-ų	-unz	-u	???
Gen	-īz (-aiz?)	-ijǭ	-īz (-aiz?)	-ijǭ	-auz	-iwǭ	-auz	-iwǭ
Dat	-ī? (-ai??)	-imaz	-ī? (-ai??)	-imaz	-iwi	-umaz	-iwi	-umaz
Instr	-ī	-imiz	-ī	-imiz	-ū	-umiz	-ū	-umiz

[3] Personal pronouns made fewer distinctions than other nominal paradigms in PGmc, as in PIE. The first- and second-person pronouns had distinctive Nom, Acc, and Dat forms for the singular, dual, and plural. The Gen was expressed by the neuter Acc singular of the possessives, which declined like strong adjectives. Other than an original Gen singular form, no additional distinctions beyond these can be reconstructed for PIE. The third-person pronouns followed the same patterns as demonstratives. In northern West Germanic, the original proximal demonstrative replaced the inherited third-person pronouns. In North and West Germanic, a new proximal demonstrative ('this, these') arose through the addition of the particle *-se* to the distal demonstrative ('that, those').

[4] Following Ringe (2006), question marks represent varying levels of uncertainty about the reconstruction: "?" represents some uncertainty, "??" indicates more uncertainty, and "???" is written in place of forms that are unreconstructable.

Table 2. Continued

	Consonant-stems			
	Masc./Fem.		Neut.	
	Sg.	Pl.	Sg.	Pl.
Nom	-Ø ~ -z (~ -s?)	-iz	-Ø	-Ø
Voc	???	-iz	-Ø	-Ø
Acc	-u	-unz	-Ø	-Ø
Gen	-iz	-ǭ	-iz	-ǭ
Dat	-i	-maz	-i	-maz
Instr	(-ē?)	-miz	(-ē?)	-miz

(Adapted from Ringe 2006:269, 272, 274, 279-280)

Due to case syncretism, nouns had four to six distinctive case forms for each number, depending on the paradigm. In the plural, the Voc always had the same form as the Nom, as in PIE.[5] The continuation of the PIE Nom/Voc/Acc syncretism resulted in only four distinctive cases for neuter nouns. Feminine stems either behaved the same as masculine stems or made only four case distinctions, with syncretic Nom/Voc/Instr forms in the singular and Nom/Voc/Acc forms in the plural. Instances of syncretism between Instr and other cases represent changes from PIE that foreshadow the eventual loss of the Instr in all Germanic languages.

Strong adjectives, demonstratives, and possessives distinguished all five cases except the Voc for masculine and feminine of both numbers. The neuter paradigms had the usual Nom/Voc/Acc syncretism. In an innovation from PIE, separate weak adjective paradigms were formed; these followed the same patterns as weak nouns of the corresponding gender.

The functions of the cases in PGmc had changed somewhat from PIE, largely due to several case mergers. The Nom and Voc retained the same functions. The Acc continued to mark DOs and motion toward (goal), including on the objects of prepositions with this meaning. It was also used in adverbial expressions indicating duration of time and extent of space. The Gen continued to mark the complement of a noun. By PGmc, it definitely had a partitive function as well. This meaning may account for its use with certain prepositions indicating motion toward, but not all the way to the goal, in contrast with the Acc. Its use with other prepositions can be explained by their grammaticalization from nouns, since nouns take Gen complements. The Loc and Abl had merged with the Dat, so in addition to the IO and other loosely connected functions

[5] Difficulty reconstructing old forms leads to some points of ambiguity: in the singular, the Voc is not securely reconstructable except for *a*-stems, but it may have been distinct from the Nom in other paradigms as well. It is unclear if the Instr was distinct from the Dat for *i*-stems.

such as inalienable possession, it also inherited the location function from the Loc and motion from (source), separation, and standard of comparison from the Abl. Since these functions were often used with prepositions, the Dat came to be the most frequent prepositional case. The Instr continued to indicate the instrument. It marked the objects of prepositions with this meaning as well as accompaniment. All surviving cases except the Voc marked the complements of certain verbs (see Ringe 2006:233-234).

2.3.2. North Germanic

In terms of case loss, the North Germanic languages developed in two divergent paths. All of which were descended from Old Norse (ON). The continental Scandinavian languages, Swedish, Danish, and Norwegian, experienced a high degree of case loss and a reduction from three grammatical genders to two, while the insular Scandinavian languages, Icelandic, and Faroese, largely retained the case and gender system of ON. The first attestations of ON were in runic inscriptions. The main division between ON dialects, when ON was first attested, was between east and west, which did not directly correspond to the later divergence between continental and insular languages.

2.3.2.1. Insular: The Development of Icelandic and Faroese

In the western branch of North Germanic, there were two main dialects of ON: Old Icelandic (OIc) and Old Norwegian (ONw). In terms of nominal inflection, they were essentially the same (see Kristoffersen 2002:911-915, 919-920). Old Icelandic (OIc) was used in Iceland around 1100 to 1350 and the middle period to 1550 (see Henriksen & van der Auwera 1994:7). Generally, OIc distinguished four cases morphologically: Nom, Acc, Gen, Dat; three genders; and two numbers: singular and plural.[6] By the time North Germanic varieties were attested, the Voc had completely merged with the Nom, and the Instr with the Dat. In both instances, the less frequent case marking was absorbed by the case marking that had a core grammatical role and was more frequent. Since the Dat and Instr suffixes were both used on the objects of prepositions, syntactic

[6] For first- and second-person pronouns, there is also a dual, creating a three-way number distinction, as in PGmc. Personal pronouns have three or four distinctive case forms. This can be considered an increase from PGmc, in which the pronouns did not have their own Gen forms. The first- and second-person singular and third-person feminine singular pronouns distinguish all four cases, while the first- and second-person dual and plural have lost the distinction between the Acc and Dat. The third-person masculine singular pronouns have the same form for Nom and Acc (see Faarlund 2004:35-36).

overlap may have played a role in their merger.[7] Thus, the Nom inherited the function of direct address, and the Dat inherited the functions of instrument and accompaniment. Otherwise, nominal inflection in OIc and the other early North Germanic languages had the same general distinctions and functions as PGmc (see Faarlund 2004:16-33).

ONw was attested from the 11th to 14th centuries. Its development in Norway continued as Middle Norwegian (MNw) from the 14th century to the mid-16th century, when Danish almost completely displaced it in the written record (see Faarlund 1994:38-39, Henriksen & van der Auwera 1994:8). Likely due to Swedish and Danish influence, when the Norwegian came to be used in writing again in 1814 (Henriksen & van der Auwera 1994:8). The modern period of Norwegian began in 1536 (Askedal 1994:268). ONw had more in common with these continental Scandinavian languages than with the rest of the western North Germanic branch. Therefore, Norwegian is discussed together with the continental languages in section 2.3.2.2 below. In addition, ONw developed into Faroese on the Faroe Islands. An intermediate stage designated Old Faroese is assumed for the 14th century to mid-16th century. It is not well-attested but would not have differed much from ONw in nominal inflection, given that the most significant changes in Faroese are relatively recent. For these reasons, Old Faroese, and likewise Middle Icelandic, which was used during the same period, are not considered as separate stages in my analysis.

The modern period is considered to have begun around 1550 for Icelandic and Faroese. Modern Icelandic retains the same distinctions in nominal inflection as OIc. It distinguishes four cases morphologically: Nom, Acc, Gen, Dat; three genders; and two numbers (see Thráinsson 1994:152-155).[8]

[7] See section 3.1.2.5 for further discussion of the motivations behind the absorption of other oblique cases by the Dat in Germanic and other IE languages.

[8] In both Icelandic and Faroese, the distinction between dual and plural has been lost, with the original dual forms of the first- and second-person pronouns supplanting the plural. The original plural forms for these persons are used as honorific forms. In Faroese, personal pronouns make the same case distinctions as in ONw, maintaining the Gen (see Thráinsson 1994:157, Barnes & Weyhe 1994:200-201).

Table 3. Changes in Noun Declension from Old Icelandic to Modern Icelandic[9]

	Strong Declension											
	< a-stems						< i-stems					
	Masc.				Fem.		Masc.				Fem.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-r > -ur	-ar	-r > -ur	-ar	-Ø	-ar	-r > -ur	-ir	-r > -ur	-ir	-Ø	-ir
Acc	-Ø	-a	-Ø	-a	-Ø	-ar	-Ø	-i	-Ø	-i	-Ø	-ir
Gen	-s	-a	-ar	-a	-ar	-a	-s	-a	-ar	-a	-ar	-a
Dat	-i	-um	-i	-um	-Ø	-um	-Ø	-um	-Ø	-um	-Ø	-um

	Strong Declension				Weak[10] Declension					
	< r-stems		Neut.		< an-stems		< ōn-stems		< an-stems	
	Fem.				Masc.		Neut.		Fem.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-Ø	-r > -ur	-Ø	-Ø	-i	-ar	-a	-u	-a	-ur
Acc	-Ø	-r > -ur	-Ø	-Ø	-a	-a	-a	-u	-u	-ur
Gen	-ar	-a	-s	-a	-a	-a	-a	-na	-u	-na
Dat	-Ø	-um	-i	-um	-a	-um	-a	-um	-u	-um

Several sound changes occurred between OIc and Icelandic. An epenthetic *u* broke up final clusters ending in *r* around the end of the 13th century (see Faarlund 1994:43). This vowel, along with many others, has undergone significant changes in phonetic realization (see Thráinsson 1994:143-147). However, none of these changes resulted in the neutralization of any case, number, or gender distinctions. A new masculine strong class has been formed in Icelandic with *-ar* in both the Gen singular and Nom plural. This combination only occurred on two nouns in OIc (see Faarlund 2004:25). Otherwise, Icelandic has retained the same declension classes as OIc.

Gender syncretism on agreement targets remains about the same in Icelandic as in OIc. For most NP-internal agreement targets (determiners and strong adjectives), all three genders are usually distinctive when the forms for the same case and number are compared. However, there is masculine/neuter syncretism in the Gen singular, and there are no gender distinctions in the Gen and Dat plural (see, for example, Faarlund 2004:33-34, 37-38, Thráinsson 1994:155-156).

[9] The grammatical information is drawn from Faarlund (1994) for Old Icelandic and Thráinsson (1994) for modern Icelandic.

[10] In reference to Germanic languages, the term "strong" is used to refer to nouns and adjectives that declined based on vowel stem paradigms, while "weak" is used for nouns and adjectives declined based on the *n*-stem paradigm. A particular noun was consistently either strong or weak, but in a Germanic innovation the strong adjectives, supplemented with more distinctive pronominal endings, became associated with indefinite syntactic contexts, while the weak endings became associated with definite contexts (Ringe 2006:169-170).

Table 4. Changes in Noun Declension from Old Norwegian to Faroese[11]

	Strong Declension					
	< a-stems					
	Masc.				Fem.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-r > -ur	-ar	-r > -ur	-ar	-Ø	-ar
Acc	-Ø	-a » -ar	-Ø	-a » -ar	-Ø	-ar
Gen	-s > -s (W), » -Ø (S)	-a > -a (W), » -ar (S)	-ar > -ar (W), » -Ø (S)	-a > -a (W), » -ar (S)	-ar > -ar (W), » -Ø (S)	-a > -a (W), » -ar (S)
Dat	-i > -i, » -Ø	-um	-i > -i, » -Ø	-um	-Ø	-um

W=Written, S=Spoken

	Strong Declension					
	< i-stems					
	Masc.				Fem.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-r > -ur	-ir	-r > -ur	-ir	-Ø	-ir
Acc	-Ø	-i » ir	-Ø	-i » ir	-Ø	-ir
Gen	-s > -s (W), » -Ø (S)	-a > -a (W), » -ir (S)	-ar > -ar (W), » -Ø (S)	-a > -a (W), » -ir (S)	-ar > -ar (W), » -Ø (S)	-a > -a (W), » -ir (S)
Dat	-Ø	-um	-Ø	-um	-Ø	-um

	Strong Declension			
	Neut.		< r-stems	
			Fem.	
	Sg.	Pl.	Sg.	Pl.
Nom	-Ø	-Ø > -Ø (W), » -r (S)	-Ø	-r > -ur
Acc	-Ø	-Ø > -Ø (W), » -r (S)	-Ø	-r > -ur
Gen	-s > -s (W), » -Ø (S)	-a > -a (W), » -r (S)	-ar	-a > -a (W), » -ur (S)
Dat	-i > -i, » -Ø	-um	-Ø	-um

	Weak Declension					
	< an-stems		< ōn-stems		< an-stems	
	Masc.		Fem.		Neut.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-i	-ar	-a	-ur	-a	-u > -u (W), » -ur (S)
Acc	-a	-a » -ar	-u	-ur	-a	-u > -u (W), » -ur (S)
Gen	-a	-a > -a (W), » -ar (S)	-u	-na » -a (W), » -ur (S)	-a	-na > -na (W), » -ur (S)
Dat	-a	-um	-u	-um	-a	-um

In Faroese, the Gen is generally only used in writing, with a few exceptions; otherwise, it maintains the same general distinctions in nominal inflection as Icelandic. Due to the loss of a productive Gen in speech, Faroese nouns tend to have only a two-way distinction between Nom and Dat; the Acc is syncretic with one of these depending on the class, except in the singular of

[11] The grammatical information is drawn from Kristoffersen (2002) for Old Norwegian and Barnes & Weyhe (1994) for modern Faroese.

masculine strong nouns, which have a three-way case distinction (see Barnes & Weyhe 1994:197-198). In terms of sound changes, the same epenthetic *u* as in Icelandic is also found in Faroese. Likewise, many vowels have undergone significant changes in phonetic realization (see Barnes & Weyhe 1994:191-193). Consonants underwent more changes than in Icelandic (ibid., 193-195). As in Icelandic, however, none of these changes resulted in the neutralization of any case, number, or gender distinctions.

To a greater extent than in Icelandic, analogical processes have brought about changes in nominal inflection, including increased case syncretism, while number syncretism has increased in written Faroese but been completely removed in the spoken language. One change that contributed to the increase in the written language was the extension of the strong Gen plural *-a* to feminine weak nouns. This ending had already been extended to masculine weak nouns in ONw, so these may have also influenced the feminine form. This is a change to the Gen, so it would not be expected if the Gen was no longer used. However, the Gen may have been used infrequently enough in speech that it was easier to remember a single ending for the Gen plural of all masculine and feminine nouns. If so, the retention of *-na* for neuter weak nouns is a little unexpected. However, if this form were also replaced by *-a*, it would have become syncretic with all singular forms as opposed to just the Nom. Thus, there was much more at stake in terms of profiling of number marking for neuter weak nouns than for feminine weak nouns, and the iconic principle overcame the principle of system adequacy.[12] Another analogical process was the leveling of the distinction between the Nom and Acc plural forms of masculine nouns. The Nom forms are now used in place of the original Acc forms, e.g., *a*-stem *hesta* » *hestar* 'horses', as they already were for all neuter classes and the plural of all feminine classes. Thus, the change in the masculine classes was most likely by analogy with the other genders and is an application of the principle of system adequacy. For weak masculine nouns, this profiled number marking by removing the Acc plural form from number syncretism. For most strong masculine classes, this change had no effect on number syncretism. However, a new class with the *-ar* ending in the Gen singular and Nom plural formed as in Icelandic, and the leveling of the Acc plural actually increased number syncretism for this class. The decrease in number syncretism for the weak

[12] Number profiling and other semantic factors in the organization of inflectional paradigms are discussed in section 3.1.2.1 below.

class and the increase for the strong class canceled out, so this process did not result in a net change in number syncretism.

In spoken Faroese, the complete removal of all number syncretism found in ONw and written Faroese resulted from two relatively general changes. The first was the loss of a productive Gen,[13] which accounts for most of the decrease in number syncretism. As part of this process, several functions of the Gen have been assumed by the Acc, and Gen forms have been leveled to the Acc. For *a*-stems with the *-ar* ending in the Gen singular, this eliminated the syncretism between this form and the Nom/Acc plural. For weak nouns, it removed the syncretism among the Gen plural and either the Acc/Gen/Dat singular, for masculine nouns, or the Nom singular, for feminine nouns. Gen forms were likely leveled to the Acc as opposed to another case due to semantic overlap, but the leveling may have occurred first in the classes where the Gen contributed to number syncretism. Evidence for the selective leveling of the Gen comes from the written language: prepositions which traditionally assigned the Gen, such as *til*, now only do so with pronouns, e.g., *til mín* 'to me', and unmodified nouns, e.g., *til strandar* 'to the shore'. When the object consists of more than just a bare noun, the noun and any modifiers occur in the Acc form instead, e.g., *til hina strondina* 'to the other shore' (see Thráinsson et al. 2004:63). With prepositions in written Faroese, the choice of Gen or Acc depends on syntactic context; therefore, it is possible that the choice of Gen or Acc also depended on morphological context, i.e., number syncretism, in the development of spoken Faroese. Without a productive Gen, the masculine strong classes that only differed in the Gen singular merged. Now the strong classes within each gender are distinguished only by the Nom/Acc plural form.

The second change was the extension of Nom/Acc plurals ending in *-r* from masculine and feminine classes to the neuter classes. This removed the syncretism with the Nom/Acc singular for strong neuter nouns. For weak neuters, it did not affect number syncretism as defined for my analysis, but number marking still became more salient. Feminine weak nouns have the same *-u* ending in the oblique singular as the weak neuter Nom/Acc plural does in written Faroese, so a potential for ambiguity exists if the gender is not known. Furthermore, this makes the grammar more transparent by creating an almost one-to-one association between forms ending in *-r* and the Nom/Acc plural. With the loss of the *-ar* Gen singular, the only exception is masculine Nom singular *-ur*. This development can be seen as a product of all three

[13] This loss of productivity is discussed further in section 3.1.2.6 below.

of the principles of selection: the *-r* form is preferred over a zero form in the marked plural by the iconic principle, the consistent use of *-r* plurals make the system more adequate, and *-r* achieves distinctive strength as an overstable marker.

Gender syncretism on NP-internal agreement targets (determiners and strong adjectives) has increased somewhat between ONw, which had the same distinctions as OIc, and Faroese. The Dat singular neuter form has been leveled to the masculine form for both determiners and strong adjectives. Determiners have also lost the distinction between masculine and feminine in the Nom singular. In the written language, the masculine Acc plural has adopted the feminine form while remaining distinct from the Nom plural, which means that only the neuter is distinctive in the Acc plural. This analogy contrasts with the change to the Acc plural for masculine nouns, but in the spoken language, the same analogy as in nouns has been applied, restoring the gender distinction (see, for example, Barnes & Weyhe 1994:199-202).

Table 5. Case Mergers in Spoken Faroese

	Old Norwegian/ Written Faroese				Spoken Faroese		
PIE	Nom	Acc	Gen	Dat	Nom	Acc	Dat
Nom	x				x		
Voc	x				x		
Acc		x				x	
Gen			x			x	
Dat				x			x
Loc				x			x
Instr				x			x
Abl				x			x

2.3.2.2. Continental: The Development of Swedish, Danish, and Norwegian

In the eastern branch of North Germanic, Old Scandinavian (OSc), consisting of the Old Swedish (OSw) and Old Danish (ODan) dialects. OSw was attested from approximately 1225 to 1375. ODan was attested from 1050 to 1340 (see Henriksen & van der Auwera 1994:5, Norde 1997a:15) Norde (2001) distinguishes "idealized" forms of OSw nominal inflection from "pessimistic" forms. The idealized forms made almost all of the same distinctions as OIc; along with Delsing's (2002:927-929) description of OSw/ODan, these are used for OSw in my analysis.

Middle Swedish (MSw) was attested from approximately 1375 to 1526 (Henriksen & van der Auwera 1994:6, Norde 1997a:15). This was a period of significant morphosyntactic change,

comparable to Middle English (ME) but notably later.[14] My study uses a combination of Norde's (2001) pessimistic forms, which are most applicable for early MSw, and Mørck's (2005:1130-1133) description of forms for MSw at various stages of development. Taken together, these MSw forms still maintained the same case, number, and gender distinctions as the idealized OSw forms, but with less consistency. For nouns, no paradigm consistently distinguished all four cases, but no two cases were completely syncretic. With the application of certain optional analogical processes, the Acc was no longer distinctive in any paradigm: it was always syncretic with the Nom and/or Dat. Throughout the MSw period, the system of morphological case marking increasingly broke down.

Table 6. Changes in Noun Declension from Old Swedish to Middle Swedish[15]

	Strong Declension					
	< a-stems				< ō-stems	
	Masc.		Neut.		Fem.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-er*, » -Ø	-ar > -æ, » -ær	-Ø	-Ø	-Ø	-ar > -æ, » -ær
Acc	-Ø	-a > -æ, » -ær	-Ø	-Ø	-Ø	-ar > -æ, » -ær
Gen	-s	-a > -æ, » -æs	-s	-a > -æ, » -s	-ar > -æ, »-ær, » -Ø, » -s	-a > -æ, » -æs
Dat	-i > -e, » -Ø	-um > -om	-i > -e, » -Ø	-um > -om	-Ø	-um > -om

	Strong Declension					
	< i-stems/u-stems				< i-stems	
	Masc.				Fem.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-er*, » -Ø	-ir > -e, » -er	-er*, » -Ø	-ir > -e, » -er	-Ø	-ir > -e, » -er
Acc	-Ø	-i > -e, » -er	-Ø	-i > -e, » -er	-Ø	-ir > -e, » -er
Gen	-ar > -æ, » -ær, » -s	-a > -æ, » -es	-s	-a > -æ, » -es	-ar > -æ, » -ær, » -Ø, » -s	-a > -æ, » -es
Dat	-i > -e, » -Ø	-um > -om	-i > -e, » -Ø	-um > -om	-Ø	-um > -om

[14] The continental Scandinavian languages also lagged the developments of the West Germanic languages in other ways, especially in diverging sufficiently to become separate languages. See Table 21 below for a comparative timeline of case loss in the Germanic languages.

[15] The grammatical information is drawn from Delsing (2002), (Mørck 2005), and Norde (2011) for OSw and MSw, Andersson (1994) for CSS.

Table 6. Continued

	Weak Declension					
	< an-stems				< ōn-stems	
	Masc.		Neut.		Fem.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-i > -e	-ar > -æ, » -ær	-a > -æ	-un > -on	-a > -æ	-ur > -o, » -or
Acc	-a > -æ	-a > -æ, » -ær	-a > -æ	-un > -on	-u > -o	-ur > -o, » -or
Gen	-a > -æ, » -s	-a > -æ, » -æs	-a > -æ, » -s	-na > -næ, » -ons	-u > -o, » -s	-na > -næ, » -os
Dat	-a > -æ	-um > -om	-a > -æ	-um > -om	-u > -o	-um > -om

Sound changes had a much more significant effect on nominal inflection in continental Scandinavian than in the insular languages. The three vowels that occurred in unstressed syllables underwent the following changes in OSw and the Scanian dialect of ODan: [ɪ] > [e], [a] > [æ], [u] > [o]. These changes still kept these vowels distinct from each other and thus did not erase any case or gender distinctions (see Delsing 2002:937). In most dialects of OSw, the loss of word final *-r* was not regular, and the possibility of different outcomes for palatal *-R*, which derived from PGmc **-z*, and alveolar *-r*, from PGmc **r*, also complicates matters (see Norde 1997a:100, Enger 2013:7-8).[16] The *-r* was still intact in the strong masculine Nom singular of indefinite forms through the whole period of OSw, e.g., *a*-stem *fisker* 'fish', but was frequently deleted in the definite forms, e.g., *a*-stem *fiskrinn* > *fiskinn* 'the fish' (see Delsing 2002:937). In Norde's pessimistic forms, the *-r* was lost from the Nom and Acc plural, e.g., feminine *ō*-stem *siangar* > *siangæ* 'beds' (2001:263). However, this varied by text and phonetic factors such as the initial sound of the following word (see Norde 1997a:101-102). Despite this variability, my investigation treats the loss of *-r* (but not the preceding vowel, as in the masculine/feminine Nom/Acc plural) as regular sound change and its retention (e.g., in the strong masculine Nom singular) as morphological blocking, i.e., when an otherwise regular sound change does not apply to a particular morpheme.

The regular loss of *-r* would have neutralized case and number distinctions in most masculine and feminine paradigms. It neutralized the number distinction in the Gen for feminine *i*-stems, e.g., singular *dygþar*, plural *dygþa* > *dygþæ* 'virtue(s)', as well as those masculine *i*-stem

[16] Palatal *-R* and alveolar *-r* were distinguished in early inscriptions of Runic Old East Norse. In inscriptions after 950, alveolar *-r* began to replace palatal *-R* in certain phonetic contexts, and *-R* is not attested after around 1100. No written distinction between these sounds was made in OSw or ODan, but the loss of *-r* was more likely when it derived from palatal *-R*, as in nominal inflections, than from alveolar *-r*, as it often was when part of the stem, e.g., *faþir* 'father' (see Norde 1997a:100-102).

and *u*-stem nouns that had the *-ar* Gen singular form.[17] For *ō*-stems, the Nom/Acc plural were involved as well, since they were already syncretic with the Gen singular. In addition, the Nom plural joined the syncretism between the Dat singular and Acc plural for masculine *i*-stems and *u*-stems, e.g., *i*-stem Dat singular *siþi*, Nom plural *siþir*, Acc *siþi* > *siþe* 'custom(s)'. For masculine weak nouns, the Nom plural joined the syncretism among the Acc/Gen/Dat singular and Acc/Gen plural, e.g., Dat singular *nakka*, Acc/Gen singular/plural *nakka*, Nom plural *nakkar* > *nakkæ* 'neck(s)'. The greatest increase in number syncretism from the loss of *-r* was in feminine weak nouns, as the Acc/Gen/Dat singular became syncretic with the Nom/Acc plural, e.g., Acc/Gen/Dat singular *haku*, Nom/Acc plural *hakur* > *hako* 'chin(s)'.

Several analogical processes have also applied in MSw, with different effects on case and number syncretism. In the Nom singular of masculine strong nouns, *-er* was either retained in full or leveled to the uninflected Acc singular form, e.g., *a*-stem *fisker* » *fisk* 'fish'. Neither option affected number syncretism for *a*-stems. For masculine *i*-stems/u-stems, the zero-ending option for the Dat singular prevented syncretism with the Nom/Acc plural. If *-er* was retained in the Nom singular and plural, however, some number syncretism still remained. In addition to this optional prevention of number syncretism, the retention of Nom singular *-er* probably played an important role in distinguishing masculine from the other genders. In contrast to strong nouns, the optional retention of *-r* in the Nom plural of masculine weak nouns and Nom/Acc plural of feminine weak nouns prevented these forms from becoming syncretic with the Acc/Gen/Dat singular; this followed the iconic principle.

Like the strong masculine Nom singular, the MSw strong feminine Gen singular was retained in full or leveled to the uninflected form used for the other singular cases, e.g., *i*-stem *dygþar* » *dygþær* or *dygþar* » *dygþ* 'virtue'. For feminine *ō*-stems, either option preserved the distinction between the Gen singular and plural, as well as the Nom/Acc plural, when these developed regularly. If the Nom/Acc plural also retained *-r*, only the leveling option avoided number syncretism. For feminine *i*-stems, only the number syncretism within the Gen forms was avoided because the Gen singular had *-ær*, while the Nom/Acc plural had *-er* instead. This same syncretism was also avoided by the same morphological blocking in the masculine *i*-stem/*u*-stem nouns that had the *-ar* Gen singular in OSw, but this form was also sometimes attested with its

[17] Masculine *i*-stems and *u*-stems had already merged in OSw, but both had subclasses that differed in the Gen singular form.

regular development to *-æ*, in contrast to the leveling for feminine strong nouns. This additional option suggests influence from other masculine Gen singular forms, particularly *-s*, which was unaffected by sound change and not leveled. An alternative and/or additional explanation is that the leveling occurred in feminine *ō*-stems first because the potential number syncretism was much higher than in the *i*-stems of either gender. Feminine *i*-stems then followed the example of the other major strong feminine declension class, while masculine *i*-stems/*u*-stems were influenced by other masculine classes as discussed above.

In addition to the optional retention of *-r* in the forms discussed above, Acc plural forms with *-r* were attested for masculine as well as feminine nouns by the end of the MSw period. Even if the retention of *-r* in other forms was the result of dialectal differences with respect to *-r* loss as a sound change, its spread to the masculine Acc plural must have been an analogical change. Acc plural forms with *-r* are not expected unless the Nom plural also has *-r*. For masculine weak nouns, this process removed another form from the syncretism with the Acc/Gen/Dat singular. For masculine *i*-stem and *u*-stem nouns, this eliminated number syncretism involving the regular development of the Dat singular, i.e., *-e*, but the Acc plural became syncretic with the Nom singular instead if this form remained *-er*. It also occurred with masculine *a*-stems but with no effect on number syncretism. A number of factors may have contributed to this extension of *-r*. First, the regular development with *-r* loss resulted in Nom/Acc plural syncretism for masculine nouns as well, providing an additional motivation for treating these two forms the same. Second, this was one of the last distinctions between the Nom and Acc, at least for strong nouns. Masculine and feminine weak nouns still retained this distinction in the singular, but strong masculine Nom singular *-er* was the only other distinctive Nom form, and this was also frequently leveled to the Acc. Thus, the spread of *-r* followed both the principle of system adequacy, in that it increased consistency across classes, and the iconic principle, in that it increased the salience of a marked form, the Acc plural.

The Dat singular had already lost its distinctiveness on strong feminine nouns in OSw, as in OIc. Over the course of the MSw period, the strong masculine and neuter Dat singular *-e* ending was also leveled to the uninflected Acc singular form, e.g., masculine *i*-stem *siþe* » *siþ* 'custom'. This process appears to have been completed earlier for masculine nouns; neuter nouns with *-e* were attested longer (see Mørck 2005:1131). For masculine *i*-stem and *u*-stem nouns, this meant that the Dat singular was no longer syncretic with the Nom or Acc plural, whether or not

these forms retain *-r*. In contrast, the leveling of the Dat singular increased number syncretism for neuter *a*-stems, since it acquired the uninflected form already used for the Nom/Acc of both numbers, e.g., *skipe* » *skip* 'ship(s)'. It is possible that number profiling motivated this leveling for the masculine *i*-stems and *u*-stems; since both the Dat singular and Nom/Acc plural are marked forms but number is more semantically relevant than case, this only followed the iconic principle to the extent that relevance as well as markedness is associated with salience. The leveling process may have then been extended to masculine and neuter *a*-stems by analogy once these were the only remaining distinctive Dat singular forms, in accordance with system adequacy.

The Gen singular form *-s* had begun spreading from *a*-stems to other masculine strong nouns in OSw. During the MSw period, this form spread to strong feminine nouns, then weak nouns. By 1450, it was an option for masculine and neuter weak nouns; by 1500, it had spread to feminine weak nouns, and finally the plural, where it was usually attached after the Nom/Acc plural form without *-r*. By Early Modern Swedish (EModSw) (1556-1732) it was used in all declension classes, indefinite and definite (see Mørck 2005:1132-1133, Norde 1997a:116-127). Of the feminine classes, *i*-stems were the first to adopt the *-s* Gen, probably due the strong parallels that still remained between this class and masculine *i*-stems. In almost all declension classes, Gen forms with *-s* were completely distinctive. When *-s* was attached to the zero plural of neuter strong nouns, however, there was no longer a number distinction in the Gen, further increasing the number syncretism in this paradigm. For other strong nouns, the *-s* form was not the only Gen singular form that avoided number syncretism. The effect on syncretism was more significant for masculine and feminine weak nouns. When masculine weak nouns had *-ær* in the Nom/Acc plural and *-æs* in the Gen plural, all number syncretism was eliminated. Even if the Nom and/or Acc plural lacked *-r*, at least the Gen singular *-s* form was no longer involved in any number syncretism, for both masculine and feminine weak nouns. Therefore, the *-s* Gen had become an overstable marker based on its distinctive strength. The merger of the masculine and feminine into a common gender can be attributed to both sound changes and analogical processes. Vowel reduction of [ɪ] > [e], and [a] > [æ] in the Nom/Acc/Gen plural forms of strong/weak masculine and strong feminine made the endings in these two genders identical. The Nom/Dat singular endings of strong masculine nouns had an optional leveling to a zero ending in MSw, which made Nom/Dat strong feminine endings syncretic. Likewise, the utilization of the

optional levelling of Nom/Acc plural *-ær, -er* and Gen plural *-æs, -es* in the strong/weak masculine and strong feminine made these two genders indistinguishable.

Gender syncretism on agreement targets increased significantly during the MSw period. In OSw, most NP-internal agreement targets (determiners and strong adjectives) made the same gender distinctions as in OIc/ONw, i.e., masculine/neuter syncretism in the Gen singular and no gender distinctions in the Gen and Dat plural. In MSw, the distinction between masculine and feminine in the Acc plural was no longer reliable due to *-r* loss. The loss of the length distinction between /n:/ and /n/ was also taking place during this period (see Kürschner & Nübling 2011:370). This eventually removed the masculine-feminine distinction from the Nom singular of the definite article. The distinctive feminine Acc singular forms of agreement targets were often leveled to the Nom form. Similarly, the uninflected form came to be used frequently for the masculine Nom and Acc singular forms of strong adjectives, as with Nom singular *-er* on strong nouns, as well as for all the Dat singular forms on agreement targets. As with nouns, the Gen singular form *-s* came to be used for all genders. As a result, masculine and feminine were only reliably distinguished in the Nom plural by the end of the MSw period (see, for example, Mørck 2005:1130-1133, 1140-1141). Some of these developments also occurred in Faroese, including the masculine-feminine neutralization in the Nom singular of the definite article and the Acc plural of all agreement targets. The latter syncretism has been tolerated in written Faroese but was distinct by analogy in speech, using the Nom/Masc form; therefore, it is unlikely that these developments were the sole motivation in MSw for the leveling of the remaining masculine-feminine distinctions in the singular. The general loss of the Nom-Acc distinction beginning around this time with strong and plural forms is a more likely cause. In fact, the use of the Nom form in the masculine Acc plural as part of this merger made the masculine and feminine distinctive in the Acc plural, at least temporarily, was important in bringing about the loss of the remaining distinctions between masculine and feminine in the core cases of agreement targets. Instead, this should be attributed to later sound changes, which are discussed below.

Table 7. Changes in Noun Declension from Old Danish to Middle Danish[18]

	Strong Declension					
	< a-stems				< ō-stems	
	Masc.		Neut.		Fem.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-er » -Ø	-ar > -æ	-Ø	-Ø, » -æ, » -ær	-Ø	-ar > -æ
Acc	-Ø	-a > -æ	-Ø	-Ø, » -æ, » -ær	-Ø	-ar > -æ
Gen	-s	-a » -æs	-s	-a » -s, » -æs	-ar » -Ø, » -s	-a » -æs
Dat	-i » -Ø	-um » -æ	-i » -Ø	-um » -Ø, » -æ, » -ær	-Ø	-um » -æ

	Strong Declension					
	< i-stems/u-stems				< i-stems	
	Masc.				Fem.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-er » -Ø	-ir > -æ, » -ær	-er » -Ø	-ir > -æ, » -ær	-Ø	-ir > -æ, » -ær
Acc	-Ø	-i > -æ, » -ær	-Ø	-i > -æ, » -ær	-Ø	-i > -æ, » -ær
Gen	-ar » -s	-a » -æs	-s	-a » -æs	-ar » -Ø, » -s	-a » -æs
Dat	-i » -Ø	-um » -æ, » -ær	-i » -Ø	-um » -æ, » -ær	-Ø	-um » -æ, » -ær

	Weak Declension					
	< an-stems				< ōn-stems	
	Masc.		Neut.		Fem.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-i > -æ	-ar > -æ, » -ær	-a > -æ	-un > -æn	-a > -æ	-ur > -æ, » -ær
Acc	-a > -æ	-a > -æ, » -ær	-a > -æ	-un > -æn	-u > -æ	-ur > -æ, » -ær
Gen	-a > -æ, » -s	-a » -æs	-a > -æ, » -s	-na » -æns	-u > -æ, » -s	-na » -æs
Dat	-a > -æ	-um » -æ, » -ær	-a > -æ	-um » -æn	-u > -æ	-um » -æ, » -ær

ODan nominal inflection is assumed to be the same as OSw. Middle Danish (MDan) was attested approximately 1100 to 1525 (Henriksen & van der Auwera 1994:6). Delsing discusses phonological change and morphological analogy during the OSc period and provides a small selection of forms representative of Danish after the changes of this period (2002:937). My analysis uses these forms to complement Mørck's (2005:1130-1133) description of MDan forms at several stages of development. Due to the merger of the Nom, Acc, and Dat, MDan only distinguished two cases: Nom and Gen. The masculine and feminine also merged into a common gender during this period, leaving only two genders.

Sound changes had an even greater effect on nominal inflection in the transition from ODan to MDan than in OSw to MSw. In the Jutlandic and Zealandic dialects of ODan, all

[18] The grammatical information is drawn from, Delsing (2002), and (Mørck 2005) for ODan and MDan, Haberland (1994) for CSDan.

unstressed vowels were reduced to [ə], spelled <æ> (Delsing 2002:937). This development was more extreme than in MSw, in which three unstressed vowels remained distinctive. Since all case endings were unstressed, this resulted in the loss of many case, gender, and number distinctions, especially in the weak declensions. For masculine weak nouns, this caused the Nom to join the other singular cases in their number syncretism with the Acc/Gen plural, e.g., Nom singular *nakki*, Dat singular *nakka*, Acc/Gen singular/plural *nakka* > *nakkæ* 'neck(s)'.

The morphological factor in *-r* loss appears to have been even clearer in MDan than in MSw. In the Nom and Acc plural, this process applied to different degrees across declension classes: consistently for masculine *a*-stems and feminine *ō*-stems but optionally for *i*-stems, *u*-stems, and weak nouns (see Delsing 2002:928-929, Mørck 2005:1131, 1133). As with MSw, the retention of *-r* is analyzed as morphological blocking of an otherwise regular sound change. When *-r* was lost, several distinctions in various paradigms were neutralized, particularly between the Nom and Acc plural. On its own, the loss of *-r* would have had more or less the same effects as in MSw, i.e., the neutralization of the number distinction in the Gen for *i*-stems and *u*-stems that had the *-ar* Gen singular form, syncretism among the Dat singular and Nom/Acc plural for masculine nouns in these two classes, syncretism among the Gen singular and Nom/Acc/Gen plural for *ō*-stems, syncretism among Acc/Gen/Dat singular and Nom/Acc/Gen plural for weak masculine nouns, and syncretism among the Acc/Gen/Dat singular and Nom/Acc plural for feminine weak nouns. In combination with vowel reduction, however, it led to the loss of additional case, gender, and number distinctions. For feminine weak nouns, the Gen plural remained distinctive, but all of the singular forms became syncretic with the Nom/Acc plural, e.g., Nom singular *haka*, Acc/Gen/Dat *haku*, Nom/Acc plural *hakur* > *hakæ* 'chin(s)'. Among masculine and feminine nouns, declension class could no longer be determined based on the Nom and Acc singular/plural forms.

Analogical processes also had an impact on MDan nominal inflection. The expected masculine Nom singular form after the sound changes above would be *-æ*, but by MDan this form had instead been leveled to the uninflected Acc singular form (see Delsing 2002:937). This was the only attested outcome for this form, in contrast to MSw, where the morphological blocking of *-r* loss sometimes occurred. The ODan strong masculine and neuter Dat singular *-i* ending was also leveled to the uninflected form. This process was complete by 1300, earlier than in MSw (see Mørck 2005:1131). Thus, the iconic principle appears to have been stronger than in

MSw with respect to the preference for uninflected forms in the unmarked singular. For masculine *i*-stem and *u*-stem nouns, the leveling of these singular case forms removed the syncretism between the Dat singular and Acc plural, as in MSw. They also prevented both the Nom and Dat singular from becoming syncretic with the regular development of the Nom plural, but also with the regular development of the Gen plural, which was not the case in MSw due to less extreme vowel reduction.[19] In contrast to MSw, the leveling of the Dat singular decreased number syncretism for neuter strong nouns; even though this form joined the existing syncretism among the Nom/Acc of both numbers, it was prevented from becoming syncretic with the Gen plural. After these changes, the Nom-Acc distinction was completely neutralized, in contrast to MSw. Although profiling of number marking and iconicity may have been factors in the leveling of the Dat singular, they were not in the leveling of the Dat plural form *-um* to the Nom/Acc plural form in all declension classes, since *-um* and its expected reflex **-æm* were completely distinctive and the Dat plural is highly marked, while the Nom/Acc plural were often involved in syncretism. More likely, this form fell out of use after the Dat had completely merged with the Nom/Acc in the singular, at least for indefinite nouns. Thus, this development was more in accordance with the principle of system adequacy.

As mentioned above, *-r* was sometimes retained in the Nom/Acc plural of *i*-stems, *u*-stems, and weak nouns. As in MSw, forms with *-r* also optionally spread to the Acc plural of masculine nouns, which did not have *-r* in ODan but had become syncretic with the regular development of the Nom. Given other analogical processes, including those applying to ODan Gen singular *-ar*, which are discussed next, the retention of *-r* did not affect number syncretism for the strong classes to which it applied and would not have for the classes to which it did not apply. For masculine and feminine weak nouns, however, it prevented or eliminated the syncretism among the Nom/Acc plural, and the singular form now shared by all cases; for feminine weak nouns, all number syncretism was thereby avoided. In addition, the *-æ(r)* plural was sometimes used with neuter *a*-stems in MDan. When it applied, this extension eliminated all number syncretism in this class. The retention and extension of *-r* followed the iconic principle; system adequacy was decreased by the uneven application of these processes across classes but increased with the elimination of a zero plural for neuter *a*-stems.

[19] However, an analogical development of the Gen plural generally supplanted the regular development, as described below.

By 1350, the Gen singular form *-s* had been generalized from *a*-stems to all declension classes, in the singular but also in the plural, where it was generally attached to the Nom/Acc plural form without *-r*. Thus, it became an overstable marker earlier than in MSw. The other strong Gen singular marker in ODan, *-ar*, had fallen out of use, but the form used for the other singular cases was sometimes used for the Gen as well, as part of a growing tendency to mark the Gen explicitly on only one element in a phrase (see Mørck 2005:1132-1133). This tendency, which is discussed further at the end of this subsection, has been described as a part of a shift from concordial case marking to phrase marking (see Norde 1997a:128-131). Either development eliminated any number syncretism that remained for masculine and feminine strong nouns after the other analogical processes described above. As in MSw, the number distinction in the Gen was neutralized for neuter strong nouns when the Gen plural was formed by attaching *-s* to the zero plural. As with other number syncretism in this class, however, this was avoided in MDan by the use of the *-æ(r)* plural, in this case as the base for *-s*. For weak nouns, the use of Gen forms with *-s* had the same effects as in MSw: it made the Gen singular distinctive regardless of plural form and eliminated all number syncretism for masculine weak nouns when *-ær* was used in the Nom/Acc/Dat plural. Thus, all number syncretism was optionally eliminated in MDan but also most case distinctions, an even more extreme development than MSw. As in MSw, sound changes and analogical processes caused the merger of the masculine and feminine genders into a common gender in MDan. Vowel reduction of unstressed vowels made the masculine and feminine forms identical. This can be seen in the Nom/Acc plural forms of strong masculine and feminine as well as weak Nom/Acc/Gen/Dat singular forms of both genders. The levelling of Nom singular *-er* and Dat singular *-i* to a zero ending of strong masculine nouns made the Nom singular strong feminine syncretic.

Gender syncretism on NP-internal agreement targets (determiners and strong adjectives) increased significantly between ODan, which had the same distinctions as OSw, and MDan, including the complete merger of the masculine and feminine. The loss of the length distinction between word-final unstressed /n:/, written as *-nn*, and /n/, written as *-n*, was complete by MDan, earlier than in MSw, and the leveling of other distinctive masculine and feminine singular forms was much more consistent; this was probably connected to the loss of the Nom-Acc distinction on nouns. The more extreme vowel reduction than in MSw neutralized the masculine-feminine distinction in the plural as well. The Dat forms were no longer used after its merger with the

Nom and Acc, and Gen singular *-s* was used for all genders, as in MSw (see, for example, Mørck 2005:1130-1133, 1140-1141, Kürschner & Nübling 2011:370). Thus, the masculine-feminine merger was already complete in MDan, in contrast to MSw. However, the neuter was still distinctive in both numbers.

As mentioned in the previous section, ONw was attested in Norway from the 11th to 14th centuries, followed by MNw from the 14th century to the mid-16th century (see Faarlund 1994:38-39). The same distinctions in nominal inflection as in ONw were retained for most of the MNw period. By the 16th century, however, the Nom and Acc had merged generally in most dialects.

Table 8. Changes in Noun Declension from Old Norwegian to Middle Norwegian[20]

	Strong Declension					
	< a-stems				< ō-stems	
	Masc.		Neut.		Fem.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-r > -e, » -er, » -Ø	-ar > -æ, » -ær	-Ø	-Ø	-Ø	-ar > -æ, » -ær
Acc	-Ø > -Ø, » -e	-a > -æ, » -ær	-Ø	-Ø	-Ø	-ar > -æ, » -ær
Gen	-s	-a > -æ, » -æs	-s	-a > -æ, » -Ø, » -s	-ar > -æ, » -ær, » -Ø, » -s, » -o	-a > -æ, » -æs
Dat	-i > -e, » -Ø	-um > -om	-i > -e, » -Ø	-um > -om	-Ø	-um > -om

	Strong Declension					
	< i-stems/u-stems				< i-stems	
	Masc.				Fem.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-r > -e, » -er, » -Ø	-ir > -e, » -er	-r > -e, » -er, » -Ø	-ir > -e, » -er	-Ø	-ir, -r > -e, » -er
Acc	-Ø > -Ø, » -e	-i > -e, » -er	-Ø > -Ø, » -e	-i > -e, » -er	-Ø	-ir, -r > -e, » -er
Gen	-ar > -æ, » -ær, » -s	-a > -æ, » -e, » -es	-s	-a > -æ, » -e, » -es	-ar > -æ, » -ær, » -Ø, » -s	-a > -æ, » -e, » -es
Dat	-Ø	-um > -om	-Ø	-um > -om	-Ø	-um > -om

	Weak Declension					
	< an-stems				< ōn-stems	
	Masc.		Neut.		Fem.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-i > -e, » -æ	-ar > -æ, » -ær	-a > -æ	-u > -o	-a > -æ, » -o, » -e	-ur > -o, » -or
Acc	-a > -æ, » -e	-a > -æ, » -ær	-a > -æ	-u > -o	-u > -o, » -æ, » -e	-ur > -o, » -or
Gen	-a > -æ, » -e, » -s	-a > -æ, » -æs	-a > -æ, » -s	-na > -næ, » -æ, » -os	-u > -o, » -æ, » -e, » -s	-na > -næ, » -æ, » -os
Dat	-a > -æ, » -e	-um > -om	-a > -æ	-um > -om	-u > -o, » -æ, » -e	-um > -om

[20] The grammatical information is drawn from (Mørck 2005) for ONw and MNw, Askedal (1994) for NNw.

Sound changes similar to those in MSw affected nominal inflection in MNw. Final clusters ending in *r* were broken up by epenthesis. This occurred around 1250 as in OSw, earlier than in the development of Icelandic and Faroese. Although this was during the ONw period, it did not have any effect on case and gender distinctions until other sound changes also applied. There was variation in the spelling of this epenthetic vowel, but *e* is assumed here, as in OSw. In many MNw dialects underwent a reduction of unstressed vowels resembling the process described above for OSw, i.e., phonetic changes occurred without the neutralization of any distinctions. In some dialects, however, unstressed /u/ merged with /i/ as *e*, but this vowel still remained distinct from unstressed /a/, written as *æ*, in contrast to the complete merger of unstressed vowels in most ODan dialects (see Mørck 2005:1129-1130). Since *o* (< /u/) still occurs in plural suffixes in Neo-Norwegian (NNw), also known as Nynorsk, substitutions of *e* for *o*, as well as for *a*, in MNw nominal inflections are considered analogical. The loss of final *-r* in MNw had some of the same complicating factors as in the development of MSw, including dialectal variation (see ibid., 1133). As with MSw and MDan, however, this investigation treats the loss of *-r* as a regular sound change and its retention or spread to new forms as analogical change.

In combination, epenthesis, vowel reduction, and *-r* loss neutralized some case and number distinctions in most masculine and feminine classes. For masculine/feminine *i*-stems with Gen singular *-ar*, the number distinction was neutralized in the Gen, e.g., feminine *i*-stem singular *bœnar*, plural *bœna* > *bœnæ* ‘prayer(s)’. For *ō*-stems, the Nom/Acc plural were already syncretic with the Gen singular, so they were also part of this syncretism. Unlike in OSw/ODan, the Dat singular of masculine *i*-stems already shared the uninflected form with the Acc singular in ONw, so it was not involved in any number syncretism before or after these sound changes. However, the Nom singular became syncretic with the Nom and Acc plural, e.g., Nom singular *vinr*, Nom plural *vinir*, Acc *vini* > *vine* ‘friend(s)’. For masculine weak nouns, the existing syncretism among the Acc/Gen/Dat singular and Acc/Gen plural expanded to include the Nom plural, e.g., Dat singular *granna*, Acc/Gen singular/plural *granna*, Nom plural *grannar* > *grannæ* ‘neighbor(s)’. As in OSw, feminine weak nouns experienced the greatest increase in number syncretism due to these sound changes with the loss of the distinction among the Acc/Gen/Dat singular and Nom/Acc plural, e.g., Acc/Gen/Dat singular *viku*, Nom/Acc plural *vikur* > *viko* ‘hook(s)’.

Many of the same analogical processes that occurred in MSw and MDan applied in MNw. Some of these were optional in MDan, and even more were in MSw, but all were in MNw. In other words, the reflexes that resulted from regular sound change were attested alongside analogical forms during at least part of the MNw period. Final *-r* was sometimes retained in the Nom singular of masculine strong nouns, as in MSw, or leveled to the uninflected Acc form, as in both MSw and MDan. Final *-r* was also optionally retained in the Nom/Acc plural of feminine nouns and the Nom plural of masculine nouns; for the latter, these forms were also optionally extended to the Acc plural, as in MSw and MDan. For masculine *i*-stems, number syncretism among the Nom singular, Nom plural, and Acc plural was avoided when the uninflected form was used for the former, or when *-er* was used only in the former or only in the later. For masculine and feminine weak nouns, the use of Nom/Acc plural forms with *-r* avoided syncretism among these forms and the Acc/Gen/Dat singular. The iconic principle likely played a role in these developments, as in MSw, but not as strongly as in MDan.

A number of different analogical processes could apply to the ONw *-ar* Gen singular form. As in MSw but not MDan, one option was the retention of final *-r*. Leveling to the uninflected form of the other singular cases was an option for feminine strong nouns, but not masculine *i*-stems, just as in MSw and MDan. In addition, *-o* was sometimes used as the Gen singular form of feminine *ō*-stems, presumably by analogy with the Acc/Gen/Dat singular form of feminine weak nouns. All of these options for the Gen singular avoided syncretism with the Gen plural. For *ō*-stems, they all eliminated syncretism with the Nom/Acc plural as well, except when *-r* was retained in both the Gen singular and Nom/Acc plural. The Gen singular form *-s* was extended from *a*-stems to other classes in the same order as in MSw. By 1350, it was an option for all strong classes; by 1500, it had also spread to weak nouns of all three genders and the plural, as an additional ending attached to the Nom/Acc plural form without *-r* (see Mørck 2005:1132-1133). Except for neuter strong nouns, where a zero-plural followed by *-s* in the Gen plural was the same as the Gen singular form, the Gen forms with *-s* were completely distinctive, with the same effects on number syncretism as in MSw. Thus, the distinctive strength of the *-s* Gen likely encouraged its spread as an overstable marker; this progressed at about the same rate as in MSw, which was slower than in MDan.

By 1500, other analogical processes not shared with MSw or MDan were also possible in the Gen plural (see Mørck 2005:1132). The *n* of the inherited feminine and neuter weak Gen

plural was generally lost, e.g., feminine *vikna* » *vikæ* ‘hooks’. As in written Faroese, this process likely occurred by analogy with the inherited Gen plural form in every other declension class. In contrast to written Faroese, however, this process applied to neuter strong nouns, where it resulted in syncretism with the regular development of the singular form shared by all cases. With respect to this development, therefore, the principle of system adequacy was stronger than the iconic principle rather than vice versa, as in written Faroese. As mentioned above, however, this syncretism could be avoided with the use of the Gen plural form *-os* instead, which may explain why there was less need to retain the distinctive *-næ* form. For strong nouns, another possibility was the leveling of the Gen plural to the Nom/Acc plural form without *-r* but without the addition of *-s*. These forms were already syncretic for masculine *a*-stems and feminine *ō*-stems as a result of *-r* loss, so this was likely an extension of that pattern. For neuter *a*-stems, the Gen plural joined the syncretism among the Nom/Acc singular and plural, all uninflected, as a result of this leveling. For feminine *i*-stems, the resulting *-e* form avoided syncretism with the regular development of the Gen singular, but this form was syncretic with the regular development of the Nom singular for masculine *i*-stems.

In ONw, the uninflected Acc form was already used for the Dat singular of masculine *i*-stems as well as strong feminine nouns. As a result, there was no number syncretism between the Dat singular and Acc plural as there was in OSw/ODan. The same process applied first to masculine *a*-stems and then to neuter *a*-stems during MNw, as it did in MSw (see Mørck 2005:1131). This had no effect on number syncretism for the former class, but for the latter, it joined the syncretism among the other uninflected forms, i.e., the Nom/Acc of both numbers and sometimes the Gen plural. Thus, profiling of number marking was most likely not the motivation for this development, but it could have occurred by analogy with masculine *i*-stems, where it did profile number. In other words, system adequacy was more important than iconicity, as with the Gen plural development discussed in the previous paragraph.

As a result of the optional extension of *-r* to the Acc plural of masculine nouns—or its loss in the Nom plural—and the optional leveling of the strong masculine singular form, the Nom and Acc were no longer reliably distinguished for masculine strong nouns in MNw, as was already the case in ONw for neuter and feminine nouns, with the exception of the singular of feminine weak nouns. By 1500, close to the end of the MNw period, the Nom-Acc distinction had also been lost in the other singular paradigms (see Mørck 2005:1131). The inherited Nom

and Acc singular forms were used interchangeably: *-e* or the uninflected form for masculine strong nouns, *-e* or *-æ* for masculine weak nouns, and *-æ* or *-o* for feminine weak nouns. In addition, the masculine singular *-e* form was sometimes extended to the singular of feminine weak nouns. It is possible that this occurred in some dialects due to further vowel reduction that neutralized unstressed *o* and *e* as *e* (see Mørck 2005:1129-1130). However, *-e* does not appear to have been used in place of *-o* in the plural of feminine or neuter weak nouns, and *o* still occurs in unstressed syllables in NNw, so the substitution of *-e* for *-o* in the singular may be purely analogical. The use of *-e* for masculine and feminine weak nouns and the use of *-æ* for feminine weak nouns eliminated any number syncretism that had not been eliminated by other analogical processes.

Gender syncretism on agreement targets increased in MNw in much the same way as in MSw. Most NP-internal agreement targets (determiners and strong adjectives) in ONw made the same gender distinctions as in OIc, OSw, and ODan. As in MSw, the distinction between masculine and feminine in the Acc plural was consistently lost in MNw. They were sometimes neutralized in the singular cases as well, but perhaps not as often as in MSw (see, for example, Mørck 2005:1130-1133, 1140-1141). In contrast to Swedish, these developments did not lead to the eventual merger of the masculine and feminine in NNw, as described below.

In the early modern period, the varieties of continental Scandinavian underwent further changes to nominal morphology. Developments that were complete in MDan and in progress in MSw and MNw were completed in most varieties during this period. In EModSw, these resulted in the loss of almost all distinctions between the masculine and feminine forms on agreement targets. Due to less extreme vowel reduction than in MDan, however, the distinction in the Nom plural would have remained. This was also leveled in favor of the feminine form, which was also the inherited masculine Acc plural form. The loss of the Nom-Acc distinction and the loss of all other masculine-feminine distinctions may have both contributed to this leveling. Thus, the merger of masculine and feminine into a common gender was complete in the dialect areas that form Contemporary Standard Swedish (CSS) and Contemporary Standard Danish (CSDan), but not everywhere in continental Scandinavia. In Norwegian dialects (as well as some Swedish dialects), the masculine-feminine distinction was lost in the plural, but it was the masculine Nom plural form that survived. In the singular of determiners and a small number of strong adjectives, however, a distinctive feminine form *-a* developed (see Mørck 2005:1133, Askedal 1994:231,

234-235). During this period, Swedish and Norwegian also caught up with Danish in terms of case loss, at least on indefinite nouns. In Gustav Vasa's Bible (1541), a conservative EModSw text, strong masculine nouns had already lost the Nom-Acc distinction, but it was maintained in the weak declension. Other significant changes involved the Dat and Gen. Indefinite Dat forms were lost except for the Dat plural ending *-om* (used for all genders), but definite forms were retained well into EModSw, the period of Swedish from 1526-1732 (see Norde 1997a:15, 28). Definite Dat forms still survive in some contemporary Swedish and Norwegian dialects (see Reinhammar 1993).

The modern period is considered to have begun during the 16th century for Swedish, Danish, and Norwegian. In the development of CSS and CSDan, all case distinctions and the distinction between masculine and feminine gender have been lost for nouns. These only distinguish two genders: common and neuter (see Norde 2001:241-242, Haberland 1994:323-325). The common gender plural forms ending in *-r* that had been lost in many varieties of MSw and MDan were restored, e.g., CSS *böcker*, CSDan *bøger* 'books' (see Norde 2001:244-245).[21] The common gender forms derive from a mix of masculine and feminine forms. However, eastern dialects of Danish still distinguish three genders; western dialects no longer have grammatical gender, instead using the common and neuter forms of attributive demonstratives and pronouns with countable and non-countable instances of nouns. CSDan also has a phrasal *-s* Gen, a construction in which *-s* appears at the end of the possessor NP rather than the head noun, e.g., *Kongen af Danmarks bolsjefabrik* 'the King of Denmark's candy factory' (see Haberland 1994:323-325). Likewise, Andersson states that NPs in CSS have a two-way distinction between what is called Basic case, resulting from the merger of Nom and Acc, e.g., *flicka* 'girl', and a phrasal *-s* Gen, e.g., *mannen på gatans åsikter* 'the man on the street's opinions' (1994:278-280, 289). Two standards exist for Norwegian: Dano-Norwegian (Bokmål), based on Danish with Norwegian elements integrated from the speech of the educated urban classes, and NNw (Nynorsk), originally based primarily on the more conservative western dialects but later incorporating more (south-)eastern forms (see Askedal 1994:219-221). In NNw, all case distinctions have been lost for nouns, but there are still three genders (ibid., 229-235). However,

[21] In CSDan, coda *-r* is generally pronounced as non-syllabic [ɐ]; the ending *-er* is pronounced as [ɐ], in contrast to *-e* as [ə] (see Haberland 1994:319-320). Thus, *-r* is not pronounced as such, but its reflex is still phonologically distinctive. A similar situation occurs in CSHG (see, for example, Eisenberg 1994:355).

many Norwegian dialects maintained the Dat case longer than any standard form of Dano-Norwegian, and certain dialects in the central and northwestern provinces of Oppland, Hedmark, Møre og Romsdal, and Trøndelag still do. In these dialects, nouns have what is called a Basic form, e.g., *kone* 'wife', and a Dat form, e.g., *konen*. These dialects maintain the same gender and number distinctions as NNw (Garbacz 2014).[22]

In the modern varieties of the continental Scandinavian languages without any morphological case distinctions, almost all former case functions are expressed with the Basic form of nouns: direct address, subject, complement of 'be', DO, IO, and object of prepositions. Adnominal possession is expressed with the phrasal *-s* Gen or prepositional constructions, and prepositional alternatives exist for some IOs, but these are not completely interchangeable (see Askedal 1994:253-254, Andersson 1994:291-296, Haberland 1994:336-339). In dialects that retain a morphological Dat, it is used for IOs as well as the complements of certain verbs and adjectives. Some prepositions exclusively take the Dat or the Basic form, while others have an opposition between Dat for location and the Basic form for motion toward (goal) (Eyþórsson et al. 2013:223-224).

[22] Pronouns in CSS, CSDan, and the two standard varieties of Norwegian generally distinguish two forms: oblique and Non-Oblique. In reference to these and other languages, oblique refers to forms that are used for all non-subject/non-possessive grammatical relations including the objects of verbs and prepositions. Thus, while oblique forms often derive from the Acc, they are also used in contexts previously associated with other case markings. For the first- and second-person, the singular oblique forms are reflexes of the Acc, while the plural oblique forms are reflexes of an already merged Acc/Dat. The third-person Oblique forms are reflexes of the Dat. Norwegian also has third-person oblique forms that are reflexes of the Acc, but these are the same as the Subject form. In addition, third-person pronouns have an invariable Possessive form derived from the Gen, but only certain third-person pronouns distinguish oblique forms from the Subject form: masculine and feminine singular, along with third-person plural in Swedish writing and formal speech, in CSDan, and in Dano-Norwegian (Andersson 1994:278-284 for Swedish, Haberland 1994:324-329 for Danish, Askedal 1994:229-234 for Norwegian).

Table 9. Case Mergers in Swedish, Danish, and Norwegian

	Old Norwegian / Old and Middle Swedish / Old Danish / Middle Norwegian				Middle Danish	
PIE	Nom	Acc	Gen	Dat	Nom	Gen
Nom	x				x	
Voc	x				x	
Acc		x			x	
Gen			x			x
Dat				x	x	
Loc				x	x	
Instr				x	x	
Abl				x	x	

It can be seen from the data given in this section that the distinction between Nom and Acc began to disappear before other case distinctions in the development of continental Scandinavian nominal inflection from OSc. The Nom completely merged with the Acc in MDan due to sound changes and leveling. Changes proceeded more slowly in Swedish, but probably in the same order: Norde points out that the Gustav Vasa's Bible (1541) had distinctive Gen and Dat forms but the distinction between Nom and Acc was limited to the weak declension. This text is from the EModSw period but is especially conservative, so it likely represents a significantly earlier stage in the development of the spoken language. Even texts from the later part of the preceding MSw period tended to have case marking on only one element in a phrase rather than concordial case (see Norde 1997a:28, 136-139). The loss of concordial case is an important step in the loss of morphological case, so the Gen and Dat were likely lost not long after. However, the order in which they were lost is less clear, especially in Danish, which lost all nominal case distinctions by the end of MDan; no traces remained longer in certain dialects, so dialectal evidence cannot be used, in contrast to Norwegian and Swedish.

Using relative medieval literature for the chronology of case loss, Berg (2015:182, 192) argues that the Gen was lost before the Dat in Norwegian. In MNw, the Dat and Acc started replacing the Gen as the object of verbs and prepositions. These uses of the Gen were likely lost by around 1400, even before the collapse of the Nom-Acc distinction during the 15th century. Definite Dat forms survive in some contemporary Norwegian dialects, while even possessive uses of Gen *-s* were supplanted by prepositional constructions in most Norwegian dialects—the phrasal *-s* Possessive was later reintroduced to Norwegian through Danish influence. Thus, the

order of case loss in Norwegian depends on the status of the *-s* form after its functions had been narrowed to marking possession. Support for the loss of the Gen even before the Acc comes from spoken Faroese, in which the Gen has been lost as a productive case, but Acc and Dat remain distinct from the Nom. Since the Dat has survived in some Swedish dialects as well, it was probably the last case lost in the development of CSS as well. The Gen form *-s* had generally ceased to function as a word-level morphological case by late MSw, but the Acc may still have been lost first, as it seems to have been in Danish.

In summary, of all morphological cases, Voc and Inst had been lost in the development from PGmc to the old attested stages of all North Germanic languages as the Voc merged with the Nom, while the Instr subsumed by the Dat. Among the North Germanic languages, Icelandic and Faroese have not lost any nominal inflection differences due to sound change, with the exception of the Gen in spoken Faroese. In all the other North Germanic languages, vowel reduction in unstressed syllables had affected nominal inflection. Phonetic vowel reduction occurred in MSw and MNw, but phonemic distinctions were mostly maintained, therefore the consequences were small; distinctions were only neutralized in combination with other sound changes. Because of extreme vowel reduction in Danish, it lost all case distinctions significantly sooner than Swedish and Norwegian, especially as the Nom-Acc distinction was retained the longest on singular weak nouns in Swedish.

Other sound changes, when combined with vowel reduction, resulted in additional neutralizations in nominal inflection. In the middle stages of continental Scandinavian languages, the final *-r* was frequently eliminated. This removed any residual distinctions between Nom and Acc plurals, and occasionally with the Gen plural. Furthermore, it significantly increased number syncretism for weak nouns, but only to a lesser extent for feminine and some masculine strong nouns. Several more analogical processes were applied. These changes were common across languages, although their consequences on nominal inflection were more noticeable in languages with more severe case loss. In MSw, MDan, and MNw, the strong masculine/neuter Dat singular form was leveled to the uninflected Acc form. The Acc, Dat, and sometimes Gen singular have been leveled to the Nom in weak paradigms as attested earliest in MNw as an optional process. In Faroese, MSw, MDan, and MNw, there was leveling between other Nom and Acc forms. The Dat plural was also leveled to the Nom/Acc plural in feminine weak nouns in MDan, which increased number syncretism. In spoken Faroese, the leveling of the Gen to the Acc throughout

all paradigms eliminated all number syncretism. Number profiling occasionally had been accomplished by the morphological blocking of *-r* in certain paradigms in MSw, MDan, and MNw.

It had also been a regular process to extend forms across paradigms, which occasionally led to a complete merging of declension classes. Based on the principles of iconicity, system adequacy, and distinctive strength, Nom/Acc plural forms that were more distinctive had extended to neuter classes and other classes with a zero plural or a plural form that was involved in number syncretism. In MDan and more frequently in the evolution of CSDan, masculine and feminine *-r* plurals began to expand to neuter classes. This had also happened in spoken Faroese. The Gen -s form of strong masculine/neuter was extended to other declension classes in most of the languages that lost all nominal case distinctions. This occurred in the beginning of MSw, MDan and MNw. Gender distinction was also reduced. Of North Germanic investigated, Swedish and Danish reduced the number of gender distinctions to just two as a result of the merger of masculine and feminine into a common gender.

2.3.3. West Germanic

The West Germanic languages considered in this section are High German, Low German, Dutch, and English. Other West Germanic languages not considered here include Yiddish, Afrikaans, and Frisian. Yiddish developed from Middle High German (MHG), so it shares its early developments with High German, and it has not diverged significantly since then in terms of nominal inflection (see Jacobs et al. 1994:388, 401-406). Afrikaans only split from Dutch in the 18th century; other than a complete loss of grammatical gender as in English, and possibly due to contact with it, nominal inflection in Afrikaans resembles that of CSDu (see Donaldson 1994:478-480, 485-486). Frisian has developed independently but is not attested as early or as well as other West Germanic languages; its nominal inflection resembles that of CSDu (see Hoekstra & Tierstma 1994:513-516). With the exception of English, the West Germanic languages considered in my analysis are referred to as the continental West Germanic languages. All West Germanic languages have experienced a high degree of case loss on nouns, with many parallels to the continental Scandinavian languages. Morphological case has been retained in most German varieties, but mainly on determiners and adjectives. Thus, these are similar to Icelandic and Faroese in terms of overall nominal distinctions, but less so with respect to nouns

in particular, which have more robust case marking in Icelandic and Faroese. In addition, English has lost grammatical gender entirely.

2.3.3.1. The Development of High German[23]

Old High German (OHG) was attested from the 8th century to the end of the 11th century. Generally, OHG fully distinguished four cases morphologically: Nom, Acc, Gen, Dat; three genders; and two numbers.[24] In addition, there was a distinctive Instr for certain singular masculine and neuter forms. Otherwise, the Instr was syncretic with the Dat (see van der Wal & Quak 1994:72, 93-95, Wright 1906:45-55). Even the limited distinctiveness of the Instr from the Dat contrasts with the North Germanic languages, in which the Instr had completely merged with the Dat by the earliest attestations.

Middle High German (MHG) was used from the 12th to 15th centuries and maintained most of the same general distinctions in nominal inflection as OHG. It distinguished four morphological cases: Nom, Acc, Gen, Dat; three genders; and two numbers.[25] The distinction between Dat and Instr had been maintained for the neuter definite article/distal demonstrative and the interrogative pronoun but lost for the masculine and on nouns and adjectives (Wright 1917, §41-53, §68-70).

[23] The grammatical information is drawn from Wright (1906) for OHG, Wright (1917) for MHG, Eisenberg (1994) for CSHG.

[24] As in ON, personal pronouns had three or four distinctive case forms. In OHG, first- and second- person pronouns distinguished all four cases, as in PGmc. The third-person masculine and feminine singular pronouns also distinguished all four cases. The third-person plural and neuter singular pronouns had the same form for Nom and Acc (see van der Wal & Quak 1994:97-99).

[25] Personal pronouns maintained almost all of the same distinctions as in OHG. However, there were several changes. In the third-person feminine singular, Gen/Dat syncretism arose due to vowel reduction, and the beginnings of Nom/Acc syncretism can be observed. In the first- and second-person plural, the beginnings of an Acc/Dat syncretism, which is also found in OS, ON, OE, etc., can also be observed. In the latter two instances of syncretism, distinctive forms still exist, but the same forms were often used for both cases involved (Wright 1917, §65).

Table 10. Changes in Noun Declension from Old High German to Middle High German

	Strong Declension					
	< a-/wa-/i-stems		< ja-stems		< a-stems	
	Masc.		Masc.		Neut.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-Ø	-a, -i > -e, -Ø	-i > -e	-a > -e	-Ø	-Ø
Acc	-Ø	-a, -i > -e, -Ø	-i > -e	-a > -e	-Ø	-Ø
Gen	-es > -es, -(e)s	-o, -eo > -e, -Ø	-es	-eo > -e	-es > -es, -s	-o > -e, -Ø
Dat	-e > -e, -Ø	-um, -im > -en, -n	-e	-im > -en	-e > -e, -Ø	-um > -en, -n
Instr	-u, -iu > -e, -Ø	-um, -im > -en, -n	-iu > -e	-im > -en	-u > -e, -Ø	-um > -en, -n

	Strong Declension			
	< ja-stems		< wa-stems	
	Neut.		Neut.	
	Sg.	Pl.	Sg.	Pl.
Nom	-i > -e	-i > -e	-o, -Ø	-o, -Ø
Acc	-i > -e	-i > -e	-o, -Ø	-o, -Ø
Gen	-es	-eo > -e	-wes > -wes, » -es	-wo > -we, » -e
Dat	-e	-im > -en	-we > -we, » -e	-wum > -wen, » -en
Instr	-iu > -e	-im > -en	-we > -we, » -e	-wum > -wen, » -en

	Strong Declension					
	< z-stems		< ō-/jō-/wō-stems		< i-stems	
	Neut.		Fem.		Fem.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-Ø	-ir > -er	-a > -e, -Ø	-ā > -e, -Ø	-Ø	-i > -e
Acc	-Ø	-ir > -er	-a > -e, -Ø	-ā > -e, -Ø	-Ø	-i > -e
Gen	-es > -es, -s	-iro > -er	-a > -e, -Ø	-ōno > -en, -n	-i > -e, » -Ø	-eo > -e
Dat	-e > -e, -Ø	-irum > -ern	-u > -e, -Ø	-ōm > -en, -n	-i > -e, » -Ø	-im > -en
Instr	-u > -e, -Ø	-irum > -ern	-u > -e, -Ø	-ōm > -en, -n	-i > -e, » -Ø	-im > -en

	Weak Declension					
	< n-stems					
	Masc.		Neut.		Fem.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-o > -e	-on > -en	-a > -e	-un > -en	-a > -e	-ūn > -en
Acc	-on > -en	-on > -en	-a > -e	-un > -en	-ūn > -en	-ūn > -en
Gen	-en	-ōno > -en	-en	-ōno > -en	-ūn > -en	-ōno > -en
Dat	-en	-ōm > -en	-en	-ōm > -en	-ūn > -en	-ōm > -en
Instr	-en	-ōm > -en	-en	-ōm > -en	-ūn > -en	-ōm > -en

Several sound changes resulted in significant changes in nominal inflection between OHG and MHG. Unstressed vowels were reduced to [ə], spelled <e> (Wright 1917, §7). This change and its results are comparable to what happened in MDan, the only North Germanic language with the total neutralization of vowel distinctions in unstressed syllables. Similar to MDan, it neutralized many case, gender, and number distinctions, including the last remaining distinctions

between the Instr and Dat. For masculine *a*-stems and *wa*-stems, which had already merged in OHG, the Dat and Instr singular became syncretic with the Nom/Acc plural, and the Gen plural joined this syncretism as well, e.g., *a*-stem Dat singular *tage*, Instr *tagu*, Nom/Acc plural *taga*, Gen *tago* > *tage* 'day(s)'. For masculine and neuter *ja*-stems, the Nom/Acc singular also joined this number syncretism, while only the Dat and Instr singular and Gen plural were involved for neuter *a*-stems and *wa*-stems. The most extreme effects on strong nouns occurred in the paradigm for feminine *ō*-stems, *jō*-stems, and *wō*-stems, which had already merged in OHG: all singular forms became syncretic with the Nom/Acc plural, e.g., *ō*-stem Nom/Acc/Gen singular *ërda*, Dat *ërdu*, Nom/Acc plural *ërdā* > *ërde* 'earth(s)'. The effects on the weak classes were similarly extreme, expanding the existing syncretism in these paradigms: only the Nom singular for all genders and Acc singular for neuters remained distinctive, but the Gen/Dat plural forms lost their distinctiveness during the transition to MHG, thereby increasing number syncretism, e.g., masculine Acc singular *boton*, Gen/Dat *boten*, Nom/Acc plural *boton* > *boten* 'messenger(s)'.

Even though vowel reduction resulted in the merger of the masculine *i*-stem endings with those of the *a*-stems and *wa*-stems by MHG, syncretism between the Dat/Instr singular and Nom/Acc/Gen plural was avoided for most nouns in this class as a result of umlaut in the plural. This process, i.e., the fronting of back vowels and diphthongs before a following /i/ or /j/, began in OHG with /a/ > /e/ and was expressed in writing for all back vowels by 1200; in the spoken language, it must have occurred before the reduction of unstressed /i/ to [ə] and the loss of /j/ (except between vowels), since umlaut was only conditioned by /i/ and /j/ (see Wright 1917, §10). Except for masculine *i*-stems that already had a front root vowel, umlaut applied in the plural case forms, since they originally contained /i/, e.g., OHG Nom/Acc plural *gesti*, Gen *gesteo/gestio*[26] > MHG *geste* 'guests', but not in the singular, since they did not generally contain /i/ by OHG, e.g., Dat singular *gaste*, Instr *gastiu/gestiu/gastu* > *gaste*). Thus, the application of umlaut in masculine *i*-stems was phonologically predictable, except in the Instr singular, which was likely influenced by the other singular forms and the *a*-stem form without /i/. Other than this potential blocking of umlaut, the merger of the Instr with the Dat in nouns can be attributed completely to regular sound change. For feminine *i*-stems, the Gen plural joined the

[26] The Gen plural form *-eo* was more common in early OHG, but this diphthong regularly developed into *-io* in the 9th century, which explains why it triggered umlaut (see Wright 1906:20-21).

syncretism between Gen/Dat singular and Nom/Acc plural, e.g., Gen/Dat singular/Nom/Acc plural *ensti*, Gen plural *ensteo* > *enste* 'favor(s)'. In contrast to masculine *i*-stems, the Gen/Dat singular forms were still *-i* in OHG, so umlaut applied to them as well, and thus did not prevent this syncretism. An analogical change optionally avoided this syncretism, however, as described below.

While the vowel reduction in MHG was shared with MDan, other sound changes did not have any parallels in the North Germanic languages. When a reduced vowel in the transition from OHG to MHG occurred in the final syllable (and in certain other contexts) after a liquid ([l] or [r]) or nasal ([m] or [n]), it was lost completely, which resulted in a further loss of distinctions in certain paradigms (Wright 1917, §9). For masculine *a*-stems, *wa*-stems, and *i*-stems, this deletion caused the Nom/Acc singular to join the syncretism among the Dat singular and Nom/Acc/Gen plural, e.g., *a*-stem Nom/Acc singular *engil*, Dat *engile*, Nom/Acc plural *engila*, Gen plural *engilo* > *engel* 'angel(s)'. Likewise, it caused the Nom/Acc singular/plural to join the syncretism between the Dat singular and Gen plural for neuter *a*-stems, e.g., Nom/Acc singular/plural *fenstar*, Dat singular *fenstare*, Gen plural *fenstaro* > *venster* 'window(s)'. For weak nouns, it neutralized the distinctiveness of the Gen plural from the other plural forms and the oblique singular forms that had become syncretic with them due to vowel reduction, e.g., masculine Gen plural *botōno* > *boten* 'messengers'. Another change applied before this deletion: final *-m* regularly changed to *-n* over the course of the 9th century (Wright 1906:43).[27] On its own, this did not affect any distinctions in nominal inflection, but in combination with the other changes, it had a significant effect, particularly on the weak declensions, where it neutralized the last remaining case distinction in the plural, the Dat, e.g., masculine Dat plural *botōm* > *boten* 'messengers'. However, it did not affect the pronominal endings used on determiners, strong adjectives, and pronouns because these had a vowel after the *m* at the time. For example, the masculine/neuter Dat singular form of the distal demonstrative/definite article underwent the following changes: OHG *demu* > MHG *dëm(e)* > Contemporary Standard High German (CSHG) *dem*.

[27] This development, which is also occurred in other West Germanic languages, probably resulted from the weak pronunciation of final nasals in unstressed syllables. Coronals are unmarked relative to other places of articulation (see Paradis & Prunet 1991). Thus, speakers, especially L1 learners, probably started assuming they were hearing *-n* instead of *-m*, and thus came to use *-n* instead in their own pronunciation.

The merging and reorganization of noun declensions played an important role in MHG. In addition to the mergers that resulted from the sound changes above, some nouns switched declension classes. One notable development was the shift of neuter *a*-stems, with a zero plural, to the neuter *z*-stem declension, which was characterized by *-er* in the plural and therefore avoided number syncretism. This class grew from under ten nouns in OHG to almost twenty during the MHG period (see Wright 1917, §47). In fact, the only other analogical changes that affected syncretism concerned umlaut. One was the optional leveling of the feminine *i*-stem Gen/Dat/Instr singular form, in which umlaut had applied in the plural, to the uninflected Nom/Acc form, in which umlaut had not applied, e.g., *enste* » *anst* 'favor'. When it applied, this process brought the singular paradigm in line with other feminine strong nouns and eliminated all number syncretism. In addition, umlaut spread to the plural forms of some masculine *a*-stems by analogy with masculine *i*-stems (see Wright 1917, §44). These changes reflect the morphologization of umlaut, i.e., its transition from a productive phonological process to a morphologically conditioned one. Over time, this process tended to increase the distinctiveness of the plural versus singular for nouns, in accordance with the iconic principle. It strengthened the association of umlaut as a plural marker, while eliminating its inconsistent use as a case marker, thereby increasing system adequacy.

Gender syncretism on agreement targets increased significantly between OHG and MHG. In OHG, most NP-internal agreement targets (determiners and strong adjectives) had Nom and Acc forms of both numbers that distinguished all three genders, but the masculine and neuter were syncretic in the Gen and Dat singular, and there were no gender distinctions in the Gen and Dat plural. The same singular distinctions were retained in MHG, while the plural Nom/Acc forms only distinguished the neuter, and even this distinction was sometimes leveled in MHG. Thus, the loss of gender distinctions in the plural was nearly complete by the end of the MHG period, but the three-gender system remained robust in the singular. Notably, there was hardly any masculine-feminine syncretism, in contrast to MSw and MDan.

The modern period is considered to have begun around 1500 for High German. CSHG maintains most of the same general distinctions in nominal inflection as MHG. It distinguishes four morphological cases: Nom, Acc, Gen, Dat; three genders; and two numbers (see Eisenberg

1994:359-367).[28] The distinction between Dat and Instr has been completely lost, and the former functions of the Instr are now expressed with prepositional constructions, most of which have a noun in the Dat. In addition, the Gen is becoming less productive in everyday language (spoken and written), as the adnominal Gen construction, e.g., *das Buch meines Vaters* 'my father's book', is replaced by the prepositional construction with *von* followed by the Dat. e.g., *das Buch von meinem Vater* 'the book of my father', the non-concordial Possessive *-s*, e.g., *(mein) Vaters Buch* '(my) father's book', and the resumptive possessive construction with a Dat possessor and a possessive determiner, e.g., *meinem Vater sein Buch* 'my father's book' (see, for example, Scott 2014, Sick 2005).

Table 11. Changes in Noun Declension from Middle High German to Contemporary Standard High German

	Strong Declension							
	Masc./Neut.		Neut./Masc.				Fem.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-Ø	-e	-Ø	-Ø	-Ø	-er	-e, -Ø	-e, -Ø » -en
Acc	-Ø	-e	-Ø	-Ø	-Ø	-er	-e, -Ø	-e, -Ø » -en
Gen	-(e)s	-e	-(e)s	-e » -Ø	-(e)s	-er	-e, -Ø	-en; -e, -n » -en
Dat	-Ø; -e » -Ø	-en	-Ø; -e » -Ø	-n	-Ø; -e » -Ø	-ern	-e, -Ø	-en; -n » -en

	Weak Declension			
	Masc.		Fem.	
	Sg.	Pl.	Sg.	Pl.
Nom	-e > -e, » -Ø	-en	-e	-en
Acc	-en > -en, » -e, -Ø	-en	-en » -e	-en
Gen	-en	-en	-en » -e	-en
Dat	-en > -en, » -e, -Ø	-en	-en » -e	-en

Sound changes between MHG and CSHG have not had much effect on nominal declension. However, a number of analogical processes have brought about significant changes in nominal inflection, including increased case and gender syncretism, while number syncretism has decreased significantly, especially when some ongoing changes are considered. The merging of paradigms had an even greater effect on nominal inflection in the transition from MHG to CSHG. Due to sound change in MHG, some masculine strong nouns had already developed a

[28] Personal pronouns maintain almost all of the same distinctions as in MHG. The Acc/Dat syncretism in the first- and second-person plural and the Nom/Acc syncretism in the third-person feminine singular have become complete (Eisenberg 1994:361).

zero ending for the Nom/Acc plural, and some neuter strong nouns had already developed the *-e* Nom/Acc plural associated with masculine strong nouns, as described above. Along with the existing overlap in the singular, this resulted in the same declension pattern for some masculine and neuter strong nouns. What began as an accidental overlap has led to a more general extension of forms from nouns of one gender to the other. Other neuter nouns adopted the *-e* Nom/Acc plural by analogy, and several masculine nouns have adopted the *-er* plural forms, which now occurs on around a hundred neuter nouns, an increase from under twenty in MHG (see Wright 1917, §47, Kürschner & Nübling 2011:365). A new mixed declension has also emerged that combines the strong Gen singular -(*e)s* and the weak plural -(*e)n* forms. This paradigm has already been adopted by over 40 nouns, including both masculine and neuter nouns. This is a clear instance of profiling number marking because it avoids the number syncretism among the -(*e)n* forms that occurs for weak nouns, strengthening the association of -(*e)n* with the plural (see Kürschner & Nübling 2011:366). This class absorbed most of the neuter weak nouns, but many masculine weak nouns remain.

Another change was the leveling of strong masculine and neuter Dat singular *-e* to the uninflected Nom/Acc form, e.g., strong masculine *tage* » *Tag* 'day'. The effects and potential motivations of the same process have already been discussed for the continental Scandinavian languages in section 2.3.2.2 above. For masculine nouns that followed the primary strong paradigm in MHG, this means the three plural forms other than the Dat are no longer syncretic with the Dat singular. As in the continental Scandinavian languages, the same change had the opposite effect on neuter strong nouns with a zero plural. Those that had not already lost the Dat singular *-e* to regular sound change joined those that had; the influence of the latter and masculine strong nouns was probably too strong to be resisted, but the *-er* plural has been extended to many more neuter nouns, thereby avoiding this number syncretism. These processes followed the principles of iconicity and system adequacy.

One analogical process was the leveling that occurred during the merger of the strong and weak feminine classes. The Nom/Acc plural form of strong feminine nouns was leveled to the Gen/Dat plural form, e.g., strong feminine Nom/Acc plural *ërde* » *Erden*, Gen/Dat *ërden* > *Erden* 'earths', as they merged with weak feminine nouns, which already had this *-en* form for all cases in the plural, e.g., *zungen* > *Zungen* 'tongues'. While the weak pattern was adopted in the plural, the strong pattern of no case distinctions, e.g., *ërde* > *Erde* 'earth', was adopted in the

singular, e.g., weak feminine Nom singular *zunge* > *Zunge*, Acc/Gen/Dat *zungen* » *Zunge* ‘tongue’. This can be analyzed as the extension of the weak Nom singular form to the other singular case forms by analogy with the strong nouns and in accordance with the iconic principle. This merger eliminated all number syncretism from feminine nouns, but also all case distinctions.

An ongoing change to weak nouns, which are always masculine, is further profiling number at the expense of case and gender. These have been the only nouns in CSHG that distinguish the Acc and Dat singular from the Nom, so it follows from both the iconic principle and principle of system adequacy that the Nom is being extended to these forms, thereby reducing number syncretism and increasing system adequacy by bringing weak nouns closer to the pattern found in other masculine nouns, which only have a distinct Gen form in the singular. With this development, weak nouns only differ from the new mixed declension in the form of the Gen singular, so these developments may be related. Earlier in the modern period, a small number of weak nouns adopted the Gen singular form *-ens*, adding *-s* to the inherited *-en* form by analogy with strong masculine nouns, e.g., MHG *namen* » CSHG *Namens* ‘name’. This change also served to distinguish the Gen singular from the plural, while retaining more case distinctions in the singular. Thus, Gen singular forms have generally maintained and sometimes even increased their salience, even when they are involved in number syncretism, as on masculine weak nouns, in contrast to other singular forms. According to the iconic principle, this indicates that the Gen singular is more marked than other singular cases. In the plural, however, the Dat retained more salience than the Gen.

Gender syncretism on agreement targets has increased slightly from MHG to CSHG. Most NP-internal agreement targets (determiners and strong adjectives) in CSHG remain reliably distinctive in the Nom and Acc singular. However, the loss of the last gender distinction in the plural is complete in CSHG.

Table 12. Case Mergers in High German

	Old High German					Middle High German / CSHG			
PIE	Nom	Acc	Gen	Dat	Instr	Nom	Acc	Gen	Dat
Nom	x					x			
Voc	x					x			
Acc		x					x		
Gen			x					x	
Dat				x					x
Loc				x					x
Instr					x				x
Abl				x					x

2.3.3.2. The Development of Low German[29]

Old Saxon (OS), also called Old Low German, was attested from the 8th century to the end of the 11th century. Generally, OS made the same morphological distinctions in nominal inflection as OHG. In addition, a distinctive Instr singular form was used for feminine long *i*-stems, probably by analogy with *a*-stems, which also appears to be the source for the masculine *i*-stem form (see van der Wal & Quak 1994:72, 93-96, Gallée 1891:55-70, Ringe 2006:269, 272).[30] Two significant sound changes occurred earlier in OS than in OHG. Phonemic vowel quantity had already been neutralized in unstressed syllables by the beginning of the OS period, while this development occurred later in OHG, in the transition to MHG. As a result, morphological distinctions relying solely on vowel length in OHG had already been neutralized in OS. For nominal inflection, the most significant effect was on feminine *ō*-stems and *jō*-stems, which had already merged in OS. These nouns had the same form for the Nom/Acc/Gen singular and Nom/Acc plural, e.g., *ō*-stem *ertha* 'earth(s)'. In OHG, the Nom/Acc plural form was still *-ā*, and this syncretism did not arise until MHG. The shift of final *-m* to *-n* appears to be reflected more frequently in OS texts than in OHG. On the other hand, final *-m* is still attested in MLG, especially in certain regions, but overall alternations between *-m* and *-n* indicate scribes were attempting to maintain a distinction that had been lost in their speech (Lasch 1914:144-145). For

[29] The grammatical information is drawn from Gallée (1891) for OS, Lasch (1914) for MLG and ModLG.

[30] In addition to the dual forms, OS personal pronoun forms varied slightly from OHG in that first- and second-person pronouns had lost the distinction between Acc and Dat, and the beginnings of Nom/Acc syncretism and Gen/Dat syncretism can be observed in the third-person feminine singular pronouns. In the latter two instances of syncretism, distinctive forms still exist, but the same forms were often used for both cases involved (see van der Wal & Quak 1994:97-99, Gallée 1891:82-83).

the most part, this development did not affect distinctions in nominal inflection until vowel reduction occurred, but as a result of this change and the neutralization of vowel length in unstressed syllables, the weak masculine Dat plural had already joined the syncretism among the Acc singular and Nom/Acc plural in OS but not OHG, e.g., OS Nom/Acc singular/Nom/Acc/Dat plural *bodon* 'messenger(s)'.

Middle Low German (MLG) was used from the 12th to 15th centuries and maintained most of the same general distinctions in nominal inflection as OS. It distinguished four morphological cases: Nom, Acc, Gen, Dat; three genders; and two numbers.[31] The Instr had been lost as a productive case marking; it only survived in a few fossilized forms, while its functions were assumed by the Dat (Lasch 1914:192-202).

Table 13. Changes in Noun Declension from Old Saxon to Middle Low German

	Strong Declension					
	< a-/long i-stems		< ja-stems/short i-stems		< long and short a-stems	
	Masc.		Masc.		Neut.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-∅	-os » -e; -a, -i > -e	-i > -e	-ios > -es; -i > -e	-∅	-∅; -u > -e
Acc	-∅	-os » -e; -a, -i > -e	-i > -e	-ios > -es; -i > -e	-∅	-∅; -u > -e
Gen	-es > -es, -s	-o, -io > -e, -∅	-es; -ies > -es	-io > -e	-es	-o > -e, -∅
Dat	-e > -e, -∅	-un, -in > -en	-ie, -i > -e	-iun > -en	-e	-un > -en
Instr	-u > -e, -∅	-un, -in > -en	-iu > -e	-iun > -en	-u > -e	-un > -en

	Strong Declension					
	< ja-stems		< ō-/jō-stems/short i-stems		< long i-stems	
	Neut.		Fem.		Fem.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-i > -e	-i > -e	-a, -i > -e	-a, -i > -e, » -en	-∅	-i > -e
Acc	-i > -e	-i > -e	-a, -i > -e	-a, -i > -e, » -en	-∅	-i > -e
Gen	-ies > -es	-io > -e	-a, -i > -e, » -en	-ono > -en	-es; -i » -∅	-io > -e
Dat	-ie > -e	-iun > -en	-u, -i > -e, » -en	-un, -ion > -en	-i > -e, » -∅	-iun > -en
Instr	-iu > -e	-iun > -en	-u, -i > -e, » -en	-un, -ion > -en	-u > -e, » -∅	-iun > -en

[31] Personal pronouns maintained almost all of the same distinctions as in OS. However, the syncretisms in the third-person feminine singular were complete (Lasch 1914:213, 216).

Table 13. Continued

	Weak Declension					
	< n-stems		< n-stems		< n-stems	
	Masc.		Neut.		Fem.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-o > -e	-on > -en	-a > -e	-un > -en	-a > -e	-un > -en
Acc	-on > -en	-on > -en	-a > -e	-un > -en	-un > -en, » -e	-un > -en
Gen	-en	-ono > -en	-en	-ono > -en	-un > -en	-ono > -en
Dat	-en	-on > -en	-en	-on > -en	-un > -en	-on > -en
Instr	-en	-on > -en	-en	-on > -en	-un > -en	-on > -en

In its development from OS, MLG underwent many of the same sound changes as MHG did from OHG. As a result, many of the same case, gender, and number distinctions were neutralized. Unstressed short vowels were all reduced to [ə], usually spelled <e> (Lasch 1914:116-117). As in MHG, the merger of the Instr with the Dat in nouns can be attributed completely to this change. For masculine *a*-stems and neuter short *a*-stems, the Dat and Instr singular became syncretic with the Nom/Acc plural (except when it had a final -*s*), and the Gen plural joined this syncretism as well, e.g., masculine *a*-stem Dat singular *dage*, Instr *dagu*, Nom/Acc plural *daga*, Gen *dago* > *dage* 'day(s)'. For masculine and neuter *ja*-stems, the Nom/Acc singular also joined this syncretism, while only the Dat and Instr singular and Gen plural were involved for neuter long *a*-stems. For feminine *(j)ō*-stems, the Dat/Instr singular joined the syncretism among the Nom/Acc/Gen singular and Nom/Acc plural, e.g., *ō*-stem Dat singular *erthu*, Nom/Acc/Gen singular/Nom/Acc plural *ertha* > *erde* 'earth(s)'. For short *i*-stems, these forms were already all syncretic in OS, but they merged with the *(j)ō*-stem forms as a result of vowel reduction. The effects of vowel reduction on the weak classes were even more extreme than on the strong classes, expanding the existing syncretism in these paradigms: only the Nom singular, as well as the Acc for neuters, and Gen plural remained distinctive, e.g., masculine Acc singular *bodon*, Gen/Dat *boden*, Nom/Acc/Dat plural *bodon*, Gen *bodono* > *boden* 'messenger(s)'.

As in MHG, umlaut in the plural forms of most masculine long *i*-stems avoided syncretism between the Dat/Instr singular and Nom/Acc/Gen plural, even though vowel reduction resulted in the merger of the endings for these nouns with those of the *a*-stems by MLG, e.g., OS Dat singular *gaste*, Instr *gastu* > MLG *gaste* 'guest' vs. Nom/Acc plural *gesti*, Gen *gestio* > *geste*. For feminine long *i*-stems, the Instr singular and Gen plural joined the syncretism among the Dat, and sometimes Gen, singular and the Nom/Acc plural. Already in OS,

however, only the plural case markings triggered umlaut, while the Gen/Dat and Instr singular had the same stem as the Nom/Acc singular, e.g., Gen/Dat singular *ansti*, Instr *anstu* > *anste* 'favor' vs. Nom/Acc plural *ensti*, Gen plural *enstio* > *enste*. Thus, number syncretism was avoided in both OS and MLG for feminine long *i*-stems, except those that did not undergo any umlaut because the stem already contained a front vowel, e.g., MLG *schrift* 'writing'. The restriction of umlaut to the plural was phonologically predictable for masculine long *i*-stems but not for feminine long *i*-stems; thus, the morphologization of umlaut in accordance with the principles of iconicity and system adequacy seems to have begun earlier in the development of Low German than High German.

The loss of reduced *e* in the final syllable was much less consistent in MLG than in MHG, where it occurred regularly after liquids and nasals. In MLG, it sometimes occurred in these environments, but not regularly. In the Gen plural of weak nouns, the final vowel was consistently lost, thereby neutralizing the last remaining case distinction in the plural, e.g., Gen plural *bodono* > *boden* 'messengers'. In the Gen and Dat singular of masculine *a*-stems, but not the Nom/Acc plural, *e* was also lost occasionally after *t/d* or another vowel. In the Gen plural of masculine and neuter *a*-stems it was lost even more frequently (but still not consistently), apparently regardless of the preceding consonant (Lasch 1914:119-122, 193). It is difficult to determine the nature of these changes. Dialectal differences may have played a role, but there also seems to be a morphological factor, as with the later loss of Dat *-e* in the development of High German. Lasch notes that *e* was retained more often when it had grammatical value, specifically in distinguishing number (ibid., 119). However, a zero ending for the Gen plural made it syncretic with the Nom/Acc singular for *a*-stems, e.g., masculine *a*-stem Nom/Acc singular *cuning* > *köninc*, Gen plural *cuningo* » *köninc* 'king(s)'. This syncretism pattern was also found in Late Proto-Slavic (LPS) on the equivalent masculine class, *(j)o*-stems; these forms, as well as other Gen plural forms, developed a zero ending due to sound change, but the iconic principle was violated, and markers from other classes and cases were often adopted on the basis of their distinctive strength as the various Slavic languages developed. It is somewhat surprising, therefore, that a similar violation of the iconic principle might arise due to analogical processes or at least not be repaired by them. With respect to Lasch's observation above, it seems that, at least in Low German, analogical processes such as morphological blocking were more likely to

profile number marking within the same case, where syntactic context is less likely to help disambiguate, than among different cases.

Other changes to nominal inflection in MLG have a much clearer morphological factor. In fact, analogical processes such as leveling had a greater effect on MLG than MHG. Some of these are similar to processes that happened later in the development of High German. Already in OS, feminine long *i*-stems often had the Gen singular form *-es*. For feminine *i*-stems in MHG, this was mostly limited to the adverbial Gen *naht(e)s* 'at night', formed by analogy with the masculine *a*-stem *tag(e)s*' in the daytime', both of which survive as adverbs in CSHG (see Wright 1917, §49). The more general process in OS has parallels with the spread of Gen singular *-(e)s* to feminine strong nouns in the transition from Old English (OE) to Early Middle English (EME) and in the development of continental Scandinavian, where *i*-stems were also the first feminine nouns to be used with this form.[32] In MLG, *-es* continued to be an option for the Gen singular of feminine long *i*-stems. The other option was a zero ending, matching the Nom/Acc singular. The Dat singular form *-e* was optionally leveled to the Nom/Acc as well. This process would only have eliminated number syncretism among the Gen/Dat singular and Nom/Acc/Gen plural for nouns without umlaut, since this process was already limited to the plural in OS. Thus, the influence of other declension classes, i.e., system adequacy, may have played a greater role than the profiling of number marking. With *-es* in the Gen singular and *-e* or a zero ending in the Dat singular, feminine long *i*-stems followed the masculine *a*-stem inflectional pattern. With a zero ending in the Gen and Dat singular, they instead followed other feminine strong nouns, including short *i*-stems, in a complete lack of case distinctions in the singular. Therefore, it appears that there were competing influences from masculine and other feminine strong nouns on feminine long *i*-stems. The spread of Gen singular *-(e)s* to new classes continued in MLG. As in the development of High German, Lasch notes that some weak nouns adopted the Gen singular form *-ens*, adding *-s* to the inherited *-en* form by analogy with strong masculine nouns, e.g., *willen* » *willens* 'will' (1914:202). Thus, the *-s* Gen was apparently on its way to becoming an overstable marker, but as a result of later developments described below, it did not achieve this status to the extent that it did in many other Germanic languages.

For masculine *a*-stems in OS, Nom/Acc plural forms occurred both with a final *-s*, as in OE, and without *-s*, as in OHG. By MLG, however, *-e*, the regular reflex of the forms without *-s*,

[32] See section 3.1.2.2 for further discussion of this process in Germanic languages and its implications.

had largely displaced *-es*, the reflex of forms with *-s*, perhaps by analogy with the majority of other strong classes. A notable exception is masculine *ja*-stems, which often retained *-es*. Lasch suggests that this retention was motivated by the form's utility (1914:195). If utility is interpreted as distinctiveness, a comparison of masculine *a*-stems and *ja*-stems supports this conclusion. For masculine *a*-stems, Nom/Acc plural *-es* would have been syncretic with the Gen singular, e.g., *dages* 'day(s)', while *-e* was syncretic with the Dat singular and Gen plural, e.g., *dage* 'day(s)', except when these forms had zero endings, which was relatively often, as described above. Thus, the shift to *-e* either slightly decreased or slightly increased syncretism, depending on other changes. For masculine *ja*-stems, on the other hand, *-es* was only syncretic with the Gen singular, e.g., *kêses* 'cheese(s)', but *-e* was syncretic with the Nom/Acc/Dat singular and the Gen plural, e.g., *kêse* 'cheese(s)', clearly an increase in syncretism, including number syncretism within the same case. Thus, *-es* may have been retained for masculine *ja*-stems as part of the tendency for Germanic and other IE languages to profile number, which can be seen as a narrower application of the iconic principle, as mentioned in the discussion of MSw in section 2.3.2.2 above. Another manifestation of this tendency, as well as the principle of system adequacy, is the spread of umlaut to the plural of *a*-stems, including neuters, e.g., Nom/Acc/Gen plural *vete* alongside *vate,* from *vat* 'barrels', early in MLG; on the other hand, plural forms of masculine long *i*-stems without umlaut are also attested in MLG, e.g., *gaste,* from *gast* 'guests' (see Lasch 1914:47, 194, 197).

The early stages of the merger among the strong and weak feminine classes can be observed in MLG. The Nom/Acc plural form of feminine *(j)ō*-stems and short *i*-stems was optionally leveled to the Gen/Dat plural form, and for some of these nouns, this *-en* form spread to the Gen/Dat singular as well. The former change eliminated all number syncretism but also all case distinctions, while the latter restored a distinction between Nom/Acc and Gen/Dat in the singular but resulted in the same degree of number syncretism as the forms inherited regularly without analogical processes. Meanwhile, the Acc singular of feminine weak nouns sometimes used the Nom form *-e* instead of *-en*, which matched all other case/number forms. This decreased number syncretism slightly. When all of these changes applied, all feminine nouns except long *i*-stems followed the same paradigm. Unlike the slightly later merger of these classes in the development of High German, the resulting paradigm was much closer to the weak paradigm and still had a significant degree of number syncretism but some case distinctions. Thus, profiling

number marking does not appear to have motivated these developments; they only followed the iconic principle in the sense that the Nom/Acc plural form, which was marked in terms of number but not case, and the Gen/Dat singular form, which was marked in case but not number, both became more salient, even though they became less distinctive due to the increased syncretism. Other analogical developments in Low German, as well as most other Germanic languages, tend to prioritize iconicity for number marking over case marking, but the Gen singular is sometimes an exception, as with this MLG development. In this way, it is similar to the ongoing change to weak nouns in CSHG that is leveling the Acc and Dat singular, but not the Gen; for feminine nouns in MLG, the Dat seems to have patterned with the Gen, probably as a result of their consistent syncretism in other singular paradigms.

In OS, gender syncretism on NP-internal agreement targets (determiners and strong adjectives) was already quite high. Determiners and strong adjectives had distinctive Acc singular forms for each gender, but only determiners had distinctive Nom singular forms. Both failed to distinguish the masculine and neuter in the Gen and Dat singular, and neither reliably distinguished gender in the plural. Thus, the loss of gender distinctions in the plural was already well underway in OS, much earlier than in the development of High German. In MLG, the Acc singular retained distinctive forms for all three genders on most NP-internal agreement targets, but only the neuter was still reliably distinctive in the Nom singular of some determiners. Thus, gender syncretism increased slightly.

The modern period is considered to have begun around 1500 for Modern Low German (ModLG). Nominal inflection has been significantly reduced from MLG. ModLG dialects tend to distinguish two morphological cases: Nom and Acc; three genders; and two numbers. The remaining case distinction relies entirely on agreement targets; nouns themselves only inflect for number.[33]

[33] Personal pronouns consistently distinguish the Nom and Acc.

Table 14. Changes in Noun Declension from Middle Low German to Modern Low German

	Strong Declension			
	Masc.		Neut.	
	Sg.	Pl.	Sg.	Pl.
Nom	-Ø, -e > -Ø	-e > -Ø, » -en; -es > -s	-Ø, -e > -Ø	-Ø, -e > -Ø; » -en, -s
Acc	-Ø, -e > -Ø	-e > -Ø, » -en; -es > -s	-Ø, -e > -Ø	-Ø, -e > -Ø; » -en, -s
Gen	-es, -s » -Ø	-e, -Ø > -Ø, » -en, -s	-es » -Ø	-e > -Ø, » -en, -s
Dat	-e > -Ø	-en > -en, » -Ø, -s	-e > -Ø	-en > -en, » -Ø, -s

	Strong Declension	
	Fem.	
	Sg.	Pl.
Nom	-Ø, -e > -Ø	-en; -e » -en, -s
Acc	-Ø, -e > -Ø	-en; -e » -en, -s
Gen	-Ø, -e > -Ø; -en, -es » -Ø	-en; -e » -en, -s
Dat	-Ø, -e > -Ø; -en » -Ø	-en > -en, » -s

	Weak Declension					
	Masc.		Neut.		Fem.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-e > -Ø	-en	-e > -Ø	-en	-e > -Ø	-en
Acc	-en » -Ø	-en	-e > -Ø	-en	-en » -Ø; -e > -Ø	-en
Gen	-en » -Ø	-en	-en » -Ø	-en	-en » -Ø	-en
Dat	-en » -Ø	-en	-en » -Ø	-en	-en » -Ø	-en

In the development of ModLG from MLG, a few sound changes had a significant effect on nominal inflection. One change was the loss of *e* in the final syllable.[34] This had begun as an apparently sporadic process in MLG, as described above, but seems to have occurred more consistently over time. However, this change was preceded or accompanied by the lengthening of short vowels in open stressed syllables in MLG, as in Middle Dutch (MDu) (see van der Wal Wal & Quak 1994:92). This lengthening only seems to have applied on ModLG noun forms ending in *-e*, and not, for example, those ending in *-en*, which suggests that it was compensatory lengthening accompanying the loss of a syllable. This change helped maintain the distinctiveness of the noun forms to which it applied, i.e., those ending with a single consonant followed by *-e*, even after this *-e* was lost, e.g., strong masculine Nom/Acc singular *dag* vs. Dat singular/Nom/Acc/Gen plural *dage* > *daag* 'day(s)'. Some noun forms that had two consonants before *-e* underwent a different change, losing the final consonant along with *-e*, e.g., strong

[34] As a result of this sound change, final *-en* is generally realized as a syllabic [n]. Some orthographies omit the *e* to reflect this, but my analysis follows those that retain it (see, for example, Reershemius 2004).

neuter Nom/Acc singular *peerd* vs. Dat singular/Nom/Acc/Gen plural *peerde* > *peer* 'horse(s)'. This change has apparently been tolerated despite resulting in a plural form that is shorter than the singular, which goes against the iconic principle. Perhaps this is because it still profiles number marking. For masculine and neuter strong nouns that had not already undergone umlaut and lost *-e* without an accompanying change, however, the remaining distinctions among all forms were eliminated, except the Gen singular and Dat plural, e.g., strong neuter Nom/Acc singular/plural *jâr*, Dat singular/Gen plural *jâre* > *johr* 'year(s)'. In addition, the differences among the strong classes within each gender were mostly eliminated by the loss of *-e*, at least in the singular. Differences remained in the plural, but they often depended on which additional changes these forms had undergone as a result of their phonological form, i.e., open-syllable lengthening, final consonant loss, or umlaut. As described below, plural forms originally associated with particular declensions were extended to others, further breaking down the original system of declension classes.

Analogical processes eliminated all number syncretism for most nouns, but all case distinctions on nouns were lost as a result. Many nouns adopted more distinctive plural forms originally associated with other declension classes. The spread of umlaut to plural forms where it was not expected based on regular sound change, which began in MLG, has continued, e.g., Nom/Acc singular *hûs* > *huus*, Nom/Acc plural *hûs* » *hüüs* 'house(s)' (see Lasch 1914:47). Umlaut also spread to new nouns in the development of High German, but neuter strong nouns were more likely to shift to the *z*-stem declension, characterized by *-er* in the plural, which also triggered umlaut, e.g., Nom/Acc/Gen plural *Häuser* 'houses'. In ModLG, however, this class did not become productive as in High German (see Lasch 1914:198). The weak plural form *-en* was extended to some strong nouns, e.g., neuter Nom/Acc/Gen[35] *lant* » *lannen* 'lands'. For masculine and neuter strong nouns, this eliminated all number syncretism but also all case distinctions in the plural, since the Dat plural already had the *-en* form. For feminine strong nouns, this process had already begun in MLG with the partial merger of feminine strong and weak nouns, as described above, but it expanded to include feminine nouns derived from long *i*-stems, e.g., Nom/Acc/Gen *tîte* » *tieden* 'times'. Similarly, the plural form *-s* < *-es*, which was mostly limited to masculine *ja*-stems in MLG, spread to other strong nouns of all genders, e.g., neuter

[35] The Gen plural form was *lande* in MLG but after the loss of *-e* it was *lant* or *lan*, if the final consonant was deleted along with *-e*.

mechdeken » *mäkens* ‘girls’. This eliminated syncretism with all singular forms except Gen *-(e)s*, which was used for all masculine and neuter strong nouns, as well as some feminine strong nouns.

Any number syncretism remaining after the spread of more distinctive plural forms was eliminated by analogical leveling, along with all case distinctions. Dat singular forms that had joined plural forms in undergoing stem changes with the loss of *-e* were leveled to the uninflected form, e.g., strong masculine *daag* » *dag* ‘day’, strong neuter *peer* » *peerd*. The Gen singular form *-(e)s* was leveled to the uninflected form, e.g., strong masculine *dages* » *dag* ‘day’. This may have been motivated by the increased productivity of *-s* as a plural form, making confusion more likely. All singular forms with *-en* were also leveled to the uninflected form, e.g., weak neuter Gen/Dat *herten* » *hart* ‘heart’. As with *-s*, the increased productivity of *-en* as a plural form may have been a motivation. The leveling of these singular forms has parallels with the rise of the mixed declension and ongoing leveling of singular weak forms in CSHG. A notable difference is that the Gen forms have not been leveled in CSHG, possibly because *-(e)s* has not become a productive plural form in the same way that it has in ModLG, so it can still be associated primarily with the Gen singular. However, the weak Gen singular form *-en* has also been retained in CSHG despite the productivity of *-en* as a plural form. In Low German but not High German, therefore, there appears to have been a shift in how the iconic principle applies to the Gen singular. Perhaps the Gen lost its markedness—and therefore its need for a salient form—as it lost its syntactic productivity in favor of other constructions, even before it was no longer distinctive on nouns, as in spoken Faroese. This is supported by the fact that the Gen also does not survive on agreement targets, where number syncretism was likely much less of a factor. Moreover, the Acc is still distinct from the Nom on masculine singular agreement targets even though this distinction has been completely lost on nouns, so it would have been possible for the Gen to survive in this way as well if there had not been a morphosyntactic motivation for its loss. The only distinctive case form that remained after all of these changes was the Dat plural form *-en*, but only on nouns that did not have this as the general plural form. Since this is one of the most common general plural forms, there was likely a lot of pressure for the Dat plural to be leveled on those nouns where it remained distinctive, especially given the lack of a distinctive Dat form in the singular. With the leveling of the Dat plural, all case distinctions were lost on nouns, e.g., Dat plural *dagen* » *daag* ‘days’.

In ModLG, there are still some gender distinctions on NP-internal agreement targets (determiners and strong adjectives). The definite article still has distinctive forms for all three genders in the Acc singular and a distinctive neuter form in the Nom singular. Strong adjectives are often uninflected, as they already generally were in the Nom singular in OS and MLG. When they are inflected, however, all three genders can be distinctive in the Acc singular but also the Nom, due to the adoption of the Acc forms. In the plural, the neuter can be distinctive because it always remains uninflected. Thus, the genders cannot be reliably distinguished in the plural, but the neuter is sometimes distinctive, an innovation that appears to reverse the complete loss of gender distinctions in the plural that began in OS and appeared complete in MLG.

Table 15. Case Mergers in Low German

	Old Saxon					Middle Low German				Modern Low German	
PIE	Nom	Acc	Gen	Dat	Instr	Nom	Acc	Gen	Dat	Nom	Acc
Nom	x					x				x	
Voc	x					x				x	
Acc		x					x				x
Gen			x					x			x
Dat				x					x		x
Loc				x					x		x
Instr					x				x		x
Abl				x					x		x

The Instr was lost in the development of MLG from OS as in MHG from OHG. The Gen was most likely the next case lost. Scott (2014:250) states that many low German dialects, including some ModLG dialects, have retained the Dat, but almost all lack a productive Gen case. Thus, the Gen was most likely lost before the Dat in dialects that only retain a distinction between Nom and Acc.

2.3.3.3. The Development of Dutch[36]

Old Dutch (ODu), also known as Old Low Franconian, was used from the 8th century to the 12th century. Generally, ODu distinguished four cases morphologically: Nom, Acc, Gen, Dat; three genders; and two numbers. The Instr merged with the Dat before ODu stage. Written evidence for ODu comes from a very limited corpus of fragmentary texts. These appear to reflect the

[36] The grammatical information is drawn from van der Wal and Quak (1994) for ODu and MDu, De Schutter (1994) for CSDu.

neutralization of phonemic vowel quantity in unstressed syllables and the shift of final *-m* to *-n*, as was generally the case for OS but not OHG. While these two changes increased syncretism for masculine weak nouns in OS, the additional merger of unstressed /u/ with /o/ in ODu increased syncretism for feminine and neuter weak nouns as well. The Dat plural became syncretic with the Nom/Acc for weak nouns of all genders. The Acc singular joined this syncretism for masculine weak nouns, as in OS. Feminine weak nouns were affected the most, however, as the Acc/Gen/Dat singular also joined the syncretism in the plural, e.g., Acc/Gen/Dat singular/Nom/Acc/Dat plural *tungon* 'tongue(s)'. Unlike other West Germanic languages of the same period, there is no evidence for a distinctive Instr in ODu.

MDu was used from the 12th to 15th centuries and maintained most of the same general distinctions in nominal inflection as ODu. It distinguished four morphological cases: Nom, Acc, Gen, Dat; three genders; and two numbers. However, Nom and Acc were never distinguished on nouns, only on determiners, adjectives, and pronouns (see van der Wal & Quak 1994:72, 75-77).[37]

Table 16. Changes in Noun Declension from Old Dutch to Middle Dutch

	Strong Declension					
	< a-/i-stems		< a-stems		< i-stems	
	Masc.		Neut.		Fem.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-Ø	-a, -i > -e; -as > -es, -s	-Ø	-Ø > -Ø, » -e, -en; -o > -e	-Ø	-i > -e
Acc	-Ø	-a, -i > -e; -as > -es, -s	-Ø	-Ø > -Ø, » -e, -en; -o > -e	-Ø	-i > -e
Gen	-is > -es, -s	-o > -e	-is > -es, -s	-o > -e	-i > -e, » -Ø	-o > -e
Dat	-i > -e	-on, -in > -en	-i > -e	-on > -en	-i > -e, » -Ø	-in > -en

[37] Personal pronouns maintained more distinctions than other paradigms in Middle Dutch. The first- and second-person pronouns, along with third-person plural and masculine singular pronouns, distinguished three case forms, with no distinction between Acc and Dat. However, the clitic form of the third-person masculine singular Acc can be distinct from the Dat, and the Acc form of the third-person plural clitic, along with the third-person neuter singular clitic, was syncretic with the Nom rather than Dat. The third-person neuter singular lacked a Gen form, except as a clitic. The third-person feminine singular only made two distinctions: Gen and Dat were consistently the same, and Acc also had this same form, except as a clitic, which had a syncretic Nom/Acc form (see van der Wal & Quak 1994:77-78).

Table 16. Continued

	Weak Declension					
	< n-stems Masc.		< n-stems Neut.		< n-/ō-stems Fem.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-o > -e	-on > -en	-a > -e	-on > -en	-a > -e	-on > -en, -a » -en
Acc	-on » -e	-on > -en	-a > -e	-on > -en	-a > -e, -on » -e	-on > -en, -a » -en
Gen	-in > -en	-ono > -en	-in > -en	-ono > -en	-on > -en, » -e	-ono > -en
Dat	-in » -e	-on > -en	-in » -e	-on > -en	-on > -en, -o > -e	-on > -en

In its development from ODu, MDu underwent many of the same sound changes as other West Germanic languages. As a result, many of the same case, gender, and number distinctions were neutralized. Unstressed vowels were all reduced to [ə], spelled <e> (see van der Wal & Quak 1994:74). For masculine *a*-stems and *i*-stems, as well as neuter *a*-stems, the Dat singular became syncretic with the Nom/Acc plural (except when it had a final *-s*), and the Gen plural joined this syncretism as well, e.g., masculine *a*-stem Dat singular *dagi*, Nom/Acc plural *daga*, Gen *dago* > *dage* 'day(s)'. In addition, this vowel reduction resulted in the merger of the masculine *a*-stem and *i*-stem paradigms, which only differed in the Nom/Acc plural in ODu. For feminine *i*-stems, the Gen plural joined the syncretism among the Dat/Gen singular and the Nom/Acc plural, e.g., Gen/Dat singular/Nom/Acc plural *dādi*, Gen plural *dādo* > *dade* 'action(s)'.

A number of analogical processes had an effect on MDu nominal inflection. As mentioned above, feminine *ō*-stems and weak nouns had already partially merged in ODu. By MDu, the merger was mostly complete; thus, this development was more advanced than in MLG, which in turn was more advanced than in MHG. The *ō*-stem Acc singular form *-e* was also adopted by weak feminine nouns, removing this form from the syncretism among Gen/Dat singular and Nom/Acc/Dat plural. The weak Nom/Acc plural form *-en* came to be used consistently for *ō*-stems as well. This change removed the syncretism among the Nom/Acc singular and plural, as well as the Gen and Dat singular, when they had the *-e* form. Although the form with *-n* was dominant for the Gen singular in ODu, this was often leveled to *-e* in MDu by analogy with the other singular forms, including the Dat, which had inherited the form without *-n* from *ō*-stems, e.g., Gen singular *erden* » *erde*, Nom/Acc/Dat singular *erde* 'earth'. At the same time, the Dat singular sometimes retained the *-en* form that it had inherited from weak nouns. When the Gen/Dat singular had the *-e* form, all number syncretism was eliminated but so were all case distinctions. A feminine class derived from *i*-stems remained distinct from the merged

class described above. Likely by analogy with the leveling of other feminine nouns, however, the Gen/Dat singular form *-e* was optionally leveled to the zero ending of the Nom/Acc singular. This eliminated all number syncretism in this class, as well as all case distinctions in the singular. Masculine and neuter weak nouns were also influenced by the strong classes, although they did not merge completely. The Dat singular forms with final *-n* were leveled to the Nom, as was the Acc singular for masculines.[38] This matched the pattern in strong masculine nouns of an opposition in the singular between the Gen and a Nom/Acc/Dat form, increasing system adequacy. It also reduced number syncretism; now only the Gen was syncretic with the *-en* plural form. As a result, the last distinction between Nom and Acc was eliminated for nouns. A very similar process is ongoing in CSHG; the same motivations, including iconicity and system adequacy, have been involved in both languages, but they appear to have triggered analogical changes much earlier in Dutch than in High German.

There were several other analogical processes affecting nominal inflection in MDu. Neuter long *a*-stems, which had a zero ending in the Nom/Acc plural in ODu, often adopted the *-e* ending used by other *a*-stems (neuter and masculine). This eliminated syncretism among the Nom/Acc singular and plural, but instead made the Nom/Acc plural join the syncretism between the Dat singular and Gen plural, e.g., Nom/Acc plural *boec* » *boece*, Dat singular/Gen plural *boece* 'book(s)'. This decreased the overall number of forms involved in number syncretism; it may also reflect a preference for number syncretism between different cases, as noted with respect to MLG in the previous subsection. However, these nouns also increasingly adopted the Nom/Acc plural form *-en* by analogy with weak nouns. This made the Nom/Acc plural syncretic with the Dat plural, e.g., Nom/Acc plural *boec* » *boecen*, Dat *boecen,* but reduced number syncretism, leaving only the syncretism between the Dat singular and Gen plural. In accordance with the iconic principle, these processes both increased the salience of plural forms; they also reduced the differences among the neuter classes, but these did not merge entirely. As in MSw, and MDan, the formal similarity among masculine and feminine forms seems to lead the eventual merger of these genders into a common gender.

There is not enough evidence for agreement targets in ODu, but they were probably similar to those in OS in terms of gender syncretism, just as MDu was very similar to MLG in this respect. By MDu, the syntactic distinction between strong and weak adjectives had been lost.

[38] Weak neuters, like all neuters in IE languages, already had Nom/Acc syncretism.

The strong forms were generally used except in the Nom/Acc singular, where weak forms were used (if the strong and weak forms were not already the same anyway). In addition, the weak masculine/neuter Gen form *-en* was used alongside the strong form *-s*, e.g., *goets*/*goeden* 'good' (see van der Wal & Quak 1994:75). Determiners and adjectives had distinctive Acc singular forms for each gender, and determiners also had a distinctive Nom singular form for the neuter. Both had distinctive Gen and Dat singular forms for the feminine, but neither distinguished gender in the plural reliably. Thus, the situation in MDu was more or less the same as in MLG, with the complete loss of gender distinctions in the plural. In addition, the Acc singular was the only core case form that distinguished the masculine and feminine, and this last distinction was later lost in most dialects, as described below.

The modern period is also considered to have begun in 1500 for Dutch. All case distinctions have been lost for nouns, but there are still two numbers (see De Schutter 1994:460-462). As in CSS and CSD, in Contemporary Standard Dutch (CSDu), the masculine and feminine have merged into a common gender in opposition to the neuter. As in English, CSS, and CSDan, pronouns still have forms for three genders, but they are assigned on a purely semantic basis, i.e., based on natural gender. However, some southern dialects of Dutch retain three grammatical genders, which are marked on articles and adjectives as well as pronouns (see De Vos & De Vogelaer 2011:248-249). As in other Germanic languages that have completely lost morphological case, the functions of direct address, subject, complement of 'be', DO, and object of prepositions are all expressed with bare nouns. Bare nouns are in partially free variation with prepositional constructions for IOs. Adnominal possession is expressed by a prepositional construction with *van* 'of, from', the *-s* Gen (limited to proper names and kinship terms), and resumptive possessive determiners, e.g., *z'n* 'his' in *de boer z'n kinderen* 'the farmer's children'. The latter construction developed in MDu, originally in combination with possessors in the Gen or Dat. It has now become the preferred option for animate possessors in spoken Dutch. The other former case functions are generally expressed with prepositional constructions (De Schutter 1994:459, 467-468, Norde 2007:56-58).

In the early modern period, the most significant sound change for nominal inflection was the regular deletion of final *-n* after [ə], although this is generally not reflected in writing (see De Schutter 1994:452). This neutralized any remaining distinctions on weak nouns, at least as long as they developed regularly. In addition, the distinction between the masculine Nom and Acc/Dat

forms was lost on determiners and strong adjectives, e.g., Nom *die goede*, Acc/Dat *dien goeden* > *de goede* 'the good', thereby eliminating the last Nom-Acc distinction. The loss of *-n* also neutralized the distinction between Nom/Acc/Gen and Dat in the plural of agreement targets, as well as strong nouns that had Nom/Acc plural *-e*, e.g., masculine *a*-stem Nom/Acc/Gen *dage*, Dat *dagen* > *dage* 'days'; only strong nouns with Nom/Acc plural *-(e)s* retained distinct Dat and Gen plural forms after this change, e.g., masculine *a*-stem Nom/Acc *duivels*, Gen *duivele*, Dat *duivelen* 'devils'.

Analogical changes eliminated the remaining number syncretism and case distinctions. The *-(e)s* form has spread from masculine *a*-stems to nouns derived from other classes. In CSDu, the plural of the vast majority of nouns is prosodically determined: *-s* for polysyllabic nouns and *-e(n)* for monosyllabic nouns (see Kürschner & Nübling 2011:375-376, De Schutter 1994:458-459). Possibly to avoid syncretism with the *-e(n)* plural, Dat singular *-e* was leveled to the zero ending of the Nom/Acc, e.g., masculine *a*-stem *dage* » *dag*. As noted by De Vos and De Vogelaer (2011:248), however, some southern (or Flemish) dialects did not undergo the deletion of final *-n* to the same extent, and these also tend to retain the distinction between masculine and feminine. In the Moerzeke dialect, for example, masculine articles and adjectives end in *-n* when the following word begins with a vowel, /b/, /t/, or /d/. In these contexts, at least, the masculine and feminine are distinctive. These forms also suggest that the Nom singular was ultimately leveled to the Acc, not vice versa, in dialects where sound change did not fully eliminate the distinction. This has parallels in continental Scandinavian and the Romance languages, as described in sections 2.3.2.2 above and 2.4 below. In Low German, the distinction between masculine and feminine agreement targets relies on the presence or absence, respectively, of *-n* in the Acc; this adds further support for the role of *-n* loss in the merger of the masculine and feminine in CSDu.

Table 17. Case Mergers in Dutch

	Old Dutch				Middle Dutch		
PIE	Nom	Acc	Gen	Dat	Nom	Gen	Dat
Nom	x				x		
Voc	x				x		
Acc		x			x		
Gen			x			x	
Dat				x			x
Loc				x			x
Instr				x			x
Abl				x			x

As mentioned above, the distinction between the Nom and Acc had been completely lost for nouns in MDu. Only masculine agreement targets maintained this distinction, until they also lost it due to sound change in most dialects. With the adoption of *-e(n)* as the most common plural form in all cases, the Dat plural was distinctive in fewer nouns (see van der Wal & Quak 1994:75), and the Dat was almost certainly the next case marking to be lost. Beginning in MDu and continuing through the modern period, the Gen used adnominally and survived the longest. The adnominal use survived the longest, at least until 18th century, but still ceased to be used productively by the early 20th century at the latest (see Scott 2014:157, 311-315).[39]

2.3.3.4. The Development of English[40]

Old English (OE) was attested from around 600 to 1100. Generally, OE fully distinguished four cases morphologically: Nom, Acc, Gen, Dat; three genders; and two numbers.[41] In addition, the distinctive PGmc Instr was retained for the singular masculine and neuter forms of demonstratives, interrogatives, and strong adjectives. Otherwise, the Instr had become syncretic with the Dat before English emerged. By Late Old English (LOE), sound changes had resulted in

[39] The gradual decline of the Gen in Dutch is also discussed in section 3.1.2.7.

[40] The grammatical information is drawn from Allen (1995), Marsden (2010), and Kemanade (1994) for OE, Irvine (2004), Moore (1928) for LOE, Allen (2008), Campbell (1959), and Minkova (1983) for EME.

[41] There was also a dual for first- and second-person pronouns, creating a three-way number distinction, as in PGmc. As in the other old Germanic languages, personal pronouns had three or four distinctive case forms. The first- and second-person pronouns distinguished at least three forms. In earlier texts, especially poetry, distinctive Acc forms are attested for the singular and plural, as in PGmc, but not the dual. Later, the Dat form came to be used instead of the Acc for all numbers. The third-person masculine singular pronouns distinguished all four cases, while third-person feminine singular pronouns had the Gen/Dat syncretism. The third-person plural and neuter singular pronouns had the same form for Nom and Acc (see Marsden 2010:357-358, Kemanade 1994:121).

a number of changes to nominal inflection. Although syncretism increased and the Dat assumed the functions of the Instr, as it had earlier in the North Germanic languages, the same general morphological distinctions in nominal inflection were retained in LOE (see Marsden 2010:xv, 1, 355-378, Kemanade 1994:110, 119-120, Irvine 2004:cxxxix-clvii).

Table 18. Changes in Noun Declension from Old English to Late Old English

	Strong Declension					
	< a-stems		< ja-stems/i-stems		< long a-stems	
	Masc.		Masc.		Neut.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-Ø	-as	-e	-as	-Ø	-Ø
Acc	-Ø	-as	-e	-as	-Ø	-Ø
Gen	-es	-a	-es	-a	-es	-a
Dat	-e	-um > -um, -an, -on	-e	-um > -um, -an, -on	-e	-um > -um, -an, -on

	Strong Declension					
	< short a-stems		< ja-stems/i-stems		< ō-stems/i-stems	
	Neut.		Neut.		Fem.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-Ø	-u > -u, -o, -a	-e	-u > -u, -o, -a	-Ø; -u > -u, -o, -a	-a, -e
Acc	-Ø	-u > -u, -o, -a	-e	-u > -u, -o, -a	-e, -Ø	-a, -e
Gen	-es	-a	-es	-a	-e	-a > -a, » -ena
Dat	-e	-um > -um, -an, -on	-e	-um > -um, -an, -on	-e	-um > -um, -an, -on

	Weak Declension					
	< n-stems		< n-stems		< n-stems	
	Masc.		Fem.		Neut.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-a	-an	-e	-an	-e	-an
Acc	-an	-an	-an	-an	-e	-an
Gen	-an	-ena	-an	-ena	-an	-ena
Dat	-an	-um > -um, -an, -on	-an	-um > -um, -an, -on	-an	-um > -um, -an, -on

Several sound changes significantly affected nominal inflection in the transition from OE to ME. Due to a series of mergers during the OE period, only a front-back distinction remained for unstressed vowels by LOE (Campbell 1959, §377-378). For neuter *ja*-stems, *i*-stems, and short *a*-stems, this resulted in syncretism among Nom/Acc and Gen plural, e.g., short *a*-stem Nom/Acc plural *scipu*, Gen *scipa* > *scipa* ‘ships’. For feminine short *ō*-stems, this made the Nom singular syncretic with the Gen plural, as well as the Nom/Acc plural when the *-a* form was used, e.g., Nom singular *giefu*, Nom/Acc/Gen plural *giefa* > *giefa* ‘gift(s)’. Perhaps to avoid this

syncretism, and certainly in keeping with the iconic principle, the Gen plural form *-ena*, an extension from weak feminine nouns, was sometimes used instead in LOE. In the earliest attested stages of the continental West Germanic languages, the weak Gen plural ending had already been adopted by most feminine strong classes except *i*-stems, so this change was actually relatively late in OE.

By LOE, word-final *-m* had changed to *-n* (Moore 1928:243), a development that occurred around the same time in other West Germanic languages. This was still not consistently reflected in the orthography, but I assume here that morphological distinctions relying on this phonological contrast were consistently neutralized. In combination with the merger of unstressed back vowels, it caused the Dat plural of weak nouns to join the syncretism among the Gen/Dat singular and Nom/Acc plural, as well as the Acc singular for masculine and feminine weak nouns, e.g., masculine Acc/Gen/Dat singular/Nom/Acc plural *naman*, Dat plural *namum* > *naman* 'name'.

Middle English (ME) was spoken from around 1100 to 1500. Early Middle English (EME) (1100-1300) represents a time of transition during which the morphological case and gender systems broke down. Many noun declensions lost productivity, and their members were absorbed by one of the few remaining classes. Before these systems were completely lost, most of the same general distinctions made by OE could be found in at least one of the remaining paradigms. In this sense, it distinguished four morphological cases: Nom, Acc, Gen, Dat; three genders; and two numbers.[42] However, Nom and Acc were never distinguished on nouns, only on determiners, strong adjectives, and pronouns. In addition, the Instr had been completely lost by EME. Gender was still generally marked on determiners, but often the gender indicated was different from the historical gender (see Burrow & Turville-Petre 2007:20-23, Kemanade

[42] Personal pronouns mostly maintained the same distinctions as in OE. The distinction between Acc and Dat in the first- and second-person had been completely lost in favor of the Dat and the same process was in progress in the third-person masculine singular (see Burrow & Turville-Petre 2007:24-25). All pronouns lost this distinction by the 14th century (see, for example, van Gelderen 2000:211-221). For the third-person feminine, neuter, and plural, this profiled the distinction between subjects and objects, since the Nom and Acc forms were syncretic, but the Dat was distinctive. In the paradigms where the change had occurred earlier, however, all three forms were distinctive, so there must have been another motivation. Earlier, the third-person feminine singular pronouns usually had two distinctive forms: Nom/Acc and Gen/Dat. However, alternative forms specific to the Nom and Gen occurred in some texts, so it had become theoretically possible to make four distinctions, more than in OE. These texts may also have a distinctive Nom form of the third-person plural, increasing the number of possible distinctions to four (ibid, 24-26).

1994:110, 120). In the First Continuation of the *Peterborough Chronicle*,[43] for example, the historically neuter noun *godspel* 'gospel' and the historically feminine noun *niht* 'night' are attested with the masculine determiners *þone* and *þes*, respectively, and there are many other similar examples (see Irvine 2004:cxliv). As mentioned below and discussed further in section 3.1.2.2 below, this confusion around gender was likely connected to the reduction and reorganization of declension classes.

Table 19. Changes in Noun Declension from Late Old English to Early Middle English

	Strong Declension				Weak Declension	
	< Masc. a-stems		< short a-stems		< n-stems	
	All Genders		Neut.		All Genders	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-Ø	-as > -es	-Ø	-Ø	-a, -e > -e	-an » -en
Acc	-Ø	-as > -es	-Ø	-Ø	-an, -e > -e	-an » -en
Gen	-es	-a > -e, » -es, -ene	-es	-a > -e, » -es, -ene	-an > -e	-ena > -ene, » -en
Dat	-e > -e, » -Ø	-um, -an, -on > -e, » -es, -Ø	-e > -e, » -Ø	-um, -an, -on > -e, » -es, -Ø	-an > -e	-um » -en

Sound changes continued to affect nominal inflection in EME. Even the front-back distinction for unstressed vowels was lost by the 11th century (Campbell 1959, §379). All vowels in unstressed syllables were spelled <e> by this stage, representing [ə] (Minkova 1983:192). This neutralized many case, gender, and number distinctions. For strong nouns, it neutralized the distinction between Dat singular and Gen plural, e.g., masculine *a*-stem Dat singular *stāne*, Gen plural *stāna* > *stane* 'stone(s)'. The paradigm derived from masculine *a*-stems also acquired number syncretism between the Gen singular and Nom/Acc plural, e.g., Gen singular *stānes*, Nom/Acc plural *stānas* > *stanes*.

Later, final *-n*, including original *-m*, was lost in unstressed syllables (see Lightfoot 1999:137, Moore 1928:244). This caused the Dat plural to join the syncretism between Dat singular and Gen plural for strong nouns, e.g., masculine *a*-stem Dat plural *stānum*/*stānan*/*stānon* > *stane*. In combination with vowel reduction, it also neutralized all case and gender distinctions for singular weak nouns, e.g., masculine Nom *nama*, Acc/Gen/Dat *naman* > *name* 'name' (see Allen 1995:206). However, as with *-r* loss in continental Scandinavian, final nasal loss was not a

[43] The First Continuation and Second Continuation of the *Peterborough Chronicle* are additional annals for the years 1121-1131 and 1132-1154, respectively, written at the end of the copy of the *Anglo-Saxon Chronicle* made at the Peterborough Abbey in 1121. The main section is considered LOE, the continuations EME (see Irvine 2004, Allen 2008:126).

completely regular sound change. It did not apply to nasals at the end of roots, as the word *seven* < OE *seofon* attests, or derivational suffixes, as in *stolen* < OE *stolen*. Despite this variability, my investigation treats the loss of final unstressed nasals as regular sound change and their retention as morphological blocking. Thus, morphological blocking restored the distinction among the Nom/Acc plural forms and the oblique singular forms of weak nouns, e.g., masculine Nom/Acc plural *naman* » *namen*, and prevented the neutralization of the Dat plural, in contrast to its regular development in strong nouns. These results were similar to the analogical leveling of singular weak forms in the continental West Germanic language, and the motivations were likely the same, even if the process differed slightly.

Other analogical processes have also applied, and the number of declension classes has been drastically reduced, with different effects on case, number, and gender syncretism. Two other analogical processes affecting number syncretism sometimes applied to strong nouns. As in other Germanic languages, the Dat singular *-e* was often leveled to the uninflected Nom/Acc singular form, but the Dat plural *-e* was often leveled as well, to the Nom/Acc plural form. If this was *-es*, then number syncretism decreased as a result, restoring the distinctiveness of the Gen plural. On nouns with an uninflected plural form, i.e., those descended from neuter *a*-stems, the same change increased number syncretism, but as in other Germanic languages these nouns could not resist this general process. The distinctiveness of the Gen plural was also retained by the extension of the weak form *-ene* to more strong nouns; previously, the weak Gen plural form had sometimes been extended to feminine strong nouns, but not masculine or neuter strong nouns. Thus, it had become something of an overstable marker, although not to the same degree as *-(e)s* in the Gen singular. Eventually, however, the preferred option became the leveling of all plural forms to the Nom/Acc, even though it led to increased syncretism with the Gen singular. In the sense that the Possessive *-'s* marker continues the strong Gen singular, this syncretism survives in Modern English (ModE).

Gender syncretism on agreement targets increased significantly between OE and EME. The masculine and neuter were neutralized in the Gen and Dat singular of most NP-internal agreement targets (determiners and strong adjectives) in OE, as well as in the Nom singular of strong adjectives. Determiners lacked gender distinctions in the plural throughout OE, while strong adjectives had lost gender distinctions in the Nom/Acc plural by LOE. Thus, the loss of gender distinctions in the plural was nearly complete by the end of the OE period, and

distinctions were limited even in the singular. The loss in the plural happened around the same time as in the development of Low German and Dutch.

However, the eventual loss of all gender distinctions in the singular as well as the plural is unique to English among the Germanic languages under investigation here.[44] This is likely connected to the extension of inflectional endings beyond their original declension classes, although the causality is unclear, as discussed in section 3.1.2.2 below. Masculine Acc *-ne* (used on agreement targets) and Gen *-(e)s*, unambiguous markers of case in the singular, spread to all genders in certain manuscripts. In other words, they were used as overstable markers based on their distinctive strength. This change in function is reflected in the distinction between *the* and *that* in ModE, which derive from the masculine and neuter Nom singular forms of the OE distal demonstrative. Similarly, as the Gen underwent functional narrowing and morphological case was lost in general, Gen *-(e)s*, which had already become an overstable marker, could be reanalyzed as a clitic/phrasal affix, another example of a syntactically more transparent function. Of course, gender distinctions still remain on pronouns in English, but they have been semanticized as in other languages that lack additional indications of gender.

By Late Middle English (LME) (1300-1500), all case and gender distinctions had been lost for nouns, adjectives, and determiners. These only distinguished two numbers. The definite article lost this distinction early, and adjectives eventually lost this distinction as well (see Burrow & Turville-Petre 2007:20-29). The phrasal *-'s* Possessive developed from the masculine/neuter singular Gen (see Allen 2002).

The modern stage of English is considered to have begun in 1500. ModE retains a number distinction on most nouns. As in other Germanic languages that have completely lost morphological case, the functions of direct address, subject, complement of 'be', DO, and object of prepositions are all expressed with bare nouns. Bare nouns and the phrasal *-'s* Possessive are in partially free variation with prepositional constructions for IOs and adnominal possession, respectively. The other former case functions are generally expressed with prepositional constructions.

[44] Afrikaans has also lost all gender distinctions (see Donaldson 1994:486).

Table 20. Case Mergers in English

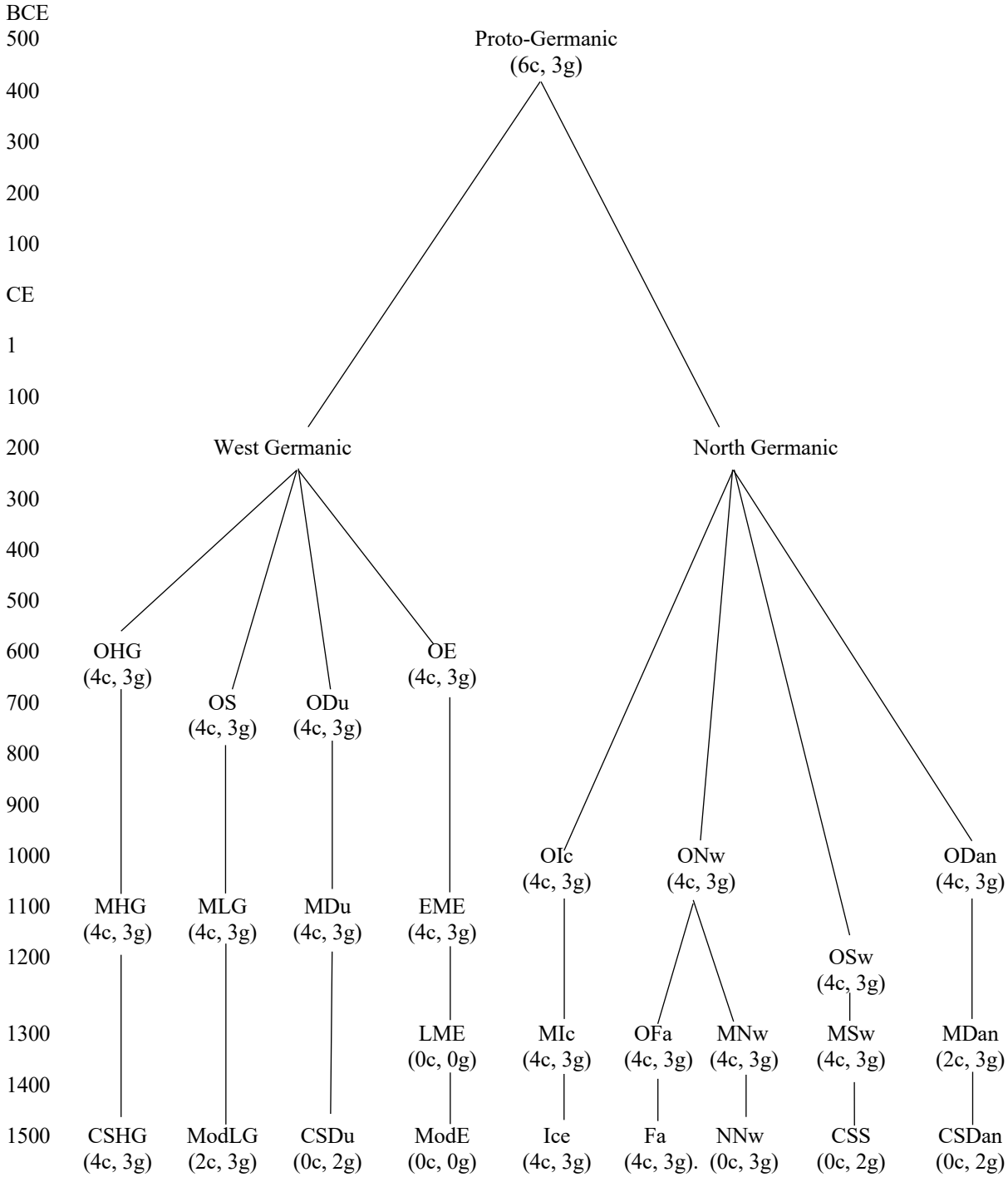
	Old English					Late Old English / Early Middle English				Middle English	
PIE	Nom	Acc	Gen	Dat	Instr	Nom	Acc	Gen	Dat	Nom	Gen
Nom	x					x				x	
Voc	x					x				x	
Acc		x					x			x	
Gen			x					x			x
Dat				x					x	x	
Loc				x					x	x	
Instr					x				x	x	
Abl				x					x	x	

In OE, the Instr still had a few distinctive forms on singular determiners, e.g., masculine/neuter Instr *þȳ*, *þī*, *þon*, Dat *þǣm*, *þām* 'the, that', adjectives, e.g., masculine/neuter Instr *dole*, Dat *dolum* 'foolish' and interrogative pronouns, e.g., Instr *hwȳ*, *hwī*, *hwon*, Dat *hwǣm*, *hwām* 'who, what', but it was already completely syncretic with the Dat on nouns, e.g., masculine *a*-stem Dat/Instr *stāne* 'stone'. In the development of ME nominal inflection from OE, the distinction between Nom and Acc was the next lost. Allen (1995:163-165) points out that the loss of a distinctive oblique form in the weak singular paradigm due to sound change left only singular feminine *ō*-stems with this distinction. However, this distinction could not be maintained on so few nouns and the Nom form was leveled to the other singular forms. Allen (2008:150-151) states that the Dat merged with the Acc by EME. As shown in Table 19, the Dat merged with the Nom/Acc in the singular and sometimes in the plural by analogy, although it regularly merged with the Gen in the plural of *a*-stems. Allen (2008:151) points out that the Gen was lost with the reanalysis of the *-(e)s* form as a clitic/phrasal affix. Evidence for both orders of case loss can be found in a comparison of two EME texts. In my investigation in chapter V, I found that Layamon's Brut retains four distinctive cases in productive use on nouns, The Owl and the Nightingale has lost the Nom-Acc distinction.

2.3.4. Summary

This section has discussed case and gender developments in the Germanic languages. The facts of these changes are summarized in Table 21 below.

Table 21. Timeline Summary of Case and Gender in Germanic Languages



All modern Germanic languages have experienced some degree of case loss since PGmc, but some have lost morphological case marking entirely while others maintain relatively robust case systems. The Instr was the first case lost in the development of all Germanic languages from PGmc. The order in which the other cases were lost varies, but generally the distinction between Nom and Acc was lost next, followed by the Dat or Gen. Likewise, some Germanic languages have completely lost grammatical gender while others still have the three genders of PIE. In those with two genders, the masculine and feminine have merged into a common gender, in opposition with the neuter.

Many of the lost case and gender distinctions can be attributed to sound change. Conversely, Icelandic and Faroese have not lost any distinctions in nominal inflection due to sound change, and the only distinction lost in these languages was the Gen, in spoken Faroese. Vowel reduction in unstressed syllables has had an effect on nominal inflection in all of the other Germanic languages. In MSw and MNw, phonetic vowel reduction occurred, but phonemic distinctions were mostly retained, so its effects were relatively minor; distinctions were only neutralized in combination with other sound changes. Although the nature of the reduction was different, similarly minor effects resulted in LOE, but by EME, as well as the middle stages of the remaining Germanic languages, all distinctions among unstressed vowels were neutralized. Vowel reduction also increased number syncretism for strong classes to some extent, as well as overlap among these classes. In the West Germanic languages, the Nom remained distinct from the oblique cases on weak nouns but increases in number syncretism on strong nouns occurred as in MDan. For feminine *(j)ō*-stems, all distinctions among the singular cases and the Nom/Acc plural were lost. In ODu and OE, the Instr had already merged with the Dat on nouns; this merger was completed in MHG and MLG due to vowel reduction.

In combination with vowel reduction, other sound changes led to additional neutralizations in nominal inflection. Final *-r* was generally lost in the middle stages of the continental Scandinavian languages. This neutralized any remaining distinctions between Nom and Acc plural, and sometimes also with the Gen plural. In addition, it increased number syncretism greatly for weak nouns and to a lesser extent for feminine (and some masculine) strong nouns. In the West Germanic languages, sound changes involving final nasals were the most significant for nominal inflection other than vowel reduction. Final *-m* was merging with final *-n* in the old stages and this process was complete in the middle stages. As a result, the Dat

plural was no longer distinct from the Nom/Acc plural for weak nouns. Except in EME, the final unstressed vowel after *n* was also lost in the weak Gen plural, so no case distinctions remained for weak nouns in the plural. In MHG, this was part of a more general process by which final unstressed vowels were lost after nasals and liquids. This increased number syncretism further for nouns ending in these consonants in their uninflected forms. Final *-n* was lost in EME and in the development of CSDu, resulting in the loss of all case distinctions in the singular of weak nouns, as in MDan. It also neutralized the distinction between the Dat and Gen plural for strong nouns, along with Nom/Acc plural forms without *-s*. Later, final *-e* was lost more generally in ModLG and LME. This may have contributed to the loss of the strong Dat singular form in both, or it may have occurred too late to make much difference; in Low German, it would have led to the loss of the number distinction for many strong nouns but stem changes such as umlaut and compensatory lengthening often accompanied this change.

Umlaut occurred as a phonological process in the early development of all Germanic languages. However, it was morphologized—in accordance with the principles of iconicity and system adequacy—to a much greater degree in some, most notably High and Low German. It was already limited to the plural in OS, and this took hold in MHG as well; in this way, some number syncretism was avoided on the nouns to which it applied. In MHG and MLG, it spread beyond the nouns where it was phonologically conditioned. It is possible that this morphologization helped case distinctions survive because it profiled number marking without leveling case forms in the singular. On the other hand, very few case distinctions survive on nouns in these languages, since the leveling often occurred anyway. Instead, case distinctions largely survive on agreement targets.

Many other analogical processes applied in the Germanic languages. These developments were often similar across languages, but their effects on nominal inflection have tended to be more drastic in languages with more extreme case loss. Almost all of these processes profiled number marking, but this was not necessarily the primary motivation for all of them; the iconic principle and/or principle of system adequacy can account for most of the changes, including those that profile number marking. Other than a minor reorganization of declension classes, Icelandic nominal inflection has not undergone any analogical processes. In all of the other languages, case forms have been leveled to other forms in the same paradigm. The strong masculine/neuter Dat singular form was leveled to the uninflected Acc form in MSw, MDan,

MNw, EME, CSHG, and CSDu. Since this change occurred in the development of CSHG even though the Dat was not lost generally, it is perhaps the best example of an analogical process motivated by profiling number marking. The weak Acc, Dat, and sometimes Gen singular have been leveled to the Nom in several languages. This process is attested earliest as an optional process in MNw and MDu; it was complete in the early modern development of these languages as well as Swedish and Low German. All of these languages have lost all case distinctions on nouns, so this could have been the leveling of some of the last remaining distinctions among the oblique cases and the Nom as these morphological distinctions were lost more generally. However, profiling number marking likely played a role as well, since the same process is also underway in CSHG, with an almost perfect resemblance to the situation in MDu. Leveling between other Nom and Acc forms occurred in Faroese, MSw, MDan, and MNw. In MDan, the Dat plural was also leveled to the Nom/Acc, a change that actually increased number syncretism. The leveling of the Gen to the Acc across all paradigms in spoken Faroese helped eliminate all number syncretism, but the primary motivation was more likely the Gen undergoing functional narrowing and losing productivity, as is happening to a lesser degree in CSHG. Number profiling sometimes appears to have been achieved by the morphological blocking of sound changes in certain forms. The loss of *-r* was optionally blocked for various forms in MSw, MDan, and MNw. In EME, *-n* loss was blocked in the plural of weak nouns, with the same result as the leveling of the singular oblique forms to the Nom in other Germanic languages.

Extensions of forms across paradigms have also been a common process, sometimes resulting in the full merger of declension classes, specifically when the system is adequately disrupted by sound changes, the remaining forms are refunctionalized using the principles of selection: iconicity, system adequacy, and distinctive strength. In many Germanic languages, Nom/Acc plural forms that are more distinctive have spread to neuter classes and other classes with a zero plural or a plural form that is otherwise involved in number syncretism, in accordance with the principles of iconicity, system adequacy, and distinctive strength. Masculine and feminine *-r* plurals spread to neuter classes starting in MDan and with increasing frequency in the development of CSDan; this has also occurred in spoken Faroese, and to a much lesser extent, in CSS. Similarly, the strong masculine *-s* plural and weak *-en* plural were extended to other classes in EME, ModLG, and CSDu. The *-s* plural was totally generalized by LME, while these are the two most common plural forms in ModLG and CSDu, with the choice depending

on the number of syllables in the noun. In the continental West Germanic languages, the strong and weak feminine classes eventually merged entirely, with a strong singular form and weak plural form. This was attested starting in ODu and already mostly complete in MDu; it was less advanced in MLG and especially MHG. The strong masculine/neuter Gen *-s* form was extended to other classes in most of the languages that lost all case distinctions on nouns, but none that retain case distinctions on nouns. This began in MSw, MDan, MNw, EME, and OS. Except in ModLG, where it lost productivity in favor of other possessive constructions, the *-s* Gen was eventually generalized as an overstable marker in these languages, sometimes even in the plural. By this point, however, it had generally been reanalyzed as a clitic or phrasal affix as all other Gen forms and functions fell out of use. These facts suggest that this reanalysis required both the extension and functional narrowing of the Gen to occur first; in turn, these were likely connected to the loss of all or most other morphological cases.

2.4. Changes in Romance Case Systems

The Romance languages considered in this section include Italian, Spanish, French, and Romanian. Other Romance languages not considered here include Portuguese, Catalan, Occitan, Rhaeto-Romance, and Sardinian. In terms of nominal inflection, these have generally undergone the same developments as languages chosen for analysis but are not attested as well. For example, Portuguese nominal inflection developed almost exactly as in Spanish, even though these two languages have diverged in certain phonological developments. Romanian retains three genders and a simplified nominal case system with three morphological cases: a combined Nom-Acc,[45] a combined Gen-Dat, and Voc. Other Romance languages have all experienced a high degree of case loss, as well as the loss of the neuter, reducing the number of grammatical genders from three to two. Since all the Romance languages are direct descendants of Latin, the Classical Latin (CL)[46] case paradigm is presented here as the proto-paradigm for them.

CL refers to the literary form of the language originally used from about 75 BCE to the 3rd century CE. Generally, CL distinguished six cases morphologically: Nom, Voc, Acc, Gen,

[45] Two case names are connected by a hyphen when they represent a single category for all nouns, e.g., "Nom-Acc". In contrast, two or more case names connected by a slash represent a syncretism of categories that are still distinctive for some nouns, e.g., "Nom/Acc". The use of a hyphen described here should not be confused with its use in a phrase such as "the Nom-Acc distinction," which is equivalent to "the distinction between Nom and Acc".

[46] The grammatical information is drawn from Bennet (1913) for CL, Grandgent (1907) for VL.

Dat, Abl; the three IE genders; and two numbers: singular and plural.[47] The PIE Instr had merged with the Abl in CL. The Voc was usually syncretic with the Nom, except for one common class of nouns. In addition, there were vestiges of the Loc, but these were almost always syncretic with either the Abl or Gen and were limited to certain types of place names and a few specific nouns, e.g., *a*-stem Gen/Loc singular *Rōmae* 'Rome', *a*-stem Dat/Abl/Loc plural *Thēbīs* 'Thebes', *ō*-stem Gen/Loc singular *Corinthī* 'Corinth', *C*-stem Abl/Loc singular *rūre* 'countryside' (see Bennet 1913:12-29).

The functions of the morphological cases in CL had changed somewhat from PIE, largely due to several case mergers. The Nom continued to mark the subject of finite verbs and the complement of 'be' and other linking verbs. The Voc was still used for direct address. The Acc continued to mark DOs and motion toward (goal), including on the objects of prepositions with this meaning. In addition, it marked the objects of some prepositions without this particular meaning, including most recently grammaticalized prepositions. It was also used in adverbial expressions indicating duration of time and extent of space. The Gen continued to mark the complement of a noun. By CL, it definitely had a partitive function as well. As the complement of verbs and adjectives, it often served this function, at least in a broad sense. The Dat continued to mark IOs and loosely connected functions such as Dat of reference, which includes uses such as the ethical Dat and inalienable possession. The Instr had completely merged with the Abl, and the Loc had been supplanted by the Abl in most contexts, so in addition to motion from (source), separation, and standard of comparison, the Abl also inherited the location function from the Loc and the instrument function from the Instr, as well as other related functions such as manner, cause, and accompaniment. The Abl was used with prepositions for some of these functions. The Loc retained its location function without a preposition, but only with certain types of place names and a few specific nouns. All surviving cases except the Voc and Loc marked the complements of certain verbs (see Bennet 1913:107-108, 120-152).

[47] Personal pronouns generally distinguished four forms, with no distinction between Voc and Nom. The first- and second-person singular pronouns had an Acc/Abl syncretism not found elsewhere in the language. The first- and second-person plural pronouns had the usual Dat/Abl syncretism.

2.4.1. Western Romance

Western Vulgar Latin (WVL) was the spoken language from the Classical period until the split into the various Romance languages was complete, around the 9th century. Grandgent (1907:4-5, 147-156) provides an account of WVL based on a combination of written sources and reconstructions based on developments attested in the Romance languages. By the beginning of the Romance period around the 6th to 7th century, WVL is believed to have already lost many of the distinctions in nominal inflection found in CL, including most of the morphological cases and one gender: the neuter. As reconstructed for this period, it distinguished two morphological case markings: Nom and Acc; three genders; and two numbers. The neuter gender was marginal, having lost much of its distinctiveness from the masculine due to sound changes; in most varieties, it was completely lost for nouns sometime in the early Romance period.

Table 22. Changes in Noun Declension from Classical Latin to Western Vulgar Latin (partially reconstructed)

	< ā-stems (also some ē-stems)		< o-stems (also some u-stems)		< o-stems	
	Fem.(/Masc.)		Masc.		Neut.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-a	-ae > -e, » -as	-us > -us, -os	-ī > -i	-um > -u	-a > -a, » -i
Voc	-a	-ae > -e, » -as	-e » -us, -os	-ī > -i	-um > -u	-a > -a, » -i
Acc	-am > -a	-ās > -as	-um > -u, -o	-ōs > -os, -us	-um > -u, -o	-a > -a, » -os, -us
Gen	-ae » -a	-ārum » -as	-ī » -u, -o	-ōrum » -os, -us	-ī » -u, -o	-ōrum » -a, -os, -us
Dat	-ae » -a	-īs » -as	-ō > -o, -u	-īs » -os, -us	-ō > -o, -u	-īs » -a, -os, -us
Abl	-ā > -a	-īs » -as	-ō > -o, -u	-īs » -os, -us	-ō > -o, -u	-īs » -a, -os, -us

	< C-/i-stems			
	Masc./Fem.		Neut.	
	Sg.	Pl.	Sg.	Pl.
Nom	-s; -Ø; -is > -es	-ēs > -es	-Ø	-a, -ia > -a, » -es
Voc	-s; -Ø; -is > -es	-ēs > -es	-Ø	-a, -ia > -a, » -es
Acc	-em, -im > -e	-ēs > -es, -īs » -es	-Ø	-a, -ia > -a, » -es
Gen	-is » -e	-um, -ium » -es	-is » -Ø	-um, -ium » -a, -es
Dat	-ī » -e	-ibus » -es	-ī » -Ø	-ibus » -a, -es
Abl	-e; -ī > » -e	-ibus » -es	-e; -ī » -Ø	-ibus » -a, -es

Phonological changes played an important role in the increase of syncretism in WVL. Final *-m* was lost in all Romance languages, except in monosyllables, and there is evidence for this change in inscriptions both before and after the Classical period. There may have been an intermediate stage in which the vowels that had preceded *-m* were nasalized, but regardless these vowels were indistinguishable from inherited final vowels by the 1st century CE (Grandgent

1907:129-130). For nouns, this had the effect of neutralizing many of the distinctions among the Acc and the Nom/Voc singular, e.g., *ā*-stem Nom/Voc *lūna*, Acc *lūnam* > *luna* 'moon' (see Pharies 2007:103). Vowel length distinctions were neutralized by the 3rd or 4th century CE, particularly in unstressed syllables, which case endings were part of for all but a few exceptional monosyllabic nouns, since final vowels were basically never stressed in polysyllabic words (Grandgent 1907:61, 75).[48] For *ā*-stems, this caused the Abl singular to join the Nom/Voc/Acc syncretism, e.g., *lūnā* > *luna*. Two other changes to unstressed vowels in the final syllable that increased syncretism: the mergers of *e* and short *i* as *e* and of *o* and short *u* as *o* (ibid., 103-104). Following Grandgent (ibid., 152-156), my analysis represents the outcome of the former merger as *e*, its reflex in most Romance languages, but the latter as interchangeable between *o* and *u*, each of which occurs as the reflex in some depending on the languages (ibid., 152-156). These changes must have occurred before the neutralization of length, since *ī* generally remained distinct from *e*. For example, masculine *o*-stem Nom plural -*ī* survived as the Contemporary Standard Italian (CSI) plural form -*i*, while consonant-stem Acc singular -*em* survived as the CSI singular form -*e*. In combination with the loss of length distinctions, the merger of *e* and short *i* resulted in syncretism among the Nom/Voc singular and Nom/Voc/Acc plural for all masculine and feminine consonant-stems and *i*-stems, e.g., masculine consonant-stem Nom/Voc singular *patris*, Nom/Voc/Acc plural *patrēs* > *patres* 'father(s)'. For some of these nouns, Nom/Voc singular became syncretic with Nom/Voc/Acc plural, e.g., masculine *i*-stem Nom/Voc/Gen singular *pānis*, Nom/Voc/Acc plural *pānēs* > *panes* 'bread(s)'. In combination with the loss of length distinctions, the merger of *o* and short *u* resulted in syncretism between the Nom singular and Acc plural for masculine *o*-stems, e.g., Nom singular *mūrus*, Acc plural *mūrōs* > *murus/muros* 'wall(s)'. In combination with the loss of final -*m* as well, it can also account for the case syncretism of Acc, Dat, and Abl singular for *o*-stems, e.g., Acc *mūrum*, Dat/Abl *mūrō* > *muru/muro*, as well as Nom/Voc/Acc singular for the neuters, e.g., *bracchium* > *bracciu/braccio* 'arm'. Similar sound changes, i.e., the loss of certain final consonants and neutralization of vowel distinctions in unstressed syllables, also occurred in many of the Germanic languages. These changes are likely the result of a shift to a rhythmic dynamic stress in both families, as discussed in section 3.1.1.1 below.

[48] See section 3.1.1.1 for a more detailed account of stress in CL.

Analogical processes reduced number syncretism almost to the level of CL, but with the eventual loss of most case distinctions. The leveling of the Gen singular to the Acc eliminated number syncretism with the Nom/Voc plural for *ā*-stems and masculine *o*-stems, and with Nom/Voc/Acc for masculine/feminine consonant-stems and *i*-stems, e.g., masculine *o*-stem CL Gen singular Nom/Acc plural *mūrī* » WVL Gen/Acc singular *muru/muro,* Nom/Acc plural *muri* 'walls'. Despite the distinctiveness even after sound changes, the alternative masculine/feminine *i*-stem Acc plural form *-īs* fell into disuse in favor of the much more frequent *-es*, e.g., *panīs* » *panes* 'breads'. Since *-īs* had avoided the previously mentioned syncretism when used on those nouns, the complete displacement of this form by *-es* increased the overall number syncretism in the language slightly. In this instance, the large difference in frequency appears to have been more important than number profiling in the survival of one form as opposed to another. In addition, the Dat singular was leveled along with the Gen for *ā*-stems, e.g., *lūnae* » *luna* 'moon', so number syncretism removed. A similar leveling of the Gen to the Acc occurred at some point in the development of many Germanic languages, as discussed in section 2.3 above: in feminine classes, the Gen singular forms that had not already become syncretic with the Acc due to sound change were leveled to the Acc. The leveled Gen forms were often syncretic with the Nom plural, as in WVL.

The *ā*-stem Acc plural form *-as* started spreading to the Nom/Voc plural in WVL, e.g., *lune* » *lunas* 'moons'. Possible causes mentioned by Alkire and Rosen include analogy with consonant- and *i*-stems, which contained most other feminine nouns and lacked a distinction between Nom/Voc and Acc in the plural, and the difficulty for the plural to retain case distinctions that have already been lost in the singular (2010:346).[49] Another explanation is that after the disruption of sound changes in this declension class, the form *-as* is refunctionalized by using the principle of system adequacy, in that it increased consistency in the paradigm, and the iconic principle, in that it increased the salience of a marked form, the Nom/Voc plural.

This last explanation is also applicable for the leveling of the Gen, Dat, and Abl plural to the Acc in every declension class once they had lost these case distinctions in the singular. This

[49] The West Germanic languages lost the distinction between Nom and Acc plural before they were attested, and sound change accounts for this neutralization later in the development of the continental Scandinavian languages. In Faroese, however, the Acc plural was leveled to the Nom for masculine nouns even though they retained this distinction in the singular. Analogy with the other genders appears to have motivated the change in Faroese, supporting the significance of analogy across declension classes in the leveling of case forms in WVL and IE languages more generally.

is supported by attestations of the Gen plural form *-oro*, which had been extended from *o*-stems to other declensions, after the disappearance of other Gen forms (see Banniard 2013:100).[50] Further support comes from Juge (2009:72-73), who notes that the loss of the Gen cannot be the direct result of sound change because the Gen plural form survives as a pronoun in some Western Romance languages, e.g., Italian *loro* 'they/their', French *leur* 'their/to them'. He notes that this pattern of retention in isolated paradigms is a common result of morphological loss cross-linguistically. Similar outcomes can be seen in many of the languages under investigation here; all of the languages that have lost case distinctions on nouns still retain at least some on pronouns. Number profiling cannot account for the leveling of the Dat singular form *-ī* for consonant- and *i*-stems, e.g., *patrī* » *patre* 'father'; this may have occurred by analogy with the other declension classes or with the *i*-stem Abl form *-ī*, which the more frequent consonant-stem form *-e* had already started replacing in CL. At some point, masculine *o*-stem Voc singular *-e*, the last distinctive Voc form, was leveled to the Nom.

Many members of the neuter *o*-stem class adopted the masculine *o*-stem forms in the plural, e.g., Nom *braccia* » Nom/Acc *bracci* 'arms'. Grandgent attributes this in part to the similarity of the singular paradigms (1907:145-146). Only the Nom/Voc differed in the singular, e.g., *murus* 'wall' vs. *bracciu/braccio* 'arm'. Alkire and Rosen note that consonant-stem neuters with Nom/Voc/Acc singular *-us*, e.g., *corpus* 'body', also came to be treated as masculine *o*-stems based on the identical Nom form, while other members of this class, e.g., *flumen* 'river', and *i*-stem neuters, e.g., *mare* 'sea', were reinterpreted as masculine or feminine nouns of the corresponding class (2010:193-195). For neuters of all classes, the Nom/Voc/Acc plural *-a* was also the same as the predominantly feminine *ā*-stem singular forms. In CL, this was only the Nom/Voc, but after the sound changes above, also the Acc and Abl. This evidently led to confusion because some neuter plurals were reanalyzed as collectives and then singular *ā*-stems, e.g., *opera* 'works > work' (see Grandgent 1907:146-147). The spread of non-neuter forms in the plural avoided this confusion of neuter plural with feminine singular. The Gen singular was also leveled by analogy with the corresponding non-neuter classes, thereby avoiding number

[50] The eventual loss of the Dat plural form *-om*/*æm* in the development of the standard continental Scandinavian languages may have also been motivated by the lack of a distinctive Dat singular form. Notably, dialects that have retained this form also have a definite Dat singular form. On the other hand, CSHG retains a distinctive Dat form only in the plural for nouns; rather than negate the role of markedness in changes to paradigms, however, this probably reflects the importance of determiners in maintaining case distinctions lost on nouns in CSHG.

syncretism with the new Nom/Voc plural. Neuter nouns that retained the inherited Nom/Voc/Acc plural *-a* experienced the same leveling of the other cases to this form, following the same general tendency of the marked plural not retaining case distinctions lost in the unmarked singular. Very few neuter consonant- and *i*-stems seem to have retained this plural form, however, so this is no longer considered an independent class after WVL.[51]

Agreement targets in CL tended to decline like the major nominal classes, with some pronominal forms on determiners. Adjectives either used *ā*-stem forms for feminine and *o*-stem forms for masculine and neuter, e.g., Nom *bonus* 'good (m.)', *bona* 'good (f.)', *bonum* 'good (n.)', or they used *i*-stem forms for all genders, e.g., Nom *fortis* 'strong (m./f.)', *forte* 'strong (n.)'. For the former class, the feminine forms were distinctive except in the Dat/Abl plural, but as with *o*-stem nouns, masculine and neuter forms were the same except in the Nom/Voc singular and Nom/Voc/Acc plural. For the latter class, the masculine and feminine forms were almost always the same, except in the Nom singular for a small subclass, e.g., *acer* 'sharp (m.)', *acris* 'sharp (f.)', *acre* 'sharp (n.)', while the neuter forms usually differed in the Nom/Voc/Acc of both numbers, although one subclass had the same Nom/Voc singular form for all three genders, e.g., *fēlīx* 'happy'. Of course, these same patterns also applied to the majority of nouns themselves. Determiners mostly used *ā*-stem and *o*-stem forms as well. However, most had the same Gen and Dat singular forms across all three genders, e.g., Gen *illīus*, Dat *illī* 'that (one)'. At the same time, many had a distinctive neuter form in the Acc as well as Nom/Voc singular, as with *i*-stem adjectives. The distinction between masculine and neuter was clearly already weaker than other gender distinctions, but it was relatively robust in the core cases. The loss of *-m* neutralized the Acc singular distinction for *i*-stem adjectives, but otherwise the same distinctions were retained in WVL. Confusion between the neuter and masculine grew during this stage despite a consistent distinction in the Nom/Voc as long as final *-s* was retained. Confusion between the neuter Nom/Voc/Acc plural and feminine singular forms also seems to have contributed to the decline of the neuter, as discussed above (Bennet 1913:34-40, 50-51). The two-gender system of masculine vs. feminine that eventually resulted from these developments in the Western Romance languages contrasts with the two-gender system of common vs. neuter

[51] In addition to these developments concerning the neuter, two relatively minor declension classes, feminine *ē*-stems and masculine *u*-stems, were predominantly absorbed by *ā*-stems, which were mostly feminine, and masculine *o*-stems, respectively. The motivations for these mergers, as well as similar mergers in Germanic and Slavic languages, are discussed in section 3.1.2.2 below.

that has developed in the continental Scandinavian languages and CSDu. These different outcomes can largely be attributed to divergent sound changes. Namely, major classes of nouns and agreement targets retained the distinction between masculine and feminine in the core cases in WVL, but not in the continental Scandinavian languages and CSDu. On the other hand, these Germanic languages have particularly salient neuter forms for agreement targets (ending *-t*) and never developed syncretism between neuter plural and feminine singular forms.

Table 23. Case Mergers in Western Romance

	Classical Latin						Western Vulgar Latin	
PIE	Nom	Voc	Acc	Gen	Dat	Abl	Nom	Acc
Nom	x						x	
Voc		x					x	
Acc			x					x
Gen				x				x
Dat					x			x
Loc						x		x
Instr						x		x
Abl						x		x

The order in which the oblique cases merged with the Acc in the development of WVL is not entirely certain. CL inherited Dat/Abl plural syncretism in all classes from PIE, and almost all Abl singular forms became syncretic with the Acc by regular sound change in WVL. The *i*-stem form that remained distinct from the Acc singular was syncretic with the Dat instead. Only neuter consonant-stems had a completely distinctive Abl singular form, but this was only one small class, especially after many of its nouns had been absorbed by other classes, and the form was unlikely to remain unleveled for long. Thus, the Abl was at least syncretic with the Acc at an early stage, and that it should have been lost first based on the order of loss in other IE languages, which tends to correlate with token frequency (see the discussion of Luraghi (2004) with respect to Greek in section 2.5.3 below). The Dat was probably lost next because sound change had already made the Dat syncretic with the Acc for singular *o*-stems. In addition, there were orthographic substitutions between final unstressed *-e* and *-i*, even when the latter was originally long, including in the Dat and Abl singular of consonant- and *i*-stems, e.g., consonant-stem Dat *luce* 'light' < CL *lucī*, but also other words, e.g., *quase* 'as if' < CL *quasī* (Grandgent 1907:103). These substitutions could reflect a more advanced merger in certain dialects, in

contrast to others in which CL -*ī* remained distinct from -*e*, as indicated, for example, by the CSI reflexes for the consonant-stem singular (-*e*) and masculine *o*-stem plural (-*i*). Alternatively, frequent substitutions in the Dat and Abl singular, but apparently not in other nominal endings with -*ī* in CL, could reflect the merger of the Dat and Abl described above. It remains uncertain, but I argue that these are also evidence for the merger of the Dat with the Acc, since the latter had become syncretic with the Abl for consonant-stems due to -*m* loss. The Dat/Abl plural were most likely leveled to the Acc together, perhaps once the Abl became strongly linked with the Acc in the singular. The leveling of the *ā*-stem Dat singular may have been motivated by number profiling, since it was previously syncretic with the Nom/Voc/Acc plural. In contrast, sound changes hardly affected the distinctiveness of the Gen; the merger of the Gen was more likely the result of number profiling, as discussed in section 2.4.1 above. Therefore, the Gen probably merged with the Acc after the Abl and Dat. This order of case loss follows the frequency of the cases more closely than the order of case loss in Germanic languages.

In CL, the Voc was already syncretic with the Nom in all paradigms except the masculine *o*-stem singular. Since the distinction between these two cases relied on a single declension class, it was probably lost relatively early. At the same time, it survived in EVL long enough to be reinforced by contact with South Slavic languages, so it may have outlasted at least the Abl in WVL as well.

In addition to their original functions, the Nom and Acc inherited the functions of the cases they absorbed in WVL. Thus, the Nom inherited the function of direct address from the Voc. The Acc inherited the functions of the oblique cases, often with the reinforcement of prepositions. The construction with *de* 'from, about, of' followed by the Acc generally supplanted the functions of the Gen and was also the most frequent replacement for Abl functions that did not already require a preposition in CL, but several other prepositions were also common for these Abl functions. The functions of the Dat were often expressed with *ad* 'to' followed by the Acc. The functions of the Loc were completely replaced by prepositional constructions, first with the Abl and then with the Acc (see Grandgent 1907:42-48).

2.4.1.1. The Development of Italian

Contemporary Standard Italian (CSI) has developed out of the Tuscan dialect, which was established as the literary language of Italy beginning in the early 14th century. All case

distinctions have been lost, but it still distinguishes two genders and two numbers. Similar to the Germanic languages that have completely lost morphological case, the functions of direct address, subject, complement of 'be', DO, and object of prepositions are all expressed with bare nouns in CSI. The other former case functions are generally expressed with prepositional constructions (see Harris 1988:18, Vincent 1988:289, 304-305). In my discussion, Pre-Italian refers to the hypothetical descendant of the Tuscan dialect of VL and the direct ancestor of Standard Italian.

Table 24. Changes in Noun Declension from Western Vulgar Latin (reconstructed) to Pre-Italian (reconstructed)

	< ā-stems		< o-stems		< o-stems		< C-/i-stems	
	Fem.(/Masc.)		Masc.		Neut.		Masc./Fem.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-a	-e, -as > -e	-us, -os > -o	-i	-u, -o > -o	-a	-∅; -s > -∅; -es > -i	-es > -i
Acc	-a	-as > -e	-u, -o > -o	-os, -us > -o	-u, -o > -o	-a	-e	-es > -i

Alkire and Rosen provide a reconstruction of Pre-Italian nominal inflection (2010:189), in which sound changes neutralized most of the remaining case distinctions. Regular sound change explains the loss of the distinction between Nom and Acc in both the singular and plural of *ā*-stems. The Acc plural ending *-as*, which had already started spreading to the Nom plural in WVL, merged with the inherited Nom *-e* via two sound changes: final [s] became [j], and then the resulting [aj] was monophthongized to [e], e.g., *lunas* > **lune* 'moons'. The *-es* ending in masculine and feminine consonant-stems and *i*-stems developed similarly: [es] > [ej] > [i], e.g., *panes* > **pani* 'bread(s)'. For masculine *o*-stems, Acc plural *-os* and Nom singular *-us* underwent similar changes: [us]/[os] > [oj] > [o], e.g., Nom singular *murus*, Acc plural *muros* > **muro* 'wall(s)'. These changes resulted in syncretism among these two forms and the Acc singular, while the Nom plural remained distinctive, e.g., **muri* 'walls'. Due to the regular merger of *-us* and *-u* as *-o*, the masculine and neuter *o*-stems became completely identical in the singular. These same changes applied to the corresponding adjective forms. Other agreement targets may have retained the distinction between masculine and neuter in the Nom singular, but they were not enough to resist the decline of the neuter and were eventually leveled to the Acc regardless. Alkire and Rosen (2010:193-195) note that some neuter nouns retained an *-a* plural, but due to the strong association between *-a* and the feminine gender, these forms either came to be used

with feminine plural agreement targets, as described in the following paragraph, or were reinterpreted as a feminine singular, e.g., *mirabilia* 'marvels' > **meraviglia* 'marvel'.

In the development of CSI, two forms underwent analogical leveling. As a result, all number syncretism was removed, but all case distinctions were lost.[52] In masculine *o*-stems, Acc plural *-o* was leveled to Nom plural *-i*, e.g., **muro* » *muri* 'walls'. This direction of leveling contrasts with the usual direction in the Western Romance languages, which supports the role of number profiling. In consonant- and *i*-stems, the Nom singular was leveled to Acc singular *-e*, e.g., **pani* > *pane* 'bread', since one of the Nom singular endings was syncretic with the plural forms. Around 30 nouns descended from neuter *o*-stems retain agreement with masculine in the singular, e.g., *uovo buono* 'good egg', but feminine in the plural, e.g., *uova buone* 'good eggs'. These are known as ambigeneric nouns and are sometimes considered a third gender category based on their unique distribution (see Alkire & Rosen 2010:194-195, Baerman et al. 2005:82-83, Kuryłowicz 1964: 212, Loporcaro 2014:7).

2.4.1.2. The Development of Spanish[53]

Early Spanish (ES), also known as Medieval Castilian, was used from the 13th to 16th centuries. All case distinctions had been lost for ES nouns and articles. These continued to distinguish two genders and two numbers (see Pharies 2007:104).[54]

[52] Although the collapse of the Nom-Acc distinction generally occurred before the loss of the Dat and Gen in Germanic languages that experienced total case loss, the merger of strong and weak feminine classes in the continental West Germanic languages achieved a similar outcome for these nouns as the leveling in the development of CSI. Generally, the strong endings were adopted in the singular, while the weak endings were adopted in the plural. For the strong nouns, this merger can also be seen as the leveling of the Nom/Acc plural to the Gen/Dat plural; for the weak nouns, it can be seen as the leveling of the oblique singular forms to the Nom.

[53] The grammatical information is drawn from Alkire and Rosen (2010), Pharies (2007), and Penny (2002) for ES.

[54] Pronouns in Medieval Castilian distinguished up to four forms: Nom, Acc, Dat, and Prepositional (Prep). Prep is the form used after prepositions, e.g., *a mí* 'to me', *para ti* 'for you'. The first- and second- person singular pronouns had three distinctive forms, with the same form for Acc and Dat. The first- and second-person plural pronouns made no distinctions until the transition to Modern Spanish, when a separate Nom form arose. The third-person pronouns had three distinctive forms, with the same form for Nom and Prep, but distinctive Acc and Dat DO and IO forms (see Pharies 2007:107-112).

Table 25. Changes in Noun Declension from Western Vulgar Latin (reconstructed) to Early Spanish

	< ā-stems		< o-stems		< o-stems	
	Fem.(/Masc.)		Masc.		Neut.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-a	-as; -e » -as	-us, -os » -o	-i » -os	-u, -o > -o	-a » -os
Acc	-a	-as	-u, -o > -o	-os, -us > -os	-u, -o > -o	-a » -os

	< C-/i-stems	
	Masc./Fem.	
	Sg.	Pl.
Nom	-s, -Ø, -es » -e, -Ø	-es
Acc	-e > -e, -Ø	-es

The loss of the remaining case distinctions in the development of ES from WVL should be attributed primarily to analogical processes in response to earlier sound changes rather than additional sound changes. Notably, final *-s* was retained in ES, in contrast to Pre-Italian. However, Penny notes that ES experienced the loss of final *-e*. Final *-e* was later restored in many phonetic contexts during the second half of the 13th century (2002:58-59). In the interim, however, the distinction between Nom and Acc singular was neutralized for some consonant-stems, and case loss was likely complete by the time *-e* was restored, so it would have been restored on the singular form regardless of its function. After sound change, Nom and Acc singular remained distinctive for non-neuter *o*-stems, *i*-stems, and some consonant-stems. However, the Nom singular was syncretic with the Acc plural for masculine *o*-stems, which came to be consistently written as *-os* in ES, e.g., Nom singular/Acc plural *murus*/*muros* > **muros* 'wall(s)', and with both the Nom and Acc plural for some nouns in the other classes, e.g., masculine *i*-stem **panes* 'bread(s)'. Thus, the leveling of the Nom singular to the Acc in these classes removed number syncretism, e.g., **muros* » *muro*, **panes* » *pan*. The remaining distinction between Nom and Acc plural was then lost in favor of the Acc, as in the singular. For *ā*-stems, this leveling had already begun in WVL. According to Barðdal and Kulikov, the last surviving case distinction was in the plural of masculine *o*-stems, e.g., Nom **anni* vs. Acc *annos* 'years'. They attribute the eventual loss of this distinction to paradigmatic leveling, i.e., the pressure of the other paradigms, which had already lost this distinction (2009:473). This was also one of the last surviving case distinctions in the development of Italian, and the same motivation likely played a role there, in addition to the elimination of number syncretism discussed above.

In contrast to Italian, however, both the Nom and Acc plural forms of masculine *o*-stems avoided number syncretism after the leveling in the singular, so the direction of leveling was likely influenced by the similarity of the Acc form to the other plural forms with *-s*, as well as the additional salience of this form compared to the Nom.

No neuter-type paradigm remained by the time of ES, in contrast to the small class of nouns in Italian that retained masculine agreement in the singular and feminine agreement in the plural; most neuter nouns that had not already become masculine by WVL did so by ES. These nouns had inherited Nom/Acc syncretism from PIE, so it is possible that they also contributed to the leveling of Nom to Acc, particularly in the singular where they generally had the same form as the Acc of the other genders. Other neuter nouns were reinterpreted as feminine singular (in form and usually meaning) based on their plural forms, e.g., *mirabilia* 'marvels' > *maravilla* 'marvel' (see Alkire & Rosen 2010:192-194). Agreement targets largely followed suit and the neuter was lost as a gender category. The only remaining vestiges of the neuter are distinctive singular forms of determiners, e.g., the definite article form *lo*, and the subject pronoun *ello* 'it'. Since there are no longer any neuter antecedents, these neutral forms have been repurposed to refer to nonspecific antecedents (see Corbett 1991:214-215, Pharies 2007:115-117).

Personal *a* (< 'to') also developed during ES as an analytic marker of definite human DOs, in addition to IOs (see Alkire & Rosen 2010:190). It was initially used to counter ambiguities and was grammaticalized in the 1600s (see Penny 2002:115-116). This development was more likely a response to case loss than a cause. It is similar in function to the borrowing of the Gen form for the Acc in Slavic languages, but it represents the extension of an analytic construction rather than a synthetic one.

The modern stage of Spanish is considered to have begun in the 16th century. Modern Spanish (ModSp) retains two genders and two numbers. Indefinite and non-human DOs, along with the functions of direct address, subject, complement of 'be', and object of prepositions are all expressed with bare nouns. The other former case functions were generally expressed with prepositional constructions. Except for the use of personal *a*, this situation is similar to the other Western Romance languages (see Green 1988:138).

In Standard Spanish, there is a redundancy in marking the /s/ plural in the NP that includes nouns, adjectives, and articles,[55] as in the example below:

(1) *Tienen muchos juegos de esos pintados en el suelo diferentes.*
They have a lot of different games like that painted on the ground.

(Poplack 1980:55)

In example (1), the plurality of the noun *juegos* is marked by final /s/ not only on itself, but also on the adjectives *muchos*, *pintados*, and *diferentes*, and on the demonstrative pronoun *esos*. In Puerto Rican Spanish, however, there is an aspiration of numerous occurrences of /s/ in syllable-final positions, and deletion of many instances of /s/ in absolute final positions. Consequently, with the lack of phonological compensation for the /s/ reduction/deletion, there is no distinction between plural and singular NPs, which makes number categories syncretic. In terms of functionality, Poplack (ibid., 57) explains that the functional hypothesis anticipates that the reduction/deletion of *-s* would not occur in environments where it would result in morphological indistinctiveness. However, the Puerto Rican's situation provides evidence to the contrary. This indicates that the functional hypothesis is inadequate for describing the phenomenon.

The reduction of word-final plural marking puts the burden of clarification primarily on the articles. Puerto Rican Spanish is trending toward increased use of the masculine article, which undergoes a stem vowel change from singular to plural e.g., *el* > *los*, even in feminine contexts (see Poplack 1980:65). This stem vowel change means that the masculine article can still clearly communicate singularity vs. plurality, even when the syllable-final /s/ is aspirated or dropped. I agree with Poplack (1980) that the functional hypothesis is insufficient to explain the situation. Morphological blocking of sound change is expected to maintain the distinction of a category. For example, there is a morphological blocking of *-r* in MSw which profiled number category.[56] Since sound change is accidental and can cause category loss, other ways to mark the lost category are utilized. Puerto Rican Spanish is similar to MHG in that a high degree of number syncretism occurred due to sound changes, but analogical processes reduced the number

[55] See Labov (1987:322), and Terrell (1975:431-32) for further discussion.

[56] See section 2.3.2.2 for more discussion.

syncretism significantly by CSHG. This shows that when number distinction is lost, language users tend to find a way to revive it. The utilization of the articles in Puerto Rican Spanish to mark number category supports number profiling hypothesis; the articles are used to compensate the loss of number category.

2.4.1.3. The Development of French[57]

Old French (OF) was used from the 9th to 14th centuries. Except for the total loss of the neuter, it retained the same distinctions in nominal inflection as WVL. It distinguished two morphological cases: Nom and Acc; two genders: masculine and feminine; and two numbers (see Alkire & Rosen 2010:188-189, Härmä 2000:611, Rickard 2003:20, 48-51, 61).

Table 26. Changes in Noun Declension from Western Vulgar Latin (reconstructed) to Old French

	< ā-stems		< o-stems		< o-stems	
	Fem.		Masc.		Neut.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-a > -e	-as > -es; -e » -es	-us, -os > -s	-i > -Ø	-u, -o » -e	-a > -e
Acc	-a > -e	-as > -es	-u, -o > -Ø	-os, -us > -s	-u, -o » -e	-a > -e

	< C-/i-stems		< C-/i-stems	
	Masc.		Fem.	
	Sg.	Pl.	Sg.	Pl.
Nom	-s, -es > -s; -Ø	-es » -Ø	-s, -es > -s; -Ø	-es > -s
Acc	-e > -Ø	-es > -s	-e > -Ø	-es > -s

In the development of OF from WVL, sound changes further increased syncretism, particularly number syncretism. Rickard describes how vowels in the final syllable underwent a drastic reduction at this stage. In this position, [a] was reduced to [ə], generally spelled <e>, while other vowels were only reduced to [ə] if they were needed to avoid an unpronounceable final consonant cluster; otherwise, they disappeared entirely. Along with the loss of the vowel in the penult when the antepenult was stressed, these changes caused almost all words to have either final stress or penultimate stress with [ə] in the final syllable (2003:13-14). Although this development shifted stress closer to the end of the word, in contrast to the Germanic languages, it

[57] The grammatical information is drawn from Alkire and Rosen (2010), and Rickard (2003) for Old French (OF), and MF (Middle French), Harris (1988) for CSF (Contemporary Standard French).

still had a similar outcome as Germanic vowel reduction, in that it tended not to preserve distinctions among inflectional endings because they were lost or reduced to [ə] as part of the process. As for the effects of this vowel reduction on nominal inflection, masculine *o*-stems were reduced to two forms: *-s* for Nom singular and Acc plural, e.g., *murs* 'wall(s)', and a zero ending for Acc singular and Nom plural, e.g., *mur* 'wall(s)'.[58] In addition, some consonant-stems and *i*-stems developed syncretism between Nom and Acc singular, e.g., masculine consonant-stem Nom *frater*, Acc *fratre* > *frere* 'brother', while others developed into an irregular class that retained this distinction due to the different reflexes of forms that differed in stress position, e.g., masculine consonant-stem Nom *báro* > *ber*, Acc *baróne* > *baron* 'warrior' (see Rickard 2003:49).

Unexpectedly, number syncretism in OF actually increased as a result of analogical processes, but some of these processes prevented an even greater increase in number syncretism, and the different declension classes became more similar. Masculine consonant- and *i*-stems adopted a zero ending in the Nom plural by analogy with masculine *o*-stems, e.g., **pains* » *pain* 'breads'. After sound changes, this had been the only form that still regularly differed among these declension classes. Thus, they had already mostly merged by OF. The spread of the *ā*-stem Acc plural form to the Nom that had begun in WVL was completed in OF, e.g., **chevre* » *chevres* 'goats'. In the singular, *ā*-stems (e.g., Nom/Acc *chevre*) still differed from feminine consonant- and *i*-stems (e.g., Nom *fins*, Acc *fin* 'end'), but all had plural forms ending in *-s* for both cases (e.g., *chevres*, *fins*).

The neuter category had been lost by the OF period, around the same time as in other Western Romance languages. Neuter *o*-stems that had not already become masculine in WVL either were reinterpreted as feminine singular (in form and usually meaning) based on their plural forms, e.g., *gaudia* 'joys' > *joie* 'joy', or became indeclinable masculine nouns with a plural meaning, e.g., *milia* > *milie* 'thousands' (see Rickard 2003:32). The latter group likely did not adopt the masculine *o*-stem forms since they completely lacked a form with *-s*. Any remaining neuter consonant- and *i*-stems developed in similar ways. Since no class of nouns with

[58] OF examples are given in standard orthography, which represents the pronunciation much more transparently than in Contemporary Standard French (CSF). These examples do not contain silent letters; the loss of final consonants except in liaison, i.e., before a syntactically connected word beginning with a vowel, was not complete until the 16th century. Vowels were nasalized before nasal consonants, but nasalized vowels did not become phonemic until the loss of coda nasal consonants and the denasalization of vowels before intervocalic nasal consonants during the Middle French (MF) period (see Rickard 2003:47, 63-65).

different gender agreement in the singular and plural developed, the loss of the neuter in French had more in common with Spanish than Italian.

In the development of Middle French (MF), which was spoken from the 14th to 17th centuries, all case distinctions were lost along with the removal of all number syncretism (see Rickard 2003:61, 68). As generally occurred in the Western Romance languages, the Acc form survived in both the singular and plural, e.g., *mur* 'wall', *murs* 'walls'. This process became widespread in the transition to MF during the 1200s; one possible motivation was the consistency of the Acc across declension classes—a zero ending in the singular and *-s* in the plural—in contrast to the synchronically opaque distribution of the Nom (see Alkire & Rosen 2010:191). The iconic principle likely also played a role: the salient *-s* form survived in the marked plural, while the zero-form survived in the unmarked singular. As mentioned above, *-s* was still pronounced consistently at this stage, which supports the primacy of analogical change over phonological change in the loss of the last case distinction in French. This leveling also eliminated the remaining differences in declension among almost all nouns of the same gender. Except for a few indeclinable and irregular nouns, there was one pattern for masculine nouns and one for feminine nouns.

In contrast to the other Western Romance languages, in which the same endings are still used for both genders on consonant- and *i*-stem nouns and adjectives, adjective declension also developed a single pattern for each gender during the MF period. A new gender distinction arose for *i*-stem adjectives with the adoption of *o*-stem forms for masculine and *ā*-stem forms for feminine. This process began in OF with the masculine, paralleling developments in nouns. However, this still left the same Acc forms for both genders, and these were the forms that were retained after case loss. Thus, the gender distinction was only consistently achieved with the spread of the *ā*-stem forms containing *-e* to feminine *i*-stems, starting with isolated examples in OF. Now only fossilized examples of *i*-stem adjective forms remain, e.g., *pas grand-chose* 'nothing much' in contrast to productive *la grande chose* 'the large thing' (see Alkire & Rosen 2010:191-192, Rickard 2003:50). These changes to adjectives, along with the parallel changes to nouns discussed above, brought about something close to a one-to-one correspondence between gender and adjective declension by late MF, which can be considered a form of gender

profiling.[59] Since profiling the gender marking appears to have played a major role in the development of French nominal inflection, it may also explain why analogy unexpectedly increased number syncretism in OF: gender marking was profiled at the expense of both case and number marking.

The modern stage of French is considered to have begun in the 17th century. CSF retains two genders and two numbers. As in other Western Romance languages, the functions of direct address, subject, complement of 'be', DO, and object of prepositions are all expressed with bare nouns in CSF. The other former case functions are generally expressed with prepositional constructions (see Harris 1988:235-237). Sound change during the modern period has neutralized the distinction between masculine and feminine forms of nouns and certain adjectives, so there is no longer a one-to-one correspondence between gender and declensions in both nouns, and adjectives. Specifically, the loss of final [ə] eventually had this effect on adjectives ending in an oral vowel. When this sound lost its syllabic value in the 16th century, secondary effects of its presence remained, so the distinction was merely shifted earlier in the word: consonant-final adjectives retained the final consonant only in the feminine forms, e.g., masculine singular *court* [ku:r], feminine singular *courte* [ku:rt] 'short', and final *-e(s)* was still reflected by a lengthened final vowel in vowel-final adjectives, e.g., masculine singular *vrai* [vrɛ], feminine singular *vraie* [vrɛ:] 'true'. Even this length distinction was lost in the 19th century, however, resulting in a new class of adjectives without gender distinctions in the spoken language. The loss of the feminine *-e* made the distinction with masculine opaque. This caused the gender assignment system less clear, making the assignment rules of gender now complicated (see Corbett 1991:134, 315, Rickard 2003:63). As a result, CSF resembles the other Western Romance languages in having some adjectives without gender distinctions, but these adjectives belong to a new phonologically-conditioned class in CSF, in contrast to the inherited class of *i*-stems in the other languages. In addition, most final consonants were lost except in liaison with a following vowel-initial word. This process began in the 13th century but was not complete until the 16th century. As a result, the *-s* plural marker is only pronounced in limited contexts. Thus, number is not always overtly marked.

[59] Of the Germanic languages discussed in section 2.3, only NNw has this kind of total association between gender and declension. Types of gender-declension relationships are considered in section 3.1.2.1 below.

2.4.2. Eastern Romance: The Development of Romanian[60]

Eastern Vulgar Latin (EVL) was the variety of VL spoken in Dacia, beginning with its Romanization in the 2nd century CE. It distinguished three genders and two numbers (Grandgent 1907:147-156). Alkire and Rosen reconstruct a three-case system with Nom, Acc, and Dat (2010:279-280). My analysis, however, assumes an earlier stage with the same six cases as CL: Nom, Voc, Acc, Gen, Dat, Abl. This allows the effects of sound change to be considered separately from analogical processes and functional mergers. These cases would have had the same functions as in CL, but with the more frequent replacement or reinforcement by prepositional constructions.

Table 27. Changes in Noun Declension from Classical Latin to Eastern Vulgar Latin (reconstructed)

	< ā-stems (also some ē-stems)		< o-stems (also some u-stems)		< o-stems	
	Fem.(/Masc.)		Masc.		Neut.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-a	-ae > -e	-us > -us, -os	-ī > -i	-um > -u, -o	-a
Voc	-a	-ae > -e	-e	-ī > -i	-um > -u, -o	-a
Acc	-am > -a	-ās > -as	-um > -u, -o	-ōs > -os, -us	-um > -u, -o	-a
Gen	-ae > -e	-ārum > -aru	-ī > -i	-ōrum > -oru, -oro	-ī > -i	-ōrum > -oru, -oro
Dat	-ae > -e	-īs > -is	-ō > -o, -u	-īs > -is	-ō > -o, -u	-īs > -is
Abl	-ā > -a	-īs > -is	-ō > -o, -u	-īs > -is	-ō > -o, -u	-īs > -is

	< C-/i-stems			
	Masc./Fem.		Neut.	
	Sg.	Pl.	Sg.	Pl.
Nom	-s; -Ø; -is > -es	-ēs > -es	-Ø	-a, -ia
Voc	-s; -Ø; -is > -es	-ēs > -es	-Ø	-a, -ia
Acc	-em, -im > -e	-ēs > -es; -īs > -es	-Ø	-a, -ia
Gen	-is > -es	-um, -ium > -u, -o > -eu, -eo	-is > -es	-um, -ium > -u, -o > -eu, -eo
Dat	-ī > -i	-ibus > -ebus, -ebos	-ī > -i	-ibus > -ebus, -ebos
Abl	-e; -ī > -e	-ibus > -ebus, -ebos	-e; -ī > -i	-ibus > -ebus, -ebos

Phonological changes in the development of EVL also led to a significant increase in syncretism. Most of the key sound changes in WVL had their beginnings, if not conclusions, well before the Romanization of Dacia in the 2nd century CE, so they also applied in EVL, with similar effects. As in WVL, the loss of final *-m* (see Grandgent 1907:129-130) neutralized the distinction

[60] The grammatical information is drawn from Alkire and Rosen (2010) for Pre-Romanian and Contemporary Standard Romanian (CSRm).

between the Acc and Abl singular for masculine and feminine consonant-stems and *i*-stems, e.g., feminine *i*-stem Acc *vulpem*, Abl *vulpe* > **vulpe* 'fox', and the distinction between the Acc and Nom/Voc singular for *ā*-stems, e.g., Acc *capram*, Nom/Voc *capra* > **capra* 'goat'. In combination with the neutralization of vowel length distinctions (see ibid., 75), the *ā*-stem Abl singular also joined the Nom/Voc/Acc singular syncretism, e.g., *caprā* > **capra*. As in WVL, *e* merged with short *i* in the final syllable, and *o* merged with short *u* in this same position as well, with similar effects on syncretism (see ibid., 103-104). My study represents the outcome of the former merger as *e*, but the latter as interchangeable between *o* and *u*, as in WVL. In combination with the loss of length distinctions, the merger of *e* and short *i* resulted in syncretism among the Gen singular and Nom/Voc/Acc plural for all masculine and feminine consonant-stems and *i*-stems, e.g., masculine consonant-stem Gen singular *flōris*, Nom/Voc/Acc plural *flōrēs* > **flores* 'flower(s)'. For some of these nouns, Nom/Voc singular were already syncretic with Gen singular, so they also joined this number syncretism, e.g., masculine *i*-stem Nom/Voc/Gen singular *panis*, Nom/Voc/Acc plural *panēs* > *panes* 'bread(s)'. In combination with the loss of length distinctions, the merger of *o* and short *u* resulted in syncretism between the Nom singular and Acc plural for masculine *o*-stems, e.g., Nom singular *lupus*, Acc plural *lupōs* > **lupus/lupos* 'wolf/wolves'. In combination with the loss of final *-m* as well, it can also account for the syncretism of Acc, Dat, and Abl singular for *o*-stems, e.g., Acc *lupum*, Dat/Abl *lupō* > **lupu/lupo*, as well as Nom/Voc for the neuters, e.g., *vīnum* > **vinu/vino* 'wine'. The merger of *e* and short *i* also occurred in other unstressed environments: in hiatus, i.e., directly before another vowel, e.g., feminine *i*-stem Gen plural *vulpium* > **vulpeu/vulpeo* 'foxes', and in the penult generally, e.g., Dat/Abl plural *vulpibus* > **vulpebus/vulpebos* (see ibid., 94, 99). These changes did not result in any immediate increases in syncretism in the VL period. In WVL, the affected forms were leveled to the Acc plural, as were the corresponding forms in the other declension classes, so sound change was likely not a major factor. However, these changes became important in combination with other changes in Pre-Romanian, as described below.

As mentioned above, my analysis assumes that most of the analogical processes occurred in the development of Pre-Romanian from EVL or in the further development of Contemporary Standard Romanian (CSRm), in response to the increased number syncretism. This allows the effects of sound change to be considered separately from analogical processes and functional mergers. As in WVL, the alternative masculine/feminine *i*-stem Acc plural form *-īs* fell into

disuse in favor of the much more frequent *-es*, e.g., *panīs* » *panes* ‘breads’, likewise resulting in a minor increase in number syncretism. One change that was likely complete by EVL, since it had already begun in CL, was the leveling of the masculine and feminine *i*-stem Acc singular form *-im* and Abl singular form *-ī* to the more frequent consonant-stem forms *-e(m)* and *-e*, respectively, e.g., Acc *secūrim*, Abl *secūrī* » *secure* ‘axe’. This would have prevented these forms from becoming syncretic with the Nom/Acc/Voc plural due to later regular sound change, as in the following hypothetical example: Acc singular *secūrim*, Abl *secūrī*, Nom/Voc/Acc plural *secūrēs* > **securi*. However, the change was more likely motivated by the low frequency of these forms that represented some of the last differences between *i*-stems and consonant-stems.

As in WVL, EVL adjectives developed in parallel to nouns. Determiners that previously distinguished neuter *-ud* from masculine *-um* in the Acc singular adopted *-u* (< *-um*) for the neuter as well by analogy with adjectives. However, most agreement targets would still have distinguished all three genders in the Nom/Voc of both numbers and the Acc plural, with the exception of *i*-stem adjectives, which already had fewer gender distinctions in CL. Some additional analogical processes brought determiners even more in line with adjectives and nouns, particularly in terms of gender distinctions. Grandgent notes that as early as the 2nd century BCE, the form *-(a)e* was used in place of Gen singular *-ius* and Dat singular *-i*, e.g., Gen *illīus*, Dat *illī* » *ill(a)e* ‘that (one)’ (1907:163). This constituted an extension from *ā*-stem nouns and adjectives to determiners (see Alkire & Rosen 2010:201). Thus, feminine became completely distinctive in the singular paradigms of determiners as well as most adjectives.

As noted by Grandgent, the corresponding masculine/neuter forms of determiners were also remodeled, but by analogy with the relative and interrogative pronouns, with which they already shared many of the same endings. This remodeling involved the insertion of *-u-*, found in Gen *cuius* and Dat *cui* ‘who/which’, e.g., Gen *illīus* » *illuius*, Dat *illī* » *illui*. Finally, a new set of feminine forms developed, probably due to the influence of these new masculine/neuter forms, e.g., Gen *ill(a)e* » *illeius*, Dat *ill(a)e* » *illei* (1907:163-164).

Table 28. Changes in Noun Declension from Eastern Vulgar Latin (reconstructed) to Pre-Romanian (reconstructed)

	< ā-stems		< o-stems		< o-stems	
	Fem.(/Masc.)		Masc.		Neut.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-a > -ă	-e	-us, -os > -ui	-i	-u, -o > -u	-a » -e, -ură
Voc	-a > -ă	-e	-e	-i	-u, -o > -u	-a » -e, -ură
Acc	-a > -ă	-as > -e	-u, -o > -u	-os, -us > -ui	-u, -o > -u	-a » -e, -ură
Gen	-e	-aru, -aro » i, -e	-i » -u	-oru, -oro » -i	-i » -u	-oru, -oro » -i, -e, -ură
Dat	-e	-is > -i, » -e	-o, -u > -u	-is > i	-o, -u > -u	-is > i, » -e, -ură
Abl	-a > -ă	-is > -i, » -e	-o, -u > -u	-is > i	-o, -u > -u	-is > -i, » -e, -ură

	< C-/i-stems		< C-/i-stems	
	Masc./Fem.		Neut.	
	Sg.	Pl.	Sg.	Pl.
Nom	-Ø, -s > -Ø; -es > -i	-es > -i	-Ø	-(ur)a > -(ur)ă; -ia » -ură
Voc	-Ø, -s, > -Ø; -es > -i	-es > -i	-Ø	-(ur)a > -(ur)ă; -ia » -ură
Acc	-e	-es > -i	-Ø	-(ur)a > -(ur)ă, -ia » -ură
Gen	-es > -i	-eu, -eo > -eu, » -i; -u, -o » -eu, -i	-is > -i	-eu, -eo > -eu, » -(ur)ă; -u, -o » -eu, -(ur)ă
Dat	-i	-ebus, -ebos > -eu, » -i	-i	-ebus, -ebos > -eu, » -(ur)ă
Abl	-e	-ebus, -ebos > -eu, » -i	-i; -e » -Ø	-ebus, -ebos > -eu, » -(ur)ă

A Pre-Romanian stage between EVL and CSRm can be reconstructed. According to the reconstruction of Pre-Romanian provided by Alkire and Rosen, almost all of the regular sound changes affecting case endings have occurred by this stage, and the number of cases was reduced to four: Nom, Voc, Acc, Gen-Dat. It still distinguished three genders and two numbers (2010:280). In addition to its original functions, the Acc inherited the functions of the Abl in Pre-Romanian. By this stage, all of these were likely reinforced by prepositions. The Gen-Dat inherited the functions of both cases. These competed with constructions involving a preposition followed by the Acc, as in the Western Romance languages.

Sound changes further increased syncretism in the development of Pre-Romanian. Alkire and Rosen propose that final *-s* developed similarly to Pre-Italian, but without the loss of [j] after [u], resulting in the form *-ui* for both the Nom singular and Acc plural of masculine *o*-stems, e.g., Nom singular/Acc plural **lupus*/**lupos* > **lupui* 'wolf/wolves', as well as the form *-i* for Dat/Abl plural of all *o*-stems, e.g., **lupis* > **lupi* (2010:280). I agree with this proposal as the most plausible account. Thus, the Dat/Abl plural form of this declension class was expected to join the syncretism with the Gen singular and masculine Nom/Voc plural, which already had the form *-i*. Likewise, this change would have caused the Dat singular to join the syncretism among

the Gen singular and Nom/Voc/Acc plural for masculine and feminine consonant-stems and *i*-stems, along with the Nom/Voc singular for some of these nouns, e.g., feminine *i*-stem Nom/Voc/Gen singular **vulpes*, Dat **vulpi*, Nom/Voc/Acc plural **vulpes* > **vulpi* 'fox(es)'. For neuter consonant-stems and *i*-stems, only the syncretism between the Gen and Dat singular resulted. In contrast, both the Gen and Dat singular appear to have been leveled to the Acc in WVL before the loss of *-s* in the development of some Western Romance languages such as Italian. However, dialects with early *-s* loss may have developed this Gen/Dat syncretism as in Pre-Romanian. Sound change often caused Gen/Dat syncretism in the Germanic languages, particularly for feminine and weak nouns, but the Acc was often involved as well, and the strong masculine/neuter Gen form with *-s* was distinctive. As a result, a Gen-Dat merger in opposition to the Acc never occurred.

By Pre-Romanian, *o* and short *u* had completely merged as *u*. In combination with the loss of final *-s* and the loss of intervocalic /b/, this change accounts for resulted in *-eu* as the expected reflex for the consonant- and *i*-stem Dat/Abl plural, e.g., **vulpebus* > **vulpeu*, as given by Alkire and Rosen (2010:264, 280). This was also the expected reflex of the *i*-stem Gen plural, e.g., **vulpeu/vulpeo* > **vulpeu* 'foxes'. Thus, the Gen/Dat syncretism that had previously been limited to singular *ā*-stems may have spread to consonant- and *i*-stems of both numbers. The merger of the Gen with the Dat could easily have resulted from these sound changes, since only *o*-stems retained the distinction in the singular, and even their Gen singular form was syncretic with the Dat plural.

Analogical processes posited for this stage involved the completion of the merger of the Gen and Dat and the beginning of the loss of distinctive oblique forms in the plural. Profiling number marking could only have been a motivation for the leveling of the *o*-stem Gen singular to the Acc/Dat/Abl singular form, e.g., masculine **lupi* » **lupu,* 'wolf'. This eliminated syncretism between the Gen singular and Nom/Voc plural for masculine members, while it prevented syncretism between the Gen singular and Dat/Abl plural for neuter members. However, the loss of this last distinctive Gen singular form was likely motivated at least in part by the pressure of the other declension classes, as mentioned in the previous paragraph, and/or a functional merger, as in the other Balkan Sprachbund languages. The leveling of the Gen plural to the Dat may have spread from consonant- and *i*-stems, where there was potential confusion between these forms, e.g., masculine *o*-stem Gen **luporu* » **lupi*. Additional pressure may have come from the

Gen/Dat syncretism in the singular for *ā*-stems, as well as *o*-stems, although it is possible that *o*-stems lost the distinction in the plural first.

These changes account for the case system reconstructed by Alkire and Rosen, but they also propose that the leveling of the Gen-Dat plural forms to the Nom/Voc/Acc began at this stage. For *o*-stems, the Gen-Dat plural had already become syncretic with the Nom. As reconstructed, the inherited Dat forms were used alongside the Nom/Voc/Acc forms in the other declension classes, e.g., feminine *ā*-stem Gen-Dat **capri* and **capre* 'goat' (2010:280).

As in Pre-Italian, the single neuter form became syncretic with the masculine singular form, while the plural neuter form became syncretic with the feminine plural. Fewer neuters seem to have adopted masculine forms in EVL than in WVL, since the class still comprises about a third of nouns. In contrast to Italian, it also remains productive with some foreign words, e.g., *fax* 'fax', and some originally masculine nouns, e.g., *populus* > *popor* 'people' (Alkire & Rosen 2010:282). At this stage, singular masculine and neuter *o*-stems only differed in the Nom. This was also true for the corresponding adjective and determiner forms. I propose that number syncretism with the Dat/Abl plural, the Gen-Dat merger, and/or the attraction of the masculine class likely motivated the leveling of the Gen singular for neuter nouns and adjectives. At the same time, neuter nouns of all classes had Nom/Voc/Acc plural forms ending in *-ă*, e.g., *o*-stem **vină* 'wines', consonant-stem **timpură* 'times'.[61] The feminine *ā*-stems also had these forms for the Nom/Acc/Abl (and sometimes Voc) singular, which motivated the use of feminine agreement targets for plural neuters. Alkire and Rosen note that the *-ă* form was subsequently replaced. Some neuter *o*-stems adopted the *ā*-stem plural form *-e*, e.g., **oasă* » *oase* 'bones'. At the same time, the neuter consonant-stem form *-ură*, a combination of stem and ending, spread to other consonant-stems and some *o*-stems (ibid., 282).

In the development of Pre-Romanian from EVL, the Abl was probably the first case lost, followed by the merger of Gen and Dat. As in WVL, the Abl became syncretic with the Acc during or soon after the EVL period due to existing syncretism and regular sound change. Leveling was only necessary for neuter consonant-stems, a relatively small class. Different sound changes from WVL meant that the Gen became syncretic with the Dat for consonant- and *i*-

[61] Except in certain environments such as the one in footnote 64 below, unstressed *a* was reduced to *ă*, pronounced similarly to [ə]. Stressed *a*, as well as both stressed and unstressed *e*, also developed into *ă* in certain environments, so this sound could be stressed or unstressed in Pre-Romanian. CSRm has inherited the same outcomes (see Alkire & Rosen 2010:258-259).

stems in the singular, and the plural forms may have been confused as well. These changes may also account for the rise of Gen/Dat syncretism for singular determiners of all genders, which could have influenced indefinite nouns in turn. Given the existing Gen/Dat syncretism in singular *ā*-stems, the leveling of the *o*-stem Gen singular form as a result of number profiling neutralized the last distinction between Gen and Dat in the singular. This distinction was eventually leveled in the plural as well, bringing the Gen-Dat merger to completion during Pre-Romanian. As in the Western Romance languages, pronouns still retain the distinction between the Nom and Acc, so this distinction was most likely the last lost for nouns, in the development of Romanian from Pre-Romanian.

Table 29. Case Mergers in Romanian

	Classical Latin / Eastern Vulgar Latin						Pre-Romanian / CSRm		
PIE	Nom	Voc	Acc	Gen	Dat	Abl	Nom-Acc	Voc	Gen-Dat
Nom	x						x		
Voc		x						x	
Acc			x				x		
Gen				x					x
Dat					x				x
Loc						x	x		
Instr						x	x		
Abl						x	x		

Romanian has been attested since the 16th century, as earlier writings by Romanians were in Bulgarian Church Slavonic (BChS). By this time, it already had its modern case system, which has lost some of the distinctions found in but also gained new distinctions (see Du Nay 1996:49). Du Nay distinguished an earlier stage, Ancient Daco-Romanian, spoken from the 12th to 15th centuries (ibid.,). There is no written evidence of this stage, however. While the Western Romance languages eventually lost all case distinctions, CSRm still distinguishes three morphological cases: a combined Nom-Acc, a combined Gen-Dat, and Voc; three genders.[62] The Voc is often replaced by the Nom-Acc and its use is normally considered archaic or unrefined. In addition to the masculine and feminine gender categories, there is also a class of ambigeneric

[62] Personal pronouns make an additional distinction between Nom and Acc. First- and second-person pronouns also have a separate Gen form, derived from the corresponding CL possessive adjectives. Possessive constructions are formed by combining these forms with linking articles, which agree in gender and number with the possessed noun, while nouns and third person pronouns use the linking articles with the Gen-Dat form.

nouns that largely continues the PIE/Latin neuter in membership, but not agreement, instead behaving like masculine in the singular and feminine in the plural, e.g., Nom-Acc definite singular *timp-ul* 'the time' (cf. masculine *lup-ul* 'the wolf') but Nom-Acc definite plural *timpuri-le* 'the times' (cf. feminine *capre-le* 'the goats') (see Mallinson 1988:398-404). Based on their unique distribution, I consider these nouns in my investigation as a distinct gender, i.e., the neuter (see Baerman et al. 2005:82-83). Other Eastern Romance varieties have developed similarly from EVL. Two of these, Aromanian and Megleno-Romanian, are also discussed briefly in my study. The Nom-Acc inherited the functions of both cases. The Gen-Dat competed with prepositions. In late Latin, *ad* was used to replace the Dat, *de* for the Gen. Since early Romance, the preposition *a* instead has been used in place of the Dat, whereas *de* continued to be used instead of the Gen (see Salvi 2011:342). In addition, new prepositions were also grammaticalized with objects in the Gen-Dat. Those grammaticalized from nouns are used with a possessive construction, i.e., linking articles and the separate Gen form of the first- and second-person pronouns, as expected for the complement of a noun. Those grammaticalized from verbal adjectives take the bare Gen-Dat; this resembles the earlier use of the Dat for the complements of certain adjectives and verbs (see Wahlström 2015:121).

Table 30. Changes in Noun Declension from Pre-Romanian (reconstructed) to Contemporary Standard Romanian

	< ā-stems		< ā-stems		< o-stems	
	Fem.		Masc.		Masc.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-ă	-e	-ă	-e » -i	-ui » -Ø	-i
Voc	-ă > -ă, » -o	-e	-ă	-e » -i	-e > -e, » -Ø	-i
Acc	-ă	-e	-ă	-e » -i	-u > -Ø	-ui » -i
Gen-Dat	-e	-e; -i » -e	-e » -ă	-i; -e » -i	-u > -Ø	-i

	< ō-stems		< C-/i-stems		< C-/i-stems	
	Neut.		Masc.		Fem.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-u > -Ø	-e; -ură » -uri	-Ø, -i » -e	-i	-Ø, -i » -e	-i
Voc	-u > -Ø	-e; -ură » -uri	-Ø, -i » -e	-i	-Ø, -i » -e, -o	-i
Acc	-u > -Ø	-e; -ură » -uri	-e	-i	-e	-i
Gen-Dat	-u > -Ø	-e; -i, -ură » -uri	-i » -e	-i; -eu » -i	-i	-i; -eu » -i

Table 30. Continued

	< C-/i-stems	
	Neut.	
	Sg.	Pl.
Nom	-Ø	-(ur)ă » -(ur)i
Voc	-Ø	-(ur)ă » -(ur)i
Acc	-Ø	-(ur)ă » -(ur)i
Gen-Dat	-i » -Ø	-eu, -(ur)ă » -(ur)i

Based on Alkire and Rosen's (2010:280) reconstruction, the high degree of number syncretism predicted for Pre-Romanian was reduced by analogical processes in the development of CSRm. At the same time, the distinction between Nom and Acc was lost, and all case distinctions were leveled for indefinite nouns, except in the feminine singular. For masculine *o*-stems, the Nom singular was leveled to the Acc, which lost *-u* by regular sound change (see Alkire & Rosen 2010:259), e.g., Nom singular **lupui* » *lup*, Acc *lupu* > *lup* 'wolf', while the Acc plural was leveled to the Nom/Voc, e.g.,**lupui* » *lupi*. Either process would have removed number syncretism, so the existing Nom/Acc syncretism in other declension classes may have been an additional factor. In the plural, the *-i* form was already shared with masculine consonant- and *i*-stems, e.g., *vulpi* 'foxes', so these may have exerted an influence as well. Furthermore, the Gen-Dat plural already shared this form with the Nom. These factors appear to have overcome the iconic principle, which would favor the more salient *-ui* form. In the singular, the last distinction between the neuter and masculine forms was lost with this leveling. These processes and their results are very similar to those that occurred in the Western Romance languages and the Germanic languages that lost all case distinctions on nouns. The main differences in Romanian are the survival of a distinctive feminine Gen-Dat singular form and the reliance on definite articles to maintain other case distinctions, as in CSHG.

For masculine consonant- and *i*-stems, the leveling of the Gen-Dat singular to the Acc, e.g., **câini* » *câine* 'dog', removed syncretism with the Nom/Voc/Acc plural. Feminine consonant- and *i*-stems did not undergo this leveling, probably due to the influence of feminine *ā*-stems, which retain a distinctive Gen-Dat singular form. On the other hand, the masculine *ā*-stem Gen-Dat singular form was leveled to *-ă*, pronounced similarly to [ə], the form already shared by the remaining cases in the singular, e.g.,**tate* » *tată* 'father'. In addition, this class adopted the *-i* plural form that all other masculine nouns had. Within this class, the Gen-Dat

already had this form, but the influence of the larger classes of masculine nouns on this smaller class was likely more important. There are no masculine *ā*-stem adjectives, but all of the other analogical processes described above applied to the corresponding adjective forms as well. Thus, gender profiling, in the form of increased predictability of gender based on declension, likely played a role in the development of Romanian inflection, as it did in French.

In addition, the Nom/Voc singular of consonant- and *i*-stems was leveled to the Acc, e.g., **vulpi* » *vulpe* 'fox'. For some of these nouns, this eliminated syncretism with the Nom/Voc/Acc plural, and this may have also been the last class in which the distinction between Nom and Acc remained. The leveling of the Gen-Dat plural to the Nom/Voc/Acc, optional in Pre-Romanian, became obligatory as the last case distinctions in the plural were lost. The neuter plural form -*ură*, which had spread from certain consonant-stems to other neuters, developed into -*uri* by analogy with other plurals in -*i*, including feminine consonant- and *i*-stems, e.g.,**vină* » **vinură* » *vinuri* (see Alkire & Rosen 2010:282).

In contrast to indefinite nouns and adjectives, determiners, including the postposed definite article and demonstratives, have retained a distinctive Gen-Dat form in both numbers and both gender paradigms.[63] Sound changes had many of the same effects as on noun and adjective inflection, but they resulted in slightly different syncretism patterns in the singular due to the unique determiner forms. In Pre-Romanian, the loss of -*s* had already brought the Gen singular forms closer to the Dat forms, e.g., masculine/neuter *illuius* > **illuiu* 'that (one)'. As noted by Alkire and Rosen, /i/ and /u/ desyllabified after vowels and were lost after consonants. Assuming desyllabification occurred first and the resulting glide behaved as a consonant, the Gen form became identical to the Dat, e.g., masculine/neuter Gen **(il)luiu* > **(il)luju* > **(il)luj*, Dat **(il)lui* > **(il)luj* (2010:259). The feminine forms developed in parallel, again gaining the Gen/Dat syncretism consistently found for *ā*-stem nouns and adjectives. Thus, these changes account for the modern Gen-Dat forms of the postposed article: masculine -*lui* (pronounced [luj]) and feminine -*ei* (pronounced [ej]).[64]

[63] Adjectives usually follow their head noun but can precede it. The definite article is attached to the end of whichever element comes first in the NP, e.g., *om-ul bun*, *bun-ul om* 'the good man' (Mallinson 1988:408).

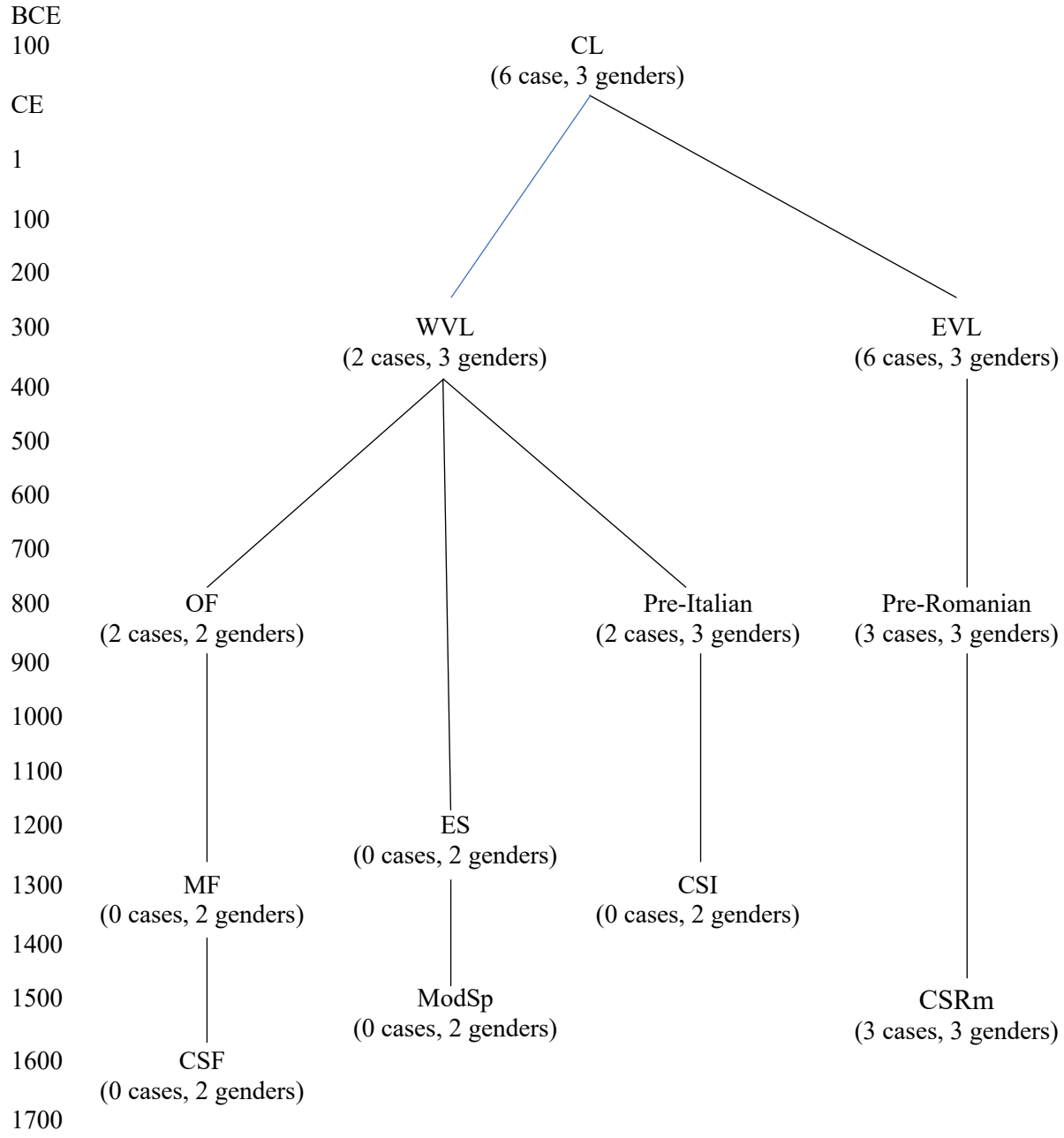
[64] In CSRm, but not Megleno-Romanian, the /l/ in this form was lost unexpectedly, probably by analogy with the Nom-Acc definite form -*a* < EVL **(il)la*, where it was lost by regular sound change. Cf. *stea* 'star' < CL *stella*, in which /l/ was also lost before -*a* (see Alkire & Rosen 2010:265, 281, 349). The vowel *a* is pronounced [a] regardless of stress, so it always contrasts with *ă*.

In addition to these differences, not all of the same analogical processes applied. Since the Gen was already syncretic with the Dat by regular sound change, and this form was syncretic not with the Nom/Voc but the Acc plural, which was itself leveled to the Nom/Voc, there was no motivation for it to be leveled to the other singular form. In the plural, determiners retain the masculine/neuter Gen form as the Gen-Dat form of all genders, e.g., the postposed article *-lor*, as well as the pre-nominal demonstratives *acestor* 'these' and *acelor* 'those', instead of losing all case distinctions as with nouns and adjectives. This form supplanted the expected feminine form, e.g., **(il)laru* » *-lor*. This process likely occurred early, since it also occurred in the Western Romance pronoun forms derived from *illorum*, the Gen plural of *ille* 'that (one)', e.g., Italian *loro* 'they/their', French *leur* 'their/to them' (see Alkire & Rosen 2010:201-203). In fact, the extension of the *o*-stem form to other declensions also occurred for nouns in WVL before these forms were lost entirely (see Banniard 2013:100). As for the retention of any distinctive Gen form in the plural, there are several possible motivations. One is that Romanian, like the Western Romance languages, derived both the definite article and the third-person pronouns from *ille*. In Western Romance, the close association of these forms was not enough for articles to maintain any of the distinctions found in pronouns, but perhaps it was in Romanian due to the additional support of the distinctive Gen-Dat on singular feminine indefinite nouns. Another factor is that determiners, in contrast to indefinite nouns, retain a distinctive Gen-Dat form in the masculine singular, which is unmarked in terms of both gender and number. Without the support of this distinction in the unmarked form, indefinite nouns of all genders likely could not maintain a distinctive Gen-Dat plural form, but determiners could.

2.4.3. Summary

This section has discussed case and gender developments in the Romance languages. The facts of these changes are summarized in Table 31 below.

Table 31. Timeline Summary of Case and Gender in Romance Languages



Romanian is the only Romance language that retains morphological case and three genders. Possible reasons for its divergent development are discussed in chapters III and IV below. Before the complete loss of morphological case, the Western Romance languages passed through a stage with a single opposition between Nom and Acc, in contrast to most Germanic languages. The two-gender system in the Western Romance languages also contrasts with those in Germanic languages. In the former, the masculine and neuter merged; in the latter, the masculine and feminine did.

As in the Germanic languages, sound changes account for many but not all of the case and gender distinctions that have been lost in the Romance languages. In both WVL and EVL, final *-m* was lost, *e* and *o* merged with short *i* and short *u*, respectively, in the final syllable, and phonemic vowel quantity was lost. The specifics of these sound changes differ from the final consonant losses and vowel reduction in the Germanic languages, but they had similar effects and likely also a similar cause, the shift to rhythmic dynamic stress. The sound changes in VL neutralized most of the distinctions among the Acc and other oblique cases in the singular, as well as the Nom for *ā*-stems. In most singular paradigms, only the Gen and Dat remained distinct from the Acc, and only the Gen for *o*-stems. These distinctive forms were leveled to the Acc in WVL but not EVL, as were the Gen and Dat/Abl plural forms. As with most instances of leveling of all forms of one case to another in the Germanic languages, these analogical processes decreased number syncretism overall, but this was not necessarily the primary motivation. Number profiling more likely motivated the leveling in the singular, where the Gen was syncretic with the Nom and sometimes Acc plural in all masculine and feminine classes. The leveling in the plural and in the neuter classes likely followed because the distinctions could not be maintained only in these less frequent and more marked paradigms. Several minor declension classes were absorbed by larger classes on the basis of gender and shared forms.

Although the Romance languages diverged after the VL period, they still underwent a number of similar sound changes and analogical processes. However, unstressed vowels underwent uniquely drastic sound changes in OF: *a* was reduced to [ə], and other vowels were lost, with certain exceptions. Although the Nom-Acc distinction survived in OF, the very limited set of forms that remained resulted in a relatively opaque system of nominal inflection. Final *-s* was retained in OF and ES, but became [j] and formed diphthongs with preceding vowels in Pre-Italian and Pre-Romanian. These diphthongs were eventually monophthongized in unstressed

syllables, except -*ui* (< -us/-os) in Pre-Romanian. As a result, the Nom-Acc distinction was neutralized in an additional plural paradigm in both languages and an additional singular paradigm in Pre-Italian. In Pre-Romanian, syncretism involving the oblique cases also increased. Likely in response, the Gen forms that remained distinct from the Dat were leveled. The leveling of the Gen/Dat plural to the Nom also began in Pre-Romanian and was completed by CSRm. Thus, all oblique case distinctions were finally lost in the plural of indefinite nouns, as had occurred much earlier in WVL. The changes to -*s* may have accelerated the loss of the Nom-Acc distinction on nouns in Italian and Romanian, but it was leveled in any paradigms where it remained in the development of all modern Romance languages regardless. The Nom form was almost always leveled to the Acc form, except for the masculine *o*-stem plural in CSI and CSRm, where the Acc form was syncretic with the singular. Thus, number profiling likely played a role at least in the direction of leveling. In OF and CSRm, declension classes were reorganized in a similar way that contrasted with the other Romance languages; masculine *a*-stems, consonant-stems, and *i*-stems diverged from feminine nouns in these classes, instead adopting a syncretism pattern that matched masculine *o*-stems. This helped make the gender-declension relationship more transparent.

2.5. Changes in the Case Systems of the Balkan Sprachbund Languages

Like the Germanic and Romance languages, the languages of the Balkan Sprachbund have experienced increased syncretism and case loss or reduction. Developments in Romanian and other Eastern Romance varieties, both before and after the formation of the Balkan Sprachbund, are discussed in section 2.4.2 above. This section addresses developments in the other major Balkan Sprachbund languages: Bulgarian, Macedonian, Greek, and Albanian. These languages are all IE, but are not closely related, except for Bulgarian and Macedonian, which are both South Slavic languages. The discussion of the Slavic Balkan Sprachbund languages begins with an overview of Proto-Slavic nominal inflection, for their Slavic historical context. The development of Bosnian-Croatian-Montenegrin-Serbian (BCMS), a group of nearly identical South Slavic languages, is included for comparison, since the standard variety of BCMS is not part of the Balkan Sprachbund. The pre-Sprachbund development of Greek and Albanian is included in their respective sections as well.

2.5.1. Proto-Slavic

Proto-Slavic (PS) is the reconstructed hypothetical ancestor of all Slavic languages. My discussion will focus on Late Proto-Slavic (LPS), which is dated to around 600-700 CE and reflects a stage during which dialectal differentiation had begun but was not yet significant. In the development of LPS from PIE, syncretism increased due to a combination of sound changes and analogical processes. In addition, a distinction arose between hard and soft stems for certain forms at this stage, i.e., between *o*-stems and *jo*-stems, as well as *ā*-stems and *jā*-stems.

Table 32. Late Proto-Slavic Noun Declension[65]

	o-stems					
	Masc.			Neut.		
	Sg.	Du.	Pl.	Sg.	Du.	Pl.
Nom	-ъ	-a	-i₂	-o	-ě₂	-a
Voc	-e/-u	-a	-i₂	-o	-ě₂	-a
Acc	-ъ	-a	-y₂	-o	-ě₂	-a
Gen	-a	-u	-ъ	-a	-u	-ъ
Dat	-u	-oma	-omъ	-u	-oma	-omъ
Loc	-ě₂	-u	-ě₂xъ	-ě₂	-u	-ě₂xъ
Instr	-omь, -ъmь	-oma	-y	-omь, -ъmь	-oma	-y

	ā-stems			u-stems			i-stems					
	Fem.(/Masc.)			Masc.			Fem.			Masc.		
	Sg.	Du.	Pl.	Sg.	Du.	Pl.	Sg.	Du.	Pl.	Sg.	Du.	Pl.
Nom	-a	-ě₂	-y₂	-ъ	-y	-ove	-ь	-i	-i	-ь	-i	-ьje
Voc	-o	-ě₂	-y₂	-u	-y	-ove	-i	-i	-i	-i	-i	-ьje
Acc	-ǫ	-ě₂	-y₂	-ъ	-y	-y	-ь	-i	-i	-ь	-i	-i
Gen	-y₂	-u	-ъ	-u	-ovu	-ovь	-i	-ьju	-ьjь	-i	-ьju	-ьjь
Dat	-ě₂	-ama	-amъ	-ovi	-ъma	-ъmъ	-i	-ьma	-ьmъ	-i	-ьma	-ьmъ
Loc	-ě₂	-u	-axъ	-u	-ovu	-ъxъ	-i	-ьju	-ьxъ	-i	-ьju	-ьxъ
Instr	-ojǫ	-ama	-ami	-ъmь	-ъma	-ъmi	-ьjǫ	-ьma	-ьmi	-ьmь	-ьma	-ьmi

[65] The reconstructed phonemic values in LPS for the non-IPA symbols are as follows: *ь is [ĭ], *ъ is [ŭ], *y is [ɨ], *ě is [æ], *ę is [ẽ], and *ǫ is [õ] (see Vakareliyska & Gyllin in press:13). The remaining symbols, including *x, correspond to their values in IPA. The *jers* *ь and *ъ ultimately derive from the PIE short (normal-grade) vowels *i and *u, respectively. The loss of *jers* in weak positions, as described in section 2.5.2.1 below, was already underway in LPS. The high mid-vowel *y derives from late PIE *ū, and the nasal vowels *ę and *ǫ derive from front and back vowels, respectively, followed by a nasal consonant. Subscripts are used to distinguish vowels with different sources and/or behavior in alternations between vowels in hard and soft endings. The vowel *ě derives from PIE *ē, while *ě₂ derives from the PIE diphthongs *ai and *oi. The vowel *i derives from PIE *ī, while *i₂ derives from the PIE diphthong *ei. The alternations between hard and soft endings, respectively, are as follows: *o ~ *e, *ъ ~ *ь, *ě₂ ~ *i₂, *y ~ *i, *y₂ ~ *ę (see Vakareliyska & Gyllin in press:8-13, 20, 31). Only the hard endings are listed in Table 32; the soft endings can be derived by applying the above alternations.

Table 32. Continued

	Consonant-stems								
	Masc.			Fem.			Neut.		
	Sg.	Du.	Pl.	Sg.	Du.	Pl.	Sg.	Du.	Pl.
Nom	-Ø	-i	-e	-Ø	-i	-i	-Ø	-ě	-a
Voc	-Ø	-i	-e	-Ø	-i	-i	-Ø	-ě	-a
Acc	-ь	-i	-i	-ь	-i	-i	-Ø	-ě	-a
Gen	-e	-u	-ъ	-e	-u	-ъ	-e	-u	-ъ
Dat	-i	-ьma	-ьmъ	-i	-ьma	-ьmъ	-i	-ьma	-ьmъ
Loc	-e	-u	-ьxъ	-e	-u	-ьxъ	-e	-u	-ьxъ
Instr	-ьmь	-ьma	-ьmi	-ьjǫ	-ьma	-ьmi	-ьmь	-ьma	-y

(Schenker 1993:87, Vakareliyska & Gyllin in press:30-34)

LPS retained all of the PIE cases except the Abl, which merged completely with the Gen as in Proto-Greek (see Schenker 1993:85, Lunt 2001:222, Sihler 1995:256). It also retained the three IE genders and numbers. Already in PIE, the Abl was syncretic with the Dat in the dual and plural, and with the Gen in the singular of all paradigms except *o*-stems. For *(j)o*-stems, it was actually the Abl singular form whose reflexes survived. In the singular, the distinction between Nom and Acc was neutralized by sound change for *i*-stems and *u*-stems. The *ā*-stem and *jā*-stem Acc plural forms spread to the Nom plural and Gen singular by analogy with feminine *i*-stems, thereby further reducing the distinctions between Nom and Acc. The Instr plural became syncretic with the Acc for *o*-stems and Nom/Voc for *jo*-stems, e.g., masculine *o*-stem Acc/Instr plural **orby* 'slaves'. The *u*-stem Voc singular *-u* spread to masculine *jo*-stems, where it was syncretic with the Dat singular, e.g., **dъdju* 'rain' (Schenker 1993:85-89). Syncretism between the *(j)ā*-stem Dat and Loc singular forms arose due to vowel contraction. Vakareliyska and Gyllin (in press:31-32) and Schenker (1993:87), among other Slavists, consider this process to have occurred in (late) PIE.[66] The Gen, Loc, and Voc singular all became syncretic for *i*-stems and *u*-stems. For *i*-stems, the Dat singular also joined this syncretism. Syncretism between the Nom and Voc singular spread from some to all consonant-stems. Finally, most consonant-stems lost the distinction between Gen and Loc singular, although some developed Dat/Loc syncretism instead.

[66] Some non-Slavist scholars such as Sihler (1995:248) for Greek and Italic, and Ringe for Germanic (2006:41-50), however, list these forms as distinctive for all PIE nouns in their reconstructions. The difference may be that they have reconstructed PIE forms that are compatible with families that split off earlier than Slavic.

The case functions in LPS had not changed much from PIE. The Nom continued to mark the subject of finite verbs and the complement of 'be' and other linking verbs. The Voc was still used for direct address. The Acc continued to mark DOs and motion toward (goal), including on the objects of prepositions with this meaning. The Dat marked the IO and was also used for other loosely connected functions such as inalienable possession, as well as the objects of a few prepositions. The Loc continued to indicate location, with and without a preposition. The Instr continued to mark the instrument. It marked the objects of prepositions indicating accompaniment and cause. The main development from PIE involved the merger of the Abl with the Gen: The Gen inherited the Abl functions of motion from (usually with a preposition), separation, and standard of comparison, in addition to its original functions, which included complement of a noun and partitive. All surviving cases except the Voc marked the complements of certain verbs (see Duridanov 1956:181-231).[67]

2.5.2. South Slavic

Old Church Slavonic (OCS), the oldest attested South Slavic language, was used from the 9th to the early 11th centuries as a written liturgical language. It was not a spoken language and was an artificial construct built on certain Macedonian dialects of LPS. Liturgical texts written after c. 1100 generally are considered to be in the later regional variants of OCS, including BChS, Russian Church Slavonic, and Serbian Church Slavonic. OCS is the closest written language to LPS. It distinguished seven cases morphologically: Nom, Voc, Acc, Gen, Dat, Loc, and Instr; the three IE genders; and the three IE numbers. A comparison of the nominal inflections attested for OCS (Lunt 2001:54, 72-73) and reconstructed for LPS (see, for example, Schenker 1993:86-87) reveals that they are in fact the same. A number of sound changes applied in the interim, but with no significant effect on noun endings. Since OCS is actually attested and is South Slavic, it is the best source of evidence for comparisons between LPS and later South Slavic languages, the vernacular languages such as Bulgarian, Macedonian, and BCMS, which do not directly descend from it.

[67] Duridanov describes the functions of the cases in OCS as part of his discussion of their later developments in MB. The case functions in LPS generally are not likely to differ much from those in OCS.

2.5.2.1. Eastern South Slavic: The Development of Bulgarian and Macedonian[68]

Late OCS and early BChS texts reflect to a significant extent an early stage in the development of spoken Bulgarian. The same morphological distinctions in nominal inflection are found in these texts as in canonical OCS texts are found in these texts, but orthographic evidence from the manuscripts indicates that a number of sound changes have applied. Therefore, I have attempted a reconstruction of the nominal inflection for a stage in which the sound changes had applied but major analogical leveling and functional mergers were still in progress. I call this hypothetical stage that follows LPS 'Pre-Middle Bulgarian' (PMB). The chronology of phonological and morphosyntactic changes is almost certainly more complex than this suggests, but this reconstruction of PMB enables a clearer analysis of the effects of sound change on the case system. In other words, this reconstruction demonstrates the extent to which case loss could be motivated by sound change alone; at the same time, it highlights the gaps that must be explained by other factors. I have reconstructed two dialectal variants of PMB: one for eastern dialects, in which the vowel raising processes described below tended to occur, and one for western dialects, in which vowel raising did not occur. These dialects also differed in the reflexes of the nasal vowels. As with the reconstruction of PMB in general, the division into two representative dialects is hypothetical; the actual dialect situation was clearly more nuanced. However, this division allows for an analysis of how the different nasal reflexes and the presence or absence of vowel raising may have changed which syncretism patterns arose. Although a distinction between Bulgarian and Macedonian would be anachronistic at this early stage, the division into two dialect groups makes it possible to consider some of the dialectal differences that contributed to the divergent development of these two languages: Bulgarian from eastern dialects, and Macedonian from western dialects.

[68] The grammatical information is drawn from Schenker (1993) for LPS; Lunt (2001), and Nandriş (1969) for OCS; Vakareliyska and Gyllin (in press) for LPS and OCS; Vakareliyska (2008), and Duridanov (1956) for BChS; Gyllin (1991), and Wahlström (2015) for MB; Sussex & Cubberley (2011), and Tomić (2006) for CSB and CSM.

Table 33. Changes in Noun Declension from Late Proto-Slavic to Eastern Dialects of Pre-Middle Bulgarian (reconstructed)

	< o-stems (inanimate)		< o-stems (animate)		< jo-stems (inanimate)	
	Masc.		Masc.		Masc.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-ъ > -∅	-i	-ъ > -∅	-i	-ь > -∅	-i
Voc	-e > -e, -i	-i	-e > -e, -i	-i	-u	-i
Acc	-ъ > -∅	-y > -i	-ъ » -a, -ŭ	-y > -i	-ь > -∅	-ę > -e, -i
Gen	-a > -a, -ŭ	-ъ > -∅	-a > -a, -ŭ	-ъ > -∅	-a > -a, -ŭ	-ь > -∅
Dat	-u	-omъ > -om, -um	-u	-omъ > -om, -um	-u	-emъ > -em, -im
Loc	-ě	-ěxъ > -ěx	-ě	-ěxъ > -ěx	-i	-ixъ > -ix
Instr	-omь > -om, -um	-y > -i, » -mi	-omь > -om, -um	-y > -i, » -mi	-emь > -em, -im	-i

	< jo-stems (animate)		< u-stems		< o-stems	
	Masc.		Masc.		Neut.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-ь > -∅	-i	-ъ > -∅	-ove > -ovi, -uve, -uvi	-o > -o, -u	-a > -a, -ŭ
Voc	-u	-i	-u	-ove > -ovi, -uve, -uvi	-o > -o, -u	-a > -a, -ŭ
Acc	-ь » -a, -ŭ	-ę > -e, -i	-ъ > -∅	-y > -i	-o > -o, -u	-a > -a, -ŭ
Gen	-a > -a, -ŭ	-ь > -∅	-u	-ovъ > -ov, -uv	-a > -a, -ŭ	-ъ > -∅
Dat	-u	-emъ > -em, -im	-ovi > -ovi, -uvi	-ъmъ > -ŭm	-u	-omъ > -om, -um
Loc	-i	-ixъ > -ix	-u	-ъxъ > -ŭx	-ě	-ěxъ > - ěx
Instr	-emь > -em, -im	-i	-ъmь > -ŭm	-ъmi > -mi	-omь > -om, -um	-y > -i, » -mi

	< jo-stems		< ā-stems		< jā-stems	
	Neut.		Fem.(/Masc.)		Fem.(/Masc.)	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-e > -e, -i	-a > -a, -ŭ	-a > -a, -ŭ	-y > -i	-a > -a, -ŭ	-ę > -e, -i
Voc	-e > -e, -i	-a > -a, -ŭ	-o > -o, -u	-y > -i	-e > -e, -i	-ę > -e, -i
Acc	-e > -e, -i	-a > -a, -ŭ	-ǫ > -ŭ	-y > -i	-ǫ > -ŭ	-ę > -e, -i
Gen	-a > -a, -ŭ	-ь > -∅	-y > -i	-ъ > -∅	-ę > -e, -i	-ь > -∅
Dat	-u	-emъ > -em, -im	-ě	-amъ > -am, -ŭm	-i	-amъ > -am, -ŭm
Loc	-i	-ixъ > -ix	-ě	-axъ > -ax, -ŭx	-i	-axъ > -ax, -ŭx
Instr	-emь > -em, -im	-i	-ojǫ > -ŭ	-ami > -ami, -ŭmi	-ejǫ > -ŭ	-ami > -ami, -ŭmi

Table 33. Continued

	< i-stems (also C-stems)	
	Fem.	
	Sg.	Pl.
Nom	-ь > -∅	-i
Voc	-i	-i
Acc	-ь > -∅	-i
Gen	-i	-ьі > -i
Dat	-i	-ьмъ > -em, -im
Loc	-i	-ьхъ > -ex, -ix
Instr	-ьjǫ > -ŭ	-ьmi > -emi, -imi

Table 34. Changes in Noun Declension from Late Proto-Slavic to Western Dialects of Pre-Middle Bulgarian (reconstructed)

	< o-stems (inanimate)		< o-stems (animate)		< jo-stems (inanimate)	
	Masc.		Masc.		Masc.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-ъ > -∅	-i	-ъ > -∅	-i	-ь > -∅	-i
Voc	-e	-i	-e > -e	-i	-u	-i
Acc	-ъ > -∅	-y > -i	-ъ » -a	-y > -i	-ь > -∅	-ę > -e
Gen	-a	-ъ > -∅	-a	-ъ > -∅	-a	-ь > -∅
Dat	-u	-omъ > -om	-u	-omъ > -om	-u	-emъ > -em
Loc	-ě	-ěxъ > -ěx	-ě	-ěxъ > -ěx	-i	-ixъ > -ix
Instr	-omь > -om	-y > -i, » -mi	-omь > -om	-y > -i, » -mi	-emь > -em	-i

	< jo-stems (animate)		< u-stems		< o-stems	
	Masc.		Masc.		Neut.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-ь > -∅	-i	-ъ > -∅	-ove	-o	-a
Voc	-u	-i	-u	-ove	-o	-a
Acc	-ь » -a	-ę > -e	-ъ > -∅	-y > -i	-o	-a
Gen	-a	-ь > -∅	-u	-ovъ > -ov	-a	-ъ > -∅
Dat	-u	-emъ > -em	-ovi	-ъmъ > -om	-u	-omъ > -om
Loc	-i	-ixъ > -ix	-u	-ъxъ > -ox	-ě	-ěxъ > - ěx
Instr	-emь > -em	-i	-ъmь > -om	-ъmi > -mi	-omь > -om	-y > -i, » -mi

Table 34. Continued

	< jo-stems		< ā-stems		< jā-stems	
	Neut.		Fem.(/Masc.)		Fem.(/Masc.)	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-e	-a	-a	-y > -i	-a	-ę > -e
Voc	-e	-a	-o	-y > -i	-e	-ę > -e
Acc	-e	-a	-ǫ > -a	-y > -i	-ǫ > -a	-ę > -e
Gen	-a	-ь > -∅	-y > -i	-ъ > -∅	-ę > -e	-ь > -∅
Dat	-u	-emъ > -em	-ě	-amъ > -am	-i	-amъ > -am
Loc	-i	-ixъ > -ix	-ě	-axъ > -ax	-i	-axъ > -ax
Instr	-emь > -em	-i	-ojǫ > -a	-ami	-ejǫ > -a	-ami

	< i-stems (also C-stems)	
	Fem.	
	Sg.	Pl.
Nom	-ь > -∅	-i
Voc	-i	-i
Acc	-ь > -∅	-i
Gen	-i	-ьi > -i
Dat	-i	-ьmъ > -em
Loc	-i	-ьxъ > -ex
Instr	-ьjǫ > -a	-ьmi > -mi

One sound change already suggested by OCS manuscripts is the loss of *jers* in weak positions. The last *jer* in a word was weak, a *jer* in the syllable before a weak *jer* was strong, but a *jer* in the syllable before a strong *jer* or any other vowel was weak (see Nandriş 1969:37-38). For all declension classes except *(j)ā*-stems and *i*-stems, this resulted in syncretism between the Instr singular and Dat plural, e.g., *o*-stem Instr singular **rabómь*, Dat plural **rabómъ* > **rabóm* 'slave(s)'. For all declension classes except *u*-stems and *i*-stems, it also gave rise to a zero ending for a marked form, the Gen plural, which resulted in the loss of the distinction between hard and soft stems in this form, e.g., *o*-stem **rábъ* > **rab*, *jo*-stem **mǫ́žь* > **mű́ž* 'men'. The same loss of distinction occurred in the other zero endings that resulted from the loss of final *jers*, i.e., in the Nom and Acc singular of masculine *(j)o*-stems and *u*-stems, except when the Acc was changed by analogy, as discussed below. For *i*-stems, the Gen plural instead became syncretic with the other seven case/number forms with *-i*, e.g., **nítьi* > **níti* 'threads'. In strong positions, the front *jer* ь lowered to *e* in PMB; in western dialects, the back *jer* ъ lowered to [o], but in eastern dialects it lowered to [ə], transcribed here as *ŭ* for continuity with the transcription of Contemporary Standard Bulgarian (CSB) (see Sussex & Cubberley 2011:113).

The regular merger of **i* and **y*, as reflected by free variation of the corresponding vowel graphemes in most BChS manuscripts, caused syncretism between Nom/Voc *-i* and Acc/Instr *-y* of the masculine *o*-stem plural paradigm, e.g., Nom/Voc *rabí*, Acc/Instr *rabý* > **rabí* (see, for example, Vakareliyska 2008:160). Another attested sound change is the reduction of the word-final V_1jV_2 sequence into V_2. This caused syncretism between the Acc and Instr of the *(j)ā*-stem singular paradigms, e.g., *ā*-stem Acc *ženǫ́*, Instr *ženójǫ* > **ženǫ́* 'woman' (Duridanov 1956:198).[69]

A sound change causing the reflexes of the LPS nasal vowels **ǫ* and **ę* to be confused in PMB would account for an additional loss of distinction between the Acc and Gen in the singular *jā*-stem paradigm. The textual evidence for this sound change is inconclusive, however. This confusion is assumed based on the inconsistent use of the Cyrillic grapheme <ѫ> for the back nasal *ǫ* and <ѧ> for the front nasal *ę* in BChS texts (see also Vakareliyska 2008, Wahlström 2015:77, citing Ivanova-Mirčeva & Haralampiev 1999:63). For example, Steinke (1968:36-37) observes the confusion of these graphemes in three apostol (Acts & Epistles) manuscripts from as early as the late 12th and early 13th centuries. The BChS Ohrid Apostol manuscript, the earliest of the three, exhibits a different distribution of the graphemes than that found in OCS: <ѧ> is used after soft consonants, <ѫ> elsewhere. The other two manuscripts, the 880 Apostol and 882 Apostol, display a relatively random use of the nasal graphemes, but one sometimes uses <e> instead of <ѧ>, the expected grapheme in OCS. Despite this apparent confusion, the major reflex of **ę* is *e*, while **ǫ* is generally realized either as *ŭ* [ə], as in CSB, or as *a*, as in Contemporary Standard Macedonian (CSM) (see Sussex & Cubberley 2011:116-117).[70] In the present investigation, these are also the assumed reflexes for the eastern and the western dialects of PMB, respectively.[71] In the western dialects, this would have made the *(j)ā*-stem Nom singular join the syncretism between Acc and Instr singular that resulted from vowel contraction, e.g., *ā*-stem Nom *žená*, Acc/Instr **ženǫ́* > **žená* 'woman'. In PMB, the merger of **ę* and **e* would have caused the Voc singular to join the syncretism among the Gen singular and

[69] The Acc/Instr homonymy in the *ā*-stem singular forms is still found in Slovene, without any effect on the general distinction between these two cases (Priestly 1993:405).

[70] The major reflex of **ě* is also *e* in CSM, but *e* or *ja* depending on the context in CSB (see Sussex & Cubberley 2011:118-121).

[71] In eastern dialects of PMB and MB through to CSB, *ŭ* [ə] has been the reflex of **ǫ* regardless of stress, so this sound can occur in unstressed as well as stressed syllables.

Nom/Voc/Acc plural for *jā*-stems, e.g., Voc singular **zemljé*, Gen singular/Nom/Voc/Acc plural **zemlję́* > **zemlję́* 'land(s)'. To account for the textual evidence in light of distinctive modern reflexes, Mirčev (1978:110-117) assumes that **ǫ* and **ę* temporarily merged in vowel quality but were able to become distinctive again because the consonants before **ę* remained palatalized. On the basis of dialectal evidence, several authors have ruled out such a temporary merger (see, for example, Wahlström 2015:78, citing Koneski 1983:40, and Ivanova-Mirčeva & Haralampiev 1999:63-68). Thus, an alternative explanation should be sought for the written confusion, but an Acc/Gen syncretism in the singular *jā*-stem paradigm could have added to the confusion caused by other mergers.

Vowel raising in unaccented syllables occurred in eastern dialects but not the westernmost dialects: /o/ > [u], /e/ > [i], and /a/ > [ə], transcribed here as *ŭ* (see Wahlström 2015:79-80). There was likely significant variation in when these processes occurred, and only a subset may have occurred in some dialects. Mirčev finds evidence for /o/ > [u] starting in 12th century BChS manuscripts, /a/ > [ə] starting in the 13th century, and /e/ > [i] from the 15th century (1978:144-147). As discussed in section 3.1.1.2 below, these processes are not all equally widespread among modern Bulgarian dialects, so this was probably true at earlier stages as well. Nevertheless, in order to keep the analysis manageable, I have included all three vowel raising processes in the reconstruction for eastern dialects of PMB and none of them for the western dialects. Where they occurred, these changes generally resulted in additional syncretism. When the final syllable was unstressed, as was often but not always the case, the Acc plural joined the existing syncretism among the Loc singular and Nom/Voc/Instr plural for masculine *jo*-stems in eastern dialects, e.g., Loc singular/Nom/Voc/Instr plural **pláči*, Acc plural **pláčę* > **pláči* 'crying'. Under the same conditions, the Nom/Voc/Acc singular joined the existing syncretism between the Loc singular and Instr plural for neuter *jo*-stems, e.g., Nom/Voc/Acc singular **známenije*, Loc singular/Instr plural **známeniji* > **známiniji* 'sign'. In fact, the earliest examples of substitutions reflecting this neutralization occurs with such neuters (see Sobolev 1991:33). When the final syllable was unstressed, neuter *o*-stems also had syncretism among the Nom/Voc/Acc and Dat singular, e.g., Nom/Voc/Acc **lě́to*, Dat **lě́tu* > **lě́tu* 'summer'. When the ending was unstressed, the *jā*-stem Voc and Gen singular and Nom/Voc/Acc plural, all already syncretic due to denasalization, became syncretic with the Dat/Loc singular as well, e.g., Voc/Gen singular/Nom/Voc/Acc plural *júnoše*, Dat/Loc singular **júnoši* > **júnuši* 'young

man/men'. In combination with denasalization, vowel raising in unaccented syllables also made the *(j)ā*-stem Nom singular join the syncretism between Acc and Instr singular that resulted from vowel contraction, e.g., *ā*-stem Nom **žíma*, Acc/Instr **zímǫ* > **zímŭ* 'winter'. As mentioned above, the development of **ǫ* into [a] in western dialects rather than [ə] as in eastern dialects meant this particular pattern of syncretism occurred for all *(j)ā*-stems in the former, as opposed to just those with unstressed endings in the latter. This appears to be the only way in which sound change resulted in greater syncretism in western dialects than in eastern dialects.

My analysis assumes that analogical processes resulting in the complete loss of certain case distinctions were not yet complete in PMB. However, some evidence for the removal of case syncretism by analogical processes is already found in manuscripts from the 12th century, namely the occasional extension of the *u*-stem Instr plural ending *-(ъ)mi* to *o*-stem masculine and neuter nouns (see Wahlström 2015:76, citing Mirčev 1978:168). As in other Slavic languages, the Acc singular of animate masculine *(j)o*-stems adopted the Gen singular form, e.g., *o*-stem **rab* » **rabá* 'slave'. This restored the distinction between Nom and Acc, which is more important for the disambiguation of arguments than the distinction between Acc and Gen. In fact, this change was already starting to appear in OCS, but it was still mostly limited to humans and more common when the Acc marked a DO than the object of a preposition. In addition, it was more common for definite DOs, and in this context inanimates occasionally had this form (see Huntley 1993:136-138). These analogical processes are attributed to PMB due to their early attestation.

Agreement targets in LPS, OCS, and the later South Slavic languages used a combination of nominal and pronominal forms. Adjectives employed short or long forms depending on whether they are definite or not; the former indicates indefiniteness while the latter indicates definiteness (Vakareliyska and Gyllin in press:37). The short form of adjectives simply followed the *(j)ā*-stem noun forms for feminine and *(j)o*-stem forms for masculine and neuter. All three genders were syncretic in the Loc singular and Gen plural. The only other form shared by masculine and feminine was the Acc plural, while the neuter was only distinct from the masculine in the Nom/Voc/Acc of both numbers. These same patterns also applied to the majority of nouns themselves. The long form of adjectives added soft-stem pronominal forms after the nominal endings. In the plural, the resulting combined forms only had a distinctive feminine form in the Nom, while the feminine was completely distinctive in the singular. Despite

different forms, the syncretism pattern between masculine and neuter was the same. Determiners used the pronominal forms (hard or soft depending on the determiner), with the same gender distinctions as the long-form adjectives (Nandriş 1969:91-93, 105-107, 113-115).

The sound changes described above for PMB should only have affected the gender distinction in two ways. First, the feminine form became syncretic with the masculine in the Nom/Voc plural of most agreement targets. Even this form should have remained distinctive for the soft-stem pronominal paradigm, except perhaps when the ending was unstressed in dialects with vowel raising. Even if this distinction sometimes remained, the masculine and feminine had essentially merged in the plural by this stage. Second, the new animate masculine Acc singular form adopted from the Gen became syncretic with the corresponding feminine form for short-form adjectives. However, this distinction remained for the inanimate masculine in all paradigms and for the animate masculine in the pronominal paradigms. In Eastern dialects, it also remained for the animate masculine of short-form adjectives when the ending was stressed. In addition to these neutralizations, however, Vakareliyska observes that a set of substitutions in the Curzon Gospel, a western Bulgarian or Macedonian BChS manuscript from the 14th century, appear to neutralize gender distinctions for certain case markings. Substitutions between <a>/<ꙗ> **(j)a* and <ѧ>/<ѫ> also occurred when the latter represented **ę*, probably due to its denasalization and lowering to the low front vowel [ä]. These look like errors in gender marking on both nouns and possessive adjectives, which used the soft-stem pronominal forms: *o*-stem masculine Gen singular (-*a*) for *jā*-stem feminine Gen singular (-*ę*) and vice versa; *jā*-stem feminine Nom/Acc plural (-*ę*) or *jo*-stem masculine Acc plural (-*ę*) for neuter Nom/Acc plural (-*a*) and vice versa (2008:161-162). However, no general distinctions in grammatical gender would have been lost, even if these substitutions reflect a merger in the underlying vernacular.

Middle Bulgarian (MB), which consisted of a spectrum of Eastern South Slavic dialects, was spoken from around the 12th to 14th centuries. Noun paradigms cannot be accurately reconstructed because there was no attested writing in the vernacular until the damascenes of the 17th century, which originated as translations of edifying (non-liturgical) religious literature written in vernacular Greek (Gyllin 1991:13-14, 46-50). However, Wahlström has provided a reconstruction of the MB case system based on the BChS corpus, the earliest vernacular texts, and evidence from modern dialects (2015:81-83). This reconstruction represents a hypothetical average over many dialects; I have modified it for my analysis to reflect the differences between

eastern and western dialects that are expected to have arisen based on the different sound changes, as suggested by the BChS manuscripts from these regions and by later dialectal data. As reconstructed, the case systems of both representative dialects distinguish four cases morphologically: Nom, Voc, Acc, Dat. These resulted from the merger of Gen and Dat, as well as the merger of Acc, Loc, and Instr. They continue to distinguish the three Slavic genders and two numbers. Of course, BChS manuscripts and modern dialects indicate more dialectal variation than reflected by the two representative dialects, and some dialects likely had case systems that differed significantly from these.

Table 35. Changes in Noun Declension from Pre-Middle Bulgarian (reconstructed) to Middle Bulgarian (reconstructed) for Eastern Dialects

	< o-stems (inanimate) (also some u-stems)		< o-stems (animate) (also some u-stems)		< jo-stems (inanimate)		< jo-stems (animate)	
	Masc.		Masc.		Masc.		Masc.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-Ø	-i	-Ø	-i	-Ø	-i	-Ø	-i
Voc	-e, -i	-i	-e, -i	-i	-u	-i	-u	-i
Acc	-Ø	-i	-a, -ŭ	-i	-Ø	-i; -e » -i	-a, -ŭ	-i; -e » -i
Gen	-a, -ŭ » -u	-Ø » -om, -um	-a, -ŭ » -u	-Ø » -om, -um	-a, -ŭ » -u	-Ø » -em, -im	-a, -ŭ » -u	-Ø » -em, -im
Dat	-u	-om, -um	-u	-om, -um	-u	-em, -im	-u	-em, -im
Loc	-ě » -Ø	-ěx » -i	-ě » -a, -ŭ	-ěx » -i	-i » -Ø	-ix » -i	-i » -a, -ŭ	-ix » -i
Instr	-om, -um » -Ø	-i; -mi » -i	-om, -um » -a, -ŭ	-i; -mi » -i	-em, -im » -Ø	-i	-em, -im » -a, -ŭ	-i

	< u-stems (also some o-stems)		< o-stems		< jo-stems	
	Masc.		Neut.		Neut.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-Ø	-ovi, -uve, -uvi	-o, -u	-a, -ŭ	-e, -i	-a, -ŭ
Voc	-u	-ovi, -uve, -uvi	-o, -u	-a, -ŭ	-e, -i	-a, -ŭ
Acc	-Ø	-i » -ovi, -uve, -uvi	-o, -u	-a, -ŭ	-e, -i	-a, -ŭ
Gen	-u	-ov, -uv » -om, -um	-a, -ŭ » -o, -u	-Ø » -a, -ŭ	-a, -ŭ » -e, -i	-Ø » -a, -ŭ
Dat	-ovi, -uvi » -u	-ŭm » -om, -um	-u > -u, » -o	-om, -um » -a, -ŭ	-u » -e, -i	-em, -im » -a, -ŭ
Loc	-u » -Ø	-ŭx » -ove, -ovi, -uve, -uvi	-ě » -o, -u	-ěx » -a, -ŭ	-i > -i, » -e	-ix » -a, -ŭ
Instr	-ŭm » -Ø	-mi » -ove, -ovi, -uve, -uvi	-om, -um » -o, -u	-i, -mi » -a, -ŭ	-em, -im » -e, -i	-i » -a, -ŭ

Table 35. Continued

	< ā-stems		< jā-stems		< i-stems	
	Fem.(/Masc.)		Fem.(/Masc.)		Fem.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-a, -ŭ	-i	-a, -ŭ	-e, -i	-Ø	-i
Voc	-o, -u	-i	-e, -i	-e, -i	-i	-i
Acc	-ŭ	-i	-ŭ	-e, -i	-Ø	-i
Gen	-i	-Ø » -am, -ŭm	-e, -i	-Ø » -am, -ŭm	-i	-i » -em, -im
Dat	-ě » -i	-am, -ŭm	-i > -i, » -e	-am, -ŭm	-i	-em, -im
Loc	-ě » -ŭ	-ax, -ŭx » -i	-i » -ŭ	-ax, -ŭx » -e, -i	-i » -Ø	-ex, -ix » -i
Instr	-ŭ	-ami, -ŭmi » -i	-ŭ	-ami, -ŭmi » -e, -i	-ŭ » -Ø	-mi » -i

Table 36. Changes in Noun Declension from Pre-Middle Bulgarian (reconstructed) to Middle Bulgarian (reconstructed) for Western Dialects

	< o-stems (inanimate) (also some u-stems)		< o-stems (animate) (also some u-stems)		< jo-stems (inanimate)	
	Masc.		Masc.		Masc.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-Ø	-i	-Ø	-i	-Ø	-i
Voc	-e	-i	-e	-i	-u	-i
Acc	-Ø	-i	-a	-i	-Ø	-e » -i
Gen	-a » -u	-Ø » -om	-a » -u	-Ø » -om	-a » -u	-Ø » -em
Dat	-u	-om	-u	-om	-u	-em
Loc	-ě » -Ø	-ěx » -i	-ě » -a	-ěx » -i	-i » -Ø	-ix » -i
Instr	-om » -Ø	-i; -mi » -i	-om » -a	-i; -mi » -i	-em » -Ø	-i

	< jo-stems (animate)		< u-stems (also some o-stems)		< o-stems	
	Masc.		Masc.		Neut.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-Ø	-i	-Ø	-ove	-o	-a
Voc	-u	-i	-u	-ove	-o	-a
Acc	-a	-e » -i	-Ø	-i » -ove	-o	-a
Gen	-a » -u	-Ø » -em	-u	-ov » -om	-a » -o	-Ø » -a
Dat	-u	-em	-ovi » -u	-om » -om	-u » -o	-om » -a
Loc	-i » -a	-ix » -i	-u » -Ø	-ox » -ove	-ě » -o	-ěx » -a
Instr	-em » -a	-i	-om » -Ø	-mi » -ove	-om » -o	-i, -mi » -a

Table 36. Continued

	< jo-stems		< ā-stems		< jā-stems	
	Neut.		Fem.(/Masc.)		Fem.(/Masc.)	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-e	-a	-a	-i	-a	-e
Voc	-e	-a	-o	-i	-e	-e
Acc	-e	-a	-a	-i	-a	-e
Gen	-a » -e	-∅ » -a	-i	-∅ » -am	-e	-∅ » -am
Dat	-u » -e	-em » -a	-ě » -i	-am	-i » -e	-am
Loc	-i » -e	-ix » -a	-ě » -a	-ax » -i	-i » -a	-ax » -e
Instr	-em » -e	-i » -a	-a	-ami » -i	-a	-ami » -e

	< i-stems	
	Fem.	
	Sg.	Pl.
Nom	-∅	-i
Voc	-i	-i
Acc	-∅	-i
Gen	-i	-i » -em
Dat	-i	-em
Loc	-i » -∅	-ex » -i
Instr	-a » -∅	-mi » -i

Sound change and analogical processes can account for the merger of Instr and Acc in MB. As mentioned above, the Nom, Acc, and Instr plural had all become syncretic for masculine *o*-stems. The extension of the *u*-stem Instr plural ending *-mi* did not catch on in the same way that similar processes did in Russian and BCMS. Another syncretism in the masculine *(j)o*-stem paradigms was between the Instr singular and Dat plural. This number syncretism may have been resolved by the leveling of the Instr singular to the Acc singular, e.g., *o*-stem virile-class[72] **rabóm* » **rabá* 'slave'. The same process may have occurred in *u*-stems and neuter *(j)o*-stems. As discussed above, the *(j)ā*-stem singular paradigms also had syncretism between Acc and Instr. Thus, most declension classes would have lost the distinction between these two cases in the singular, and some would have in the plural as well. The remaining plural forms might then have been leveled because the distinction was no longer retained in the singular or even all plural forms.

[72] The virile was a certain subset of animate: it marked a healthy free male adult human.

Sound change can only have played a direct role in the merger of the Loc and Acc in eastern dialects of MB, where they were neutralized in the singular of neuter *jo*-stems with unstressed endings, as mentioned above. However, the Loc singular was syncretic with plural forms in several paradigms. In LPS, the Loc singular was already syncretic with the Nom/Voc/Acc plural for feminine *i*-stems and the Nom/Voc/Instr plural for masculine *jo*-stems. In PMB, it had also become syncretic with the Gen plural for the former class, and, when endings were unstressed in eastern dialects, with the Acc plural for the latter class, and the Nom/Voc/Acc plural for *jā*-stems. All of this number syncretism may have been eliminated by the leveling of the Loc to Acc singular, e.g., masculine *jo*-stem inanimate **pláči* » **plač*. At the same time, the Acc plural was leveled to the Nom for masculine *jo*-stems, the last class which retained this distinction in the plural; in eastern dialects of PMB, even this distinction only remained when the endings were stressed, so the leveling could be seen as an extension of unstressed pattern. As for the leveling of the Loc, the choice of the Acc rather than another case such as the Dat, as in BCMS (see section 2.5.2.2 below), may have been motivated by the existing neutralization of Loc and Acc for neuter *jo*-stems, at least in eastern dialects of MB, and by the fact that the Dat was already syncretic with the Loc for *jā*-stems, so it also participated in that number syncretism. However, these processes only account for the neutralization of the Loc and Acc singular in four declension classes, so the neutralization must have then spread to other declension classes. While this is possible, a functional merger between the Acc and Loc may have played a more important role. Duridanov notes that the Acc came to be used on the objects of all prepositions, regardless of their original case (1956:184-187, 197-201, 210-211, 236). In addition, the Loc forms that survive in a few modern dialects such as the Rhodopian dialect of Tihomir are used to express both location and goal, which was originally an Acc function (Stojkov 1968:40). A functional merger could have also played a role in the merger of the Acc and Instr, since the Acc began to replace prepositional uses of the Instr in 12th century BChS manuscripts, before its forms were lost (see Sobolev 1991:31-32). An earlier instrument-accompaniment merger via the spread of the prepositional construction *sъ* 'with' followed by the Instr from the accompaniment function to the instrument function, which was originally expressed with a bare Instr form, meant that all functions could be expressed with prepositional constructions and, by extension, the Acc (Duridanov 1956:197-198).

The merger of Gen and Dat can also be partially attributed to sound change and number profiling. As mentioned above, the *jā*-stem Gen and Dat singular were syncretic in eastern dialects of MB when the ending was unstressed. Hard stem endings were often substituted for soft stem endings in BChS manuscripts, which suggests that hard and soft stems were merging in MB (see Duridanov 1956:222). If *ā*-stem Gen singular *-i* spread to *jā*-stems, then the Gen and Dat singular would have always been syncretic for this class, even when the ending was stressed, e.g., Gen **zemljé* » **zemljí*, Dat **zemljí* 'land'. If *jā*-stem Dat singular *-i* unexpectedly replaced the *ā*-stem form *-ě* instead of the reverse substitution, then these two cases would have been syncretic for *ā*-stems as well, e.g., *ā*-stem Dat **ženě́* » **žení*, Gen **žení* 'woman'. The Dat is generally assumed as the source for the Gen-Dat case in dialects that retain it, so the use of *-i* instead of *-e* (< **-ě*) in some dialects suggests at least one of these replacements occurred. Otherwise, the *ā*-stem Gen would have to be the source instead of the Dat. It is also possible that vowel raising resulted in [i] from unstressed /ě/ as well as /e/, which means sound change alone could result in the syncretism of the *ā*-stem Gen and Dat singular in eastern dialects. However, this process lacks evidence (Wahlström 2015:48). In addition, it is less phonetically natural than the other forms of vowel raising because of the distance between the low vowel [ě] and high vowel [i]. For these reasons, my analysis assumes that raising did not apply to /ě/. Even if it did apply in some MB dialects, another motivation would be needed for the merger in the western dialects, which did not undergo vowel raising.

For masculine *(j)o*-stems, the Gen plural had been syncretic with Nom/Acc singular since LPS.[73] The removal of this number syncretism may have motivated the leveling of the Gen plural to the Dat plural in these paradigms, e.g., *o*-stem **rab* » **rabóm* 'slaves'. Even after these changes, however, Gen and Dat would have remained distinct in most paradigms, in contrast to the development of Pre-Romanian, in which sound change and number profiling can account for the Gen/Dat syncretism in most paradigms. Thus, a functional merger may be the best explanation in MB, since there is evidence of overlap in functions. Duridanov notes that the use of the Dat instead of the Gen to indicate adnominal possession also already occurred in OCS but became more frequent in BChS manuscripts. Unlike the symmetrical substitutions between different case forms in these later manuscripts, the Gen did not generally replace the Dat

[73] As mentioned above, the Acc singular of virile masculine *(j)o*-stems started adopting the Gen singular form in OCS, instead of remaining syncretic with the Nom singular and Gen plural.

(1956:202-203, 218-219). This suggests that sound changes did not play a role, but rather the loss of productivity of the Gen at the expense of the Dat. The influence of other Balkan Sprachbund languages that had already undergone a Gen-Dat merger may have played a role in this process, as discussed in section 3.2.2 below.

The mergers discussed above would have eliminated all case distinctions for singular neuter *o*-stems with unstressed endings in eastern dialects of MB. Unlike other declension classes, the syncretism between Dat and Nom/Voc/Acc meant that the leveling of the Gen to the Dat, e.g., **lě́tŭ* » **lě́tu* 'summer', would have also removed its distinction with the Nom/Voc/Acc. The same leveling of all case distinctions would then have been extended to the plural and the neuter *jo*-stem paradigm. Based on the high number of case substitutions for neuter nouns relative to their frequency in the manuscripts, as observed by Duridanov (1956) and Steinke (1968), among others, Wahlström argues that these were the first paradigms to lose all case distinctions in the development of Bulgarian, perhaps because case was not needed as much for the disambiguation of grammatical relations (2015:187). Since the additional syncretism in unstressed syllables did not occur in western dialects, these may have resisted these changes for longer, while the development of CSM was influenced by CSB.

The leveling of the Instr and Loc to the Acc would actually be expected to increase the distinctiveness of the neuter in MB, since the neuter was distinct from the masculine in the Acc but not Loc or Instr. If case distinctions were lost on neuter agreement targets in response to their loss on nouns, this would make the neuter completely distinctive. In any case, the neuter was not in any danger of being lost. It is possible that the masculine and feminine had merged in the plural. In western dialects, however, their syncretism in the soft-stem Acc plural should have been eliminated by the leveling of the masculine form to the Nom.

The standard Bulgarian literary language developed out of vernacular texts, particularly the damascenes, beginning in the 17th century. It was later codified in the 19th century, as was Macedonian (Gyllin 1991:44-45, 105). Until these 19th-century efforts toward a separate Macedonian linguistic identity, both were considered varieties of the same language (Bulgarian). In contrast to the other contemporary Slavic languages, almost all case distinctions have been lost in CSB and CSM nouns. CSB and CSM distinguish two morphological cases, Nom and Voc; the three Slavic genders; and two numbers, singular and plural, having lost the dual, like most of

the Slavic languages.[74] Distinctive Voc forms occur for many singular animate nouns, especially those referring to humans: masculine *-e*, *-u*, or *-o* (depending on the stem), e.g., *Bože moj!* 'My God!' and *Učitel(j)u!* 'Teacher!'; and feminine *-o* or *-e*, e.g., *Sestro!* 'Sister!' and *Elice!* 'Elica!' (Tomić 2006:49-63, 86-94). The feminine Voc of proper names is insulting in Bulgarian, however, so the Nom forms are used instead. As for functions other than direct address, the Nom form of nouns is used without a preposition for subject, complement of 'be', DO, and object of prepositions. The other former case functions are generally expressed with prepositional constructions, as in the Western Romance languages.

CSB and CSM have generally lost gender agreement in the plural. Adjectives distinguish all three genders in the singular but have a single plural form. A distinction between feminine and neuter plural remains for the postposed definite article, e.g., feminine *knigi-te* 'the books' and neuter *sela-ta* 'the villages', while the choice of form with masculine plural nouns depends on their final vowel, e.g., CSB/CSM *studenti-te* 'the students' and CSB *xora-ta* 'the people'. Likewise, singular masculine nouns generally take the distinctive masculine form, e.g., CSB *student-ŭt*, CSM *student-ot* 'the student', but they take the article associated with another gender if they end in the vowel associated with it, e.g., the CSB masculine singular nouns *vladika-ta* 'the bishop' (cf. feminine *glava-ta* 'the head') and *djado-to* 'the grandpa' (cf. neuter *pismo-to* 'the letter'). Noun modifiers precede their head noun as a rule, and the leftmost constituent in the NP takes the article, but with modifiers it is always determined morphologically rather than phonologically as with nouns (Tomić 2006:56, 89-93). There is no indication that any gender category is under threat in CSB or CSM, although the high number of consonant-final English loanwords, which are rendered as masculine, is greatly increasing the proportion of masculine nouns in the language (see Vakareliyska 2018:358).

[74] In CSB, personal pronouns distinguish Nom and oblique forms, e.g., third-person masculine singular Nom *toj*, oblique *nego*, as well as Acc and Dat clitic forms, e.g., third-person masculine/neuter singular Acc *go*, Dat *mu*. For first- and second-person plural, the same clitic form is used for Acc and Dat, e.g., first-person plural *ni*. In CSM, personal pronouns distinguish three forms: Nom, Acc, and Dat. The Acc and Dat forms have full and clitic forms, e.g., third-person masculine/neuter singular Acc full *nego*, clitic *go*, Dat full *nemu*, clitic *mu*. For first- and second-person singular, the same full form is used for Acc and Dat, e.g., first-person singular *mene*, but the clitic forms are still distinctive, e.g., first-person singular Acc *me*, Dat *mi*.

Table 37. Case Mergers in Bulgarian/Macedonian

	Old Church Slavonic / Eastern Dialects of Pre-Middle Bulgarian / Western Dialects of Pre-Middle Bulgarian							Middle Bulgarian for Eastern Dialects / Middle Bulgarian for Western Dialects				Bulgarian / Macedonian	
PIE	Nom	Voc	Acc	Gen	Dat	Loc	Instr	Nom	Voc	Acc	Dat	Nom	Voc
Nom	x							x				x	
Voc		x							x				x
Acc			x							x		x	
Gen				x							x	x	
Dat					x						x	x	
Loc						x				x		x	
Instr							x			x		x	
Abl				x							x	x	

Instr and Loc were the first cases lost in the development of MB (Mirčev 1958:258). No Instr or Loc case forms were still in productive use in the 17th century damascenes (Gyllin 1991:81). The relative order in which these two cases merged with the Acc is uncertain. Meyer claims that the Loc disappeared before the Instr (1920:47). It is true that BChS manuscripts of the 12th to 14th centuries provide relatively few examples of replacement of the Instr by the Acc compared to the Loc, so it is quite possible that this process began later than the Loc, but by no means certain. In fact, some of the prepositional constructions used in place of the Instr were already attested in OCS (Duridanov 1956:182-195, 201). In addition, if syncretism due to regular sound change and leveling due to number syncretism are considered important factors in the process, it is more likely that the Instr was lost first. After sound changes had applied, i.e., as reconstructed for PMB, syncretism between Acc and Instr occurred in singular *(j)ā*-stems, e.g., western *ā*-stem Nom/Acc/Instr **žená* 'woman', and in plural masculine *o*-stems, e.g., Nom/Voc/Acc/Instr **rabí* 'slave'. In eastern dialects, the Acc and Instr were also syncretic when the final syllable was unstressed in plural masculine *jo*-stems, e.g., Nom/Voc/Acc/Instr **pláči* 'crying'. On the other hand, the only syncretism between Acc and Loc after sound changes was in singular neuter *jo*-stems when the ending was unstressed in these dialects, e.g., Nom/Voc/Acc/Loc **známiniji* 'sign'. Moreover, number profiling may have motivated the leveling of the Instr to the Acc for an additional seven declension classes in the singular and three in the plural, while it motivated the leveling of the Loc for at most four additional declension classes in the singular. Thus, the distinction between Instr and Acc was neutralized by regular sound change and by leveling as part of number profiling in far more paradigms than the distinction between Loc and Acc. The

early loss of the Loc and Instr has parallels in the Germanic languages, Latin, and Greek. As discussed with reference to Greek in section 2.5.3 below, this tendency across IE languages can probably be attributed to their relatively low frequency.

The Acc replaced prepositional uses of the Gen, as well as the Instr and Loc, early in the vernaculars, while other uses of the Gen were mostly replaced by the Dat or analytic constructions before 1400 (Gyllin 1991:77, 80). This suggests that the Gen merged with the Dat after the Instr and Loc had been lost, as does the frequency of these cases in other IE languages. After the sound changes reconstructed for PMB, syncretism between Gen and Dat occurred in one declension class in the singular, along with an additional one in eastern dialects when the syllable was unstressed. Number profiling may have motivated the leveling of the Gen for five declension classes in the plural. Thus, based on syncretism after sound change and leveling as part of number profiling, the distinction between Gen and Dat was slightly less robust than the Loc-Acc distinction, but more robust than the Instr-Acc distinction. In combination with other evidence, the Gen was still most likely lost after the Loc.

In the development of CSB and CSM, the Gen-Dat and Nom both merged with the Acc, but the order is uncertain. On the one hand, Gyllin notes that analytical constructions with the preposition *na* 'on, at, of, to' were competing with the Gen-Dat in the damascenes. The Gen-Dat survived longest in the IO function, but even this function likely came to be limited to human masculine singular referents in the vernaculars (1991:78-79). Thus, it is unclear whether the Gen-Dat should still be considered productive at this later stage. On the other hand, Mirčev argues that Acc forms began to replace Nom forms after the Acc had become the general prepositional case (1958:263). Gyllin observes that most Nom forms were already syncretic with the Acc in MB. As for the classes which retained this distinction in the singular, the *(j)ā*-stems lost it by the 18th century in most dialects, while the animate *(j)o*-stems still had a distinctive Acc form derived from the Gen in the writings of Paisij and probably also some 18th century dialects. Ultimately, the Nom forms won out as the default in these declension classes, but this may have been a later development (1991:80). As a core, high frequency case, the Acc is expected (e.g., by Luraghi 2004) to be lost no sooner than the Gen-Dat, and the modern dialects appear to support this. Tomić notes that northern Macedonian dialects, which are transitional to BCMS, tend to retain the Acc but not Gen-Dat. The same is true of western Macedonian dialects, in which Gen-Dat forms fell out of use around the beginning of the 20th century except for some

proper names and kinship terms in certain dialects (2006:53-54). The Rhodopi Bulgarian dialects continued to use Acc, Dat, and Loc forms at least into 20th century, although they may be archaic for current speakers (ibid., 88). However, there is no evidence for dialects with a distinctive Gen-Dat but not Acc, so the Gen-Dat was more likely lost first in the dialects on which the standard language was based.

2.5.2.2. Western South Slavic: The Development of Bosnian-Croatian-Montenegrin-Serbian[75]

This section considers the development of BCMS,[76] whose standard variety is not part of the Balkan Sprachbund, lacking most of the characteristic grammatical features, many of which do not directly relate to case marking.[77] A comparison of changes to nominal inflection in BCMS, which involved some simplification but minimal case loss, allows for a better idea of which developments in Bulgarian and Macedonian can be attributed to the influence of the Balkan Sprachbund and which might have happened independently. Many of the sound changes that applied in these languages were similar, but several increased syncretism further in the development of Bulgarian and Macedonian than in the development of BCMS. These can partially account for the different outcomes, but the more significant differences were in the analogical processes that occurred, especially the functional mergers. In Bulgarian and Macedonian, these correspond to those in other Balkan Sprachbund languages, while BCMS has undergone a different functional merger.

Vernacular Bosnian, Croatian, Montenegrin and Serbian texts are first attested in the 13th century, the most important of which were legal texts. These were still heavily influenced by Church Slavonic, but more vernacular elements started appearing in the 14th century (see Svane 1958:17, 24). Based on Svane (1958), Belić (1962), and Leskien (1976), I have reconstructed the nominal inflection of Middle Serbian (MSrb) during the 14th and 15th centuries. Since it is the

[75] The grammatical information is drawn from Svane (1958), Belić (1962), and Leskien (1976) for MSrb, Browne (1993) for BCMS.

[76] This also includes Croatian since Croatia was predominantly Roman Catholic, so most of its medieval writings were in Latin. However, there are some medieval Croatian Church Slavonic writings in Glagolitic script. Bosnian, Croatian, Montenegrin, and Serbian are the standard varieties of four very closely related dialects of what used to be known as "Serbocroatian" (see Brown & Alt 2004:10, Browne 1993:307).

[77] See section 3.2.2 for a discussion of these grammatical features and their implications for case loss.

basis for the modern Serbian standard, an ekavian Neo-Štokavian dialect is assumed, i.e., one with /e/ as the reflex of **ě*, *što* for 'what', and accent retraction (see Browne 1993:307-308).[78] At this stage, the same general morphological distinctions in nominal inflection were made as in LPS, but the dual was in the process of merging with the plural, and case syncretism had increased significantly.

Table 38. Changes in Noun Declension from Late Proto-Slavic to Middle Serbian (reconstructed)

	< o-stems (inanimate) (also some i-stems, u-stems, C-stems)		< o-stems (animate) (also some i-stems, u-stems, C-stems)		< jo-stems (inanimate)	
	Masc.		Masc.		Masc.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-ъ > -∅	-i	-ъ > -∅	-i	-ь > -∅	-i
Voc	-e	-i	-e	-i	-u	-i
Acc	-ъ > -∅	-y » -e	-ъ » -a	-y » -e	-ь > -∅	-ę > -e
Gen	-a	-ъ > -∅, » -ā, -ī, -ov	-a	-ъ > -∅, » -ā, -ī, -ov	-a	-ь > -∅, » -ā, -ev
Dat	-u	-omъ > -om, » -oma, -i, -im, -ima	-u	-omъ > -om, » -oma, -i, -im, -ima	-u	-emъ > -em, » -ema, -i, -im, -ima
Loc	-ě > -e, » -u, -i	-ěxъ > -ex, » -ix	-ě > -e, » -u, -i	-ěxъ > -ex, » -ix	-i > -i, » -e, -u	-ixъ > -ix, » -ex
Instr	-omь > -om	-y > -i, » -mi, -imi, -im, -om, -oma, -ima	-omь > -om	-y > -i, » -mi, -imi, -im, -om, -oma, -ima	-emь > -em	-i > -i, » -mi, -imi, -im, -em, -ema, -ima

[78] The modern Bosnian and Croatian standards are also Neo-Štokavian, but they are (i)jekavian, with /ije/ as the reflex of **ě* in long syllables and /je/ in short syllables. Old Štokavian dialects have not undergone accent retraction; they are not an earlier stage of Neo-Štokavian. The names of the dialects are named based on the question word 'what'. Other dialects spoken primarily in Croatia include Čakavian, with *ča* for 'what' and /e/ and/or /i/ as the reflex of **ě*, and Kajkavian, with *kaj* for 'what' and /e/ as the usual reflex of **ě* (see Brown & Alt 2004:9*,* Browne 1993:307-309). In addition to phonological and lexical differences, these dialects have not undergone the same morphological changes as Neo-Štokavian; notably, they retain distinctive Dat, Loc, and Instr plural forms as in Slovene. Both dialects have more in common with Slovene than Štokavian, and Kajkavian is particularly close to Slovene (see Browne 1993:382-386, Priestly 1993:388, 399-406, Thomason 2009:62).

Table 38. Continued

	< jo-stems (animate)		< u-stems		< o-stems (also C-stems)	
	Masc.		Masc.		Neut.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-ь > -∅	-i	-ъ > -∅	-ove » -ovi	-o, -∅	-a
Voc	-u	-i	-u » -e	-ove » -ovi	-o, -∅	-a
Acc	-ь » -a	-ę > -e	-ъ > -∅, » -a	-y » -e	-o, -∅	-a
Gen	-a	-ь > -∅, » -ā, -ev	-u » -a	-ovъ > -ov	-a	-ъ > -∅, » -ā
Dat	-u	-emъ > -em, » -ema, -i, -im, -ima	-ovi » -u	-ъmъ » -om, -oma, -i, -im, -ima	-u	-omъ > -om, » -oma, -i, -im, -ima
Loc	-i > -i, » -e, -u	-ixъ > -ix, » -ex	-u	-ъxъ » -ex	-ě > -e, » -u, -i	-ěxъ > -ex, » -ix
Instr	-emь > -em	-i > -i, » -mi, -imi, -im, -em, -ema, -ima	-ъmь » -om	-ъmi > -mi, » -imi, -im, -om, -oma, -ima	-omь > -om	-y > -i, » -mi, -imi, -im, -om, -oma, -ima

	< jo-stems		< ā-stems (also some C-stems)		< jā-stems (also some C-stems)	
	Neut.		Fem.(/Masc.)		Fem.(/Masc)	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-e	-a	-a	-y » -e	-a	-ę > -e
Voc	-e	-a	-o	-y » -e	-e > -e, » -o	-ę > -e
Acc	-e	-a	-ǫ > -u	-y » -e	-ǫ > -u	-ę > -e
Gen	-a	-ь > -∅, » -ā	-y » -e	ъ > -∅, » -ā	-ę > -e	-ь > -∅, » -ā
Dat	-u	-emъ > -em, » -ema, -i, -im, -ima	-ě » -i	-amъ > -am, » -ama, -ami	-i	-amъ > -am, » -ama, -ami
Loc	-i > -i, » -e, -u	-ixъ > -ix, » -ex	-ě » -i	-axъ > -ax	-i	-axъ > -ax
Instr	-emь > -em	-i > -i, » -mi, -imi, -im, -em, -ema, -ima	-ojǫ > -u	-ami > -ami, » -ama, -am	-ejǫ > -u	-ami > -ami, » -ama, -am

	< i-stems (also some C-stems)	
	Fem.	
	Sg.	Pl.
Nom	-ь > -∅	-i
Voc	-i	-i
Acc	-ь > -∅	-i
Gen	-i	-ьi > -ī
Dat	-i	-ьmъ > -am, » -im, -imi, -ma, -ima
Loc	-i	-ьxъ > -ax, » -ix
Instr	-ьjǫ > -ju, -u	-ьmi > -mi, » -imi, -im, -ma, -ima

MSrb nominal inflection was affected by some of the same sound changes as the other South Slavic languages. As in other Slavic languages, the *jers* were regularly lost in weak positions except when this would result in certain consonant clusters (see Reinhart 2014:1297-1298, Browne 1993:309). As in PMB, this resulted in number syncretism between the Instr singular and Dat plural for all declension classes except *(j)ā*-stems and *i*-stems, e.g., *o*-stem Instr singular **gradomь*, Dat plural **gradomъ* > **gradom* 'city/cities'. Likewise, the Gen plural was reduced to a zero ending for all classes except *u*-stems and *i*-stems. As a result, the distinction between hard and soft stems was also lost in this form, e.g., *o*-stem **gradъ* > **grad*, *jo*-stem **mǫžь* > **muž* 'men', resulting in the use of only one *jer* letter, the front *jer* ь, after letters representing both etymologically hard and soft consonants. The same loss of distinction occurred in the other zero endings that resulted from the loss of final *jers*, i.e., in the Nom and Acc singular of masculine *(j)o*-stems and *u*-stems, except when the Acc was changed by analogy, as discussed below. Unlike in PMB, the *i*-stem Gen plural appears to have regularly developed into *-ī*, keeping it distinct from the seven to eight case/number forms with *-i*, e.g., **niti* vs. Gen plural **nitьi* > **nitī* 'thread(s)'. Thus, this is an example of a sound change that resulted in less syncretism in MSrb than in PMB. The merger of **i* and **y* as **i* occurred in MSrb, as in the rest of South Slavic (see Sussex & Cubberley 2006:43). This would have resulted in the same syncretism between Nom/Voc *-i* and Acc/Instr *-y* of the masculine *o*-stem plural paradigm as in PMB, e.g., Nom/Voc **gradi*, Acc/Instr **grady* > **gradi*.

As in PMB, the nasal vowels **ǫ* and **ę* were denasalized in the development of MSrb. However, their reflexes, [u] and [e], respectively, remained distinct from each other (see Reinhart 2014:1297-1298, Browne 1993:309). Thus, the merger of **ę* and **e* caused the Voc singular to join the syncretism between the Gen singular and Nom/Voc/Acc plural for *jā*-stems just as in PMB, e.g., Voc singular **zemlje*, Gen singular/Nom/Voc/Acc plural **zemlję* > **zemlje* 'land(s)', but *(j)ā*-stem Acc *-u* remained distinctive, so Nom/Acc singular syncretism did not increase. V_1jV_2 sequences were sometimes contracted in MSrb. However, the result of the contraction was a long vowel, so it remained distinct from a single V_2, e.g., the feminine Nom singular long form adjective **dobraja* > **dobrā* vs. the short form **dobra* 'good' (see Sussex & Cubberley 2006:135-136). In addition, no vowel raising in unaccented syllables occurred. Thus, syncretism increased due to sound changes in BCMS, but significantly less than in PMB.

By the 14th century, masculine *i*-stems, *u*-stems, consonant-stems, and *jo*-stems had merged with masculine *o*-stems to form one general class for almost all masculine nouns, although the merger was not complete (see Svane 1958:25, Browne 1993:320). These mergers did not simply consist of the extension of all *o*-stem endings to the other declension classes involved. As Browne explains, the *jo*-stem forms that began with [e] usually retained this vowel after palatals and certain other consonants, e.g., Instr singular **mužem* 'man', but some *jo*-stems eventually adopted *o*-stem forms with [o], e.g., **padežom* 'case'. Most *jo*-stems retained Voc singular *-u*, which also spread to some *o*-stems ending in velars. The *jo*-stem Acc plural form *-e* was retained rather than the *o*-stem form that had become syncretic with Nom/Voc *-i* due to sound change (1993:315, 319-320). The *u*-stem Gen plural form *-ov* had spread to some masculine *o*-stems, primarily monosyllables, but also disyllables, e.g., *zakonovь* 'laws' (MS 176/21, Nr. 159, 1367, cited by Svane 1958:72-73). It also spread to some masculine *jo*-stems as *-ev*, e.g., *kraljevь* 'kings' (MS 201/11, Nr. 190, 1382, cited by ibid.). The Instr plural form *-mi*, from *u*-stem *-ъmi* and *i*-stem *-ьmi*, was sometimes used on *(j)o*-stems as well, including neuters (see Belić 1962:64, Leskien 1976, §731). Each form that was retained from one of the smaller classes was more distinctive than the *o*-stem form, so the principle of distinctive strength accounts for the survival of these overstable markers.

Finally, the three Loc singular forms of the merged classes, i.e., *o*-stem *-e*, *u*-stem *-u,* and *jo*-stem/*i*-stem *-i*, appear to have been used somewhat interchangeably during this period; *-u* also spread to neuters even though there were no neuter *u*-stems (see Turbić-Hadžagić & Petrović 2011:78, Malić 1988:108). Each of these forms was involved in a different syncretism: *-e* with the Acc plural for masculine nouns, *-u* with the Dat singular for masculine and neuter nouns, and *-i* with the Nom/Voc plural for masculine nouns, as well as the inherited *(j)o*-stem Instr plural for masculine and neuter nouns. There may have been similar variation between the *o*-stem Loc plural form *-ex* and the *jo*-stem form *-ix*, but *-ex* was dominant until the end of the 15th century, and instances of *-ix* for original *jo*-stems were likely due to OCS influence (see Svane 1958:75, Johnson 1972:335).[79] Other than the forms mentioned above, the *o*-stem forms were adopted. For the Nom plural, it is likely that *-ov-* was kept, but followed by *-i* instead of *-e*, as in BCMS. In the combined masculine class, the Acc singular of animate nouns also adopted the Gen singular

[79] As mentioned above, my discussion assumes an ekavian dialect. In ikavian dialects, the reflex of the LPS *o*-stem Loc plural *-ěxъ* would have merged with the *jo*-stem form *-ix* anyway.

form. The motivation was very likely the same as in MB and other Slavic languages, favoring the Nom-Acc distinction over the Acc-Gen distinction for animate nouns.

The *ā*-stem and *jā*-stem declension classes had also merged in MSrb. To an even greater extent than with the *(j)o*-stems, the *jā*-stem forms replaced the *ā*-stem forms. The Voc singular was the only form that was not already shared by the two classes or derived from the *jā*-stems. The *ā*-stem Voc singular *-o* avoided number syncretism with the Nom/Voc/Acc plural, which may have motivated its retention and spread, e.g., **zemlje* » *zemljo* 'land'. However, some *jā*-stems retained Voc singular *-e* despite the number syncretism (see Browne 1993:321). Otherwise, the resulting *(j)ā*-stem paradigm had the same number syncretism as its antecedents, i.e., among the Gen singular and Nom/Voc/Acc plural. The tendency for the more distinctive form to survive when two classes merge, as opposed to simply the most frequent, appears to have been slightly stronger in MSrb than in MB, but it is not clear that this made a difference in the outcomes regarding case loss. In addition to this merger, some feminine consonant-stems came to be treated as *(j)ā*-stems, others as *i*-stems (see ibid., 322).

In addition to the inherited Gen plural forms, i.e., *-ov* from *u*-stems, *-ī* from *i*-stems (which also survives in BCMS, e.g., *kostī* 'bones'), and zero endings from *(j)o*-stems and *(j)ā*-stems, the new form *-ā* had begun to appear by the first half of the 14th century (Svane 1958:72-74). There are a number of competing theories about the origin of this new form, most of which involve some kind of analogy with *-ī*. One possibility is that the final *jers* of the *(j)o*-stems and *(j)ā*-stems were not lost as expected in a weak position, but were somehow treated as strong *jers*, perhaps because they had attracted the accent. Strong *jers* became *-a* in MSrb, and this could have been lengthened to *-ā* by analogy with *-ī* (see Vondrák 1904, Leskien 1976, §726, Mladenović 2008:136ff.). One piece of evidence cited in favor of this origin is that a Gen plural form written as <ьь> was common in Serbian Church Slavonic manuscripts, e.g., the *o*-stem <градьь> *grad(ā)* 'cities'. This appears to show the lengthening of a strong *jer* to *ī*, but Belić (1962:72, 83) has argued that <ьь> was a purely orthographic convention to distinguish the Gen plural from the *(j)o*-stem Nom and inanimate Acc singular, all of which had zero endings after the fall of the *jers*. Instead, he argues that *-ā* first arose in *(j)ā*-stems by analogy with variation in the Dat/Instr plural, for which the forms *-am*, *-ami*, and *-ama* could all be used in MSrb, as described in more detail below. If speakers interpreted this form as a base of *-am* plus an optional *-i* or *-a*, then by analogy they may have interpreted the Gen plural as a zero ending plus an

optional -*ī* (borrowed from *i*-stems) or -*ā* (ibid., 82-83). However, the Dat/Instr dual form -*ama* only started to be used interchangeably with the plural forms -*am* and -*ami* in 15th century manuscripts, so these forms were an unlikely source of analogy for a form that arose in the 14th century (see Svane 1958:79). It is more likely that -*ā* arose in *(j)ā*-stems because the feminine *i*-stem Gen, Dat, Loc, and Instr plural forms all contained *i* by this time and the *(j)ā*-stem Dat, Loc, and Instr plural forms all contained *a*, so -*ā* was supplied for the Gen plural to balance the paradigms, with a long vowel by analogy with -*ī* (see Johnson 1972:342, 349ff.). A source in the *(j)o*-stems is also possible; these had -*a* in the Gen singular, while feminine *i*-stems had -*i*, so -*ā* would be the analogous plural form to *i*-stem -*ī* (see Krispel 2013:65). However, it seems more likely that -*ā* spread to *(j)o*-stems from *(j)ā*-stems. Further evidence that *i*-stems contributed to the rise of the long Gen plural forms comes from a similar development in CSR: the Gen plural form -*ej* has spread from *i*-stems to many nouns in other declension classes which inherited a zero ending, especially *(j)ā*-stems, e.g., *jā*-stem *statej* 'articles'.

As mentioned above, Dat/Instr dual forms were being used with a plural meaning by the 15th century, indicating that the semantic distinction between the dual and plural was lost around this time. Since the same dual form could be used for Dat or Instr, the original plural forms came to be seen as interchangeable as well. For *(j)ā*-stems, the three forms involved, i.e., Dat plural -*am*, Instr plural -*ami*, and Dat/Instr dual -*ama* all contained *a*, so this development was relatively straightforward. For the masculine and neuter classes, the situation was complicated by vowel differences between the forms. The inherited *o*-stem forms were Dat plural -*om*, Instr plural -*i*, and Dat/Instr dual -*oma*. The *jo*-stem forms were analogous but with *e* instead of *o*. As mentioned above, the Instr plural form -*mi* from *u*-stems and *i*-stems was also used. Additional forms arose through the mixing of these forms: -*im* from -*i* and -*om*/-*em*, -*imi* from -*i* and -*mi*, and, by the end of the 15th century, -*ima* from -*i* and -*oma*/-*ema*. Feminine *i*-stems had the inherited Dat plural -*am*, Instr plural -*mi*, and Dat/Instr dual -*ma*, but forms with *i* were used as well, i.e., -*im*, -*imi*, and -*ima* (see Belić 1962:644ff., Matasović 2008, §230). As -*ima* became more common, -*oma*/-*ema* fell into disuse (see Leskien 1976, §731). After the MSrb period, -*ima* became the dominant form, as described below. A parallel development involving the spread of dual forms to the plural during the loss of the dual as a productive number category was the adoption of irregular plural forms of some paired body parts in BCMS and other Slavic languages from the dual, e.g., the neuter *o*-stem *okči* 'eyes' with -*i* and palatalization, instead of -

a (see Sussex & Cubberley 2006:225). These dual forms likely survived because they were used more frequently than the plural forms (see Meiser 1992:207-208). However, the Dat/Instr dual form was almost certainly used less frequently than the Dat and Instr plural forms for the vast majority of nouns, so other factors must have been involved, as discussed below.

Agreement targets in MSrb maintained most of the same gender distinctions as in LPS and OCS. Long and short adjective forms both remained in use, but long Gen, Dat, and Loc forms had begun to spread to the short paradigms (see Reinhart 2014:1299). In the Loc singular of short-form adjectives, this restored the distinction between masculine/neuter (*-om*) and feminine (*-oj*); previously, the hard-stem form *-ě* and soft-stem form *-i* were used for all genders. Otherwise, adjectives, along with determiners, developed more or less as expected based on the changes to nouns, with contraction of the long forms to long vowels, as in the example above. In contrast to MB, the distinction between masculine and feminine was retained in the Nom/Voc plural due to the adoption of soft-stem forms, as in nouns.

The standard literary BCMS was developed from Neo-Štokavian dialects in the 19th century. BCMS distinguishes six cases morphologically: Nom, Voc, Acc, Gen, Dat, Instr; three genders; and two numbers, singular and plural. The dual has been lost, and, unlike the situation in other Slavic languages, the Loc has almost entirely merged with the Dat. The location function is now expressed by prepositions with Dat objects. However, certain dialects maintain distinctive Loc forms (see Tomić 2006:108-110, Browne 1993:306-308, 318-323).

Table 39. Changes in Noun Declension from Middle Serbian to BCMS

	< o-stems (inanimate) (also some u-stems)		< o-stems (animate) (also some u-stems)		< jo-stems (inanimate)	
	Masc.		Masc.		Masc.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-Ø > -Ø, » -o	-i; -ovi	-Ø > -Ø, » -o	-i; -ovi	-Ø > -Ø, » -e	-i > -i, » -evi
Voc	-e	-i; -ovi	-e	-i; -ovi	-u	-i > -i, » -evi
Acc	-Ø > -Ø, » -o	-e > -e, » -ove	-a	-e > -e, » -ove	-Ø > -Ø, » -e	-e > -e, » -eve
Gen	-a	-ā, -ī; -Ø » -ā, -ī; -ov » -ovā	-a	-ā, -ī; -Ø » -ā, -ī; -ov » -ovā	-a	-ā; -Ø » -ā; -ev » -evā
Dat	-u	-ima; -om, -oma, -i, -im » -ima, -ovima	-u	-ima; -om, -oma, -i, -im » -ima, -ovima	-u	-ima; -em, -ema, -i, -im » -ima, -evima
Loc	-u; -e, -i » -u	-ex, -ix » -ima, -ovima	-u; -e, -i » -u	-ex, -ix » -ima, -ovima	u; -i, -e » -u	-ix, -ex » -ima, -evima
Instr	-om	-ima; -i, -mi, imi, -im, -om, -oma » -ima, -ovima	-om	-ima; -i, -mi, -imi, -im, -om, -oma » -ima, -ovima	-em	-ima; -i, -mi, -imi, -im, -em, -ema » -ima, -evima

	< jo-stems (animate)		< o-stems		< jo-stems	
	Masc.		Neut.		Neut.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-Ø > -Ø, » -e	-i > -i, » -ev	-o, -Ø	-a	-e	-a
Voc	-u	-i > -i, » -evi	-o, -Ø	-a	-e	-a
Acc	-a	-e > -e, » -eve	-o, -Ø	-a	-e	-a
Gen	-a	-ā; -Ø » -ā; -ev » -evā	-a	-ā; -Ø » -ā	-a	-ā; -Ø » -ā
Dat	-u	-ima; -em, -ema, -i, -im » -ima, -evima	-u	-ima; -om, -oma, -i, -im » -ima	-u	-ima; -em, -ema, -i, -im » -ima
Loc	-u; -i, -e » -u	-ix, -ex » -ima, -evima	-u; -i, -e » -u	-ex, -ix » -ima	-u; -i, -e » -u	-ix, -ex » -ima
Instr	-em	-ima; -i, -mi, -imi, -im, -em, -ema » -ima, -evima	-om	-ima; -i, -mi, -imi, -im, -om, -oma » -ima	-em	-ima; -i, -mi, -imi, -im, -em, -ema » -ima

	< ā-/jā-stems		< i-stems	
	Fem.(/Masc.)		Fem.	
	Sg.	Pl.	Sg.	Pl.
Nom	-a	-e	-Ø	-i
Voc	-o, -e	-e	-i	-i
Acc	-u	-e	-Ø	-i
Gen	-e » -ē	-ā; -Ø » -ā, -ī, -ū	-i	-ī > -ī, » -ijū
Dat	-i	-ama; -am, -ami » -ama	-i	-ima; -am, -im, -imi, -ma » -ima
Loc	-i	-ax » -ama	-i > -i, » -ju, -u	-ax, -ix » -ima
Instr	-u » -ōm	-ama; -ami, -am » -ama	-ju > -ju, » -i; -u > -u, » -i	-ima; -mi, -imi, -im, -ma » -ima

A high degree of variation occurred in MSrb nominal inflection. The variants that have been retained in BCMS are generally among those that involve the least number syncretism, as was the case in the development of MB. Some case distinctions were lost as a result, but far fewer than in MB. In the Gen plural, only the long vowel forms survived; these avoided all syncretism by virtue of their length, in contrast to the zero ending Gen plural forms, which were syncretic with the Nom singular for most masculine nouns, as well as the Acc singular for inanimate masculine nouns. For all declension classes except feminine *i*-stems, which already had *-ī*, *-ā* became the most common Gen plural form, e.g., *ā*-stem *ribā* 'fish'. The *-ī* form has also been extended from *i*-stems to some *(j)ā*-stems and some masculine nouns that were not originally *i*-stems. At the same time, feminine *i*-stems sometimes adopted *-ijū* for the Gen plural, the lengthened form of the earlier Gen/Loc dual *-iju*, e.g., *kostijū* alongside *kostī* 'bones'. This form may have also influenced the use of *-ū* as a third alternative for *(j)ā*-stems, the other feminine class; *-u* was Gen/Loc dual for *(j)o*-stems and *(j)ā*-stems but was only lengthened and adopted in the plural for *(j)ā*-stems.

The *u*-stem *-ov* form did not survive as a Gen plural form on its own, as it did in other Slavic languages such as CSR. Instead, the *u*-stem pattern with *-ov-* in the Nom/Voc and Gen plural spread to all plural forms of most monosyllabic, and some disyllabic, masculine *o*-stems (and as *-ev-* to original *jo*-stems) but with the same endings as *(j)*o-stems after *-ov-*/*-ev-* instead of the inherited *u*-stem endings, e.g., Nom/Voc **gradi* » *gradovi* 'cities', not **gradove* (see Browne 1993:315, 319-320). In the Gen plural, *-ā* was added to *-ov*, resulting in *-ovā*; this reflects a strong paradigmatic pressure for the Gen plural to have a long vowel, with *-ā* as the default. The spread of *-ov-*, both within the plural paradigm and across declension classes, is a prime example of number profiling using an overstable marker.

With the lengthening of the Gen singular to *-ē*, e.g., **zemlje* » *zemljē* 'land', likely by analogy with the Gen plural forms, all number syncretism was eliminated for *ā*-stems. In the development of Bulgarian and Macedonian, more drastic changes avoided this syncretism, i.e., the replacement of the Gen-Dat with analytical *na*-constructions. The long vowel for the *(j)ā*-stem Instr singular is possibly the result of contraction, but the *-m* must have been added by analogy with *o*-stems. This cannot have been motivated by syncretism because *-ō* is already a distinctive form; a more likely motivation is consistency across declension classes, which falls under the principle of system adequacy.

In the Loc singular of the masculine and neuter classes, only *-u* survived. Originally from *u*-stems, this form was syncretic with the Dat singular, but avoided the number syncretism associated with the other forms, *-e* and *-i* (see Browne 1993:320). In addition to number profiling, another paradigmatic motivation is parallelism with the feminine classes, which already had Dat/Loc syncretism (see Janda 1996:331).

In the Dat/Instr plural, two forms inherited from the Dat/Instr dual, *-ama* and *-ima*, were retained for *(j)ā*-stems and all other declension classes, respectively. In contrast to some of their competitors, these forms avoided all number syncretism. Their adoption resolved the syncretism between the Instr singular and the original Dat plural forms for the masculine and neuter classes, e.g., *o*-stem Instr singular *prozorom,* Dat plural **prozorom* » *prozorima* 'window(s)'. This contrasted with the leveling of the Instr to Acc singular in MB. The use of these forms also restored the distinction between the Nom/Voc and Instr plural for masculine *(j)o*-stems, which may have been more important than the distinction between Dat and Instr plural. Another possible motivation for the retention of only the dual forms is that they remained more phonetically salient than the inherited Dat plural forms, which were reduced to one syllable by the loss of final *jers*. The Dat/Instr plural is highly marked, so the most salient form available would be preferred by the iconic principle. This also supports MacWhinney's (1985:205) finding that a strong cue renders direct mapping and thus facilitates learning. The use of the Instr plural form in place of the Dat would have accomplished the same thing for *(j)ā*-stems, e.g., **zemljam* » *zemljami*, but the dual form had the added benefit of consistency across declension classes. In addition, the preference for the dual form may come back to the fact that it was already used for both the Dat and Instr when it adopted a plural meaning in MSrb, whereas the competing Dat and Instr forms may have still had stronger associations with one case or the other. This development is an example of the redistribution of forms associated with different numbers according to case as two number categories merged, since plural forms survive elsewhere in the plural. A much more common development is the retention of different case forms according to number when a case merger occurs, as with the merger of the Dat and Instr in AG described in the following subsection (see Meiser 1992:203-204).

Over the course of the 17th century, the inherited Loc plural forms were gradually replaced by the Dat/Instr plural forms in use at the time. These included inherited Dat and Instr plural forms as well as the syncretic dual form (see Leskien 1976, §732). Thus, the dual forms

had not yet won out at this stage, but they probably did not long after the Loc joined this syncretism and lost its status as a distinctive case. A potential phonological motivation for the merger of the Loc plural with the Dat/Instr is the loss of *-x* in many central Štokavian dialects (see Brozović & Ivić 1988:25, Popović 1960:428f.). This process likely occurred from the end of the 16th century to the mid 18th century, so the timing is plausible (see Belić 1960:106). Specifically, the loss of *-x* would have made the *jo*-stem Loc plural *-i* syncretic with its original Instr plural form, e.g., Instr **konji*, Loc **konjix* > *konji* 'horses'. One issue with this argument is that *-x* was retained in the pronominal Gen plural forms, which were syncretic with the Loc before it joined the Dat/Instr syncretism instead, so either its loss was not regular in the dialects from which standard BCMS developed, or it was restored in these forms based on dialects that had not lost *-x*. A potential morphological motivation is the existing Dat/Loc syncretism in the singular of all classes (see Leskien 1976, §440). For *(j)ā*-stems and *i*-stems, this already existed in LPS, and it is more likely that the masculine and neuter classes developed this syncretism first in the singular because of markedness: the collapse of a distinction in the marked plural is often tolerated without any effect on the singular, as with the Dat and Instr in BCMS, but the collapse of a distinction in the unmarked singular often leads to the total loss of that distinction. Given the saliency of plural markers in the Dat/Instr plural, the loss of the distinction in the singular does not make the number category syncretic.

The merger of the Dat and Loc in BCMS contrasts with the Gen-Dat and Acc-Loc mergers in MB, both characteristic of the Balkan Sprachbund languages. Both PMB and BCMS inherited Dat/Loc syncretism in the singular *(j)ā*-stem and *i*-stem paradigms. Perhaps due to the lack of vowel raising, the Loc singular was never leveled to the Acc in BCMS, or if a functional merger preceded the formal merger of these cases in MB, this functional merger simply did not occur in MSrb, perhaps due to a lack of the external factors associated with the Balkan Sprachbund. Without any motivation to level the Loc to the Acc, it was instead leveled to the case with which it already shared some syncretism. However, other Slavic languages inherited the same Dat/Loc syncretism for *(j)ā*-stems and *i*-stems, as well as *u*-stem Loc *-u*, but did not undergo a full merger of the Dat and Loc, so the additional analogical processes mentioned above and/or semantic overlap must have been a factor.[80] In addition, feminine *i*-stems have

[80] The semantic overlap of the Dat and Loc that may have enabled their functional merger in Hittite, AG, and potentially BCMS is discussed in section 3.1.2.5.

developed Loc/Instr singular syncretism. One or both inherited forms can be used for both cases, depending on the noun. If the Loc form is used, it is syncretic with the Dat as well but also the Voc and Gen singular and the Nom/Voc/Acc plural. Thus, the spread of the Instr singular form to the Loc decreases number syncretism.

Gender syncretism on agreement targets has increased somewhat in BCMS compared to MSrb. The long Gen,[81] Dat, and Instr adjective forms have been adopted in the short paradigms. As a result, gender distinctions on plural agreement targets only remain in the Nom/Voc, which have distinctive forms for all three genders, and the Acc, which has a distinctive neuter form but masculine/feminine syncretism (see Browne 1993:323-329). These core case forms appear to be the most important for maintaining gender distinctions, however, so the increased syncretism is still unlikely to result in the total loss of gender agreement in plural, as has happened in CSB, CSM, and other Slavic languages such as CSR.

Table 40. Case Mergers in Serbian

	Old Church Slavonic / Middle Serbian							BCMS					
PIE	Nom	Voc	Acc	Gen	Dat	Loc	Instr	Nom	Voc	Acc	Gen	Dat	Instr
Nom	x							x					
Voc		x							x				
Acc			x							x			
Gen				x							x		
Dat					x							x	
Loc						x						x	
Instr							x						x
Abl				x							x		

So far, changes in the nominal inflection of the major Romance and Slavic languages in the Balkan Sprachbund have been discussed, as well as members of both families outside of the Balkan Sprachbund, i.e., Western Romance languages in section 2.4.1 above and BCMS in this section. Potential explanations of the differences in case loss among the members of these two families inside and outside the Balkan Sprachbund are considered in chapter III. The next two subsections discuss Greek and Albanian, respectively, two more Balkan Sprachbund languages—each comprising its own branch of IE—that have undergone similar case mergers as

[81] The masculine animate Acc singular adopted the long form along with the Gen singular, with which it is syncretic.

the Balkan Sprachbund languages discussed above, but without the total case loss of Bulgarian and Macedonian.

2.5.3. Greek[82]

Mycenaean Greek (MyG) was written from 1400 to 1200 BCE, making it the oldest attested form of Greek and therefore the closest approximation to Proto-Greek. The MyG forms in Table 41 are interpretations of the written MyG evidence. Linear B, the syllabic writing system used for MyG from 1400 to 1200 BC, failed to capture many of the phonetic distinctions between case forms, but authors have filled in the details based on later stages of Greek and the comparative method (see, for example, Thompson 2010:189, 193, Trąba 2018:115). No Voc forms were attested due to the nature of the surviving texts (see Thompson 2010:193). However, these assumed to be the same as Ancient Greek (AG) when the AG forms directly continue the reconstructed PIE forms, as in all plural Voc forms, which are identical to the Nom, and the singular forms for *o*-stems, masculine *ā*-stems, and some consonant and other stems (see Sihler 1995:256, 268, 283-284).[83]

As in Slavic, the PIE Abl had already merged with the Gen in MyG. In MyG, the Gen forms were retained; in the plural they took over the functions of the Abl from the Dat/Abl form (see Sihler 1995:251). Some authors such as Luján (2014) posit a distinctive Loc in MyG. According to Sihler, both the Loc and Instr were at least partially distinctive in MyG but underwent a functional merger with the Dat by AG, with a mixture of case forms surviving (1995:251). However, the more common opinion is that the Loc had already lost productivity by this stage (see Trąba 2018:115). On the other hand, fossilized Loc forms for a limited set of place names and some other nouns survived at least into AG, e.g., *oíkoi* ‘at home’ in contrast with Dat *oíkōi* (see Fykias & Katsikadeli 2012:255). Some distinctive Instr forms were certainly still used in MyG, but it is possible that these were used with the same functions as the Dat (see Trąba 2018:115). This would mean that it had already lost productivity, like the Loc (see Adrados 2005). Following Trąba (2018:115-116), the present investigation treats Instr but not Loc as a

[82] Grammatical information is drawn from Sihler (1995), Horrocks (2010), Adrados (2005), Thompson (2010), Mussies (1971).

[83] In AG, feminine *ā*-stems and some consonant stems have adopted the Nom singular form in place of the inherited Voc form (see Sihler 1995:268, 283-284).

distinctive case in MyG. The Dat forms given in Table 41 are a mixture of original Dat and Loc forms. The plural forms all derive from the Loc, while the singular forms derive from the Dat or a combination of Dat and Loc. Both *-ei*, originally Dat, and *-i*, originally Loc, are found for singular consonant-stems (see Sihler 1995:251, 253). In addition, the expected Loc singular forms for *o*-stems (*-oi*) and *ā*-stems (*-ai*) differed from the Dat only in vowel length, so they would not have been distinguished in written MyG, but most scholars assume that only Dat forms are represented in the written record (see Thompson 2010:193).

AG was used from the 8th to 3rd centuries BCE. It can refer to the whole set of Greek dialects used at the time, but my study focuses on the Attic dialect, the main basis for the formation of the Koine (see Trąba 2018:117-118). It distinguished five cases morphologically: Nom, Voc, Acc, Gen, Dat; the three IE genders; and three numbers: singular, dual, and plural. The dual made the same case distinctions in all noun paradigms: Nom/Voc/Acc and Gen/Dat (see Weiss 2010:105-106).

Table 41. Changes in Noun Declension from Mycenaean Greek to Ancient Greek

	< o-stems		< o-stems		< ā-stems	
	Masc.		Neut.		Fem.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-os	-oi	-on	-a	-ā > -ā, -ē	-ai
Voc	-e	-oi	-on	-a	-a » -ā, -ē	-ai
Acc	-on	-ons > -ous	-on	-a	-ān > -ān, -ēn	-āns > -ās
Gen	-ojo > -ou	-ōn	-ojo > -ou	-ōn	-ās > -ās, -ēs	-āōn > -ôn
Dat	-ōi	-oihi » -ois	-ōi	-oihi » -ois	-āi > -āi, -ēi	-āhi » -ais
Instr	-ō » -ōi	-ois	-ō » -ōi	-ois	-ā » -āi, -ēi	-ap^{h}i » -ais

	< ā-stems		< C-/i-/u-stems		< C-/u-stems	
	Masc.		Masc./Fem.		Neut.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-ās > -ās, -ēs	-ai	-Ø, -s	-es	-Ø	-a > -a, -ē
Voc	-a	-ai	-Ø, -s	-es	-Ø	-a > -a, -ē
Acc	-ān > -ān, -ēn	-āns > -ās	-a, -n	-as	-Ø	-a > -a, -ē
Gen	-āo » -ou	-āōn > -ôn	-os	-ōn	-os	-ōn
Dat	-āi > -āi, -ēi	-āhi » -ais	-ei, -i	-si > -si, » -sin	-ei, -i	-si > -si, » -sin
Instr	-ā » -āi, -ēi	-ap^{h}i » -ais	-ē » -ei, -i	-p^{h}i » -si, -sin	-ē » -ei, -i	-p^{h}i » -si, -sin

A number of sound changes applied to case markers between MyG and AG. Most of these sound changes did not increase syncretism, but the simplification of final *-ns* to *-s* with compensatory

lengthening made the *ā*-stem Acc plural form syncretic with either the Gen singular (for feminine nouns, e.g., Gen singular **khó:rās*, Acc plural **khó:rāns* > *khó:rās* 'country/countries') or the Nom singular (for masculine nouns, e.g., Nom singular **ktitās*, Acc plural **ktítāns* > *ktítās* 'inhabitant(s)'). In addition, a series of sound changes may have resulted in syncretism between the Instr singular and Nom/Voc/Acc plural for neuter *s*-stems and *u*-stems. Silher posits that the combination of stem and ending had the form **-es-a* for the Nom/Voc/Acc plural of *s*-stems in Proto-Greek and **-w-a* for *u*-stems. Due to the loss of initial and intervocalic [s] (see Thompson 2010:191), the MyG *s*-stem form would have been **-eha*, e.g., **géneha* 'races'. With the addition of [e] by analogy with masculine/feminine *u*-stems, which had the form **-ewes*, in all Attic-Ionic dialects, the *u*-stem form would have been **-ewa*, e.g., **ástewa* 'towns'. Due to the subsequent loss of [w] and intervocalic [h], the Homeric form for both stems was *-ea*, e.g., *génea* and *ástea*. In Attic, this was contracted to *-ē*, e.g., *génē* and *ástē* (1995:307, 326). With the same loss of these stem consonants, the Instr singular form of these nouns would have been syncretic with the Nom/Voc/Acc plural, but this form most likely already lost productivity by the time these sound changes were complete.

If the merger of the Dat and Instr was not complete by this stage, it could have been motivated by the avoidance of number syncretism in this one form, but this is unlikely. No other number syncretism occurred in MyG. This was more likely a functional merger (see Sihler 1995:264), as in the early merger of the Instr with the Dat in the Germanic languages, with the Abl in the development of Classical Latin (CL), and with the Acc in MB. In AG, the Dat form survived in most paradigms, but the Instr was retained instead in the plural of *o*-stems, and the *ā*-stem plural form developed by analogy with the *o*-stem Instr form. Other analogical changes occurred in the development of AG, but they did not result in case or gender loss. Changes to masculine *ā*-stems increased parallelism with masculine *o*-stems. Silher observes that both feminine and masculine *ā*-stems had already converged with masculine *o*-stems, with most forms differing only in the theme vowel. The Nom and Gen singular were the only *ā*-stem forms that did not have a close parallel in masculine *o*-stems. Thus, the masculine *ā*-stem Nom singular acquired an *-s*, e.g., **ktítā* » *ktítās* (> *ktítēs* in Attic-Ionic) 'inhabitant', possibly already in MyG, but this is unclear due to the writing system. The masculine *ā*-stem Gen singular form *-ao* had already been influenced by the *o*-stem form *-ojo* in MyG. The expected reflexes of these forms in AG were *-ō* (pronounced [ɔ:] at this stage) and *-ou* (pronounced [o:]), respectively. Thus, these

forms had only a small difference in tongue height, which was neutralized in favor of the *o*-stem form, e.g., **ktítao* > **ktítō* » *ktítou.*The leveling of the feminine *ā*-stem Voc singular to the Nom further differentiated this class from masculine *ā*-stems, which generally retained a distinctive Voc singular form (1995:268, 274-275). This last change was probably due to the influence of the distinctive masculine *o*-stem Voc singular. The addition of *-n* to Dat plural *-si* was mostly limited to Attic-Ionic and was probably by analogy with the Dat form of the first- and second-person plural pronouns. It later spread to various verb forms with similar endings. Over time this *-n* came to be used primarily before vowels and pauses, but this was never completely consistent (ibid., 232-233).

By the time the functional mergers discussed above were complete in AG, the cases had the following functions. The Nom and Voc retained the same functions. The Acc continued to mark DOs and motion toward (goal), including on the objects of prepositions with this meaning. It was also used in adverbial expressions indicating duration of time and extent of space. The Abl had merged with the Gen, so it inherited the Abl functions of motion from (source), separation, and standard of comparison, in addition to marking the complement of a noun and expressing partitivity. As the complement of verbs and adjectives, it often served this function, at least in a broad sense. The Loc and Instr had merged with the Dat, so in addition to marking IOs and loosely connected functions such as ethical Dat and inalienable possession, the Dat inherited the location function (usually with a preposition) from the Loc and the instrument function from the Instr, as well as other related functions such as manner and accompaniment. All surviving cases except the Voc marked the complements of certain verbs (see Morwood 2002:10-22, Trąba 2018:117-120).

Agreement targets in the different stages of Greek tended to decline like the major nominal classes, with some pronominal forms on determiners. They are not well-attested in MyG, so the discussion here focuses on AG. Many adjectives used *ā*-stem forms for feminine and *o*-stem or *u*-stem forms for masculine and neuter. For these classes, the feminine forms were usually distinctive except in the Gen plural, while masculine and neuter forms were distinctive in the Nom/Voc/Acc of both numbers, the Acc plural, and sometimes the Acc singular. Though rarer, compound adjectives used the same *o*-stem forms for feminine as well as masculine, and some other adjectives used consonant-stem forms, for which the masculine and feminine forms were also the same, while the neuter forms differed in the Nom/Voc/Acc of both numbers (Sihler

1995:348-355). Determiners mostly used *ā*-stem and *o*-stem forms as well, but the neuter was always distinctive in the Nom/Voc/Acc of both numbers. The distinction between masculine and neuter was slightly weaker than other gender distinctions, but less so than in CL.

Koine Greek (KG) was used from the 4th century BCE to around the 4th or 5th centuries CE. Generally, like AG, KG distinguished five cases morphologically: Nom, Voc, Acc, Gen, Dat; the three IE genders; and two numbers: singular and plural, having lost the dual in its formation (see Mussies 1971:101-105).[84] One of the major sources of written evidence for KG, particularly in a more vernacular form, is Biblical Greek, i.e., the New Testament, as well as the earlier Septuagint translation of the Hebrew Old Testament. Other significant sources for vernacular KG include documents written on papyrus such as private letters (see Horrocks 2010:106, 114-115, 147, Stolk 2015:92).

Table 42. Changes in Noun Declension from Ancient Greek to Koine Greek

	< o-stems		< o-stems		< ā-stems	
	Masc.		Neut.		Fem.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-os	-oi > -oi, -ü	-on	-a	-ā > -a; -ē > -e	-ai > -ε
Voc	-e > -ε	-oi > -oi, -ü	-on	-a	-ā > -a; -ē > -e	-ai > -ε
Acc	-on	-ous > -us	-on	-a	-ān > -an; -ēn > -en	-ās > -as
Gen	-ou > -u	-ōn > -on	-ou > -u	-ōn > -on	-ās > -as; -ēs > -es	-ôn > -ón
Dat	-ōi > -o	-ois > -ois, -üs	-ōi > -o	-ois > -ois, -üs	-āi > -a; -ēi > -e	-ais > -εs

	< ā-stems		< C-stems		< C-stems	
	Masc.		Masc./Fem.		Neut.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-ās > -as; -ēs > -es	-ai > -ε	-∅, -s	-es > -εs	-∅	-a
Voc	-a	-ai > -ε	-∅, -s	-es > -εs	-∅	-a
Acc	-ān > -an; -ēn > -en	-ās > -as	-a > -a, » -an	-as	-∅	-a
Gen	-ou > -u	-ôn > -ón	-os	-ōn > -on	-os	-ōn > -on
Dat	-āi > -a; -ēi > -e	-ais > -εs	-i	-si, -sin	-i	-si, -sin

[84] When the koineization process occurred, the Attic dialect still had the dual, but the Ionic dialect had already lost this distinction, so the dual was not retained in KG.

Table 42. Continued

	< i-stems	
	Fem.	
	Sg.	Pl.
Nom	-is	-eis > -is
Voc	-i	-eis > -is
Acc	-in	-eis > -is
Gen	-eōs > -eos	-eōn > -eon
Dat	-ei > -i	-si, -sin

No case loss occurred in the development of KG from AG, but sound changes increased syncretism. The loss of *-i* from long diphthongs (see, for example, Horrocks 2010:116) resulted in syncretism between Nom/Voc and Dat singular for feminine *ā*-stems, e.g., Nom/Voc *khó:rā*, Dat *khó:rāi* > *khó:rā* 'country'. The later neutralization of vowel length (see, for example, Adrados 2005:193) caused the Gen plural to become syncretic with the Acc singular for *o*-stems, e.g., Acc singular *lū́kon*, Gen plural *lū́kōn* > *lū́kon* 'wolf/wolves'.[85] In combination, these changes also made the Dat singular syncretic with the Voc for some masculine *ā*-stems, e.g., Voc *neanía*, Dat *neaníāi* > *nɛanía* 'young man'. By the Hellenistic period (323–31 BCE), several diphthongs had undergone sound changes as a result of monophthongization: [ai] and [e] > [ɛ], [oi] > [ü], and [ei] > [i] (ibid.). This last change resulted in syncretism between Nom singular and Nom/Voc/Acc plural for *i*-stems, e.g., Nom singular *pólis*, Nom/Voc/Acc plural *póleis* > *pólis* 'city/cities'.

A large part of the simplification in KG involved the regularization and/or loss of productivity of minor declension classes, including *u*-stems of all genders and masculine *i*-stems. One analogical change that began to occur later in the Koine period was the spread of Acc *-an* to consonant-stems by analogy with *ā*-stems, e.g., *phū́laka* » *phū́lakan* 'guard' (see Horrocks 2010:28, Mussies 1971:104,). This was the first of several changes that led to the eventual merger of these two stem classes, as described below. Agreement targets continued to make the same distinctions as in AG.

Byzantine Greek was used from the 5th century to the fall of the Byzantine Empire in 1453, when the modern period is considered to have begun for Greek. Byzantine Greek was a transitional period for nominal inflection. For the most part, KG continued to be used in writing,

[85] For neuter *o*-stems, the Nom/Voc/Acc were already syncretic.

so the changes in the spoken language from KG to Modern Greek (ModGr) are only sporadically attested (see Trąba 2018:120-125). For this reason, Byzantine Greek is not considered as a separate stage in my analysis, but developments that must have occurred during this period are discussed. By ModGr, a few of the distinctions found in KG had been lost, including the Dat. It distinguishes four cases morphologically: Nom, Voc, Acc, Gen; three genders; and two numbers: singular and plural. The Voc is usually syncretic with the Nom, except for masculine nouns. Two forms of the language competed for acceptance as the standard after Greece gained independence from the Ottoman Empire in 1829: Katharevousa, which was highly archaicizing, and Demotic, i.e., the vernacular. Demotic has prevailed but with significant Katharevousa influence in the standard form (see Adrados 2005:291-297).

Table 43. Changes in Noun Declension from Koine Greek to Modern Greek

	< o-stems		< o-stems		< ā-stems	
	Masc.		Neut.		Fem.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-os	-oi, -ü > -i	-on > -o	-a	-a; -e > -i	-ε » -es
Voc	-ε > -e	-oi, -ü > -i	-on > -o	-a	-a; -e > -i	-ε » -es
Acc	-on > -o	-us	-on > -o	-a	-an > -a; -en > -i	-as » -es
Gen	-u	-on*	-u	-on*	-as; -es > -is	-ón*
Dat	-o	-ois, -üs » -us	-o	-ois, -üs » -a	-a; -e > -i	-εs > -es

	< ā-stems		< C-stems		< C-stems	
	Masc.		Masc.		Fem.	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
Nom	-as; -es > -is	-ε » -es	-s, -∅ » -as	-εs > -es	-s, -∅ » -a	-εs > -es
Voc	-a > -a, » -i	-ε » -es	-s, -∅ » -a	-εs > -es	-s, -∅ » -a	-εs > -es
Acc	-an > -a; -en > -i	-as » -es	-a, -an > -a	-as » -es	-a, -an > -a	-as » -es
Gen	-u » -a, -i	-ón*	-os » -a	-on*	-os » -as	-on*
Dat	-a; -e > -i	-εs > -es	-i » -a	-si, -sin » -es	-i » -a	-si, -sin » -es

Table 43. Continued

	< C-stems		< i-stems	
	Neut.		Fem.	
	Sg.	Pl.	Sg.	Pl.
Nom	-Ø	-a	-is » -i	-is
Voc	-Ø	-a	-i	-is
Acc	-Ø	-a	-in > -i	-is
Gen	-os > -os, » -u	-on*	-eos > -eos, » -is	-eon*
Dat	-i » -Ø	-si, -sin » -a	-i	-si, -sin » -is

Syncretism increased significantly as a result of sound changes during the Byzantine period. Final *-n* was already weak by the end of the Koine period (Horrocks 2010:116).[86] Over the course of the early to mid Byzantine period, final *-n* was regularly lost, with particularly significant effects on the singular paradigms of vowel stems (Adrados 2005:238, 242). For *o*-stems, this resulted in syncretism between Dat and Acc singular, e.g., Acc *lű́kon*, Dat *lű́ko* > *líko* 'wolf'.[87] For *ā*-stems and *i*-stems, the Acc joined the Voc/Dat singular syncretism, e.g., Voc/Dat *khóra*, Acc *khóran* > *khóra* 'country;' Voc/Dat *póli*, Acc *pólin* > *póli* 'city'.[88] The regular raising of [e] to [i] (while [ɛ] raised to [e]) made the Voc, Acc, and Dat forms in one of the two feminine *ā*-stem singular paradigms the same as the *i*-stem forms, e.g., *time* > *timi* 'honor'. These overlapping forms led to the merger of these two declension classes in the singular, as described below. The loss of final *-n* and vowel changes, including the earlier neutralization of vowel length in KG, are similar to sound changes that preceded major case loss in the Germanic and Romance languages. These changes are likely the result of a shift to a rhythmic dynamic stress in all three families, as discussed in section 3.1.1.1 below.

Analogical processes also had a significant effect on the development of Modern Greek nominal inflection. Several of these processes decreased number syncretism. One was the retention of *-n* in the Gen plural forms of all declension classes, an instance of morphological blocking that eliminated syncretism with the Acc singular in the *o*-stem paradigms, as well as the Nom/Voc for neuter *o*-stems.

[86] This may explain why Acc singular *-an* started being used for consonant stems as well as *ā*-stems: if *-an* was pronounced as *-a* at least some of the time, then scribes may have written *-an* when they pronounced an Acc as *-a* regardless of the historical spelling.

[87] For neuter *o*-stems, the Nom/Voc/Acc were already syncretic.

[88] For feminine *ā*-stems, the Nom was also part of this syncretism.

This probably happened by analogy with determiners and clitic pronouns, which retain *-n* before vowels and plosives but also in other conditions if it is necessary to distinguish them from forms without *-n* (see Horrocks 2010:275). Another factor that may have contributed to the retention of *-n* in the Gen plural is that this form had marginal status in Demotic as spoken before standardization occurred, but it was revived due to the influence of Katharevousa (ibid., 285, 462-463). As part of this restoration, regular sound changes like the loss of *-n* may have been effectively reversed.

Several declension class mergers simplified the overall nominal inflection system. Case syncretism increased as part of these processes, but number syncretism decreased. As mentioned above, singular *i*-stems merged with one of the two feminine *ā*-stem paradigms after sound changes led to three identical forms. This motivated the remaining forms to change by analogy. The Nom singular became *-i*, e.g., *pólis* » *póli* 'city', which is syncretic with the Voc/Acc/Dat singular as in the corresponding *ā*-stem paradigm, but this meant that it was no longer syncretic with the Nom/Voc/Acc plural, e.g., *pólis*. The Gen singular sometimes became *-is*, which is in fact syncretic with the Nom/Voc/Acc plural, but the original form *-eos* was often retained, or else restored, due to the influence of Katharevousa (see Horrocks 2010:288, 462-463). The distinctiveness of the original form and the avoidance of number syncretism might explain why it resisted the merger along with the plural forms.

The most significant merger was that of the masculine and feminine consonant-stems and *ā*-stems. As mentioned above, this likely began with the overlap in the Acc singular forms. Horrocks notes that the *ā*-stem singular forms spread to consonant-stems by analogy with the Acc, e.g., Gen *elpídos* » *elpídas* 'hope', while the consonant-stem plural forms spread to *ā*-stems, e.g., Nom/Voc *khórɛ*, Acc *khóras* » *khóres* 'countries'. In the singular, this simplified the paradigm by eliminating stem allomorphy among the Nom and the other case forms, e.g., Nom *elpís* » *elpída*. The consonant-stem Acc plural *-as* had already been leveled to the Nom/Voc *-es* by analogy with various minor declension classes like *i*-stems and *u*-stems, which already had strong parallels with the consonant-stems. The only difference that remains between the two declension classes is the position of the accent in the Gen plural form, e.g., *ā*-stem *khorón* vs. consonant-stem *elpídon* (2010:286-288). It is also possible that the *ā*-stem Dat plural form *-ɛs* was retained, since this would make it syncretic with the consonant-stem Nom/Voc/Acc plural

-es after the raising of [ɛ] to [e]. Even if this did not occur, only consonant-stems had a distinctive Dat singular form after the sound changes described above. Thus, this merger left only neuter consonant-stems with a distinctive Dat singular form. Once only one relatively small declension class had a distinctive Dat singular, and possibly only this class and the *o*-stems had a distinctive Dat plural, it is unsurprising that a distinctive Dat case could not survive.

Along with the merger of consonant-stems and *ā*-stems, the masculine nouns in these classes became less like masculine *o*-stems and more like other consonant/*ā*-stems. Masculine *ā*-stems previously had *-a* in the Voc singular regardless of the vowel in the other singular cases, but they adopted the feminine pattern in which the Voc is syncretic with the Acc/Dat singular, e.g., *políta* » *polítis* 'citizen'. The masculine consonant/*a*-stem Gen singular was also leveled to this form, e.g., *polítu* » *polítis*. Horrocks considers this an extension of the pattern already used for male names, including hypocoristics and foreign names, in earlier stages of the vernacular (2010:179-180). These developments reversed earlier similarities that had developed between masculine *ā*-stems and *o*-stems and contrast with the partial or complete mergers of different declension classes of the same gender in the Germanic languages, OF, MB, and MSrb. This might be explained by the fact that the *ā*-stem paradigm was originally derived from the *o*-stem paradigm, and they shared many parallels in their early development, in contrast to the more distantly related consonant-stems. Once consonant-stems and *ā*-stems began to merge, however, the large number of consonant-stem masculine nouns probably became the more natural source of analogy for the smaller class of masculine *ā*-stems. Changes such as these may be driven by some of the same underlying causes as case loss, as discussed in section 3.1.2.2 below.

Sound change had only minor effects on gender syncretism. As mentioned above, the loss of *-n* was blocked for determiners in certain phonetic contexts and when necessary to distinguish from forms without *-n*, i.e., the neuter. The spread of the masculine Nom plural form *i* of the definite article to the feminine was primarily the result of sound change, specifically the desyllabification of both masculine *i* and feminine *ɛ* to [j] before vowels. The influence of the feminine Nom singular form *i* may have also played a role in solidifying the syncretism of the Nom plural forms in all phonetic contexts. On the other hand, analogical processes during the Byzantine period increased the number of adjectives with distinctive feminine forms: compound *o*-stem adjectives adopted *ā*-stem forms for the feminine, and many consonant-stem adjectives were replaced by or remodeled as *o*-stem/*ā*-stem adjectives (see Horrocks 2010:289-291). Thus,

the feminine has become more distinctive overall, and the neuter has retained approximately the same status as in AG; all three genders are robust in Modern Greek.

Table 44. Case Mergers in Greek

	Mycenaean Greek						Ancient Greek				
PIE	Nom	Voc	Acc	Gen	Dat	Instr	Nom	Voc	Acc	Gen	Dat
Nom	x						x				
Voc		x						x			
Acc			x						x		
Gen				x						x	
Dat					x						x
Loc					x						x
Instr						x					x
Abl				x						x	

	Koine Greek					Modern Greek			
PIE	Nom	Voc	Acc	Gen	Dat	Nom	Voc	Acc	Gen
Nom	x					x			
Voc		x					x		
Acc			x					x	
Gen				x					x
Dat					x			x	x
Loc					x			x	
Instr					x			x	
Abl				x					x

In addition to the changes discussed so far, the Gen has been losing productivity to the Acc in spoken Greek. The fall into disuse and subsequent restoration of the Gen plural in the standard language has already been discussed above. Trąba argues that this has also been happening in the singular. While the Gen singular form remains distinctive in most declension classes, it has been restricted to the possessive function. In speech that has not been influenced by the standard, all prepositions now take the Acc; thus, the distinction between Acc and Gen meanings of prepositions has been neutralized. Particularly for northern dialects, in which the bare Acc is used for IOs, oblique may be a better label than Acc for this case (2018:126-128). Trąba finds the motivation for this loss of productivity in the existing syncretism, its polysemous character, which has led to its replacement by various prepositional constructions, which convey these functions more transparently, and its lower frequency (ibid., 129). Trąba compares the spoken

Greek Gen, which retains more functions such as IO for Gen clitic pronouns, to possessive *-s* in English (ibid., 128). However, the Gen is still expressed with different fusional markers depending on declension class, which suggests that the situation is much closer to that in CSHG, for which the standard may have also helped support a Gen case with limited functionality (see section 2.3.3.1 above), than in English, for which *-s* is analyzed as a clitic or phrasal affix. The Nom and Voc retain their original functions (see ibid.).

The order that the cases were lost in the development of Greek corresponds to their frequency of use: Instr was the least frequent, followed by the Dat (see Luraghi 2004:366-367). The Instr merged with the Dat in the development of AG. The Dat was lost sometime during the Byzantine period, although a wide range of dates have been proposed. Horrocks observes restrictions in the uses of the Dat and replacement by PPs and the Acc in texts as early as the 2nd century BCE. Overlap between the functions of the Dat and Gen began later, starting with clitic pronouns indicating possession, then spreading to ethical uses, and finally IOs. These uses were subsequently extended to nouns (2010:95, 107-108, 116). In Egyptian varieties, confusion with the Gen was particularly strong due to the neutralization of /o/ and /u/ in unstressed final syllables, as attested in Egyptian letters from the 2nd through 5th centuries CE, although these also contain PPs instead of adverbial uses and Acc as the prepositional case (ibid., 116-117, 179, 184-185). However, written Koine in 6th century Byzantium still regularly has the Dat in expected contexts, even in dialogue representing more colloquial speech (ibid., 245-247, 253-255). In his study of the decline of the Dat, Humbert argues that it was not completely lost in native speech until the 10th century (1930). Thus, the loss of the Dat was likely gradual and was completed much earlier in some regions than others. Another uncertainty is how the Dat merged with other cases. The Dat appears to have undergone a formal merger with the Acc, but its functions were assumed by both the Acc and Gen in the standard language. The Acc had become the only prepositional case in spoken varieties by the late Byzantine period, but the Gen came to be used for IOs in southern dialects, including the standard, in which a prepositional construction with *se* ‘to’ followed by the Acc can also be used (see Horrocks 2010:284-285, Trąba

2018:131).[89] Northern dialects have adopted the Acc for IOs as well, perhaps due to the syncretism with the Dat, as opposed to the syntactically motivated Gen (Humbert 1930:200). With the possible exception of Albanian, the merger of the Gen and Dat began earlier in Greek than in any other Balkan Sprachbund language. However, the merger was likely completed later in Greek than in Albanian and possibly Pre-Romanian. Regardless, Bulgarian and Macedonian were the last major Balkan Sprachbund languages to undergo this merger.

2.5.4. Albanian[90]

Proto-Albanian was used from the 1st century CE to 18th century (Orel 2000:xii). Albanian has been attested since the 15th century. This is relatively late for the first attested stage of a language; due to the lack of an earlier written ancestor, the reconstruction of Proto-Albanian is more speculative than that for the other proto-languages discussed in this section.[91] Contemporary Standard Albanian (CSA), spoken from 18th century, based on the Tosk dialect, maintains all of the same general case distinctions as Proto-Albanian. Orel (2000:232-240) provides a reconstruction of Proto-Albanian nominal inflection. This system is only reflected in the indefinite forms of CSA. Generally, Proto-Albanian distinguished five cases morphologically: Nom, Voc, Acc, a combined Gen-Dat, and Abl; the three IE genders; and two numbers: singular and plural. The IE Instr and Abl were lost without a morphological trace; the Albanian Abl derives instead from the Loc. The Gen-Dat also derives from the Loc[92] in the singular but from the Gen in the plural. Thus, the Gen-Dat and Abl were syncretic in the singular but not the plural. Similarly, Nom and Acc only remained distinctive in the *o*-stem plural paradigm. Both of these configurations are typologically uncommon. The latter has been leveled

[89] In these varieties, the Gen is still used to mark the complements of certain verbs, e.g., *mniázo* 'resemble'. However, many verbs that took a Dat or Gen complement in AG now take an Acc DO, e.g., *voïtháo* 'help', which previously assigned the Dat, and *akúo* 'listen to', which previously assigned the Gen to animate objects (see Trąba 2018:129-131).

[90] Grammatical information is drawn from Orel (2000) for Proto-Albanian, Newmark (1982) for Contemporary Standard Albanian (CSA).

[91] In some ways, the evidence for Albanian is similar to Romanian, which is only attested as a separate language since the 16th century. However, Romanian still has CL and, to a lesser extent, VL as written ancestors that help inform what developments occurred during its unattested period.

[92] For *a*-stems, the source was the Dat/Loc form, already syncretic in PIE. Otherwise, the Gen and Dat had already merged in the development of Proto-Albanian from PIE. Thus, Albanian was probably the earliest Balkan Sprachbund language to undergo a complete merger of the Gen and Dat.

as expected based on markedness,[93] and the distinction between Nom and Acc is only retained in the singular definite paradigms. The definite Abl plural form has also been leveled to the Gen-Dat in the past few centuries (see Demiraj 1993:145).

The functions of the cases in Albanian have changed from PIE, largely due to several case mergers. The Nom is primarily used to mark the subject of finite verbs and the complement of 'be' and other linking verbs, but it also marks the objects of a few prepositions. The Voc is used for direct address. The Acc continues to mark DOs, but apparently not motion toward (goal); it marks the objects of some prepositions, but these mostly indicate location. The original Instr functions of instrument, manner, and accompaniment are expressed with the preposition *me* 'with' followed by the Acc. The Gen-Dat inherited Gen functions including complement of a noun and partitivity, and Dat functions including IO. The Abl, which is usually used with a preposition, has inherited the location function from the Loc; it also indicates motion from (source) and cause, original Abl functions. In addition, it can mark the complement of a noun, but in contrast to the Gen-Dat, the noun must be uninflected, i.e., indefinite Nom/Acc singular, and no linking article is used (see Çanta 2017:230-234, Demiraj 1993:107, Newmark et al. 1982:135-139).

Table 45. Changes in Noun Declension from Proto-Albanian (reconstructed) to Contemporary Standard Albanian

	< o-stems		< ā-stems	
	Masc./Neut./Ambigeneric		Fem.	
	Sg.	Pl.	Sg.	Pl.
Nom	-a > -∅, -ë	-ai > -e, -∅, » -a; -ō > -e, -ë, » -a	-ā > -ë	-ā > -ë, -a
Voc	??? » -o	-ai > -e, -∅, » -a; -ō > -e, -ë, » -a	??? » -o	-ā > -ë, -a
Acc	-a > -∅, -ë	-a > -∅, -ë, » -e, -a	-ā > -ë	-ā > -ë, -a
Gen-Dat	-ei > -i, -u	-ō » -eve, -ve, -ave, -ëve	-āi > -e	??? » -ëve, -ave
Abl	-ei > -i, -u	-aisu > -esh, » -sh, -ash, -ësh	-āi > -e	??? » -ësh, -ash

In addition to the *o*-stems and *ā*-stems, the two main declension classes that have been retained in CSA, Orel (2000) has found evidence that Proto-Albanian had *i*-stems and a few remaining consonant-stems. Almost all consonant-stems had been redistributed to other declension classes,

[93] Generally, languages retain more distinctions (e.g., among cases) on unmarked categories (e.g., singular) than marked categories (e.g., plural). See section 3.1.2.1 below for further discussion of how markedness interacts with morphological change.

and *i*-stems were redistributed during Proto-Albanian, mostly to *o*-stems. Gender assignment was reorganized during Proto-Albanian based entirely on nouns' synchronic declension class rather than their inherited IE gender: masculine/neuter *o*-stems and feminine *ā*-stems. The *ā*-stem paradigm was already heavily influenced by *o*-stems in Proto-Albanian, and analogy between these two primary declension classes increased over time.

The plurals of some IE neuter *o*-stems were reanalyzed as singular feminine *ā*-stems, a process that also occurred in Romance languages. A new class with masculine singular and feminine plural forms and agreement developed. This pattern is similar to that found in CSRm, and Orel attributes it to Romanian influence. This class was defined by the semantics of the noun (definite mass) rather than form or inherited gender (Orel 2000:216-223, 232). However, because all of the *ā*-stem plural forms are also found on *o*-stems, this can now be considered a class of *o*-stems from a synchronic viewpoint. Some scholars use the label *neuter* for this class of ambigeneric nouns (see Baerman et al. 2005:82-83, Orel 2000:223). Although many IE neuters joined the ambigeneric class, others retain distinctive neuter agreement in the Nom/Acc singular of the definite and linking articles, in contrast to CSRm and CSI. In addition to a small number of inherited neuters, neuter agreement can be found for nouns substantivized from participles and adjectives. Due to their semantics, the majority of these nouns are only used in the singular (see Newmark 1982:133-134, 148). Thus, the neuter and ambigeneric classes are distinct from each other at least in some singular agreement forms, and my study considers them as two separate gender categories, while acknowledging that the latter is in the process of displacing the former.

As reconstructed, Proto-Albanian had slightly higher levels of number syncretism than CSA. The Nom/Acc singular and plural of *ā*-stems all shared an *-ā* ending in Proto-Albanian, as did the Voc plural, which was always syncretic with the Nom plural as in PIE. Similarly, the Nom/Acc singular and Acc plural of *o*-stems all shared an *-a* ending. In both classes, however, the singular and plural forms were often distinguished by different stress positions. There were two common accentual patterns. One had fixed stress on the stem, so the forms mentioned above were syncretic. In the other there was no number syncretism, since the stem was stressed in the singular, while the ending was stressed in the plural (see Orel 2000:126-127).

Orel notes that some Proto-Albanian forms cannot be reconstructed because the CSA forms have been borrowed from other languages and/or extended from other paradigms. The Voc singular form (*-o*) of both declension classes was borrowed from South Slavic *ā*-stems, e.g., *ā*-

stem *Neno*! ‘mother!’ and *o*-stem *Burro*! ‘husband!’. This development is similar to what happened in Romanian, and likewise supports Balkan Sprachbund influence, but it was further extended from *ā*-stems to *o*-stems. In addition, the *ā*-stem Gen-Dat and Abl plural forms have been extended from the *o*-stem paradigm without any trace of the original *ā*-stem forms. The first vowel in these forms have been modeled on the *ā*-stem Nom/Voc/Acc plural, e.g., Nom/Voc/Acc *grima*, Gen-Dat *grimave*, Abl *grimash* ‘crumbs’ (2000:233, 239).

In the development of CSA from Proto-Albanian, sound changes led to an increase in the number of alternative forms within the main two declension classes but did not increase number syncretism overall. Many vowels underwent different regular sound changes depending on whether they were stressed or unstressed in Proto-Albanian. At some point after this differentiation, stress shifted away from the final syllable. Final *-ā* became *-a* when stressed but *-ë* when unstressed (ibid., 123, 127, 238-239). For *ā*-stems, this maintained the same level of number syncretism: nouns that had distinguished the Nom/Acc singular from the Nom/Voc/Acc plural based on stress in Proto-Albanian now have *-ë* in the former and *-a* in the latter, e.g., Nom/Acc singular **grímā* > *grimë*, Nom/Voc/Acc plural **grimá́* > *grima* ‘crumb(s)’; nouns that had previously had fixed stem stress now have *-ë* in all of these forms, e.g., **bárā* > *barrë* ‘load(s)’.

The effect of sound changes on number syncretism was more complex for *o*-stems. Final *-a* became *-ë* when stressed but was lost when unstressed. In addition, unstressed final *-ō* became *-ë*, and unstressed final *-ai* was lost, but stressed final *-ō* and *-ai* merged as *-e* (see Orel 2000:233-236). Together, these sound changes resulted in different syncretism patterns among *o*-stems in the development of CSA, depending on the Nom/Voc plural form and accentual pattern of the noun in Proto-Albanian. When the ending had only been stressed in the plural, there continued to be no number syncretism, e.g., Nom/Acc singular **gárda* > *gardh*, Nom/Voc plural **gardṓ/*gardái* > *gardhe*, Acc **gardá* > **gardhë* ‘fence(s)’. When the stress had been fixed on the stem, the Nom/Acc singular and Acc plural remained syncretic, now with a zero ending, e.g., **édzja* > **esh* ‘hedgehog(s)’, **swésura* > *vjehërr* ‘father(s)-in-law’. When the Nom/Voc plural had been *-ō*, it remained distinctive, e.g., **édzjō* > *eshë*, but when it had been *-ai*, it joined the syncretism, e.g., **swésurai* > *vjehërr*. In contrast to *ā*-stems, a small group of *o*-stems had been stressed on the ending in the singular (ibid., 233). When the plural was also stressed, the Nom/Acc singular and Acc plural remained syncretic, now as *-ë*, e.g., **burá* > **burrë*

'husband(s)'. When the plural had been unstressed with *-ō* in the Nom/Voc, this form became syncretic with the Nom/Acc singular as *-ë*, while the Acc plural developed a distinctive zero ending, e.g., Nom/Acc singular **brumá* > *brumë*, Nom/Voc plural **brúmō* > *brumë*, Acc **brúma* > **brum* 'dough(s)'. When the plural had been unstressed with *-ai* in the Nom/Voc, this form instead developed the same zero ending as the Acc plural, avoiding number syncretism, e.g., Nom/Acc singular **delá* > *djalë*, Nom/Voc plural **delmai* > *djem*, Acc **delma* > *djem* 'boy(s)'.[94] These different outcomes for unstressed vowels are similar in some respects to developments such as vowel raising in Bulgarian and vowel reduction/deletion in Germanic languages and French. Although vowels in all case forms were affected in Albanian, neutralizations only occurred among the Nom, Voc, and Acc. Since these forms were often syncretic already in Proto-Albanian, syncretism patterns changed, but drastic increases like those in the other languages did not occur.

Analogical processes had a greater effect on distinctions in nominal inflection than sound change did, reducing number syncretism, in addition to making the declension classes more similar. These included the spread of Voc singular, Gen-Dat plural, and Abl plural forms between classes, as described above, plus several other changes described here. The *o*-stem Gen-Dat plural is itself the combination of the regular development (*-e*) and the *u*-stem form *-ve*, e.g., *gardheve* 'fences' (see Orel 2000:237). As in the South Slavic languages, this *u*-stem form was probably extended because of its distinctiveness; the inherited form alone would have been syncretic with one of the Nom/Voc plural forms, as in the reconstruction of Proto-Albanian. As in the *ā*-stems, the first vowel in this form and the Abl plural has been changed by analogy with the Nom/Voc/Acc plural for some nouns, e.g., Nom/Voc/Acc *eshë*, Gen-Dat *eshëve*, Abl *eshësh* 'hedgehogs'. Orel notes two other analogical processes that affected *o*-stems. First, many Acc plural forms have changed to match the Nom/Voc plural forms, e.g., **gardhë* » *gardhe*, by analogy with the forms that were already syncretic, e.g., Nom/Voc/Acc *vjehërr* 'fathers-in-law'. Second, the *ā*-stem Nom/Voc/Acc plural form *-a* has replaced inherited *-e* for some *o*-stems, e.g., **are* » *ara* 'bears' (ibid., 235-236). The former change decreased number syncretism because the inherited Acc plural forms were syncretic with the Nom/Acc singular more often than the Nom plural forms were. Although the latter increased parallelism between the declension classes, it had little or no impact on number syncretism, since the extended *ā*-stem

[94] This particular noun also has a stem-change in the plural.

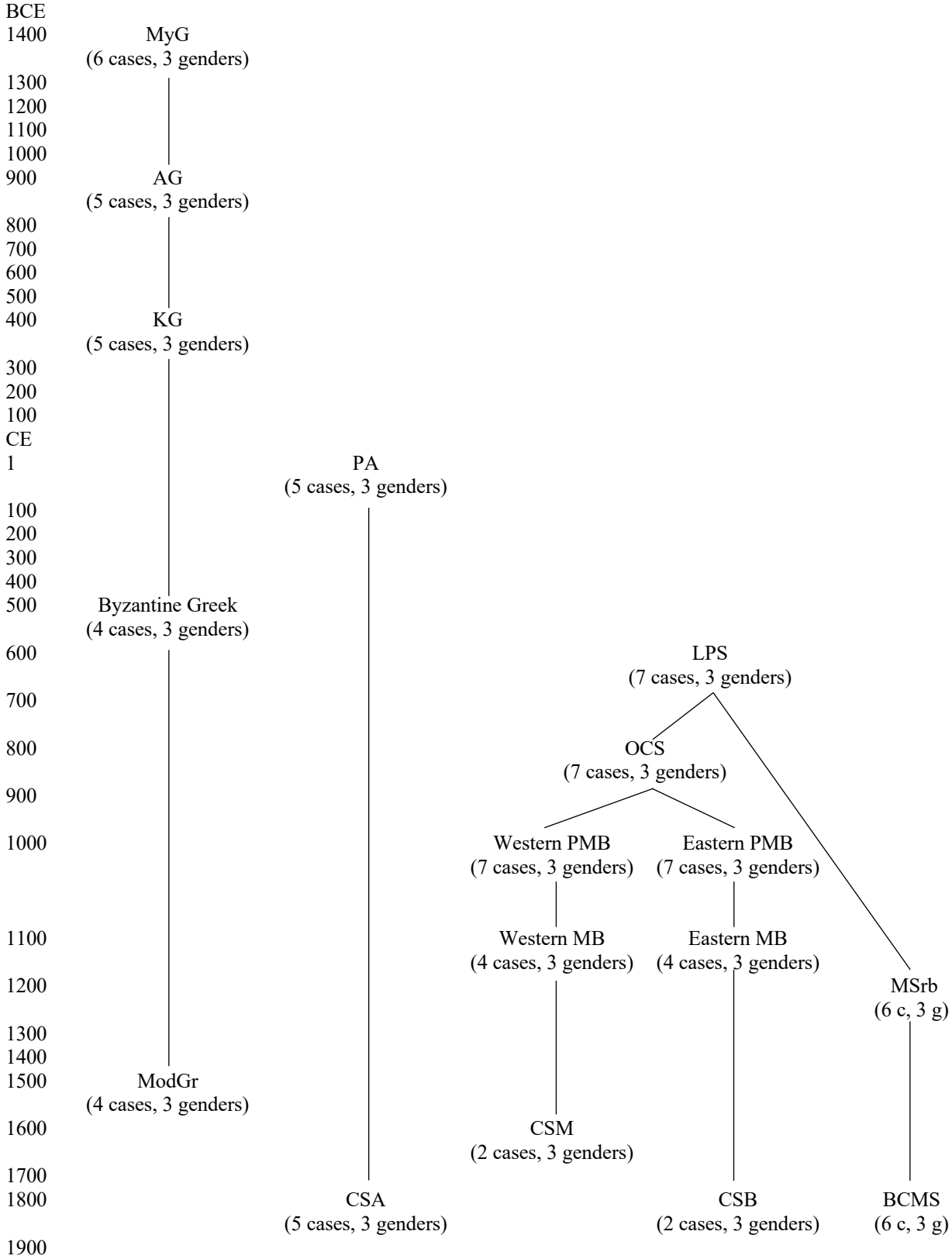
form seems to have only replaced *-e*, which was already distinct from the Nom/Acc singular. In addition, even when the same endings are used for the Nom/Acc singular and Nom/Voc/Acc plural, stem changes including umlaut and palatalization often profile number marking, as umlaut does in many Germanic languages. The Nom/Voc plural endings regularly triggered the umlaut of certain stem vowels and the palatalization of certain stem-final consonants. Although these processes sometimes reduced number syncretism in the paradigms of the nouns to which applied, they did not eliminate number syncretism until the resulting stem changes were extended to the Acc plural as well (see ibid., 72-77, 140-146, 236).

Agreement targets in CSA tend to distinguish the feminine from the masculine and neuter in the singular. The only distinctive neuter forms are the Nom/Acc singular of the postposed definite article, and the Nom of the linking article used with adjectives. Otherwise, it is syncretic with the masculine in the singular. The masculine and feminine are consistently distinguished for singular definite nouns, but not for the linking articles. These have distinctive masculine and feminine forms in the Nom singular. Definite linking articles also have distinctive Gen-Dat/Abl singular forms, but neither definite nor indefinite linking articles distinguish gender in the Acc or any plural form. Likewise, the same postposed definite article forms are used in the plural regardless of gender. Only adjective endings distinguish masculine and feminine in the plural (Wahlström 2015:100-101). As mentioned above, ambigeneric nouns have masculine agreement in the singular and feminine agreement in the plural. Fewer gender distinctions in the plural is not surprising, but the retention of a gender category based on only a few distinctive forms is. In fact, the neuter's lack of distinctiveness may be connected to its displacement by the ambigeneric pattern. As mentioned above, Romanian influence likely played a role; it had developed the ambigeneric pattern following a similar loss of distinctiveness of the neuter, so the adoption of this pattern accomplished the same thing in both languages.

2.5.5. Summary

This section has discussed case and gender developments in the Balkan Sprachbund languages. The facts of these changes are summarized in Table 46 below.

Table 46. Timeline Summary of Case and Gender in Balkan Sprachbund Languages



The Balkan Sprachbund languages all share certain developments, i.e., the early loss of less frequent cases such as the Abl, Instr, and/or Loc, and the merger of the Gen and Dat. However, they differ in the extent to which further case loss has occurred. Different degrees of case loss can also be observed between some of the Balkan Sprachbund languages and other languages in their respective families: Romanian is the only Romance language that retains morphological case, while Bulgarian and Macedonian are the only Slavic languages that have lost case completely (except the Voc). While there are also slight differences in how gender has developed among the languages within each of these groups, its development has been much more uniform than in Germanic.

Sound changes have affected nominal inflection in all Balkan Sprachbund languages, but in different ways and to different degrees. Phonemic vowel quantity, at least on the final syllable, was lost at some point in all of their developments. It was lost regardless of position in EVL (as in WVL) and KG, and in the development of CSA from Proto-Albanian. It was lost on final syllables in the development of LPS, although it redeveloped in some Slavic languages, including MSrb.[95] Germanic languages had also lost vowel length distinctions on unstressed syllables (including the final syllable) before their earliest attested stages, with the exception of High German, in which they were lost as part of the vowel reduction that occurred in the development of MHG from OHG. The loss of phonemic vowel quantity before the attested stages is beyond the scope of this investigation, but syncretism most likely increased somewhat as it did in EVL and KG.

The loss of *jers* in weak position and denasalization of **ę* and **ǫ* in Slavic languages, as well as the merger of **i* and **y* in South Slavic languages greatly increased number syncretism in PMB and MSrb, including among the Nom/Acc singular and Gen plural and between the Instr singular and Dat plural. Case syncretism among the Nom, Acc, and Instr within each number also increased as a result, but less so in MSrb than PMB because of the different reflex of **ǫ* and analogical extension of soft stem forms. In eastern dialects of PMB, vowel raising in unstressed syllables further increased syncretism involving oblique cases and between the Nom and Acc plural; this process had similar causes and effects as vowel reduction in Germanic languages and the vowel mergers in VL. Overall, these sound changes in the South Slavic languages are

[95] See section 3.1.1.2 below for a more in-depth discussion of this and other prosodic developments in the Slavic languages.

somewhat comparable in their effects to those in VL and Germanic languages, as were the vowel mergers and loss of final *-n* that occurred starting in KG; these changes in Greek greatly increased syncretism involving the Dat, as well as the overlap among certain declension classes. Vowel changes also occurred in the development of CSA, including different outcomes for unstressed vowels that resemble vowel reduction and deletion processes in other languages, but the overall effect on syncretism was limited. Many of these sound changes can be attributed to the development of rhythmic dynamic stress, but not those that occurred in all South Slavic languages, since some of these, including BCMS, have retained a pitch accent.

The Balkan Sprachbund languages have all undergone certain functional mergers, most notably of the Gen and Dat. However, other analogical processes varied significantly depending on the extent to which case loss on indefinite nouns occurred; more leveling was associated with more case loss, while the analogical preservation and extension of distinctive forms was associated with case retention. The borrowing of the Gen singular form for the Acc and the extension of the *u*-stem Instr plural *-mi* to other masculine classes occurred in both PMB and MSrb, reducing syncretism between each form and the Nom of the same number, as well as certain forms of the other number. Similar changes also occurred in other Slavic languages, and MSrb was not part of the Balkan Sprachbund, so membership in it could not have been a factor. However, the changes in MB were very different from those that began in MSrb and have been completed in BCMS, so contact with other Balkan Sprachbund languages may have played a role in the former. In MB, the Instr and Loc were leveled to the Acc, similar to how oblique cases that remained distinctive after sound changes were leveled to the Acc in WVL, but on the basis of less existing syncretism, particularly for the Loc. In addition, Gen forms were leveled to the Dat as they were in Pre-Romanian, but again based on less existing syncretism. Number syncretism decreased as a result of these processes, but functional mergers were likely a more important factor than number profiling. In BCMS, on the other hand, the Loc almost completely merged with the Dat, and the Dat/Instr dual form was adopted for the Dat/Loc/Instr plural. Extensions that profiled number also applied, including the spread of distinctive Gen plural forms, in contrast to the leveling of the Gen in MB.

A functional merger of the Instr and Dat occurred in AG, with mostly Dat forms surviving but some Instr and analogical formations. In the development of ModGr, the Dat forms that remained distinctive, one in the singular and most in the plural, were leveled to the Acc, but

its functions were assumed by a combination of the Acc and the Gen. In addition, the loss of *-n* was blocked morphologically in the Gen plural, as it was in the weak plural in EME. A more transparent gender-declension relationship developed in ModGr through a reorganization of declension classes closely resembling the developments in CSRm and OF. In the development of CSA, a new distinctive Voc singular form was borrowed from South Slavic *ā*-stems, distinctive oblique plural forms were extended, and subclasses with the same vowel in all plural forms developed. This last process increased transparency but leveled the few remaining indefinite Nom-Acc distinctions; definite nouns still retain this distinction, as in High and Low German. In contrast, CSRm only retains this distinction on pronouns, like CSB and CSM. A number of potential reasons behind the differences established in this section are discussed in chapter III below; these include differences in prosodic development, different outcomes of case variation, and the effects of various contact situations.

CHAPTER III.
MOTIVATIONS FOR CASE AND GENDER LOSS

A number of different causes have been proposed for the loss of morphological case and grammatical gender in the Germanic, Romance, and Balkan Sprachbund languages. These causes can be classified as internal (e.g., sound change), or external (i.e., language contact). A combination of multiple internal and external motivations is probably needed to account for all instances of case and gender loss across these languages, as in proposals by authors such as Barðdal (2009), Duridanov (1956), and Wahlström (2015). However, Rapaport (2016) makes a convincing argument that prosodic change can account for all of the remaining developments concerning the forms and functions of nominal inflection. This is the first argument considered in section 3.1, followed by other internal motivations, while external motivations are the focus of section 3.2. Due to the proposed connections between many of the internal and external factors, the division between the content of the sections is not strict.

Although their accounts are limited to the Balkan Sprachbund languages, Wahlström (2015) and Duridanov (1956) make arguments that cover many different factors, internal and external. Therefore, a brief introduction to their overall arguments are presented here with the aim of providing context for their specific arguments in the various subsections. In his account of the loss of case inflection in Bulgarian and Macedonian, Wahlström argues that mutual bilingualism caused individual linguistic features to diffuse among the Balkan Sprachbund languages. These features include several functional mergers: the merger of goal and location, weakening the Loc, the merger of instrument and accompaniment, weakening the Instr, and the merger of recipient and possessor, the most notable shared feature of Balkan Sprachbund case systems. They also include clitic doubling and other means for marking grammatical relations and information structure that are characterized by explicit analytism. In addition, Wahlström argues that in the development of Bulgarian and Macedonian, a high number of second language (L2) speakers prevented the restoration of case distinctions obscured or lost due to sound change, in contrast to other Slavic languages, which underwent some of the same sound changes and resulting syncretism. He concludes that multiple factors are needed to account for case loss in Bulgarian and Macedonian, including different types of contact (2015:185-190). As Wahlström (2015:20) points out, Duridanov's (1956) arguments are colored by the lens of the Soviet

philosophy of dialectical materialism, under which he was forced to work in order to publish his work in communist Bulgaria, and which attributes change to the internal conflict between contradictions inherent in every object and phenomenon. His discussions of the textual evidence serve as a backdrop for later research, but further textual evidence has come to light and significant developments have been made across many relevant fields of linguistics.

The following subsections evaluate arguments made to account for case and gender loss in the Germanic, Romance, and Balkan Sprachbund languages. Section 3.1.1.1 discusses argument about the driving force of prosody. In section 3.1.1.2, this argument is extended to the Slavic languages, with a particular focus on vowel reduction, an outcome of prosodic change that can affect nominal inflection to various degrees. The ease of production is another factor that can cause case loss discussed in section 3.1.1.3. Section 3.1.1.4 considers orthographic substations in BChS.

The remaining internal motivations concern morphosyntactic and semantic factors, some of which connect to external factors as well. First, the organization of paradigms and the merging of declension classes are considered in sections 3.1.2.1 and 3.1.2.2, respectively. Next, the role of grammaticalization, particularly of definite articles, is considered in section 3.1.2.3; Wahlström's typological investigation of this factor has implications outside of the Balkan Sprachbund. Functional mergers are another important aspect of Wahlström's account, particularly as one of the ways that contact can lead to case reduction. The paths these developments can take are discussed in section 3.1.2.5, but first case variation is considered in section 3.1.2.4 as a precursor to both functional mergers and functional narrowing, which is the focus of section 3.1.2.6. For Barðdal (2009), the functional narrowing is the most important mechanism by which contact can induce case loss. In addition to functional mergers, Wahlström (2015) instead sees the rise of analytism as an important connection between contact and case loss. This common development in the Germanic, Romance, and Balkan Sprachbund languages is considered in section 3.1.2.7. The impacts of various types of contact are discussed in section 3.2, with a particular focus on structural convergence, one of Wahlström's proposed developments due to contact that can bring about or accelerate the paths of development involving case variation, functional mergers, functional narrowing, and especially analytism, which is otherwise difficult to justify as anything other than an effect of other morphosyntactic

changes. For this reason, the subsections in 3.1 concerning these developments necessarily include references to external factors.

As part of the evaluation of these various arguments, and of the importance of particular factors in a broader sense, I consider how they can be applied to other Germanic, Romance, and Balkan Sprachbund languages beyond those considered by the authors. Sometimes this entails adding analysis of my own and mentioning how they fit into my more general analysis.

3.1. Internal Motivations

This section is divided broadly into phonological motivations and morphosyntactic/semantic motivations. The former motivations involve form without reference to function. The latter motivations relate to the connection between form and function. This connection involves a complex interplay among morphology, syntax, and semantics.

3.1.1. Phonological Motivations

Specific sound changes and their effects on nominal inflection in the Germanic, Romance, and Balkan Sprachbund languages have been discussed in chapter II above. This section considers two broader factors that are closely connected to each other and to many of these specific changes: prosodic change and vowel reduction. Prosodic change may be a factor in many of the segmental changes discussed above, but perhaps more importantly, it can explain how some case endings were lost even when no specific sound changes applied. Vowel reduction is a process commonly associated with the type of dynamic stress found in many of the languages under consideration. This process appears to be involved in many lost case and gender distinctions, but vowel reduction can occur without case or gender loss and vice versa. Thus, it is important to consider the types of vowel reduction and other factors such as stress position that mediate its effect on case marking. Orthographic substitutions are considered in the third subsection; these are one of the main sources of evidence for case variation and mergers, whether due to sound change or functional overlap, but they may arise due to unrelated scribal errors. Thus, the nature of this evidence must be considered carefully.

3.1.1.1. Prosodic Change

Prosodic change is a crucial factor in the loss of case and gender inflections in the languages under investigation. Building on previous research, Rapaport (2016) argues convincingly that the word-final case markings used by Germanic, Romance, and Hellenic languages were particularly vulnerable to the apocope which resulted from standard prosodic change. These families all underwent a shift from the PIE pitch accent to dynamic stress in the centuries preceding case loss, i.e., accent came to be predominantly marked by intensity and duration rather than pitch. Bybee et al. (1998:277) argue that this diachronic change is cross-linguistically common. Rapaport (2016) describes the ambiguous intermediate stages that native listeners had to navigate as their languages transitioned from pitch accent to dynamic stress.

The dynamic stress refers to an emphasis on loudness (Allen 1973:74). In dynamic stress languages, every accent has some element of intensity, duration, and pitch on the main stressed syllable. Some languages place an emphasis on pitch, while others place an emphasis on intensity or loudness (Bybee et al. 1998:267, Allen 1973:6). Rapaport (2016:2) states that while dynamic stress may involve elements of duration and tone contour, the defining characteristic is that it is based on loudness (intensity or amplitude) in the native perception. He supplements that model with the argument that diachronic prosody, working with phonological erosion, is primarily responsible for the loss of case endings in certain IE languages, and that this case loss occurs at a predictable point in these languages.

The PIE morphological cases were distinguished by suffixes that typically had one syllable. The loss of phonemic vowel length and the transition to rhythmic stress adjustment as well as the frequent erosion of unstressed syllables often led the suffixes to be light. Moreover, based on the reanalysis of new metrical rules, these suffixes became extrametrical. Consequently, without the role of sound change, unstressed final syllables were susceptible to vowel reduction as well as the tendency to be neutralized or lost (Rapaport 2016:4).

A fixed initial accent developed in PGmc and Proto-Italic. This may have occurred at least in part because these daughter languages predominantly adopted thematic, i.e., vowel stem, nouns, which were consistently accented on either the root or the theme vowel across all case forms, whereas many athematic, i.e., consonant stem, nouns were accented on the endings of the oblique cases (Rappaport 2016:22-23). Thus, an initial accent would have become more common and eventually extended to all words. Ringe observes that this process must have occurred

around the application of Verner's Law in PGmc.[96] Because Verner's Law describes a shift from voiceless to voiced consonants following an unstressed syllable, stress must obviously be a factor in the language at this point. The existence of stress implies that the language must have been at least partially dynamic, rather than entirely pitch-based, but at the same time this process depended on the inherited variation in accent position. Thus, stress must not have become completely fixed until after Verner's Law (2006:105). Rapaport dates this shift to around 500 BCE. Not long after, around 400 BCE, the AG pitch accent became dynamic in the development of KG, although it generally maintained the same syllable position. The accent remained fixed on the initial syllable until around 250 BCE in Old Latin, but a shift away from the first syllable can be observed in the works of Plautus around 200 BCE. Since this shift appears to have resulted from weight sensitivity, i.e., the effect of vowel length and coda consonants on the accentual pattern, a weight-sensitive dynamic stress can be reconstructed starting around this time (2016:39, 58).

In terms of Hayes' (1995) metrical stress theory, all of the languages above had developed a rhythmic dynamic stress with feet based on the moraic trochee. In this type of foot structure, a foot is composed of two moras, i.e., units of weight, with stress on the first. A light syllable, with only a short vowel in the rhyme, counts as one mora, while a heavy syllable, with a long vowel and/or coda consonant, counts as two. One foot of each word receives primary stress, while the others have secondary stress. These languages did not have the exact same foot structure, however. OE, representative of at least the old West Germanic languages, parsed feet from left to right, with primary stress on the leftmost foot. In contrast, CL and KG parsed feet from right to left, with primary stress on the rightmost foot. CL required a word-final extrametrical syllable. In OE and other Germanic languages, each moraic trochee was optionally followed by an extra mora (see Dresher & Lahiri 1991).[97] This system, termed a Germanic foot, differed slightly from extrametricality as in CL, however, since these extra moras were found within words, not just word-finally. There was no extrametricality in KG (see Lahiri et al. 1999).

[96] Verner's Law describes the process by which voiceless fricatives were voiced when they were not adjacent to a voiceless sound and the preceding vowel was unstressed. This process applied to the reflexes of PIE voiceless stops that had become fricatives by Grimm's Law and to PIE *s (Ringe 2006:102).

[97] A light syllable followed by a heavy syllable (but not vice versa) still functioned like a moraic trochee even though there were three moras. Thus, a Germanic foot could have up to four moras, but only for this particular pattern.

The weight-sensitivity and strong alternating rhythm associated with a foot structure based on moraic trochees set these languages up for further prosodic changes that threatened case distinctions. Vowel reduction is a common phenomenon associated with dynamic stress in which unstressed vowels are reduced in duration and undergo changes in quality, usually becoming more central (Rapaport 2016:40-41). At the same time, raising is also common, specifically low vowels (Kapatsinski et al. 2020:26).

The opposite process often applies to stressed vowels: they tend to be lengthened and become more peripheral in quality (Rapaport 2016:40-41). Vowel reduction beginning in OE is described in section 2.3.3.4 above. Although it is not apparent in writing, phonetic vowel reduction was also significant in CL: most short, unstressed vowels were more centralized than their long and/or stressed counterparts, i.e., [ɪ ɛ ɔ ʊ] vs. [i(:) e(:) o(:) u(:)] (see Allen 1973). Vowel reduction further reinforces the alternating pattern of rhythmic dynamic stress as described by the Iambic-Trochaic Law: a short-long-short-long alternation is perceived as iambic, while a loud-soft alternation is perceived as trochaic (Woodrow 1909). Notably, slight differences in duration are actually perceived as a loudness alternation, resulting in a trochaic perception, as in the languages under discussion (Rapaport 2016:41-46). A positive-feedback loop results, explaining the unidirectional and relatively rapid shift from pitch accent to dynamic stress in these languages (ibid., 42).[98]

While one may at first think that, as a trochaic language adds duration to each louder syllable, an iambic pattern may be heard instead, creating confusion, Woodrow (1909: 41-46) indicates the contrary. Slightly increased duration is perceived as loudness instead, and, up until the point when duration differences become obvious, subtle long-short alternations are perceived as trochaic, not iambic. Rapaport suggests that this bias in favor of trochaic loud-soft groupings reinforces emerging trochaic perception of dynamic stress and increases learners' production of loudness/duration on accented syllables, further reinforcing vowel reduction on other syllables (2016:42). As phonemically short, stressed vowels were lengthened, and phonemically long, unstressed vowels were reduced, phonemic vowel quantity was obscured, and its functional load decreased. Length became associated more and more closely with stress until phonemic vowel quantity was completely neutralized (Rapaport 2016:42-44).

[98] Vowel reduction is discussed in more detail in the following subsection.

There is a relationship between predictability of speech and language redundancy. The more predictable a portion of speech, whether due to context or frequency, the more it tends to be shorter or reduced. This process is called language redundancy which refers to the predictability that comes from the structure of the language. Its relationship with duration is anticipated by the idea that the greater a language's redundancy of word, syllable, or phoneme, the lower its acoustic redundancy, where acoustic redundancy refers to the possibility of recognition based solely on acoustic features. In this view, duration mirrors acoustic redundancy since longer portions are arguably more conspicuous and recognized. Spectral parameters, such as formant frequencies (*F1/F2*) in the center segment of a vowel, can also be used as indicators of acoustic redundancy because they are linked to vowel distinctiveness and thus probability of recognition (Aylett & Turk 2006:3048). In terms of *F1/F2* effects, the vowels would show higher centralization as language redundancy increased, and that the path of *F1/F2* changes would differ depending on vowel type (ibid, 3055).

Together, vowel reduction and the subsequent loss of phonemic vowel quantity explain the VL vowel mergers discussed in section 2.4.1 above: after the loss of phonemic length distinctions, reduced *i*, pronounced [ɪ], was closer to *e* than unreduced *i*, and reduced *u*, pronounced [ʊ], was closer to *o* than unreduced *u*. In fact, vowel reduction is attested around the same time as the loss of phonemic vowel quantity in VL, early KG, and the Middle Germanic languages. Like the shift to dynamic stress, these fast and unidirectional changes are explained by a positive-feedback loop: vowels produced with a short duration, regardless of their original phonemic quantity, are perceived as unstressed, and unstressed vowels are likely perceived as short, leading to further length changes via vowel reduction (Rapaport 2016:42-44).

Hayes (1995), Lahiri et al. (1999), and others argue that foot structure is a fundamental phenomenon that determines the formation of phonological rules rather than an emergent phenomenon resulting from independently-motivated phonological rules. Evidence comes from the conspiratorial behavior of processes that enforce foot structure compliance. Some of these processes can be observed synchronically in English dialects as they adopt loanwords. In Present Day English (PDE), i.e., contemporary varieties of English, standard and otherwise, foot structure is the same as in CL: moraic trochees parsed from right to left, with an extrametrical final syllable. In many varieties of PDE, including standard varieties, the French loan *police* does not conform to this foot structure: it has a light syllable followed by a stressed heavy syllable, so

it is not moraic and does not have an extrametrical syllable. In Glasgow Scots and some Urban American dialects, however, *police* is stressed on the first syllable rather than the second. With this change and the associated lengthening of the open stressed syllable, it conforms to the PDE foot structure. Also in Glasgow Scots, *company* is pronounced [ˈkʌmp.nɪ], with syncope of a medial extrametrical syllable bringing it into compliance (Hayes 1995). Of course, not all non-compliant words have changed to conform to the PDE foot structure, and some words are still distinguished primarily by stress position, so this foot structure is not as absolute as it apparently was in CL. Nevertheless, the general trend since ME has been for more nouns to conform to this structure, especially in non-standard varieties. Therefore, the claim for foot structures in English and other languages with lexical stress still appears valid.

Many sound changes in OE and EME enforce compliance with the Germanic foot structure. One change that most likely played a critical role in case loss was final schwa deletion, which can often be attributed to foot structure compliance. For example, *hōlie* 'holy' had a heavy syllable followed by two light syllables, which is one mora too many for a Germanic foot. Deletion of the final schwa reduced it to the three-mora maximum for a Germanic foot. In contrast, *manie* 'many' did not lose its final schwa at this stage because all three of its syllables were light and it already fit a Germanic foot (Rapaport 2016:25, 30-33).

Lass (1992) argues that the influx of Latinate loanwords in the 11th century contributed to the reanalysis of foot structure in ME from the Germanic foot system to the modern system with moraic trochees and an extrametrical final syllable. This involved a change in parsing direction and primary stress placement as well, but stress did not actually change for most words, since they only had one or two syllables. When a stress change occurred, a period of variation between the old and new forms is often attested (ibid., 89). Further quantity adjustments followed this reanalysis, gradually resolving non-compliant feet. The loss of remaining final schwas may have been one such result. For example, *sune* 'sun' had two light syllables, which formed a moraic trochee and therefore an acceptable Germanic foot. After the requirement for a final extrametrical syllable arose, however, the loss of the final schwa yielded *sun*, a single heavy syllable and acceptable moraic trochee (Rapaport 2016:54, 57).[99]

Other quantity adjustments resulted in noun paradigms with unpredictable alternations between long and short vowels in LOE and EME. The lengthening of the first vowel in disyllabic

[99] The requirement for an extrametrical final syllable does not apply to monosyllabic words.

words, attested from the 11th century, was another response to the extrametricality requirement. For example, *water* had a light syllable followed by a heavy syllable. Reanalysis of the second syllable as extrametrical left a single light syllable. Vowel lengthening resulted in a heavy syllable, an acceptable moraic trochee. This change did not apply to the Gen singular form *wateres*, however, because the case ending served as the extrametrical syllable, leaving two light syllables to form a moraic trochee. The shortening of an initial stressed vowel in trisyllabic words also occurred, allowing a light second syllable to form a moraic trochee with the first, since only one syllable can be extrametrical. For example, the first vowel was originally long in all forms of *brōdor* 'brother'. The Nom singular form remained acceptable under both the old and new foot structure, but its Gen singular form *brōdores* changed to *broderes* to conform. Both of these quantity adjustments resulted in vowel length alternations within paradigms, whose inconsistency may have confused learners (Rapaport 2016:55-57) and been leveled as a result, as with the confusing distribution of *-s* in OF discussed in section 2.4.1.3 above.

Other Germanic languages, including German, Dutch, and continental Scandinavian, have undergone similar shifts in foot structure and the resulting quantity adjustments (Rapaport 2016:54). The modern West Germanic languages have the same basic foot structure as PDE, but sometimes with differences in extrametricality (ibid., 30). CSHG and CSDu seem to have more of a tendency for primary stress on the leftmost foot than PDE, but all three languages have some derivational suffixes that attract primary stress and some that do not (see Eisenberg 1994:357-358, De Schutter 1994:444-445, 454-455). CSHG and ModGr retain some case distinctions, but they are on the path to being lost. The utilization of determiners that carry case information and their agreement with the nominal case makes the noun declension survive even though the endings have undergone a high degree of case syncretism (Rapaport 2016:5). The contemporary continental Scandinavian languages have a similar foot structure to the West Germanic languages, but they have also undergone additional prosodic developments. In Swedish and Norwegian, a new pitch accent has developed, with two contrasting tones, but the accent position remains the same (see Askedal 1994:227-229, Andersson 1994:274-277). In many varieties of Danish, a different prosodic phenomenon associated with stressed syllables has arisen. This phenomenon, *stød*, involves an incomplete glottal stop during the stressed vowel. Stressed syllables with *stød* contrast with those pronounced normally, but stress position is not affected (see Haberland 1994:313-314, 317-318, 321).

In contrast, Icelandic and Faroese are more conservative in their prosody: stress is not weight-sensitive, with syllabic trochees parsed left to right and primary stress on the leftmost foot (Lahiri et al. 1999:366). The correlation between the conservative prosody and case marking of these two languages adds further support to the importance of prosodic change in case loss. Likewise, the correlation between their conservative prosody and well-established isolation, discussed further in section 3.2 below, supports the role of contact in prosodic change (see Rapaport 2016:14-15). In fact, the slightly more innovative nature of Faroese compared to Icelandic holds for prosody as well as nominal inflection. Faroese has more exceptions to the foot structure, namely words with non-initial primary stress. These are mostly loanwords, but, in addition, not all prefixes are stressed (see Barnes & Weyhe 1994:196).

A significant portion of VL case loss can be attributed to the vowel mergers resulting from vowel reduction and neutralization of phonemic vowel quantity, along with the loss of final *-m* and, in some dialects, *-s*. In fact, the loss of final consonants may also be motivated by prosody, since the reduction of final unstressed syllables makes consonants harder to distinguish as well as vowels. In addition, a number of other prosodic developments may have contributed to Romance case loss. The CL requirement for an extrametrical final syllable was eliminated in many dialects, leading to the reduction and loss of this syllable (Rapaport 2016:22, 65). Effects of this change are clearly observable in the development of French, including the loss of most vowels in the final syllable in OF, as described in section 2.4.1.3 above, and the eventual loss of most final consonants.

Rapaport also addresses the retention of robust case systems in other IE languages. Generally, these fail to undergo one of the major prosodic developments that drove case loss in Germanic, Romance, and Hellenic languages. Armenian still retains seven morphological cases. It developed dynamic stress, but this was fixed on the final syllable. Thus, case endings were consistently stressed and never lost distinctiveness, although the forms are quite innovative. Lithuanian and Latvian also retain seven IE cases. These Baltic languages developed a mix of dynamic and pitch accents, but they did not acquire a rhythmic foot pattern because heavy syllables attracted pitch accents, not dynamic stress. Thus, phonemic vowel quantity was actually reinforced by the accent rather than neutralized, and case endings remain distinctive (2016:68-69). Albanian's stress falls on the penultimate syllable of suffixed words but on the final syllable when not suffixed (Newmark et al. 1982:16).

Finnish and other Finno-Ugric languages such as Hungarian are often cited as counterexamples to phonology-first accounts of case loss because they have actually been adding morphological cases over time despite an apparent fixed accent on the first syllable. Finnish has 14 case endings, whereas Hungarian may have 18 case endings although their proto language, Proto-Finno-Ugric, has nine case endings. However, Arnhold (2014) argues that Finnish has at most a very weak dynamic stress, and it is certainly not the kind of rhythmic stress found in Germanic and Romance languages. Instead, he finds phrase-level tones to be the most salient aspect of Finnish prosody. Vowel reduction is improbable as a result. Long and short vowels have not merged. Therefore, it has not undergone the developments that appear to be triggered by dynamic stress in the IE languages discussed above and retains phonemic vowel (and consonant) quantity. As seen above, this distinction is highly correlated with the fate of the case systems among the IE languages under discussion: languages that lose this distinction tend to lose cases, while those that retain it do not. Another difference is that Finnish is postpositional. Arnhold (2014) found that the postposition shares its pitch contour with the noun, making it more difficult for the speakers he tested to distinguish compound nouns from nouns followed by a postposition. In addition, vowel harmony extends through the entire NP or postpositional phrase. These facts establish a pathway for postpositions to be reanalyzed as case endings. This example illustrates a common means by which languages develop case endings. Since the IE languages under investigation all have prepositions but case suffixes, this opportunity for reanalysis does not exist; a system of case marking with a mix of prefixes and suffixes is strongly dispreferred by the principle of system adequacy (see Wurzel 1984:87, Meiser 1992:205-206). On the other hand, some phonological erosion of case endings has occurred, e.g., the shortening of the Allative from *-lle* to *-l* in spoken Finnish, but this and other endings still remain distinctive, and the tendency is not accelerated by the strong-weak alternation usually associated with dynamic stress (see Rapaport 2016:72).

3.1.1.2. Vowel Reduction

In pitch accented languages, no vowel reduction occurs. However, in dynamic stress languages, vowel reduction is a common occurrence. As established in the preceding section and chapter II, phonetic vowel reduction in various forms has played a critical role in the loss of case and gender distinctions across the Germanic, Romance, and some Balkan Sprachbund languages.

Pettersson and Wood (1987:261) distinguish two types of vowel reduction: phonological vowel reduction is the positional neutralization of two or more vowel phonemes. Phonetic vowel reduction is the different phonetic realization of unstressed vowels, usually involving centralization compared to their stressed equivalents. Both types of vowel reduction may include centralization, but the key distinction is that phonological reduction decreases the number of phonemic contrasts in that position, while phonetic reduction does not.

Phonemic vowel quantity can be seen in minimal pairs of vowels with the same quality but considered separate phonemes due to their difference in duration e.g. Attic Greek *lē.go* vs. *lĕ.go*, or Latin *mā.lus* vs. *mă.lus*. Through the process of lengthening accented short vowels and reducing unaccented long vowels, quantity adjustments often fused previously—different vowels are being perceived as more similar, until the distinction between a long vowel and an accented vowel became moot (Rapaport 2016:42). Duration was then mostly equivalent to stress. As a result, quantity neutralization left vowels prone to mergers (Allen, 1973). Because reducing vowels leads to vowel quantity ceasing to matter, it would be uncommon for a language to maintain both vowel reduction and vowel quantity concurrently for long. Once vowel reduction became prevalent, it soon prompts loss of phonemic vowel quantity. This can be seen in Early Koine Greek, Classical Latin, and the Middle Germanic languages (Rapaport 2016:43).

Phonetic reduction is the result of articulatory undershoot: the shorter duration of unstressed syllables makes it more difficult to hit articulatory targets. Sometimes phonetic reduction is subsequently phonologized, resulting in positional neutralization. Phonetic vowel reduction is a regular feature of languages with a dynamic stress with duration as a primary acoustic correlate (see Barnes 2006:29-30). Barnes observes that, In Romance and Slavic languages, vowel height is the most commonly neutralized feature in phonological reduction, followed by nasality and quantity (ibid., 20). While Barnes (2006) does not use a typologically unbiased sample, Kapatsinski et al. (2020:29, 31) use an unbiased sample of languages and find neutralization in the front-back dimension to be more common, then the height dimension is the next common reduction.

The prosody of Slavic languages may account, at least in part, for the different outcomes regarding vowel reduction (only found in East Slavic and Bulgarian) and case loss (which only occurred in Bulgarian and Macedonian). LPS had a pitch accent and phonemic vowel quantity for most vowels, but not in word-final position. However, its prosody had undergone numerous

changes from PIE, so these features were not inherited in the same sense as they were in AG, for example. Accented syllables had either a rising tone (acute) or a non-rising tone (circumflex). A new rising tone (neo-acute) subsequently arose due to the retraction of the accent from a weak *jer* to the preceding syllable. This tone differed from old acute, resulting in an unstable three-tone system, which was resolved in different ways depending on the individual Slavic language. In one group, including Czech, East Slavic, Bulgarian, and Macedonian, the old and new acutes merged, and the tonal contrast developed into one of vowel quantity, including on the final syllable: acutes became long, circumflexes short. In East Slavic, Bulgarian, and Macedonian, however, this quantity distinction was subsequently lost, sometimes influencing the position of the developing dynamic stress. In another group, including Slovak, Polish, and BCMS, the old acute and circumflex merged instead as a falling tone in opposition with the rising neo-acute. This contrast developed into one of quantity in the West Slavic languages of this group, as it did in Czech, while the accent was retracted by a syllable in some varieties of BCMS, which limited the freedom of the accent and strengthened its pitch-based character. A similar retraction in Macedonian may have led to the fixed stress system, as one of the non-final positions was generalized: antepenultimate in CSM, but penultimate in southern and southwestern dialects (Sussex & Cubberley 2011:39, 133, 151-154).

Thus, some Slavic languages have retained a pitch accent, while others have developed a dynamic stress. In the West Slavic languages, vowel quantity was strengthened as a phonemic feature, in part by extensive vowel contraction, while stress became fixed. Phonemic vowel quantity was lost in 16th century Polish and restricted in Slovak. Like Czech, however, these languages retain a relatively weak fixed stress. Primary stress is fixed on the initial syllable in most of these languages but on the penult in Polish. Polish prosody is not as different from the other West Slavic languages as this would suggest, however, because the penult is still initial for disyllabic words, and longer words have a secondary stress on the initial syllable, just as longer words in the other languages have a secondary stress on the penult. Unstressed vowels are not subject to reduction, and pitch is the main acoustic cue (Short 1993:538, Sussex & Cubberley 2011:152, 179, 189, Rothstein 1993:692). Thus, the prosody of the West Slavic languages resembles that of Finnish more than the Germanic languages. The fixed positions for both primary and secondary stress indicate that there is no strong-weak rhythmic alternation, explaining the lack of vowel reduction.

The dynamic stress is more energetic in the Slavic languages where it is free. Sussex and Cubberly point out that, of these, Ukrainian has the weakest stress, while CSR and Belarusian, the other two East Slavic languages, have the strongest; CSB, a South Slavic language, falls somewhere inbetween. Phonetic vowel quality differs between stressed and unstressed vowels in all of these languages, with more extreme reduction accompanying stronger stress. Energy and length are important acoustic cues (2011:179). Even in Ukrainian, phonetic vowel reduction occurs: unstressed vowels tend to be more centralized. Reduction is verging on phonemic neutralization for the vowels /ɪ/ and /ɛ/, which are both realized as [e] when unstressed (Buk et al. 2008:65, 68).

In terms of phonetic characteristics, CSB lexical accent is realized through a combination of intensity, pitch, duration, and vowel quality. Unstressed vowels are approximately 35% shorter than a corresponding stressed vowel in the same position (Tilkov & Bojadžiev 2013:152, 155). Historically, this difference in duration was most likely critical in the development of first phonetic and then phonological vowel reduction.

Vowel reduction can only contribute to the confusion and loss of nominal inflection if these are regularly unstressed. In CSB, approximately 8% of feminine nouns are stressed on the *-a* ending, e.g., *vodá* ‘water’. Approximately 10% of monosyllabic masculine nouns have stressed plural suffixes, e.g., *mъž* ‘man’, *mъžé* ‘men’, and only a few exceptional disyllabic nouns do, e.g., *vjátъr* ‘wind’, *vetrové* ‘winds’. The neuter plural suffix *-a* is regularly stressed, but the singular *-o* and *-e* endings are not, e.g., *dŭ́no* ‘bottom’, *dъná* ‘bottoms’ (Patseva 2017:10-11, 17-18). As with these surviving differences in stress between singular and plural suffixes, accent position in earlier Slavic varieties, as represented by OCS, sometimes varied among forms within a paradigm, even for the same number, so the plural is not entirely representative of all the earlier case endings in MB. However, the low frequency of stressed endings in CSB still suggests that the endings were largely unstressed during the period when case loss occurred. Patseva also notes that definite articles are most often unstressed: approximately 10% of monosyllabic masculine nouns take a stressed article in the singular, e.g., *mъž* ‘man’, *mъž-ŭ́t*, *mъž-á* ‘the man’. Feminine nouns with a final consonant, from the old IE *i*-stem declension, also take a stressed article in the singular, e.g., *mládost* ‘youth’, *mladost-tá* ‘the youth’. This is a relatively large group due to the productivity of the *-ost* suffix (ibid., 10). Given that the articles were demonstratives grammaticalized as clitics, I find it likely that the stressed feminine articles

were a later development. As for the masculine forms, all that remains is an epenthetic vowel (for *-a*), and sometimes the *-t* of the Proto-Slavic demonstrative **tъ*.[100] Thus, the original case endings on the articles were most likely never stressed, and those on the masculine forms were lost entirely.[101] Therefore, these were less likely to help preserve case distinctions than the preposed articles in Greek and German, and potentially even the postposed articles in Scandinavian.

CSR has not experienced significant case loss, even though dynamic stress and vowel reduction are stronger than in CSB. At first glance, this suggests that the retention or loss of case cannot be attributed primarily to prosodic factors as Rapaport (2016) argues. However, Shapiro (1986:185-189) argues that mobile stress patterns, in which the endings are stressed at least for all of the case forms of one number and sometimes all forms but those with a zero ending, are unmarked, while fixed stress patterns are marked. This is at least partially supported by frequency data: among the most frequent nouns, mobile stress is more common than fixed stress. However, when nouns with mobile stress become infrequent, they tend to have a fixed stress. It is safe to conclude, therefore, that in the development of Russian from Proto-Slavic, children have been consistently and frequently exposed to stressed case endings as they acquired the language, even as dynamic stress and vowel reduction arose. This does not seem to be the case for the development of Bulgarian. On the other hand, vowel reduction still seems to have had some effect on children's acquisition of gender in CSR. Smoczyńska (1985:644-648) found that Polish children have already acquired gender distinctions at the age of two. Popova (1973:269-271) found that it takes Russian children significantly longer to acquire a similar system. She attributes this delay in part to difficulty determining the declension class of nouns due to the neutralization of unstressed *-o* and *-a*, which comprise the Nom singular forms of the most common classes of neuter and feminine nouns and agreement targets. Corbett (1991:83) states the high number of hypocoristic male names, e.g., *Kolja* from *Nikolaj*, whose grammatical

[100] In written CSB, the two forms of the masculine singular article have a distribution that apparently corresponds to former case distinctions: the form *-ŭt* corresponds to the Nom, while the form *-a* (pronounced *-ŭ*) is used elsewhere (see Mladenova 2009:411-412). This distinction is not generally observed in speech, however: native speakers generally do not pronounce the final *-t*, and when they do, it is often in a non-Nom context, i.e., a hypercorrection. In addition to this synchronic evidence, there is strong historical evidence that it is an artificial literary feature introduced in the 19th century without a basis in any dialects extant at the time (Mladenova 2007:70-77). As a comparison, CSM does not have this distinction: the same forms (*-ot*, *-ta*, *-to*, *-te*) are used in all syntactic contexts.

[101] Issues concerning the relative chronology of case loss and the grammaticalization of definite articles and how these processes may have interacted are discussed in section 3.1.2.3 below.

gender is typically feminine, adds further difficulty. This delay in acquisition has not led to the loss of any gender category in CSR, but it appears to be contributing to the pressure on the neuter, as mentioned in section 3.1.2.6 below.

If vowel reduction resulting from dynamic stress contributed to case loss in the development of Bulgarian, the next question is how case loss occurred in Macedonian. As mentioned above, most Macedonian varieties, including CSM, have fixed non-final stress, either antepenultimate as in CSM or penultimate. The stress is weaker in CSM than CSB, with less difference between stressed and unstressed syllables in duration and other cues (see Sussex & Cubberley 2011:151, 179). Phonological vowel reduction does not occur in CSM. Except for the fact that the stress is on a different syllable, its prosody seems generally comparable to that of Polish, which Stieber has suggested may have stronger stress than Czech due to mobility across forms of a paradigm that differ in number of syllables (1969:65). Polish, of course, has retained a robust case system, while CSM has not. In both languages, nominal inflections are never stressed, and in CSM the articles cannot be stressed because they are not antepenultimate, but without vowel reduction this is unlikely to matter. Therefore, I support the proposal, discussed in section 3.2.1 below, that dialect contact allowed morphosyntactic changes, such as increased reliance on analytic alternatives in response to the loss of case distinctions, to spread even to those varieties where the phonological basis for those changes did not actually occur.

3.1.1.3. Ease of Articulation

Ease of production articulation is a way to produce the intended meaning with less effort. Van Trijp (2013:13, 17, 25) considers the changes to the High German definite articles in terms of usage-based factors. He argues that syncretism arose in a way that made articulation easier for the speaker, perception easier for the listener, and processing easier for both, while still allowing for the disambiguation of utterances. Because syncretic forms are more ambiguous, an apparent expectation is that they will require additional means than forms with an obvious form-meaning mapping. Speech is thought to represent a compromise between pronunciation economy and intelligibility on the other. That is, it is considered that speakers favor forms that involve the least amount of articulatory effort while being distinct enough from other forms in the language. Higher syncretism does not always imply greater ambiguity for the listener. Syncretism causes fewer paradigms that are easier to absorb, and the language user can properly rely on other clues.

Shrier (1965: 436) argues that the power of the dialect as a means of communication does not appear to be affected by the breakdown, merger, or reorganization of morphological components. Communication across dialect speakers is not impeded as long as there is a systematic structure underlying the current different inflectional endings. Any uncertainty between speakers of various dialects when a specific form denotes IO for one speaker but DO for another is cleared up by further conversation and context.

Predictable and frequent forms tend to be reduced or eliminated. Producing case-marking increases the quantity of signal a speaker must encode and articulate, increasing the speaker's production effort. Reduction, i.e., the employment of less strong signals, can have no negative implications and will be less likely to lead to negative responses in sceneries where the actual meaning is unambiguously interpretable. However, in other situations, speakers will expect to receive indirect negative feedback, e.g., a listener's expression of misunderstanding or an inability to respond in accordance with the underlying meaning. This feedback is thought to have an impact on following productions (Kurumada & Jaeger 2015:3, 4). There is an assumption that the emergence of reduction effects is associated with word frequency. It is faster to retrieve more frequent words to avoid high activation levels. There is a link between word frequency and contextual predictability. That is, because high frequency words occur more regularly in everyday speech than low frequency words, any occurrence of a high frequency word is predictable in the usual context than a low frequency word (Pierrehumbert 2001:5). There is evidence that forms that are frequently high are likely to be phonetically shortened (Sumner et al. 2013). Since longer forms are less frequent, they tend to be a burden on short memory retrieval which violate the economy principle, and thus, they are replaced by frequent shorter ones (Ramat 1992:556-557, Jendraschek 2009:16-17). Economy is a crucial principle to be undertaken by language users to facilitate efficient communication.

With a series of experiments in examining the optional case marking of Nom and Acc in Japanese, Kurumada and Jaeger (2015:3, 12) found that even when the animacy of both arguments was held constant, there was a tendency to induce case-marking for unexpected grammatical function. Furthermore, their experiments show that speakers' decision to yield case-marking is influenced by the general validity of the planned grammatical function. In languages where case-marking is optional, semantic features of arguments, such as animacy, have an impact on the assignment of case-marking, especially when both subject-object-verb and object-

subject-verb ordering are acceptable in Japanese. With a similar study on Korean speakers, Lee (2006) found that speakers were more likely to produce case-marking on animate arguments of Nom and Acc case markings. This further supports the importance of animacy in deciphering the grammatical roles to make the utterance clear for the listeners, even when that requires increasing the speaker's effort in producing case marking.

3.1.1.4. Orthographic Substitutions

In many studies including this one, orthographic substitutions are taken as evidence for the neutralization of case endings due to sound change and subsequent case mergers. However, other interpretations of these substitutions are possible. Lunt (1965:305-312) argues that orthographic substitutions in BChS manuscripts should not necessarily be interpreted as the result of case loss or, in some instances, even vowel mergers. Scribal errors can have many causes. As a comparison, the East Slavs made errors when copying liturgical poetry even though their case system survived, which suggests that similar errors in BChS liturgical poetry might have other causes as well. However, many of the apparent errors in BChS manuscripts are not in fact similar: they are often systematic, in contrast to the sporadic substitutions in the East Slavic manuscripts. Thus, Lunt is skeptical about claims of case loss in 14th-century MB, even to a three-case system.

This argument brings into questions the reconstructions used in the present analysis, but it is still possible that such a three-case system occurred later in the MB period. If it is true that no case loss occurred until after the 14th century, this could mean it occurred during the Ottoman period, after intense contact with the other Balkan Sprachbund languages began, including but not limited to Turkish.[102] Lunt (1965) does not directly argue that later case loss was motivated by contact, but this conclusion can be drawn from a later date for proposed sound changes and neutralization of case endings. In the years since this article was published, however, more careful investigation of manuscripts has been done. For example, Vakareliyska (2008) bases her arguments about orthographic substitutions in the Curzon Gospel, a 14th-century BChS manuscript, on this and two other BChS manuscripts with a shared antigraph and still finds support for the loss of case distinctions in a particular dialect underlying one of them, due to

[102] The sociolinguistic situation in the Balkan Sprachbund during Ottoman rule is discussed further in sections 3.2.2 and 3.2.3 below.

sound changes. Based on Vakareliyska (2008) and other evidence, the present analysis continues to assert the reduction to a system of three morphological cases in MB and the importance of early sound changes in this process, while acknowledging that phonological change along cannot account for all of the case loss in Bulgarian and Macedonian. At the same time, it should be acknowledged that any particular sporadic example of a non-canonical case use in a BChS manuscript may be a scribal error that occurred for reasons other than neutralization of case endings in the scribe's dialect.

3.1.2. Morphosyntactic and Semantic Motivations

The motivations considered in this section relate to the complex interplay among morphology, syntax, and semantics in changes to nominal inflection. A major unresolved issue concerns which of these domains is the starting point for these changes. Through a Granger-causality analysis of a corpus of Icelandic texts from 1150 to 2008, Moscoso del Prado (2014) demonstrates that morphological changes triggered syntactic changes in the history of the language. This supports previous proposals (e.g., Lightfoot 2002) that grammatical change starts in words and spreads to the rest of the grammar. The first two subsections below, on paradigm organization and the merging of declension classes, consider arguments that generally hold to this view. The remaining subsections, however, consider arguments that generally oppose this view, although some are still compatible with form-driven change. While they may eventually be proven incorrect, or at least incomplete, they are too numerous to ignore. The final four subsections discuss motivations that involve different aspects or stages in several related paths of development: case variation, functional mergers, functional narrowing, and the replacement of synthetic structures by analytical structures, The ease of production is also discussed.

3.1.2.1. Paradigm Organization

The paradigm is a key aspect of nominal inflection. Changes in the number of distinctive case and gender categories necessarily interact with paradigm organization, likely in both directions. This section considers several factors that play a role in paradigm organization, including markedness, relevance, analogy, and syncretism.

Tiersma (1982:832) explores the concept of morphological markedness, which he bases on the work of Jakobson (1932, 1939) and Greenberg (1966:19). Tiersma's (1982:832)

application of the term ‘markedness’ refers to categories that specifically mark for particular features, contrasting themselves with an ‘unmarked’ member of the same category. Markedness plays an important role in morphological changes. Singular is considered basic, so it is unmarked in relation to the plural. In many IE languages, the existence of a marker in the plural signifies that this category is marked in opposition to the unmarked one, the singular.

Bybee (1985) considers several other factors in morphological expression. One of her major hypotheses is that a morpheme’s degree of fusion is correlated with its degree of semantic relevance, i.e., the extent to which the content of one semantic element modifies another. She predicts highly relevant categories to be expressed lexically or inflectionally most of the time, while irrelevant categories are only expressed syntactically. Relevance is dependent on cognitive and cultural salience, and the expression of categories varies systematically by language. Generality is also a factor in the type of morphological expression: inflectional categories must be general enough to apply to all words in a lexical category (ibid., 4, 13-14, 17). Bybee argues for this hypothesis based on a cross-linguistic survey of verbal categories (ibid., 20-47). When applied to the expression of grammatical relations, case inflection is only expected when the meaning is sufficiently relevant but still general. If the meaning becomes too specific, inflected forms may be fossilized, e.g., as adverbs. If the meaning loses relevance, affixes may be reanalyzed as clitics or replaced by periphrastic constructions. All of these changes to case inflection are attested in the development of Germanic, Romance, and Balkan Sprachbund languages, and my study will consider their connection to relevance.

One indicator of relevance is proximity to the root. Greenberg (1963:112) found that number is almost always expressed between the noun base and case expression, if both categories are expressed on the same side of the base. This order of morphemes indicates that number is more relevant to nouns than case, which Bybee also explains in terms of the effects of these categories: number has a direct effect on the referent of the noun, while case only affects the relations of the referent to other elements in the clause. In IE languages, case and number are usually fused together into a single portmanteau expression, so there is no difference in their order. However, Bybee considers a morpheme expressed by stem modification to be closer to the base than an affix (1985:34). This correctly predicts that umlaut is more often morphologized to express number than case in the Germanic languages. It is also a potential motivation for the leveling of the Nom-Acc distinction in the singular of OF nouns, where it was expressed through

a stem alternation, as described in section 2.4.1.3 above. In many instances, only the Acc form, which shared its stem with the plural forms, survived, e.g., CSF singular *baron* 'baron' < Acc singular *baron*, plural *barons* < Acc plural *barons* (cf. OF Nom singular *ber*). In others, only the Nom form survived, with the plural by analogy, e.g., CSF singular *sœur* 'sister' < Nom singular *suer*, plural *sœurs* (cf. OF Acc singular *seror*, Nom/Acc plural *serors*) (see Rickard 2003:49-50). The latter group all denote people and occur frequently as agents; this likely explains why the Nom form was adopted for these nouns (see Meiser 1992:207-208). When both forms survive, they have either become different lexemes or the alternation only expresses a number distinction, e.g., CSF singular *œil* 'eye' < Acc singular/Nom plural *ueil*, plural *yeux* < Acc plural/Nom singular *ialz*. Thus, these OF stem alternations can only survive as a morphological expression for number, the most relevant category. Order of acquisition might be another indicator of relevance. Gender agreement is often acquired later than number agreement, e.g., in Hebrew (Berman 1985:273) and French (Clark 1985:699). However, Hooper (1980:176-177) found the opposite order in Portuguese and Latvian.

An important aspect in the organization of paradigms is the relation between the basic form and the forms derived from it. The basic form can be determined from child language and historical data: when children start to acquire a paradigm, they initially use only the basic form, and any overregularization errors involve forms built from the basic form; historically, analogical changes usually involve derived forms changing to become more similar to basic forms (Bybee 1985:50-51). This direction of change corresponds to Mańczak's (1980:284-285) argument that the most frequent forms in discourse tend to remain unchanged and cause other forms to change. Zero expression, i.e., the use of an uninflected form, is another indicator of the basic member of a paradigm. This observation can be traced back to Jakobson (1939), who noted a significant correlation between the use of uninflected forms and semantically unmarked members of a category. Zero expression arises diachronically when markers are only used for the marked member of a category, but uninflected forms are also actively created by speakers during restructuring (Bybee 1985:54-55). For example, a new preterite marker developed from the third singular marker in Old Provençal through reanalysis of this basic form as uninflected; the other person/number markers were then added to the third singular form instead of the original base (Bybee & Brewer 1980:210). This type of phenomenon originates in the acquisition process: if

the learner never recognized the narrower function of the basic form, it is assumed to be an uninflected form and other forms may be restructured accordingly (Bybee 1985:56).

Bybee and Brewer (1980:214-217, 222) connect the choice of basic form to autonomy, the probability that a word has its own lexical representation. A word is more autonomous if it represents fewer semantic categories, has high frequency, and/or is morphophonemically irregular. Less autonomous forms are derived from more autonomous basic forms, but frequent paradigms can contain multiple autonomous forms, as in suppletive paradigms. These criteria may explain why pronouns, with their high frequency and irregular, suppletive paradigms, resist case loss: they do not rely on the productivity of case marking because their forms are all autonomous.

Kürschner and Nübling (2011:357, 359) argue that gender and declension class are often linked in Germanic languages to profile number, which is a highly relevant category for nouns. In these languages, both gender and declension class systems are an inherent lexical feature of nouns. They consider gender to function as a means to track references and, at least in CSHG, in NP framing constructions, as the most reliable feature in an NP due to its invariability. Dye et al. (2017:7-12) find that gender in CSHG also reduces nouns' entropy, i.e., the uncertainty about which word will occur in a particular context, redistributing it so that entropy remains more constant over discourse. This is consistent with Zipf's (1949) Principle of Least Effort, i.e., that humans tend to behave in ways that minimize the average rate of work. Kürschner and Nübling (2011:357-360) note that the function of declension classes is less clear. In PGmc, gender and declension class are each partially predictable based on the other. In the later Germanic languages, these two systems are linked in various relationships, often with one improving the memorability of the other. If one loses functionality, however, the usefulness in memorization is lost, so the link functions to profile number instead.

Kürschner and Nübling (2011:361-366) point out that a trend in the history of German and other Germanic languages has been the combination of declension class markers with markers of number and case, which serve as host categories. In this way, class markers have become hidden but also more stable, since they cannot be lost unless the host grammatical categories are lost. Further developments usually involve the association of class markers with more relevant categories, i.e., from case to number. Class markers tend to disappear from the unmarked values of these host categories, however, which explains why both the old class

markers and umlaut mostly survive only in the plural. Thus, number is profiled while case marking is reduced. Gender also profiles number, but in a different way. In CSHG, for example, this has been achieved by restricting gender distinctions to the singular. Thus, gender and declension class are complementary in their profiling of number.

Kastovsky (2000:720-722) makes a similar proposal with respect to the loss of case and gender in English, arguing that distinctions in number were prioritized over case, and even more so gender, as the system of stem classes was reduced to a single default system. Based on these findings, the concept of number profiling is at the core of my quantitative analysis in chapter IV. However, the specifics of how analogical change brings about this profiling are still lacking in the literature on Germanic languages. Corbett attributes the decline of gender systems primarily to the attrition of formal markers (1991:315).

Based on an investigation of gender and declension in a sample of standard and dialectal Germanic varieties, Kürschner and Nübling also propose a typology of gender-declension relationships. One type of relationship is the total association of gender and declension in a one-to-one correspondence, as in NNw and some Alsatian dialects of German. This only occurs if all three genders are maintained because this system is more complex and therefore benefits the most from the memory reduction gained by linking the systems. At the other end of the typology is the total dissociation of these systems, as in CSDan, CSDu, and ModLG in East Friesland. This occurs when the number of genders and declensions are reduced and both are conditioned by transparent features like semantics, derivational affixes, or phonological shape. However, these systems still profile number marking when unlinked. In between these extremes are two more complex types that nevertheless improve memorability and serve to profile number. marking. In languages like CSS and the Swiss German dialect of Fribourg, there are a high number of declensions, but each is associated with only one gender. Languages such as CSHG, Faroese, Icelandic, Luxembourgish, and Yiddish have higher complexity and retain the partial predictability found in PGmc (2011:377-381). This typology is useful in explaining some of the later changes to gender systems.

In English, gender distinctions only survive on personal pronouns, for which grammatical gender has been replaced by natural gender assigned purely on a semantic basis (De Vos & De Vogalaer 2011:245). The number of grammatical genders in CSDu has been reduced to two, with a common gender category as in CSS and CSD (Kürschner & Nübling 2011:375). As in other

Germanic languages with two grammatical genders, pronouns still retain three genders, but this system is undergoing a semanticization process based on individuation: traditionally masculine pronouns are used for highly individuated nouns, e.g., concrete count nouns, while the traditionally neuter pronoun is used for low-individuation nouns, e.g., abstract mass nouns (Audring 2006). De Vos and De Vogelaer investigate the semanticization of gender in Southern Dutch, where the process is not as advanced as in the Standard, northern variety. Southern varieties have retained three grammatical genders, but in the dialect under investigation, Moerzeke Dutch, the masculine is only consistently distinguished from the feminine on pronouns and the indefinite article; the masculine forms of adjectives and the definite article are distinguished from the feminine by an *-n*, but only in certain phonological environments that apply about 40% of the time. The authors propose that a system of three grammatical genders relies on gender marking within the NP; when these distinctions become too opaque, semanticization occurs. They support this hypothesis with data from child language acquisition research and an apparent time investigation of three age groups of adults. The results indicate that natural gender has become more relevant for animate nouns, and many deviations from grammatical gender are semantically motivated for inanimate nouns. Since other deviations are not, however, the breakdown in grammatical gender must have begun first and motivated the later semanticization. Combined with decreased exposure to dialectal forms due to a tendency for parents and caretakers to suppress them in child-directed speech, the child language data suggests that grammatical gender cannot be fully or successfully acquired due to a lack of exposure. The authors consider these results to support the idea that semanticization is a form of morphological regularization (2011:247-249, 255-256, 264-265). A similar semanticization in terms of individuation has occurred in western Danish dialects; the common and neuter forms of determiners and pronouns are now used with count and mass nouns, respectively. Notably, these dialects have a single preposed definite article without gender distinctions, like English, so gender marking within the NP (see Haugen 1976:371, Haberland 1994:323-325).

Syncretism has been an important factor in the development of nominal inflection in Germanic, Romance, and Balkan Sprachbund languages. As Baerman et al. (2005) point out, the introduction of the term syncretism to linguistics is generally credited to Pott (1836), who used it to refer to the diachronic neutralization of inflectional distinctions through formal or functional mergers. Baerman et al. (2005:4) explain that Hjelmslev (1935-7:60) and Jakobson (1936:67)

introduced a synchronic perspective of syncretism within the structuralist framework. Focusing on the morphology-syntax interface, Baerman et al. (2005:2-4, 111) view syncretism as a disparity between syntax and morphology, or the failure to mark a distinction that is morphosyntactically relevant. Interactions between features are an important aspect of syncretism. For example, Nom and Acc are syncretic in the plural but not singular in many Germanic declension classes. Brown (1998:190) analyzes such a relationship as a feature dependency, i.e., a number feature can condition the presence or absence of case distinctions, but not vice versa. Based on their cross-linguistic study of syncretism, including many non-IE languages, Baerman et al. propose the following hierarchy: gender syncretism > case syncretism > number syncretism. This means that as long as a language has all of these morphological categories, it will not have number syncretism without case syncretism or case syncretism without gender syncretism. The authors note that this hierarchy should be restricted to inflectional features, so the generalization about gender syncretism would apply to agreement targets but not nouns. In addition, there are counter-examples to these generalizations (2005:7, 113-114). Nevertheless, this hierarchy captures important generalizations that apply in the development of the Germanic languages, and it corresponds to the relevance hierarchy proposed by Bybee (1985). This also supports Bybee and Brewer's (1980:225) observation that more semantically similar forms will also have more similarity in forms, resulting in paradigm leveling.

Some patterns of syncretism should be noted. One pattern is uninflectedness, in which some lexical classes lack formal differences in terms of case, gender, and number. This causes an inability to relate syntactic contents. Uninflectedness is only unique in a specific set in lexicon as other lexical classes still show inflections (Baerman et al. 2005:30). Another is neutralization which occurs when feature values are irrelevant to distinguish syntactic context and this general loss is found everywhere in the language. Lack of syntactic significance, or the complete absence of distinctions for one set of syntactic objects in a particular syntactic context, is not enough, because other syntactic objects may nevertheless recognize those distinctions (ibid., 30).

Baerman et al. focus their investigation on systematic syncretism, excluding obvious instances of homophony that arises accidentally through sound change. However, they also note that accidental homophony can be reanalyzed as morphological. If this occurs, the pattern can be extended from one stem class to others, possibly all, just as with syncretism that is a cross-

linguistically common characteristic of feature structure. At the same time, sound-change patterns unique to a particular language must compete with other more cross-linguistically generalized patterns in the language so the former are less common (2005:9-10, 169-170). One possible consequence of this distinction on Germanic, Romance, and Balkan Sprachbund nominal inflection is that new instances of syncretism due to sound change are more likely to be leveled by analogy than well-established patterns of syncretism. Conversely, sound changes that happen to extend an established pattern or bring about a new pattern in a sufficient number of declension classes are more likely to be reanalyzed as a systematic morphological pattern than those that do not.

Baerman et al. introduce several types of patterns found when syncretism occurs. Two are particularly relevant for the present investigation. The first is simple syncretism, in which two or more cells have the same form but different values of the same feature. The second, the polarity effect, is when the same form is found for entries that are seemingly unrelated, with no common semantic features (2005:13-17). A common form for Nom and Acc singular is an example of simple syncretism, as is a common form for Nom singular and Nom plural. The first example differs only in case, while the second differs only in number. In contrast, a common form for Gen singular and Nom plural is a polarity effect. The shaded cells in the following table are an example of number syncretism:

Table 47. Koine Greek *i*-Stems Feminine Noun Declension

	Sg.	Pl.
Nom	-is	-is
Voc	-i	-is
Acc	-in	-is
Gen	-eos	-eon
Dat	-i	-si, -sin

Baerman et al. note that polarity effects are much less likely to be morphologically and semantically systematic than simple syncretism, and their representation shows problems for many morphological models of syncretism (ibid., 104-111). If these issues of representation and systematicity are held to be indicative of the cognitive reality of speakers, this may explain why polarity effects are historically less stable than simple syncretism in the Germanic, Romance, and Balkan Sprachbund languages.

Gender syncretism is very common cross-linguistically, particularly in languages with gender marking that differs morphologically between the singular and non-singular numbers (Baerman et al. 2005:82). Almost all IE languages with grammatical gender fall in this category;[103] as expected, therefore, a number of the languages under investigation have gender syncretism. Corbett (1991:7) explains that the key point of examining gender is agreement. He distinguishes between controller gender, i.e., the categories of gender that the nouns contain. and target gender, i.e., the marking of gender on adjectives, verbs, and so on. Both can coincide, but they do not match in some languages. Corbett (2013) examines 257 languages and found that that of the languages, there are 145 ungendered languages. The minimum number of genders, if exists, is two, and this number of genders is common across languages. The less common number of genders is three. The even less common number of genders is four. A language can have five or more genders, but this system occurs in 24 languages. (Corbett 1991:161, 167) points out that some languages can have subgenders in their gender system. For example, Russian has three genders, and each gender is further divided into animate and inanimate.

Corbett discusses two patterns of gender syncretism across number: convergent systems, as in CSHG, have fewer gender distinctions in the plural, while crossed systems, as in CSRm and CSA, have different groupings of genders in the singular and plural (1991:155-156). Creating a new gender category occurred in Romanian, Italian and Albanian. The morphological marking of the novel gender is not completely new. As Corbett (1991:313) points out, new gender category can arise from available morphological criteria existed in the language. Nouns that take different agreement in the singular and plural are sometimes also labeled *ambigeneric*, which implies a combination of categories, but Baerman et al. consider these to have a distinct gender (2005:82-83). As shown in the following table, in Romanian the ambigeneric gender category uses the masculine singular and feminine plural:

[103] Certain Western Romance languages could be considered exceptions. In CSF, *-e* marks the feminine in the singular and plural of adjectives, although this form is phonetically realized in different ways depending on the word. However, determiners still have gender syncretism in the plural. In Spanish, singular and plural forms of some nouns and agreement targets both contain *o* for masculine and *a* for feminine. However, other nouns and agreement targets lack these markers; they either do not have distinctive gender marking or have a masculine singular form without *o*, e.g., the definite article, with masculine singular *el*.

Table 48. Contemporary Standard Romanian Gender System (Simplified)

	Sg.	ambigeneric	Pl.
Masc.	-∅		-i
Fem.	-ă		-e

Greenberg's Universal 37 states: 'A language never has more gender categories in nonsingular numbers than in the singular' (1963:112). For this statement to be meaningful, *gender categories* must be interpreted as those that are distinguished on agreement targets (see Corbett 1991:156). However, the term *gender categories* is normally used in the present study to refer to those on nouns.[104] Universal 37 is consistent with the unmarked status of the singular. Masculine and neuter *o*-stems, the most common classes of these genders for nouns and agreement targets, were already mostly syncretic in the singular in CL and likely merged completely in the singular in the development of Romanian. However, the neuter plural was still distinct from the masculine and feminine. This system would have had more gender distinctions in the plural, violating Universal 37. By adopting feminine agreement in the plural, however, probably influenced by the resemblance between the neuter plural and the feminine singular, three gender categories were retained without violating the universal. As discussed further in section 3.2.3 below, contact may explain why Romanian conformed to the universal in this way and not through the loss of the neuter category, as in the Western Romance languages.

This subsection has focused on the internal organization of paradigms but has also considered the external organization of paradigms, particularly with respect to how they correspond to gender categories. The following subsection explores the interaction among different paradigms further.

3.1.2.2. Merging of Declension Classes

The merging of declension classes is a development that has often accompanied case and gender loss in the Germanic, Romance, and Balkan Sprachbund languages; it has also occurred in IE languages that have not undergone significant case or gender loss, e.g., CSR. Like the loss of

[104] The example of the neuter in CSRm highlights this distinction in terminology: on agreement targets it is marked by a combination of the masculine and feminine categories, so it is ambigeneric, but it is considered its own category on nouns because of the unique combination of masculine agreement in the singular and feminine in the plural.

case and gender categories, a reduction of declension classes can be considered morphological simplification. Some ways these processes may be connected are discussed below.

A connectionist network is one potential model for the acquisition of declension classes. Hare et al. (1995) implemented two connectionist networks to model the acquisition of English past tense inflection, specifically six inflection classes based on those in OE: five strong classes with ablaut, each associated with a particular rime pattern, and the weak class with the *-ed* suffix. The first model was tasked with learning to categorize nonce verbs into these six classes, while the second, more complex model output the actual past tense forms. In both models, the training input was evenly distributed between the six classes. Then their ability to generalize from the training data was tested with additional verbs. Even though the weak class was much smaller than in OE, where it accounted for around 75% of verbs, the models still learned this class as the default category. Based on these results, Hare et al. conclude that the critical factor in the development of a default class is not its size but the structure of the non-default classes. A default category can arise as long as there are well-defined non-default categories and/or sufficient variation in default category members across the rest of the input space. This input space is defined by the relevant dimensions for the classification of forms; these can be phonological, as in these models, but also morphological, syntactic, or semantic. These two conditions can be satisfied to varying degrees in natural languages, resulting in a continuum of default-like behavior (1995:24-26).

In CL, PGmc, and Proto-Slavic, default declension classes were already developing. In CL, *o*-stems were the default for masculine and neuter nouns, *ā*-stems for feminine nouns (see Alkire & Rosen 2010:186-187). The same is true in Proto-Slavic (see Lunt 2001:53). PGmc *a*-stems were much larger than other classes, with more masculine than neuters, and *ō*-stems were the largest feminine class (see Ringe 2006:269). Over time, these classes expanded, absorbing smaller, less productive classes and further increasing their default-like behavior. This can be partly attributed to sound changes that obscured the distinctions among declension classes for certain forms. Already in CL, *u*-stems shared two of the most frequent forms with *o*-stems: Nom singular *-us* and Acc singular *-um*. Thus, it is not surprising that *u*-stems were subsumed by *o*-stems in VL (see Alkire & Rosen 2010:185-186). The same merger occurred in Slavic languages, sometimes with the adoption of certain more distinctive *u*-stem endings such as Gen plural *-ov*, as discussed in sections 2.5 above and 3.1.2.4 below. Forms such as these that survive despite

originally belonging to a small class have been described as overstable markers. Based on the principle of distinctive strength, these markers often spread to new classes and are more likely to be retained in a functional merger (see Wurzel 1984:139, 209, Meiser 1992:206-207). This principle explains how overstable markers can come from smaller classes as well as larger, default classes. Sometimes the merger of declensions might begin with or be accelerated by the extension of one or more overstable markers.

Gender mergers have also played an important role in these Germanic, Slavic, and VL. Declension classes composed of nouns primarily of the same gender are much more likely to merge than those that are not. Even without much formal similarity, shared gender can allow a minor class to be absorbed, as with the CL *ē*-stems. Many of these became *a*-stems in VL despite no overlapping forms, e.g., Nom singular *facies* 'face' > *facia* (see Alkire & Rosen 2010:187). At the same time, the loss of gender distinctions appears to be facilitated by formal similarity, as with the loss of the neuter beginning in WVL. The default *o*-stem class only differed in a few endings between masculine and neuter, particularly in varieties with the loss of final *-s*, and the remaining distinctions were eventually leveled or fossilized in the Western Romance languages, as described in section 2.3.3 above. In MSw, MDan, and MDu, similar forms between masculine and feminine contributed to a full merger of these two genders into a common gender, as discussed in sections 2.3.2.2 and 2.3.3.3 above.

In the old Germanic languages, sound changes eliminated many of the distinctions among the declension classes, particularly in the singular. In the transition from OE to ME, masculine *a*-stems expanded as the default, and by LME this was the only remaining productive class and grammatical gender had been lost. The *-(e)s* plural was originally associated with this class, and the *-(e)s* Gen singular was associated with *a*-stems generally. Jespersen (1894:170) notes that the *-(e)s* plural had already spread to other strong masculine classes in OE, i.e., *i*-stems and *u*-stems, and that the spread of the *-(e)s* Gen was related but actually more rapid. In a corpus linguistics study of ME texts, Newman (1999) found that the *-(e)s* plural spread more quickly in the Northern and Midland dialects than in the South and Southeast, where it was not complete until the end of the 14th century. The neuter *a*-stems were generally among the first to adopt this ending, as a result of already sharing the remaining forms. Feminine strong classes usually followed, in slightly different orders depending on the dialect. Weak nouns were among the last to adopt the *-(e)s* plural. In fact, in the South and especially the Southeast, the weak plural *-en*

even spread to some strong nouns, rivaling the *-(e)s* plural for default status, e.g., in EME texts such as Layamon's Brut (see Burrow & Turville-Petre 2007:21-22, 97).

For *a*-stems, a common plural form meant that the masculine and neuter paradigms were no longer distinguishable. Along with the spread of the *-(e)s* Gen singular, feminine strong nouns and weak nouns of all genders also became more or less indistinguishable from *a*-stems, resulting in the modern system with a single default class and a few unproductive exceptions. The shift to semantic gender assignment for personal pronouns began in LOE, before the distinctions between declension classes had been completely lost (see Baron 1971:118-119). It is unclear what kind of causal relationship existed between the loss of gender distinctions on agreement targets, the semanticization of pronominal gender agreement, and the expansion of the more frequent declension class. However, given that grammatical gender must be marked overtly on agreement targets but not nouns, I propose that the loss of gender agreement was more likely a significant factor in the expansion of the default *-(e)s* forms than vice versa. This is supported by the chronology of semanticization and the changes to declension, along with the loss of NP-internal gender agreement motivating a similar semanticization in Moerzeke Dutch, as proposed by De Vos & De Vogalaer (2011).

The extension of *-(e)s* as the default Gen marker likely enabled its reanalysis as a clitic (or phrasal affix, according to Nevis (1985), Allen (2008:183), and others). One of the criteria that Zwicky and Pullum (1983) propose to distinguish clitics from inflectional affixes concerns host sensitivity: affixes but not clitics are limited to specific syntactic categories, e.g., nouns, and to subclasses within the category, e.g., nominal declension classes. Thus, as argued by Carstairs (1987:153-155), the spread of *-(e)s* to all nouns was a necessary step in its reanalysis as a clitic. This reanalysis resulted in the loss of what was probably the last morphological case marking in EME. Thus, the merging of declension classes is closely connected to the loss of both grammatical gender and morphological case, at least in the development of English.

3.1.2.3. The Grammaticalization of Definite Articles

The grammaticalization of demonstratives into definite articles adds the additional morphological distinction of definiteness. This development of definite articles from demonstratives is common cross-linguistically (Dahl 2009:243, Heine & Kuteva 2004:4, among others). The development and presence of overt definiteness appears to have a connection with

case marking and the loss thereof. When articles are consistently suffixed to nouns, as in the North Germanic languages, they can also be considered additional inflections.

The suffixed definite article is found in almost all Scandinavian varieties, in which it was grammaticalized from the determiner *(h)inn* 'that, the'. The standard view holds that this occurred via constructions in which this determiner linked an adjective to a preceding noun, e.g., OSw *maþr inn gamli* 'the old man' (Delsing 2002:938).[105] An alternative hypothesis is that demonstrative uses of *(h)inn* were sufficient to allow for its grammaticalization as a definite article, since demonstratives were often postnominal (see, for example, Nygaard 1906:33ff.).[106] However, it has been lost in the contemporary Danish dialects spoken in Western and Southern Jutland for all case markings except as a vestige in the Gen form. Instead, these dialects use a prenominal article, e.g., *æ hus,* as opposed to CSD *huset* 'the house' (Delsing 2002:938). Thus, the article in these dialects more closely resembles the prenominal article that is characteristic of the West Germanic languages, and in fact may be influenced by the close proximity of these dialects to West Germanic varieties. As for the history of the Scandinavian articles, in ODan texts from Jutland the suffixed definite article was already unusual in non-Gen forms. The suffixed definite article was uncommon in ODan and OSw runic inscriptions, as well as in ON poetry (Delsing 2002:338, citing Hansen 1927:119-150), which Delsing (2002:938-939) attributes to its late development, around the 13th century, rather than to the style of surviving texts.

Like the West Germanic languages, the Western Romance languages all have prenominal articles. Alkire and Rosen discuss the grammaticalization of the definite article from the demonstrative paradigm *ille* 'that' in adjectival uses beginning in early VL. They propose a possible intermediate step in which forms of *ille* meaning 'the one(s)' were followed by a modifier. Possible instances are attested as early as Plautus (3rd to 2nd century BC), e.g., *mendicus atque ille opulentissimus* 'the beggar and the very wealthy man'. The 4th century

[105] This type of construction has close parallels to constructions in some of the Balkan Sprachbund languages, i.e., the repetition of the definite article between a noun and a following adjective in Greek, and the use of linking articles between nouns and adjectives in Romanian and Albanian. See section 3.2.2 for a more detailed discussion of these Balkan Sprachbund constructions.

[106] The indefinite article forms in CSS, i.e., common gender *en*, neuter *ett*, and CSD, i.e., common *en*, neuter *et*, resemble the definite article, i.e., CSS *-(e)n*, *-(e)t* and CSD *-en*, *-et* (Andersson 1994:280-283, Haberland 1994:323-326). However, the indefinite article forms are also used as the numeral 'one', from which they are derived. This is more obvious in Faroese, which has *ein* 'one, a' for the masculine Nom/Acc, and in Neo-Norwegian, which has masculine *en*, feminine *ei*, neuter *eit* (Barnes & Weyhe 1994:203, Askedal 1994:237).

travel narrative Peregrinatio Aetheriae, written in an informal colloquial style, contains far more instances of the demonstratives *ille* and *ipse* than expected in CL, and context usually suggests 'the' as a better translation for these than 'that', e.g., *illos sanctos monachos, qui ibi manebant* 'the holy monks who resided there'. By the 8th century, *ille* had been completely grammaticalized into a definite article, as in Regula Chrodegangi, a text of monastic rules, e.g., *illi seniores illas cappas quas reddere debent non commutent* 'the elders are not to sell the cloaks which they have to turn in' (2010:203-205). Grandgent (1907:164) places the beginning of the free use of *ille* and *ipse* as definite articles in the 4th century. The latter lost out to the former, except in the varieties of Sardinia, the Balearic Islands, and parts of the nearby Spanish and French coasts.

A characteristic feature of the Balkan Sprachbund languages is a postpositive definite article, as in CSB, CSM, Romanian, and Albanian, although Greek has a preposed article (Tomić 2006:59-60, 91-93, 127-128, 187, 205).[107] In all of these languages, the definite article, in its various gender and case forms, was grammaticalized from demonstratives. In fact, with the exception of Romanian, in which the article was grammaticalized from *ille* as in the Western Romance languages, and some forms in Albanian, the same IE demonstrative paradigm masculine Nom singular **so* 'that' was the source of the article. The OCS demonstratives *sь* 'this' and *tъ* 'that' occurred postpositively in BChS. However, in CSB and CSM, *tъ* became the definite article rather than *sь*, which died out in Slavic languages (see Sussex & Cubberley 2011:270-271, Vakareliyska and Gyllin in press:111, Wahlström 2015:44). The Greek determiner *ho* 'that, the' generally still had a demonstrative force in MyG and Homeric Greek, but by the time the Attic dialect of AG was attested, it had been grammaticalized into a definite article (see Silher 1995:388-389). Orel notes that the IE demonstrative paradigm **is* was the source for the definite masculine and feminine Nom singular forms in Albanian. The Gen-Dat and Abl plural forms are taken from the indefinite declension. The remaining forms derive from IE **so*, although phonological changes have obscured the demonstrative endings. For example, the Nom plural and masculine Gen-Dat/Abl singular form *-t* reflects only the initial **t* of the corresponding forms of **so* (2000:246-247).

[107] Turkish does not have a definite article derived from a demonstrative like the other Balkan Sprachbund languages, but it makes a distinction in definiteness for DOs: the Acc suffix is only used for definite objects; otherwise, the Absolute form is used (Lewis 2000:34).

This Balkan feature most likely developed and spread through contact among the Balkan Sprachbund languages, with Romanian and Albanian as the most likely sources (Bubenik 2006:198-199). Since the grammaticalization process was already well underway by the Romanization of Dacia in the 2nd century CE, the Romanian postpositive articles share the same early stages of development from *ille* as the preposed articles in the Western Romance languages. It is not clear, however, how it came to be postposed in Romanian. Lindstedt (2014:176-177) argues that the postposed definite article, characteristic of the Balkan Sprachbund, spread through contact among these languages, with Albanian as the source for the placement after the noun. Toponymic evidence from the Roman period, presented by Hamp (1982), suggests a postpositive article in the ancestor of Albanian. Internal reconstruction dating its development before borrowings from Latin (Lindstedt 2014:177, citing Voronina 1976) also supports Albanian as the source. However, *ille* is also attested after the noun in late Latin texts, as in the 4th century Vulgate, e.g., *homini illi* 'to that man' and *homo ille* 'that man' (Posner 1996:128-129). Thus, this placement could have developed independently in Romanian, but I find Lindstedt's claim of mutual interference more plausible. I support the hypothesis that contact with Albanian reinforced this alternative placement in Romanian, while the more frequent position before the noun predominated in the Western Romance languages, where there was no external influence. In Bulgarian and Macedonian, the definite articles as such did not appear in the written record until the first vernacular texts, the damascenes, which originated in the 17th century as translations of non-liturgical religious literature written in vernacular Greek (Gyllin 1991:46-52). Their apparent late development relative to other Balkan Sprachbund languages suggests that contact could have played a role. However, the demonstratives *sъ* 'this' and *tъ* 'that', as well as adjectives, were already generally post-nominal in Proto-Slavic, as attested in OCS, so contact was not necessary for the placement of the article in Bulgarian and Macedonian.

Typological evidence suggests a relationship between definite articles and case loss. In a statistical analysis based on data from the *World Atlas of Language Structures* (Haspelmath et al. [ed.] 2013), Wahlström finds a typological tendency for a lower probability of the presence of a definite article in languages that have a higher number of morphological cases. Conversely, the lower the number of cases found in a language, the higher the probability of there being definite articles. While there may be a positive correlation between the loss of case and the

grammaticalization of the definite article, this is still not evidence for causation. On the other hand, the article itself often carries or protects case marking, even if other case marking is lost in the process (2015:173-177).[108]

Mladenova (2009) discusses various theories about the connection between the emergence of definite articles and case loss. Her focus is on Bulgarian, but she also considers parallels with Germanic and Romance languages. One argument that she makes is that articles perpetuate case distinctions. Philippi (1997: 63-64) makes this same argument with respect to the Germanic languages, in which definite NPs tend to express more case distinctions than bare nouns. This is particularly noticeable in CSHG, e.g., *Tag* 'day' can be Nom, Acc, or Dat singular, but *der Tag* 'the day' can only be Nom singular. Wahlström makes a similar argument for the Balkan Sprachbund languages: regardless of the article's position, more case distinctions are retained on definite NPs than bare nouns (2015:115-116, 170-171). In Romanian, for example, the bare noun *băieți* 'boys' can be Nom-Acc or Gen-Dat, while the definite Nom/Acc plural form *băieți-i* is distinct from the Gen-Dat plural *băieți-lor*. Postpositive articles may help preserve case distinctions by adding additional phonological material that can persist after sound changes have erased less salient case markings. As evidence for this mechanism in continental Scandinavian, Beito (1957:77) notes that the Dat forms that have survived in some Norwegian dialects always derive from definite forms, while the more reduction-prone indefinite forms have been lost.

In Bulgarian and Macedonian, however, the facts do not seem to support this type of relationship. The damascenes represent a stage in which definite articles occur and case distinctions have been reduced but not completely lost (Gyllin 1991:50-52). Thus, they serve as important evidence for the interaction between case markings and articles. In the 17th-century Tixonravov damascene, the distinction between Nom and Gen-Acc for animate masculine singular nouns is retained with and without the article (Mladenova 2009:412-13, citing Demina 1971:137, 151, 229, 326). In later damascenes, the article forms underwent reduction before the nominal case inflection (Mladenova 2009:413, citing Trifonov 1931:23). In the 18th-century Svištov damascene, for example, the reduced Gen-Acc form *care-tukŭ* 'the emperor' occurs as well as the more conservative form *carę-togo*/*carě-togo*, and the reduced Dat form *zmeju-tom*

[108] Russian has Gen indefinite DO only in negative and with non-count nouns. Otherwise, Acc covers indefinite and definite DO.

'the dragon' occurs in addition to the conservative *zmeju-tomu* (Mladenova 2009:413, citing Miletič 1923:16-27). Notably, forms such as **car-togo* and **zmej-tomu*, in which only the nominal case inflection is omitted, are not attested (Trifonov 1931:23). This contrasts with continental Scandinavian languages, in which the nominal case inflections, known as internal inflections, tended to be lost before the definite suffixes. Beginning in the earliest OSw manuscripts, internal inflections were sometimes omitted, and this tendency only grew stronger over time. For example, the indefinite feminine *i*-stem Gen singular *færþ-ar* 'journey' generally had the corresponding definite form *færþ-innar* as opposed to a form like *ferð-ar-innar*, attested in OIc (Norde 1997a:107-109). Thus, some factor appears to have made the nominal case inflections more resistant to reduction than the definite article forms in Bulgarian, while the opposite occurred in continental Scandinavian. It is possible that this is somehow connected to the difference in morphosyntactic status between the definite forms in continental Scandinavian, which can be considered nominal suffixes, and in Bulgarian, which cannot because they are clitics which are attached to the first word in the NP, whether it is a noun, an adjective, or another modifier.

This grammaticalization process has also been cited as a source of case loss, e.g., in Bulgarian and Macedonian. The impetus for case loss on nouns in these languages is the development of the definite article: since it was inflected for case, the case endings on nouns became less important (Wahlström 2015:170, citing Tiraspol'skij 1980:73-74). This argument may apply to a language like CSHG, which has prenominal articles with a much higher degree of case marking than nouns. With respect to Bulgarian and Macedonian, however, this is unconvincing for a number of reasons. First, it does not explain why the case markings on the article were subsequently lost, in contrast to CSHG. In addition, Wahlström notes that the postpositive articles in the Balkan Sprachbund languages combine with the nominal case inflections (2015:170). Thus, in contrast to languages with prenominal articles, the fates of the two sets of case markings are more closely aligned such that one cannot simply make the other redundant. Finally, there is uncertainty about the timing of these developments. The vowel mergers in MB that were most likely responsible for the initial confusion of case forms in BChS manuscripts probably occurred before definite articles were fully grammaticalized, but neither advanced case loss nor definite articles appeared in the written record until the eventual appearance of the first vernacular texts, the 17th-century damascenes (Gyllin 1991:13, 50-52).

Mladenova also considers a functional argument for the replacement of case marking by definite articles. Both can perform discourse functions, but on closer examination these functions were not quite the same in the relevant period of written Bulgarian, i.e., the 17th-century damascenes. The primary definiteness-related function conveyed by case marking in Slavic languages is inclusiveness, while identifiability is an extension of this in some contexts but cannot be expressed consistently. For the objects of some verbs in CSR, for example, the Acc and Gen are used in opposition: the Acc indicates that the action of the verb applies to the whole object, i.e., it has an inclusive function, in contrast to the partitive function of the Gen. However, the Acc does not necessarily imply that the object refers to something previously mentioned in the discourse or otherwise identifiable from context. The use of the Acc in these contexts often corresponds to the use of the definite article in CSB, but not always (2009:414-416). An investigation into the early attestations of the definite article in 17th-century Bulgarian shows that the article was initially used with an identifiability function, and that its expansion to generics, with which it has an inclusive function, only came later (Mladenova 2007:93-184). Thus, the Bulgarian definite article did not initially compete functionally with case marking, so their grammaticalization was not the result of case loss (Mladenova 2009:416-417). Although articles may have competed with case marking later in their development, meaning there was potentially functional continuity, this hardly implies that the grammaticalization of definite articles was a motivation for case loss.

Instead, Mladenova argues that a higher order factor affected both case loss and the grammaticalization of the definite article, resulting in parallels in their development (2009:425). Mladenova (2009:421-22) points out that Stojkov's (1970) investigation of Bulgarian dialects reveals that case distinctions tended to be lost for the plural before the singular, the neuter before the masculine and feminine, and inanimates before animates, with masculine personal names and kinship terms retaining case distinctions the longest. Mladenova has found a similar order for the spread of the definite article by comparing how often the article was used with these categories in damascenes and CSB. In three homilies of the Tixonravov damascene, the article was used in 68% of positions expected based on CSB for plural nouns, and 70% of expected positions for feminine and neuter nouns, but only 13% for masculine singular nouns. Likewise, the article was used in 61% of expected positions for inanimate nouns, but only 13% for animate nouns

(2007:131-137). Lower coverage of expected positions implies a later development, since changes had to occur to reach the full distribution of the article in CSB.

3.1.2.4. Case Variation

Before a morphological case distinction is lost entirely, it usually passes through a stage of case variation, although this stage is not always well-attested (see Luraghi 1987:356). Thus, case variation is an important topic in the investigation of case loss.

Building on work by authors such as Meiser (1992), Luraghi (2004), and Baerman et al. (2005), Kulikov (2013) discusses the sources, outcomes, and parameters of case variation and competition in IE languages, as well as some non-IE languages. He interprets the notion of case broadly, so variation between morphological cases and analytic constructions is also considered. Case variation occurs when two or more cases are interchangeable for a given context, without a change in meaning, or at least in the types of meanings canonically associated with cases. In CSHG, for example, both Gen and Dat can be used after prepositions such as *entlang* 'along'. The frequency of each case may vary depending on register and other pragmatic factors, but the meaning is essentially the same. In contrast, case opposition occurs when two or more cases have contrasting meanings in a given context. For example, Acc and Dat are both used after prepositions such as *auf* 'on, onto', but the meanings contrast: Acc is used when motion is involved. There are also instances that fall on a continuum between free case variation and non-free case opposition. In Slavic, for example, the borrowing of the Gen form to mark the Acc has developed out of an alternation between Gen and Acc depending on animacy, originally for the masculine singular (stemming from the Proto-Slavic "virile" form, as attested in OCS), and later for the masculine/feminine plural in some Slavic languages. This difference is meaningful but involves a case-external meaning rather than one concerning the relationship of the NP to the verb (2013:53-58).

Writing during the Communist period in Bulgaria, Duridanov (1956) approached the breakdown of the Bulgarian case system from a perspective of dialectical materialism; nevertheless, despite significant differences in theoretical framework and terminology, he considered many of the same sources and outcomes of case variation that Kulikov presented much later (2013). Earlier scholars had tended to provide one-sided explanations. Duridanov (1956:156) points out, some, such as Jagić (1894) and Weigand (1888), focused only on external

conditions; others overestimated the role of specific aspects, e.g., phonetic reduction according to Miletič (1890), and displacement of bare cases by prepositional constructions according to Meyer (1920); a third group, including Vondrák (1912), was content with assuming a general, abstract influence of other languages without evidence. Instead, dialectical materialism attributes change to the internal conflict between contradictions inherent in every object and phenomenon. Some contemporaries of Duridanov that he cites, such as Lomtev (1953), tried to argue that language change is often motivated by a drive to supplement a language's shortcomings in expressing the intended meaning of its speakers. This need certainly exists, but it cannot account for all developments, e.g., the replacement of the Dat, Loc, and Instr plural forms by the dual forms *-ima* and *-ama* in BCMS. In that instance, only internal structural conditions are relevant, since the original endings were still available to express those case categories. Duridanov instead saw the contradiction between existing and emerging elements, in other words, variation, as the driving force in language change (1956:156-161). Even though Duridanov used the required communist approach, he makes some significant points that give his analysis value.

Several types of internal contradictions can be found in the nominal system used in late OCS and BChS manuscripts. Under certain conditions, these may have undermined internal laws. One type is the variation between synonymous bare case forms and prepositional constructions, either reinforcing the same case or using a different case. For example, in the Savvina Kniga, an OCS manuscript from the 11th century, a bare partitive Gen is used: *vsěkъ vasъ* 'every one of you' (Luke 14:33). In the same verse in the Codex Marianus, from the early 11th century, the preposition *otъ* 'from' reinforces the Gen: *vьsěkъ otъ vasъ*. The availability of these redundant expressions contradicts the general tendency for well-defined meanings and syntactic functions (Duridanov 1956:161-164). Another type of contradiction results from the homonymy of case endings due to sound change. For example, *o*-stem Instr singular *-omь* and Dat plural *-omъ* both became *-om* with the loss of final *jers*. This contradicts the tendency to use different forms for different functions (ibid., 164). The third type was the coexistence of productive and unproductive declension classes. Duridanov asserted that the assimilation of unproductive classes by productive classes is generally explained as a consequence of the loss of semantic distinctions among the underlying stem suffixes. However, more recent work by Harmon and Kapatsinski (2017) indicates that the causative relationship between semantic indistinction and the assimilation of unproductive forms by their more productive counterparts

may go the other way; that is, as productive forms become more productive, subsuming other forms, their increased frequency may extend their semantic applicability to new uses.

Of these three types of internal contradictions, Duridanov considered the first to be the most important factor in case loss, since some critical sound changes, such as those to the nasal vowels, are a relatively late development (around the 12th century), while clear examples of cases losing syntactic productivity to other cases and prepositional constructions are found in the 11th century (ibid., 164-165).

These contradictions can be resolved in several different ways. One option is the strengthening of the case system with forms from unproductive classes to eliminate syncretism in BChS, e.g., the adoption of *u*-stem Gen plural *-ov* by *o*-stems as these classes merged so it can work together with the prepositional system to express the necessary diversity of semantic relationships between nouns and other words. However, this option violates the tendency that productive ending can be extended to other declension classes. This occurrence of using an ending from an unproductive class is rare but it is used based on this high distinctiveness (see Albright 2008:149, Hock 2021). When a form is more unique, it is easier to be recognizable by the listener (van Trijp 2013:20). This implies the gradual reduction of synonymous expressions. Another option is the weakening of the case system in favor of prepositional constructions and the Nom-Acc form. Slavic languages have a general tendency for the first option, but the internal structure of Bulgarian developed elements of the second, as the functions of the Acc were extended, particularly in combination with prepositions (Duridanov 1956:165-168). For example, the preposition *na* ‘on’ followed by the Nom/Acc singular form *krovъ* ‘shelter’ is used with a form of the verb *nadějęti sę* ‘to trust in’, in the Sinai Psalter, an 11th-century OCS manuscript: *na krovъ krilu tvoeju nadějǫtъ sję* ‘they trust in the shelter of your wings’ (35.8, cited by Duridanov 1956:166). This verb canonically takes a Dat object, as in *silě* ‘strength’ in another line of the same manuscript: *nadějǫštei sję silě svoei* ‘trusting in their strength’ (48.7; ibid.). The functional expansion of the Acc in OCS can already be observed in manuscripts dated to the 9th to 10th centuries, but those manuscripts probably are dated too early, as is the habit of Bulgarian historical linguists. Moreover, the Nom and Acc were already syncretic in most forms, so the focus of grammatical expression gradually shifted from inflections to prepositions (ibid., 169-171). Distinctive Acc forms were leveled to the Nom as a result of this shift in focus. For example, the *a*-stem Nom form *nezъloba* ‘acacia’ is used after the preposition *vъ* ‘in’ in the Sinai

Psalter (77.72; ibid., 169). Other scholars, however, attribute this substitution to regular sound changes, and I agree with this explanation, as described in section 2.5.2 above. Duridanov cited the failure of appositions to agree as further evidence that speakers had a weakened feeling for inflections. Similar early developments in VL and similar outcomes in the Romance languages serve as further evidence that the functional expansion of the Acc can be an important step in case loss (ibid., 170). The displacement of case marking by prepositional constructions is discussed further in section 3.1.2.7 below.

A third option is a shift in the focus of case marking from nouns to articles. In Greek, articles have retained certain distinctions that were lost on nouns due to sound change, e.g., the Nom-Acc distinction for feminine nouns. This allowed the case categories to survive while syncretism on nouns was retained. High German has undergone the same route of development (Duridanov 1956:171-172). Of course, this option was not available for most Slavic languages, but it is a characteristic of the Balkan Sprachbund languages. Bulgarian and Macedonian developed a definite article, but it may have come too late, when the analytic tendency in marking grammatical relations was already too strong.

Given these three different outcomes despite similar structural conditions, Duridanov argued that external conditions can influence the resolution of internal structural contradictions. For Bulgarian, contact with the other Balkan Sprachbund languages led to structural convergence (1956:171-173). Historical data suggests intensive linguistic contact between the Bulgarian Slavic speakers and other groups in the Balkans began well before the 14th century (Georgiev 1952:81). If contact began early enough, it could explain the loss of Instr and Loc via structural convergence: the semantic relations expressed by these morphological cases were already expressed exclusively with prepositional constructions in Pre-Romanian and Byzantine Greek. In addition, populations speaking these two languages were assimilated within the Bulgarian territory. These L2 speakers may have had particular difficulty acquiring the redundant use of case marking and prepositional constructions, leading to the retention only of the latter. Native speakers would not have accepted the development of completely new analytical constructions by L2 speakers, but given two alternatives, a large number of L2 speakers could have been the deciding factor (Duridanov 1956:173-174). The nature of contact in the Balkan Sprachbund and the potential effects of L2 speakers mentioned here are discussed further in section 3.2 below.

Case variation has several potential sources, all of which are also potential motivations for case loss. Kulikov (2013:59-65) considers the same four sources as Meiser (1992:191-196), i.e., paradigmatic inconsistencies, sound change, semantic overlap, and adstrate influence. The first three can be traced back further to Wackernagel (1920:302-304). Kulikov maintains the view that the most natural cause of variation is overlapping functions of the cases (2013).

Harmon and Kapatsinski (2017:35-38) demonstrate in their recent series of studies, noted above, that when forms have sufficiently similar functions, a speaker may simply choose the form which is the most accessible in their lexicon, even if that form is not the most well-suited, potentially leading to the broadening of a form's use through repeated extension. The broadening of forms tends to correlate with increasing frequency, while narrowing correlates with decreased frequency. The over-extension of forms happens when a speaker uses a more accessible form in a new context over another form that has a similar semantic equivalent. Every form that is prone to an extension in production and can be observed by the listener can drive language change. Over time when the form is used again, the new uses can eventually increase among other members of the community. Harmon and Kapatsinski (2017: 30) show that the results of their form-to-meaning mapping task were consistent with the prediction that the extension of frequent forms in production can coexist with the entrenchment of frequent forms in comprehension.

Kulikov points out that semantic overlap can result from different perspectives of the same participant. In ancient Indo-Iranian languages, for example, the passive agent was commonly marked with the Instr, conceptualized as the Manipulator, but the Gen was also used, mostly for pronouns, with a Possessor conceptualization (2013:58-59). Overlapping functions are the focus of Luraghi (1987), as discussed in detail in the following subsection. Another source mentioned by Kulikov is paradigmatic pressure and analogy, a topic in section 3.1.2.1 above. If two cases are syncretic in some declension classes, this can support variation between these cases even in classes with distinctive forms. Other factors are usually involved as well. For example, the Abl was already the least independent case in PIE: only the *o*-stem singular paradigm had a distinctive form. Along with semantic overlap, this could have caused variation between Gen and Abl, as attested in Vedic, followed by a complete merger, as in Greek and Slavic.

Phonological changes, the focus of section 3.1.1 above, are another source of variation, since they can lead to the partial case syncretism that in turn results in paradigmatic pressure and

analogical processes. Alternatively, phonetic erosion can lead directly to variation on the way to the full merger of two or more cases. For example, Arabic lost its three-case system after the weakening, neutralization, and loss of final vowels. In the Middle Arabic of South Palestinian Christian texts from the 8th to 10th centuries CE, case forms were still used but with variation: Nom, Acc, and Gen could all be used after prepositions that were only followed by the Gen in Classical Arabic (2013:61-63). In the following examples of this variation from Gruber-Miller (1990:244ff.), the case endings are always followed by a possessive suffix, so the vowels are not word-final:

(2) *mʕ ʔb-ii-hmaa*
with father-GEN-their
'…with their father'

(3) *mʕ ʔb-aa-hmaa*
with father-ACC-their
'…with their father'

(4) *ʕlaa ʔx-uu-h*
against brother-NOM-his
'against his brother'

Thus, variation seems to have occurred in this non-final context because in the potentially more common word-final position, the phonetic distinctions between the case forms had already been neutralized or at least weakened. In chapter II above, many instances of case loss are attributed to these internal sources. Paradigmatic pressure is often the best available explanation for the leveling of the last remaining distinctive form(s) of a case after sound changes have neutralized most distinctions. In other instances, there are still too many distinctive forms after sound changes have applied and a functional merger due to semantic overlap is the best explanation.

Kulikov argues that a substrate or adstrate can be a source of case variation: a variation pattern in one language can spread to another by contact or influence the productivity of an existing pattern (see Barðdal 2009). Kulikov attributes the marking of some DOs with the Gen in

Slavic to intensive contact with Finno-Ugric speakers to the north and northeast of the (Balto-)Slavic homeland. The Slavic Gen is the closest equivalent to the Finnic Partitive, which is used in opposition with the Acc to mark DOs, depending on aspectual features of the verb. Under negation, the Partitive is obligatory. Therefore, he proposes that bilingual speakers, particularly native Finnic speakers, used both the Acc and Gen in affirmative sentences, perhaps still based on aspect or indiscriminately, but preferred the Gen in negative sentences. This pattern then spread to the Slavic varieties away from the contact zone, leading to the Gen of negation and the adoption of the Gen form for the Acc of animate nouns (2013:57, 64-65).

For the latter development to have taken place, the originally aspectual opposition must have passed through a period of some degree of flexibility in its conformity to its original semantics before a new opposition arose based on animacy. The (free) variation often does not continue for long as the uses of two cases become associated with some extra meanings. Historically, this additional semantic difference—definite/indefinite, individual/generic, aspectual meanings, modes of action …etc.—has not traditionally been indicated in the range of meanings represented by cases. This process triggers to evolve case opposition from (free) case variation (Kulikov 2013:68). Analogous to case variation, Kapatsinski (2009:168-170) examines the use of the Russian adversative conjunction distribution, i.e., *no, da,* and *odnako* and finds that their uses are based on semantic, syntactic, and stylistic motivations. In terms of syntax, the conjunction *no* is preferred in postcopular complements. The conjunction *da* is preferred when the conjuncts are verb phrases or clauses. The conjunction *odnako* is preferred when the conjuncts are nouns or adjectives. In terms of semantics, *no* is preferred when the second conjunct avoids the circumstance expressed by the first conjunct from continuing the action. Stylistically, *odnako* is a trait of formal writing, whereas *da* is not. Therefore, in the situation of case variation, the variation is not always free.

A source of case variation in contact is often hard to verify, but Kulikov's (2013) argument above appears to meet Thomason's (2008b) criteria for such a claim, which are discussed in section 3.2.2 below. This section also considers contact as a source for other developments, such as the variation and subsequent merger of the Gen and Dat, which may have spread among the Balkan Sprachbund languages after arising in one or more of them due to other factors. Another potential source of case variation not mentioned by Kulikov (2013) is dialect mixture. The Jutes, Angles, Saxons, and Frisians who initially settled southern and eastern

England and southeastern Scotland all spoke different West Germanic dialects (Trudgill 2010:5-6). Nielsen argues that the mixture of these dialects led to greater variability in OE than in other Germanic languages of that time, citing the multiple roots used more or less interchangeably for 'first', 'be' and 'which of two' (1998:78-79). As for case variation, the same preposition could sometimes take multiple cases without any apparent difference in meaning, sometimes even in the same sentence. There is even at least one instance of one preposition with two coordinated objects in different cases: *ðurh drycræft oððe ðurh runstafum* 'through sorcery or through staff', where *drycræft* is Acc, while *runstafum* is Dat (Mitchell 1985:§1177(2)). In addition, different prepositions with the same meaning or function often took different cases. Some of this unexpected case variation may have resulted from the confusion of two dialectal usages that were conflated (ibid., §1222). Mitchell considers these types of confusion to factor into both the reduction in which cases are used with prepositions and the more general loss of morphological case (ibid., §1177(2), §1222). The early appearance of this confusion lends support to its role in bringing about case loss. However, if the written language lagged behind the spoken language during this period, as suggested by authors such as Tristram (2004:113, 202), then such confusion in case marking would be expected as an indication that the writer was trying to express a distinction he no longer used in his own speech. The potential role of dialect contact in case loss is discussed further in section 3.2.1 below.

Kulikov discusses several ways that case variation can develop over time. Besides staying the same, it can disappear or be reinforced into case opposition. One way that case variation can disappear is through the functional merger of the cases involved. This occurs when the effects of phonetic changes and/or semantic overlap are not compensated by analogy or other stabilizing processes. In most Germanic and Romance languages, for example, case distinctions lost to sound change were generally not restored by analogical processes, but they were in most Slavic languages, as discussed in section II above. When a merger occurs, the forms can derive from both cases. For example, the one distinctive Abl form in PIE survived as the Slavic *o*-stem Gen singular form when the Abl merged with the Gen (2013:65-67). Functional mergers are the outcome that authors such as Luraghi (1987) and Meiser (1992) focus on; they are discussed further in the following section.

Another way in which variation can disappear is when one case ousts the other from the overlapping context, thereby decreasing the functional weight of the latter. For example, a

prepositional construction with *ad* 'to' and the Acc competed with the Dat to mark IOs in Medieval Latin texts. The prepositional construction became more frequent until it ousted the Dat (Kulikov 2013:67-68). Presumably, this reflected similar variation in the spoken language, but it is not clear that nouns still had a distinctive Dat form at this stage. Thus, the prepositional construction may have ousted a bare IO rather than the Dat. Kulikov interprets an example with *ad* followed by an Acc adjective and a Dat noun as evidence for this variation (ibid., 68). However, the Dat noun is still the object of *ad*, so this example of apparent case variation within the same NP is more likely to reflect the loss of formal case distinctions and dominance of the prepositional construction than actual variation with a morphological Dat. This outcome of case variation can be described as functional narrowing and is discussed further in section 3.1.2.6 below.

Kulikov proposes that the reinforcement of case variation to case opposition is more common when there are minor functional differences. This (almost) free variation can develop into an association with additional meanings, usually case-external meanings such as definiteness and aspect. In Spanish, for example, the *ad*-construction spread from IOs to some DOs in an opposition with the bare construction based on specificity and individuation of the referent (2013:68-69). The development of a case opposition means that the path towards case loss has been avoided, at least as long as this opposition remains stable.

Kulikov discusses six parameters of case variation; their relative ranking can determine the outcome of case competition. The first four are syntactic constraints. The Identifiability constraint is the tendency to use the same form for the same function. It resists innovations in syntax, in that functions change less, and in morphology, in that forms spread to other cases less often (2013:69-70, 77). This constraint was previously identified by Duridanov as the general tendency for well-defined meanings and syntactic functions. It is violated by the first type of contradictions he discussed (1956:164). The Markedness principle, previously identified by Meiser (1992:204-205) as the iconic principle, concerns formal (un)markedness, i.e., salience, and can be framed in terms of economy. It favors the use of the less salient form for the more basic function when variation occurs, but it also means that the more salient form gains priority when there is variation in less dominant functions. As stated by Kiparsky (2012:22), this process can be seen as a form of grammaticalization since the endings indicate new functions.

A zero-morph is the least salient expression of case meaning on the continuum of formal markedness, a morphological case is more salient, and a morphological case with an adposition is the most salient (Kulikov 2013:70-71). This principle can account for the redistribution of forms when a functional merger occurs, i.e., the survival of one case's forms in some paradigms and another's in other paradigms depending on their salience. One declension class might have a more salient form for one case, while another declension class has a more salient form for the other case. Also common is different outcomes by number: the less salient case forms are retained in the unmarked singular, while the more salient forms are retained in the marked plural (see Meiser 1992:202-204). The Distinguishability constraint is the tendency to use different forms for different functions (Kulikov 2013:71-72). Duridanov also identified this tendency; contradictions to it are the second type he discussed (1956:164). Specifically, it should be possible to distinguish the two arguments of a transitive predicate. The Primary Argument Immunity Principle (PAIP) states that only the case used for primary arguments, i.e., the Nom in Nom-Acc languages, can be exclusively expressed with zero allomorphs (2013:72-73).

Kulikov also considers a morphological parameter and a functional parameter. The former depends on paradigmatic independence (P-independence), a measure of how many declension classes have a distinctive form for a given case. That is, cases with higher P-independence are more likely to survive, while cases with lower P-independence are more likely to be replaced. P-independence was almost certainly a motivation for the absorption of the paradigmatically weaker Abl by the Gen in Greek and Slavic. The latter, semantic specificity, favors the case with a narrower range of functions (2013:74-75). Thus, semantic specificity promotes the retention of case distinctions, particularly those involving peripheral cases, or even the creation of new distinctions, as in the second Loc in CSR. Kulikov mentions that the second Loc developed from variation between the *u*-stem Loc form *-u* and the *o*-stem form *-ě* in the 13th century, as these declension classes were in the process of merging. The merger was complete after the 14th century, but variation between the Loc forms continued on nouns that denote locations until the 19th century. By the 20th century, the variation had been reinforced into an opposition with the second Loc limited to true locative usages, e.g., with the prepositions *v* 'in' and *na* 'on'. The second Loc only had an independent form for about 150 substantives, but despite its low type frequency and therefore low P-independence, it survived because of semantic specificity (ibid., 76). This process in the development of CSR resembles refunctionalization, as

described by Meiser (1992:201-202), but with the new functions emerging from variation that originally arose due to a merger of declension classes rather than a case merger.

Kulikov proposes that these six parameters interact to determine the outcome of case variation. Since outcomes vary significantly across languages despite similar conditions at an earlier stage, a universal hierarchy of parameters cannot be achieved; instead, the hierarchy must vary by language and over time. However, a few generalizations can be made. Identifiability, Formal Markedness, and Distinguishability are the most fundamental constraints. They are broad in scope, with some role in most instances of case variation. The other constraints are narrower in scope: the PAIP only applies to the core cases, while P-independence and semantic specificity rarely apply to these cases. Historical developments can increase the relevance of these parameters, however. For example, the PAIP became more important than Identifiability and Distinguishability in OF after the core case forms were rapidly eroded. In CL, both the Nom and Acc generally had overt case forms; the Acc had a zero ending for some neuters, but so did the Nom. Thus, the PAIP was satisfied. Later sound changes left OF with zero endings for the Acc singular in all declension classes, while the Nom was either a zero ending or *-s* depending on the declension. At least in the singular, therefore, this situation violated the PAIP, which apparently gave rise to variation between the Nom and Acc forms in subject position during the MF period before the distinction was ultimately lost in favor of the Acc (2013:73-74, 77-79). Both Identifiability and Distinguishability would have been violated by the leveling of the Nom to the Acc, since the form for the subject function changed and subjects were no longer morphologically distinguishable from DOs. Thus, the PAIP may have worked to bring about its collapse, possibly together with the conceptual complexity of the OF declension system, the motivation proposed by van Reenen and Schøsler (2000:337). Just because the importance of the PAIP increased relative to Identifiability and Distinguishability, however, their relative ranking did not necessarily change. The PAIP could have been highly ranked all along, but its effects only surfaced once it was violated by phonological developments.

Returning to the Medieval Latin example, Kulikov proposes that Formal Markedness was eventually favored over Identifiability for the marking of IOs as the *ad*-construction ousted the bare Dat. Identifiability had worked to resist a new construction for this function, but the need for a more marked form became stronger (2013:79). Although Dat forms were still employed in these texts, it is quite possible that the WVL of the time already lacked a distinctive Dat form. In

this case, Distinguishability in the broader sense would have also played a role because the bare IO would not have been distinguishable from the DO. In other words, the more marked prepositional construction was the only option that was still distinctive. A similar variation between the Gen and *dē* 'from, about' followed by the Abl is already attested in CL; the prepositional construction eventually won out and its reflexes are the default expression for possessors in modern Western Romance languages (see Pharies 2007:102-103, Wahlström 2015:107). The same parameters were most likely involved in this development as with IOs.

As for Slavic, Kulikov posits that the Nom/Acc syncretism was in accordance with Identifiability but not Distinguishability. After the loss of final *jers*, it was also in accordance with the PAIP in the sense that the masculine Nom singular had zero marking. The borrowing of the Gen form for the Acc of some nouns but not others violated Identifiability in favor of Distinguishability. This change may be connected to the increase in Gen marking of the DO regardless of animacy under negation and in other contexts (2013:78-79). In another sense, however, zero marking for both Nom and Acc violated the PAIP because only the Nom should be exclusively zero marked. The borrowing of the Gen form for the Acc of masculine singular animate nouns (and of plural animate, or at least human, nouns in most Slavic languages) meant that it was no longer exclusively zero marked, satisfying this interpretation of the PAIP. It should be noted that this line of reasoning only applies to masculine singular nouns, but these are unmarked in both gender and number, so they can play a significant role in more general changes. In addition, it ignores the simultaneous development of zero marking in the Gen plural due to the loss of final *jers*. In fact, this zero marking was prone to replacement, e.g., when *o*-stems adopted the *u*-stem Gen plural form *-ov*, as described above. This and other substitutions may also reflect the influence of the PAIP, as well as the Markedness principle.

Kulikov's (2013) account of case variation has potential implications for many instances of case loss considered in my study. Some of the factors discussed in this section are also discussed elsewhere in my study, particularly in sections 3.1.2.5, 3.1.2.6, 3.1.2.7, sometimes in slightly different terms. The following three sections discuss the main outcomes of case variation. Unlike in the discussion above, a distinction between variation and outcomes involving only case marking and those involving an analytic construction is made in the present study where possible. The former outcomes are a functional merger or functional narrowing; the latter are instances of analytism.

3.1.2.5. Functional Mergers

A functional merger is one possible outcome of case variation. Sometimes mergers occur with no attested variation at an earlier stage, but I assume in my study that case variation precedes a functional merger, and more generally, that linguistic variation is required for language change to occur, even if this variation is not discussed directly.

Luraghi (1987) argues that there are two different ways in which a functional merger can occur, with different implications for the development of the case system. Partial synonymy between two cases arises when they can be used for the same semantic and/or syntactic function. This synonymy can lead to a full merger if there is too much redundancy and if unacceptable ambiguity does not result. Semantic syncretism results when the two cases overlap in semantic value, regardless of syntactic overlap, while syntactic syncretism results when the two cases occur in the same syntactic position, regardless of semantic overlap. Syntactic overlap corresponds to Kulikov's (2013) concepts of case opposition, while there is still a semantic distinction or case variation. Semantic syncretism results in systems whose case markers convey semantic information, while lexical features convey syntactic information. On the other hand, syntactic syncretism results in systems whose case markers convey syntactic position or grammatical relations, while lexical features help indicate thematic relations. The quantity and thematic relations of a verb's arguments are generally inherent information in its lexical entry. Likewise, the lexical features of nouns often determine the semantic and syntactic functions they can fill. For example, toponyms are more likely than personal names to be used in a locative function and are more likely to be adjuncts than arguments. Animacy is particularly important in determining the level of ambiguity that is acceptable. Both types of syncretism can occur in the same language, but Luraghi argues that languages predominantly undergo one or the other (1987:356-359). This use of the term syncretism differs from how it is generally used in the rest of the present investigation, i.e., homophony of forms in a particular paradigm.

Luraghi discusses AG and Hittite as examples of languages with semantic syncretism. The Gen and Abl had already merged in the earliest attestations of Greek. Luraghi proposes that corresponding notions of source formed the semantic overlap between these cases that eventually resulted in their merger. Specifically, both subjective and objective uses of the Gen imply that its referent is the source of the action implied by the head noun, while the Abl canonically marks the source of motion, and by metaphorical extension, the cause. A similar process occurred for

Slavic, although Abl functions are usually expressed with a preposition followed by the Gen as opposed to a bare Gen (1987:362-363). The variation preceding this merger is discussed in the previous section.

The merger of the Dat and Loc is well-attested in Hittite and was mostly complete in MyG, although it was functionally restricted in that a preposition was usually required in uses that replaced the Loc. Luraghi argues that the semantic overlap between the Dat and Loc arose due to uses of the Dat that indicate concrete or figurative proximity, e.g., Dat of possession. In these uses, the Dat is used with animate referents, while the Loc is predominantly used with inanimate referents, such as toponyms. Thus, they are in complementary distribution, so ambiguity does not result from their merger (1987:363-364). The much later merger of Dat and Loc in BCMS may have been motivated in part by the same semantic overlap. The Instr also merged with the Dat and Loc in the development of AG. Luraghi finds the semantic overlap between the Loc and Instr in the conceptualization of location as means and vice versa, e.g., the largely synonymous English expressions *to travel by car* and *to travel in a car*. Since the Instr was mostly used with inanimates until the later development of the Instr of agent, it could merge with the Dat as well as Loc without increasing ambiguity (ibid., 365). The syncretism among the Dat, Loc, and Instr plural in BCMS may have been facilitated by the same semantic overlap and interaction with animacy. The striking parallels in the development of these cases in AG and BCMS adds support to Luraghi's argument; besides the similarities discussed above, the developments in both languages occurred without widespread homophony due to sound changes. On the other hand, Hewson (2006:278) makes an argument for the merger of the Instr with the Dat in Greek and Germanic that does not rely on the complementary distribution of animacy: means can be conceived as goal-oriented in that it allows for completing the verbal activity, just as the Dat often functions as the goal.[109]

Luraghi (1987:366, 367) discusses CL and Germanic as examples of languages with syntactic syncretism. The Abl and Instr had merged in CL, and the Loc had almost entirely merged with the Abl as well. Luraghi proposes that these three cases merged because they were all used as adjuncts. When semantic opacity is resolved based on contextual and/or lexical cues, the main functions of all three can occur without a preposition. Otherwise, the Abl without a preposition is limited to the instrumental function, as well as the closely related functions of

[109] The merger of the Dat with the Gen rather than the Loc is characteristic of the Balkan Sprachbund.

manner and cause. Due to this functional load, the Abl did not become the unmarked prepositional case; the Acc did instead. In Germanic, the Loc, Abl, and later Instr all merged with the Dat. Luraghi argues that this was possible because Germanic made a primary distinction between the central arguments, i.e., the subject and DO, and all other actants, including both the IO and adjuncts. In contrast, CL distinguished arguments from adjuncts, and the Dat remained distinctive. In Germanic, the use of prepositions with the Dat for adjuncts but not arguments accomplished this secondary distinction, and also led to the Dat becoming the unmarked prepositional case. These instances of syntactic syncretism caused semantic opacity, which was partially resolved by the increased the use of prepositions. These effects can be seen as a motivation for subsequent reductions in case morphology that are much greater than in languages with semantic syncretism (1987:366-368). As with Greek, Hewson (2006:278) argues that the conception of both Instr and Dat as goal-oriented allowed for their merger in Germanic. This alternative explanation, as argued by Luraghi (1987), would mean that semantic and syntactic syncretism are not as neatly separated.

Contrary to the reconstruction of Proto-Germanic provided in section 2.3.1 above (see Ringe 2006:199, 269-274), the Loc and Abl may have been subsumed in part by the Dat and in part by the Instr at first. Anderson (1958) argues that the Instr forms that remained distinct from the Dat in OE were actually derived from a mix of Instr, Loc, and Abl forms. In contrast to OHG and OS, OE nouns did not regularly have a distinctive Instr form. He considers the PIE Loc singular form *-*ēi* to be the most likely origin for the OE *a*-stem Dat/Instr singular form -*e*,[110] as well as the demonstrative Instr forms *þȳ* 'that' and *þȳs* 'this', and the interrogative pronoun/adverb form *hwȳ* 'why'. Meanwhile, the PIE Abl singular form *-*ēd* was probably the source of the distinctive Instr singular form -*e* for masculine and neuter strong adjectives, as well as the Instr form *þē* 'that'. If these sources are correct, none of the above forms are cognate with the OHG/OS *a*-stem Instr singular form -*u*, the demonstrative Instr form *diu/thiu* 'that' or the interrogative pronoun/adverb form *hwiu* 'why'. In OE, only the interrogative adverb *h(w)ū* 'how' and some exceptional nominal Instr singular forms in -*um* appear to be derived from the PIE Instr. Anderson also questions the assumed Instr origin of OHG -*u* because it would require PIE *-*ō* instead of *-*o*. In fact, the PIE thematic Instr singular form is now generally reconstructed as

[110] The PIE Loc singular form *-*ēi* is also the source, via PGmc *-*ai*, of the OHG/OS *a*-stem Dat singular form -*e*, while the Instr form is distinctive in these languages (see Ringe 2006:200).

*-*oh*$_1$, which became *-*ō* with the loss of laryngeals and compensatory lengthening (see Sihler 1995:248, Fortson 2004:113, Ringe 2006:15, 41). In every Germanic language, the Loc, Abl, and Instr all eventually merged with the Dat, but if Anderson (1958) is correct, the situation in PGmc may have involved more variation and overlap among these cases than reconstructions generally suggest. The inherited case forms may have combined to form functional categories in different ways depending on the Germanic dialect. The ultimate loss of the Instr in OHG, OS, and especially in OE, can then be considered a delayed outcome of this confusion during the PGmc period.

Whether a language is prone to semantic syncretism or syntactic syncretism, a by-product of having distinct prepositions for different cases, appears to have a significant effect on how its case system develops, but Luraghi does not discuss what motivates a particular language to follow one of the two patterns. In fact, she notes that both types can be found in the same language. In CL, the syntactic syncretism of the Abl and Instr may have begun as semantic syncretism. In Homeric Greek, the ending *-phi* was used for functions of the Loc, Abl, and Instr, contrary to the semantic syncretism between the Abl and Gen discussed above. However, this semantically opaque ending fell into disuse (1987). Despite these exceptions, both languages subsequently developed in a way that restored them to their preferred type: semantic opacity increased in CL but decreased in AG. Perhaps another way to frame this dichotomy, then, is that some languages tolerate a greater degree of semantic opacity in case marking than others. Even in those that develop opaque case markers, however, other changes such as the increased use of prepositions decrease the overall semantic opacity. Thus, syntactic syncretism seems to be associated with analytic tendencies, but the directionality of any causation is unclear. Analytism is discussed further in section 3.1.2.7 below.

The Romance and Germanic languages primarily experienced syntactic syncretism, and most members of these families eventually lost all case marking on nouns. Meanwhile, Greek primarily experienced semantic syncretism and retains the case system. In Slavic, only the Gen and Abl merged in Proto-Slavic (see Schenker 1993:85, Lunt 2001:222), as in AG, so it seems to belong in the group with semantic syncretism. Except for Bulgarian and Macedonian, which are in the Balkan Sprachbund, Slavic languages retain a robust case system, and later mergers, e.g., between the Dat and Loc in BCMS, also appear to be semantically motivated, providing further support for Luraghi's proposal.

Faroese and CSHG have experienced some reduction in case marking, particularly on nouns themselves. Case marking has become more syntactically predictable in the development of High German, particularly as the arguments of verbs (see Barðdal 2009:141-142). In both High German and Faroese, prepositional constructions have supplanted certain case functions, particularly of the Gen. These developments suggest that the effects of syntactic syncretism are still playing a role, albeit a less extreme one than in English, Dutch, and the continental Scandinavian languages. On the other hand, Barðdal demonstrates that Icelandic largely retains the variety of semantically-motivated case options for each syntactic position found in earlier Germanic languages, and changes that have occurred have not significantly increased semantic opacity (ibid., 135-140, 142-147). Thus, it appears that Icelandic is resisting further syntactic syncretism.

The loss of the Dat in Byzantine Greek, part of the Balkan Sprachbund, exhibits characteristics of semantic syncretism, as expected, but also syntactic syncretism. The Acc began replacing prepositional uses of the Dat in the Koine period; by the late Byzantine period, it had become the only prepositional case in spoken varieties (see Horrocks 2010:108, 284-285). These developments are characteristic of syntactic syncretism. The replacement of other Dat functions with the Gen has features of both types. Humbert (1930:166-171) observes substitutions between Dat and Gen personal pronouns in the Egyptian Greek papyri, which date from 300 BCE to 800 CE. He has identified examples of a Gen pronoun that can be interpreted as an adnominal possessor or a recipient. These two interpretations are possible for the first-person singular Gen pronoun *mu* in the following example from the 2nd century CE:

(4) *agórasón* *mu* *tò* *méros* *tû* *eleōnos*
buy-2SG.AOR.IMP.ACT 1SG.GEN the-ACC.SG part-ACC.SG the-GEN.SG olive.grove-GEN.SG
'buy my part of the olive grove' / 'buy me part of the olive grove'

Horrocks (2007:628-629) shows that this overlap in position was a syntactic basis for the replacement of the Dat by the Gen and also considers semantic motivations. In New Testament Greek, in texts of the Gospel of John specifically, Gianollo (2010:113) has found a different syntactic overlap with less difference in semantics: the Dat external possession construction and an extraposed Gen.

Largely due to the low frequency of inalienable possessives in the Egyptian Greek papyri, however, external possession constructions with the Dat are rare in these texts, and prenominal Gen pronouns do not have the characteristic properties of external possession constructions. Therefore, Stolk (2015:92, 95-117) seeks an alternative semantic motivation for the replacement of Dat pronouns by the Gen in the papyri. She finds that Gen constructions with ambiguity between denoting a possessor and a malefactive source, benefactive, or goal are the most likely starting point for the extension of the Gen into Dat functions. These are mostly attested between the 1st and 3rd century CE. The final stage, at least for pronouns, would then be uses of the Gen as benefactive, goal and addressee without any adjacent noun allowing for a possession interpretation. These are rare until the late Roman period, i.e., the 3rd and 4th centuries, and the Byzantine period, which lasted until the 7th century in Egypt. This distribution allows for a speculative chronological progression, but more research is needed to confirm the results. Given the apparent variation among different KG texts with respect to the path by which Gen pronouns replaced the Dat, it is difficult to determine exactly how the loss of the Dat actually proceeded in spoken KG. Regardless of the exact path, it appears that both semantic and syntactic overlap played a role, bringing the strict dichotomy between the two proposed by Luraghi (1987) into question.

The functional merger of the Gen and Dat also occurred in Romanian, Bulgarian, and Macedonian. The process in these three languages differed from the developments in Greek in its directionality, since the Dat generally subsumed the functions of the Gen instead of vice versa, but its motivation is no easier to classify as primarily semantic or syntactic. It also likely differed among these languages in terms of specifics. While the primary functions of the Gen were ultimately supplanted by the Dat in the development of MB, Duridanov observes that its secondary functions started being replaced by the Nom-Acc form in 13th century BChS manuscripts. One of these, the Gen of separation, is attested both with and without a preposition in OCS, but the use of prepositional constructions, particularly with *otъ* 'from', had expanded significantly by the 13th century. There are numerous examples from the 14th century of these constructions with the Nom-Acc form, which suggests that this function of the Gen was no longer productive in MB. For example, in the Plovdiv Gospel, a BChS manuscript from the 14th century, the preposition *otъ* is followed by the Nom singular form *prokaza* 'leprosy': *čisti sǫ otъ prokaza* 'clean yourself of leprosy' (Matthew 8:3, cited by Duridanov 1956:205). The expected

Gen singular form *prokazy,* still after the preposition *otъ,* is found in the Codex Zographensis, one of the two late 10th-century or early 11th-century OCS manuscripts of the four Gospels. This function was inherited from the Abl when these cases merged in the development of Proto-Slavic from PIE (see Schenker 1993:85, Lunt 2001:222), which may have made it more susceptible to replacement by other constructions. Duridanov even considers this extra functional load on the Gen to be a factor in its loss more generally, with supporting evidence from a parallel development in Greek: the secondary functions that the Greek Dat had acquired from the Instr and Loc were the first to lose productivity, and this case ultimately merged into the Gen-Dat as in MB. In other Slavic languages, the Gen of separation also came to be limited to prepositional constructions, but without replacement by other case forms, and the Gen survived as a distinctive case (1956:201-205).

Likely due to its semantic connection with the separation function, the partitive Gen also lost productivity around the same time despite being a core function. It was replaced by both the bare Nom-Acc form and this form with the preposition *otъ* 'from'. As one of many examples, in the Zagreb Octoechos (60a, cited by Duridanov 1956:213), a BChS manuscript from the 13th century, the Nom plural form *grěxy* 'sins' is used with the noun *množestvo* 'a multitude', instead of the expected Gen plural form *grěxъ*. Likewise, in the Banica Gospel (John 4:39; ibid., 214), a western BChS manuscript from the 13th or 14th century, the preposition *otъ* followed by the Nom/Acc plural form *samarěny* 'Samaritans' is used with the adjective *mnozy* 'many'. The expected Gen plural form *samarěnъ* is found in the corresponding verse in the Codex Zographensis, also with the preposition *otъ*. In turn, the Gen retained a partitive sense as the complement of some verbs in OCS but not others. This function is an example of case assignment by verbs, including but not limited to DOs. This case assignment with a partitive sense also lost productivity in favor of the Acc beginning in 13th-century manuscripts. For example, in the Strumica Apostle (1 Corinthians 1:22, cited by Duridanov 1956:216), a 13th-century BChS manuscript, the Acc singular form *znamenie* 'a sign' is the object of the verb *prosętь* 'they ask for'. The expected Gen singular form *znamenija* is found in the corresponding verse in the Šišitovac Apostol, a 14th-century Serbian Church Slavonic manuscript. The broader use of the Gen for DOs under negation also became less consistent in BChS manuscripts. Other Slavic languages have also undergone similar developments: CSR, BCMS, and colloquial Polish

have variation between Gen and Acc marking on negated objects as long as there is no partitive sense, and Czech requires the Acc in these contexts (ibid., 212-218).

Duridanov argues that the primary adnominal function of the Gen also started to lose productivity in the 13th century. Like the functions described above, it was sometimes supplanted by the Nom-Acc, with or without the preposition *otъ* 'from' (1956:218-220). For example, in the Prague Gospel, a BChS manuscript from the 14th or 15th century, the Nom plural form *xlěby* 'heads of grain' is used in an adnominal construction: *xlěby prědъloženie sъněstъ* 'to eat the offering of heads of grain'. Unlike other functions, however, the adnominal Gen was also replaced by the Dat, which already had a possessive function (ibid., 218).

At least in some texts, some types of adnominal constructions were more likely to be replaced than others. In an investigation of the Dioptra, a 14th-century BChS translation of a Greek manuscript written in 1095, Fuchsbauer (2018) finds that subjective uses of the Gen in Greek were overwhelmingly translated with a Gen marker, while objective uses were mostly translated with a Dat marker. The objective Gen has more semantic overlap with the Dat because both involve affectedness by the verbal action, while the subjective Gen is closer to a possessive or partitive function. Thus, the translator tended to imitate the Greek original despite the presumed loss of the Gen in the vernacular, except when the meaning was close enough to the core meaning of the Dat. In the *Life of St. Parasceva*, an original BChS text from the late 14th century, there was an even stronger tendency to use the adnominal Gen; the Dat still occurred in some objective uses, but at least some of these could also be interpreted as a Dat of benefit. Thus, the use of the Gen appears to reflect a more general archaizing trend by the MB literary reformers, perhaps still influenced by Greek in a general way, rather than overly literal translation from Greek on a case-by-case basis. Alternatively, the distribution of the Gen and Dat in these texts could reflect an intermediate stage in the loss of the Gen, in which uses involving affectedness had already been replaced by the Dat but others had not.

Duridanov asserts that this replacement for the Gen came to dominate in most dialects, even as at least some of the functions of the Dat were being supplanted by the preposition *na* followed by the Acc.[111] In fact, other replacements for the Gen were likely limited to certain dialects. Duridanov posits that this use of *otъ* was connected to the use of the analogous

[111] The development of this construction, which ultimately displaced possessive uses as well, is discussed further in section 3.1.2.7 below.

preposition *apó* ‘from’ to indicate the possessor in Northern Greek dialects (1956:219-220). In these dialects, the Acc is also used with this preposition (Wahlström 2015:156, citing Asenova 2002:93). However, Duridanov concedes that the direction of structural borrowing is not clear, since other Slavic varieties have developed similar constructions, e.g., Czech *synek od meho bratra* ‘the son of my brother’. Another possible source is Romanian, based on the occurrence of *ot*-constructions as well as the adnominal Dat in the Wallachian Letters, written as a continuation of the BChS literary tradition in Wallachia and Moldavia during the late 14th and early 15th centuries, and therefore with potential influence from the local Romanian varieties. This is further supported by the use of *de* ‘from’ in similar constructions in VL (1956:219). As discussed in section 3.2.3 below, however, these constructions are quite limited in CSRm, but they seem to have been more productive in earlier stages of the language. Thus, Romanian cannot be ruled out as the source for these constructions in Bulgarian.

Replacement of the adnominal Gen by the bare Nom-Acc was also dialectally limited. Unlike the *ot*-constructions, however, which are attested in the Cserged Prayers,[112] and survive in certain modern Bulgarian dialects (Sofia, Central, and Southwestern), the bare Nom-Acc was gradually supplanted by the adnominal Dat in these dialects, based on the evidence of fossilized forms (Duridanov 1956:219-220). Moreover, it is possible that substitutions of the Acc for the Gen were actually a result of phonological confusion rather than a loss in syntactic productivity. If there was confusion between the nasal vowels, it could have led to substitutions of the *(j)ā*-stem Acc singular form *-ǫ* for the *jā*-stem Gen singular form *-ę*. However, this still would not account for the use of *-ǫ* instead of the *ā*-stem Gen form *-y*, given that there is much less evidence for the substitution of soft stem forms for hard stem forms than vice versa in the development of Bulgarian (ibid., 1956:222-223). As discussed in section 2.5.2.1 above, apparent confusion of *ǫ* and *ę* may have been purely orthographical in nature, adding further doubt to this explanation.

The merger of both the Loc and Instr with the Acc similarly has characteristics of both semantic and syntactic syncretism. On the one hand, variation between the Acc and Loc occurred

[112] These texts consist of three 19th-century manuscripts written in the Bulgarian dialect of the Cserged Bulgarians who moved to Transylvania in the 13th century. They are translations of German Lutheran liturgical texts probably dating to 1680 at the earliest, but their language is considered to be highly conservative due to the geographical isolation of the Cserged Bulgarians from other Slavic varieties. They are also valuable because they do not continue the conservative Greek-influenced BChS literary tradition, given that they are not Eastern Orthodox Christian texts (Wahlström 2015:39, citing Miletič 1987 [1896-1900]:101-111, 116).

in ancient Indo-Iranian languages such as Sanskrit and Pali. For example, certain verbs of motion could take an Acc object denoting goal or a Loc object denoting location after motion. These two functions are conceptually similar, so the variation can be attributed to semantic overlap (Kulikov 2013:60-61). The same conceptual link, already attested in OCS, likely played a role in the merger of these cases in MB. As one of many examples, in the Codex Zographensis the preposition *vъ* 'in' is followed by the Loc singular form *cěsarьstvii* 'kingdom' even though *prideši* 'come' is a motion verb: *egda prideši vъ cěsarьstvii tvoemь* 'when you come into your kingdom' (Luke 23:42). The expected Acc singular form *cěsarьstvie* is found in the corresponding verse in the Codex Assemanianus, an OCS lectionary gospel manuscript from the early 11th century. Also, in the Codex Zographensis, the preposition *vъ* is followed by the Nom/Acc singular form *domъ* 'house' even though the verb *sǫtъ* 'are' does not signify motion: *vъ domъ otьca moego obitěli mnogy sǫtъ* 'in the house of my father there are many lodgings' (John 14:2). The expected Loc singular form *domu* is found in the corresponding verse in the Codex Marianus, as well as in Codex Assemanius, Dobrejšo Gospel, Book of Sava, and other manuscripts (see Duridanov 1956:185-186). Likewise, there is semantic overlap between the Loc and Instr, as discussed above. On the other hand, all three of these cases marked the objects of prepositions in OCS, particularly after the instrument-accompaniment merger allowed the core functions of the Instr to be expressed with a prepositional construction (see Duridanov 1956:197-198). Eventually the Acc became the only prepositional case, which is expected to result from syntactic syncretism but not semantic syncretism. Thus, based on Duridanov's account, the Balkan Sprachbund languages appear to be characterized by a complex interplay between semantic and syntactic syncretism rather than neatly fitting into type.

Overall, Luraghi's (1987) proposal regarding functional mergers has the potential to account for the different outcomes regarding case loss among languages of the same family. It has a number of issues, however. Baerman et al. state that one is that syncretism among oblique cases is only common in IE languages; non-IE languages are much more likely to have syncretism between Nom and Acc or between one of these core cases and an oblique case (2005:40, 52). Given that semantic and syntactic overlap among oblique cases also occurs in non-IE languages, as well as in IE languages that lack these functional mergers, another factor must be involved. Rapaport argues that accent and ablaut patterns in PIE clearly distinguished the oblique cases from the core cases, but not from each other, encouraging oblique-oblique

syncretism. More specifically, only one of the four patterns for athematic nouns had a consistent accent on the root; the other three stressed the oblique cases on a later syllable than the core cases (2016:10-11).

In addition, an explanation is still needed for why a language initially becomes prone to syncretism. One possibility is that syntactic overlap results in syncretism because the case endings involved occurred in the same phrasal contexts and were therefore prone to the same elision due to stress lapse (Rapaport 2016:66). This type of deletion may explain the earliest instances of schwa deletion in OE, which occurred in hiatus and metrically weak positions according to Lass (1992) and Minkova (1991), in a phrase such as *sune and mone* 'sun and moon', with the structure / ˈσ σ σ ˈσ σ /, changed to *sun and mone*, with the structure / ˈσ σ ˈσ σ /, which removed the stress lapse (ibid., 56-57). Elision of this kind cannot occur if the final syllable is stressed and is expected to be most common in words with initial stress. Therefore, the IE daughter language groups with initial stress, including PGmc and Proto-Italic, would have been more prone to this elision and hence syntactic syncretism. In contrast, daughter languages that retained the mobile accent, including frequently accented oblique endings, would have undergone this process much less frequently. Thus, a prosodic distinction would neatly account for the different paths languages take with respect to functional mergers.

3.1.2.6. Functional Narrowing

The ousting of one case from one or more contexts is another possible outcome of case variation in those contexts. My study generally refers to this process as functional narrowing, highlighting its connection to functional mergers: when a case has lost all of its functions to another case, the result is a functional merger. Thus, functional narrowing very often describes the process by which case variation eventually results in a functional merger. However, a case may still survive in some of its functions, i.e., thematic or grammatical relations.

Functional narrowing has been observed across the Germanic family; the Dat and Gen seem to be the most often affected in the periods under consideration. For example, Delsing discusses the loss of productivity of the Gen case marking in Scandinavian. The Gen had three uses in the OSc period: adnominal, e.g., *mansins hus* 'the man's house'; complements of verbs, prepositions, and adjectives, e.g., *til mansins* 'to the man'; and in partitive and adverbial constructions, *annars dags*' on another day'. All three uses are attested in earlier OSw, but non-

adnominal uses were lost around 1250-1350 (2002:938). By this time, only personal pronouns retained the Gen after prepositions, but even they competed with the Acc. For example, in the Revelations of Saint Birgitta, a MSw text from the late 14th century, the Gen pronoun *min* 'me' appears as the object of *til* 'to', but in another line of the same text the Acc form *mik* does instead (see Norde 1997a:161-162). Thus, the Gen was sometimes replaced by the Acc, and as the Gen's functional load weakened; prepositional constructions also replaced certain functions of the Gen (Delsing 2002:938). The functional narrowing of the Gen also made it possible for the Gen singular form *-s* to develop into the clitic *-s* in the modern continental Scandinavian languages, which allows for the group Gen construction (see also Norde 1997a:81).

Norde (1997a:32) states that Hansen applied a similar argument to the loss of both Gen and Dat in Scandinavian and also provided additional semantic motivation for the loss of productivity, namely that the Gen and Dat were lost in the Scandinavian languages because verbs and prepositions lost the ability to assign non-Acc case marking as a result of the obscuration of the semantic differences between objects marked for these cases (1956:184). This view posits that these semantic and syntactic changes led to phonetic reduction, not the other way around (ibid., 192). As discussed in section 3.1.2.3 above, however, certain Norwegian dialects retain Dat forms that derive from definite forms, but never from indefinite forms; Beito argues that this can only be explained by phonological changes (1957:77). Rapaport offers prosodic change as a potential phonological mechanism by which only definite endings survive, despite the lack of regular sound changes affecting the indefinite endings. As discussed in section 3.1.1.1 above, he has attributed the loss of final syllables in other Germanic languages such as English to a strong alternating trochaic rhythm as well as more specific prosodic constraints (2016:66-67). Therefore, at least for the loss of the Dat, functional narrowing seems to have been both a result and an indication of case loss, but not a motivating factor.

Functional narrowing also occurred during the breakdown of the OE case system in the transition to ME.[113] Texts from the transition period provide a record of decreasing functional weight in progress. Specifically, a select few EME texts still have verbs and prepositions that assign either the Dat or the Gen (see, for example, Van Gelderen 2000:212-218). In the only manuscript of The History of the Holy Rood Tree*,* a 12th-century EME text, verbs that previously assigned the Dat sometimes have Dat pronoun objects, but the typically Dat noun

[113] Chapter V is concerned with the functional narrowing of selected OE and EME texts.

endings *-e* and *-um* never occur after verbs. Verbs in this text no longer assign the Dat consistently, but prepositions still tend to do so where expected. In both manuscripts that contain the Katherine Group, five 13th-century texts written in the same West Midlands dialect, there is no distinction between Acc and Dat pronouns; a few verbs assign the Gen but not the Dat. In both manuscripts of Layamon's Brut, another 13th-century text from the West Midlands, some verbs can still assign the Dat or Gen. Dat and Gen objects are less common in the later Otho manuscript of Brut than in the earlier Cotton Caligula (ibid., 214-218). In the following examples provided by Van Gelderen, *abiden* 'await' has the Gen object *þeos wederes* 'the weather' in the Cotton Caligula manuscript but the Acc object of *weder* 'weather' in the Otho:

(5) *þere læi þa uerde. þeos weder-es abid-en*
there lay the army. the.GEN.SG weather-GEN.SG wait-3PL
'there lay the army waiting for good weather' (Cotton Caligula, II. 14093–4)

(6) *ferde. weder a-...*
army. weather.ACC.SG
'army. ... weather'. (Otho, *idem*)

However, Dat and Gen objects are still attested in the Otho manuscript. In the following example, *his domes* 'his judgment' is the Gen object of *forȝe* from *forȝeten* 'forgot':

(7) *forȝe ... þare his dom-es*
forget.PST.3SG ... there his judgement-GEN.SG
'forgot there his judgment'. (Otho, I. 13506)

The differences between these two manuscripts demonstrate the loss of syntactic productivity in progress. In a corpus study, I have found similar developments in other EME texts. This study traces the functional narrowing of case usage from Beowulf, a representative OE text, through the Cotton Caligula manuscript of Brut, and The Owl and the Nightingale. These texts reveal various stages in the functional narrowing of the Gen and Dat in favor of the Acc. This study is discussed in more detail in section 5.2 below.

CSHG has been experiencing a functional narrowing of the Gen similar to that in OSc and EME. Adverbial uses have become fossilized, e.g., *keineswegs* 'by no means' (lit. 'of no way') and time expressions such as *nachts* 'at night'. The use of the Gen is highly restricted in CSHG. Only around five verbs that take a Gen object are still in productive use (see Barðdal 2009:142). Verbs that once took both an Acc or Dat and a Gen object began to assign Acc to the second object instead of Gen in the 13th century (Seefranz-Montag 1983:171-189). The Dat has also been replacing the Gen on the complements of prepositions such as *wegen* 'because of' and *laut* 'according to' in both speech and writing. For *laut*, this may be by analogy with three synonymous adverbial prepositions derived from a noun or participle that assign the Dat: *gemäß*, *entsprechend*, and *zufolge* (Sick 2005:214). In more general terms, prepositional constructions with the Dat have a much higher type frequency than those with the Gen. The use of Dat also provides distinctiveness from the Nom/Acc for unmodified strong nouns in the plural, while the Gen does not (see ibid., 215). This suggests a tendency in the CSHG case system towards emphasizing the two-way distinction between the Nom and an oblique form, at the expense of distinctions among the oblique cases. The end result of such a tendency can be observed in ModLG, which retains only a two-way distinction between Nom and oblique, as discussed in section 2.3.3.2 above. In terms of Kulikov's (2013) parameters, Distinguishability has become stronger than Identifiability.

Other High German case markings have also become more restricted in their use. Maling (2002:31) has found that only around 100 verbs still take a single Dat object, as opposed to around 750 for Icelandic. Most verbs with a Dat or Acc syntactic subject, often called impersonal constructions, stopped being used or were attracted to the more common Nom subject construction. This followed a period of variation in subject marking among these three case markings during MHG (Seefranz-Montag 1983:162-163). Barðdal attributes this variation and eventual substitution to the similar meanings of constructions with Dat and Acc subjects. On the other hand, CSHG has retained a Dat passive construction, e.g., *Ihm wurde geholfen* 'he was helped' where *ihm* is Dat (2009:142). Thus, while many low-frequency constructions have been lost or reduced to even lower frequency, the four case markings of CSHG are not entirely confined to distinctive syntactic environments.

The Gen in Faroese has experienced an extreme loss of productivity, even in the written language. The Gen has been replaced in several different ways depending on the specific

function. As mentioned in section 2.3.2.1 above, the Acc has replaced the Gen for modified nominal objects of prepositions, even in written Faroese. The Acc is also used to mark possession with kinship terms, in contrast to Dutch and High German, which maintain *-s* Gen forms for kinship terms, even when a broader adnominal Gen has been or is being lost. In spoken Faroese, a variant of the *-s* Gen, originally only found in the masculine and neuter singular, has the form *-sa* before consonants and *-sar* before vowels and is used with names and nouns used as names, regardless of gender, e.g., *Tummas á Dómarakontórinumsa bilur* 'Thomas at the legal office's car' (see Allen 2008:47, 55-56). The various ways in which the Gen has been replaced, in combination with its surviving reflexes, suggest that its disappearance has syntactic and possibly morphological motivations, but not phonological.

None of Kulikov's (2013) parameters appears to convincingly motivate the ousting of other cases from various syntactic positions by the Acc after a period of variation. In addition to Identifiability, Distinguishability would have favored the Gen or Dat over the Acc because the Acc was more likely to be syncretic with the Nom. While sound changes lowered the P-independence of all cases in all of the Germanic languages that experienced this functional narrowing, the Acc tended to have the lowest P-independence of all. Likewise, the choice of Gen or Dat was based on the semantics of the verb, so they would have been supported by semantic specificity. Therefore, other motivations must be sought to explain the attested outcomes.

One theory on the collapse of Latin nominal inflection holds that syntactic factors were the primary motivation for the extension of the Acc to other cases. Herman observes that the Tabulae Defixionum, short inscriptions of curses from Hadrumetum in Romanized Africa, contain forms that appear to be Acc before intransitive verbs. Syntactically, these can be interpreted as a subject or an Acc of enumeration, a use in lists of objects. The functional alternation of the Nom and Acc appears to have begun with such nouns in late 2nd century CE Romanized Africa. Forms such as Nom *-us* and (apparently) Acc *-u* were not interchangeable in other positions traditionally associated with the Acc, e.g., DO. Based on these facts, Herman argues that the extension of the Acc began with the Nom due to the use of the Acc in detached structures, which are unmarked contexts, and was accelerated, not caused, by phonetic changes. He admits, however, that is not clear whether the loss of *-s* more generally was influenced by the replacement of Nom forms ending in *-s* with Acc forms or was the cause of these substitutions, as proposed by the phonology-first approach (1987:102, 106).

Assuming that this variation between Nom and Acc occurred before the regular loss of *-s*, the outcome is difficult to predict with Kulikov's (2013) parameters. Both Identifiability and Distinguishability would have favored the Nom over the Acc in this position. The PAIP may have favored the Acc if it was treated (more) like a zero form while Nom had a more marked form. It is true that many forms in the *o*-stem paradigm would have contained *-u* after the merger with *-o* but forms such as the Gen singular/Nom plural did not, so it seems like a stretch to treat *-u* as part of the stem, making the Acc a zero form. Otherwise, none of the parameters would favor the outcome Herman proposes.

Sornicola considers this question to be unanswerable due to the nature of historical evidence. Another issue is whether this case variation was restricted to Africa, which may have lost length oppositions earlier than other parts of the Empire, resulting in syncretism between the Acc and other case forms (2011:34-35). Of course, this would imply that sound changes were important factors in these syntactic developments. A larger problem with this syntax-first approach is that later Romance languages, particularly OF, are attested with a two-case system that distinguishes Nom and Acc but no other cases. This system could not result if the Acc was extended to the Nom before the other cases. Even if syntactic factors did play a role in the loss of case distinctions, the extension of the Acc to functions associated with the oblique cases, such as object of prepositions and IO, appears to be a much more likely first step, as in Greek, Bulgarian, and Macedonian.

One hypothesis accounts for functional narrowing with a usage-based constructional approach. The basic unit of this model is constructions, which are form-meaning correspondences. The meaning of a construction can be derivable from its parts or not. In a usage-based approach, the frequency of a construction, especially its type frequency, is the basis for its status (Barðdal 2009:135). Semantic coherence also plays a role in the productivity of different argument structures: more productive constructions occur with a higher number of verbs but have less semantic consistency among these verbs. Low-frequency constructions may still attract new verbs if they are very similar semantically (ibid., 14-15). In Icelandic, for example, the Acc syntactic subjects of certain verbs have started being replaced by the Dat, e.g., *Mig langar > Mer langar* 'I long' where *mig* is Acc and *mer* is Dat. Both of these constructions are much lower in type frequency but semantically much more similar and coherent than Nom subject constructions, so the verbs with Acc subjects have adopted the slightly more frequent Dat

construction (ibid., 21). Based on a corpus study, Luraghi (2004:366) found in MyG that the Acc is more frequent than the Gen, followed by the Dat. The lowest frequency of morphological cases is the Instr. Assuming this case-frequency hierarchy is true for other languages, this can explain why the Acc absorbs other oblique morphological cases over time.

Suttle and Goldberg (2011) investigate three factors of syntactic productivity: type frequency, variability, and similarity. The variability of a construction is a measure of the semantic range of its attested instances; thus, it represents the inverse of Barðdal's (2008) semantic coherence. The similarity of a new instance of a construction to existing instances is a measure of how close it is in meaning. It can be calculated two ways: summed similarity, based on proximity to all instances, or maximum similarity, based only on the closest instance. This study primarily uses maximum similarity. Participants in this experimental study were presented with sets of sentences in a fictitious language with an English verb and definite articles but a nonce particle and nouns; they rated the chance of another English verb (the target) appearing in the same construction. Type frequency varied depending on the number of distinct verbs in each set: one, three, or six. Variability was either low, when the stimuli verbs came from the same semantic class, or high, when they came from three different semantic classes. Similarity was high, when the target verb came from the same semantic class as one of the stimuli, medium, when the target came from a related class, or low, when the target came from an unrelated class. All three factors were found to have an effect on productivity, but this varied significantly depending on their interaction. Type frequency had a stronger effect with high variability. Variability had a positive effect with medium similarity, indicating that a construction is more acceptable in a new class when it is attested in multiple classes that are similar to that class. However, it had a negative effect with high similarity, since low variability when the target is already in the same class as at least one attestation means exposure to more verbs in that class, and no effect with low similarity, since no level of variability can overcome the fact that no attestation is similar to the target. To account for the results above, Suttle and Goldberg introduce the concept of coverage, i.e., the density of a construction's attestations across the category formed by these attestations and the target (2011:1254-1256). These factors of productivity should be applicable to case constructions as well.

Perek (2016) uses distributional semantics to investigate diachronic changes in productivity. Distributional semantics is an application of the observation that words with similar

meanings have similar syntactic distributions. This approach is predominantly implemented using vector-space models, in which words are associated with arrays of co-occurrence counts. A co-occurrence matrix is built by comparing the collocates of a set of words, with each word as a row and each column as a correlate. Similarity between the rows then approximates semantic similarity. Similar words according to distributional semantics are not necessarily synonyms; they can also have other semantic relations such as antonymy, co-hyponymy, hypernymy. This is not a problem for modeling syntactic productivity, however, because the importance of various semantic relations has been demonstrated for syntactic productivity (see, for example, Zeschel 2012).

As a case study, Perek investigates the use of the *hell*-construction, i.e., "V *the hell out of* NP," from the 1930s, when its popularity began to increase, until 2009, the end date of the data source, the Corpus of Historical American English (2016:11-26). This construction likely developed from semantic bleaching of an exorcism scenario, i.e., *beat the devil out of*, in the late 19th century, with *hell* replacing *devil* by metonymy, since the two concepts are closely associated (Hoeksema & Napoli 2008). A cluster analysis of the historical spread of this construction to new verbs shows that the more initial members in a semantic cluster, the faster it grows. The two clusters with highest initial type frequency were forceful actions and psych-verbs. These were the two most productive clusters, while others did not reach a critical mass of members. The higher variability within the psych-class allowed it to be even more productive than the tighter cluster of forceful actions (2016:19-23). This study shows that the importance of semantic similarity in triggering morphological case mergers.

Using a text corpus, Barðdal has found that the Acc occurs in 3% of Icelandic subject constructions, the Dat in 10%, and the Nom in 85%. Thus, the lowest-frequency construction (Acc) is being attracted by another low-frequency construction (Dat)—not by the highest-frequency construction (Nom)—due to the semantic similarity (and restrictedness) of Acc and Dat subject constructions, in contrast to the semantically open Nom subject construction (ibid., 21-22). These principles account for developments in Germanic languages: those that maintain morphological case, i.e., Icelandic, Faroese, and CSHG, still underwent changes in case-assignment constructions, losing low type frequency constructions as those verbs were attracted by high type frequency constructions with semantic overlap; those that lost morphological case, i.e., English, Dutch, and continental Scandinavian, took this even further, eventually merging all

argument structures that differed only in case marking (ibid., 14-15).[114] In other words, they underwent syntactic syncretism, as Luraghi (1987) proposes is characteristic of Germanic languages. A rapid increase in vocabulary due to borrowing in the latter group of Germanic languages may have caused their more extreme development, as most new verbs were attracted by the high type frequency constructions, leaving low type frequency constructions with an even lower proportion of verbs (Barðdal 2009:138-139).

It is possible that the influx of loanwords also affected the frequency of gender categories in English, Dutch, and continental Scandinavian, since these languages have all lost at least one gender category. In Russian, for example, the proportion of masculine nouns has increased at the expense of the neuter because far more loanwords are treated as masculine (see Corbett 1991:81-82). Similarly, as mentioned in section 2.5.2.1 above, English loanwords with final consonant are treated as masculine in Slavic (Vakareliyska 2018:358). Likewise, the masculine in French has expanded from 51% of nouns to 61% since the 17th century, largely due to English loans, which are mostly treated as masculine (Corbett 1991:82). These changes have not caused the loss of the neuter in CSR or feminine in CSF, but these categories could be lost in the future, particularly the former. In CSR, only 13% of nouns are neuter and new nouns have become masculine or feminine. In addition, the neutralization of unstressed *o* and *a* means several common neuter and feminine forms on nouns and agreement targets are often indistinguishable (ibid., 316-317), as mentioned in section 3.1.1.2 above. Similar changes in frequency may have contributed to the merger of the masculine and feminine in continental Scandinavian and Dutch, as well as the total loss of grammatical gender in English. This may have contributed to its eventual loss, likely in combination with an increase in syncretism between the masculine and feminine due to sound change and analogical processes.

The usage-based constructional approach satisfactorily accounts for functional narrowing in the Germanic languages and its connection to the general loss of morphological case. It is

[114] Classical Arabic may represent the stage of a language that has syntactically-predictable case marking: the Nom subject and Acc object could also be reliably distinguished by word order, and all prepositions took the Gen, which was also used for possession constructions. While it is unclear if this system was the result of functional narrowing, the subsequent loss of case marking adds further support to the idea that case marking is more susceptible to total loss once it is redundant with syntactic marking of grammatical relations. As mentioned in section 3.1.2.4 above, the loss of final vowels in the development of Arabic eliminated case marking in most contexts. Since case marking was already redundant, there was no motivation for it to be restored by analogy, e.g., to the few contexts where it survived the sound changes. Instead, case variation temporarily arose in those contexts until all case marking was lost.

grounded in concrete and measurable variables, i.e., constructions and their type frequencies, as opposed to an abstract distinction between two types of case. Furthermore, it explains how semantic and syntactic motivations work together with contact to bring about case loss: languages with a higher degree of vocabulary replacement are more likely to have cases that undergo functional narrowing, which results in case loss. However, it still does not explain why some languages will resist this.

Many of the languages I discuss in this study experienced a significant influx of L2 speakers around the same time that the loss of productivity and other functions appeared to accelerate. MLG speakers's use of continental Scandinavian varieties as an L2 beginning in 1250 could be connected to the functional narrowing of the Dat and Gen around this time. Likewise, a large shift of British Celtic speakers to OE may have occurred at an early enough date to trigger the decline in the productivity of these cases, and the later contact with ON speakers could have contributed as well. The extension of the Acc at the expense of other cases in the Latin of late 2nd century CE Romanized Africa also occurred not long after speakers in this region shifted to Latin. Based on the transition from Libyan and Punic names to Roman names in inscriptions in this region, Varvaro concludes that Latin was overtaking local languages during the 1st century CE (2011:14-17). Intense contact between MB and other Balkan languages, particularly Pre-Romanian, had begun by the 12th century at the latest. The 12th century is also when written evidence from BChS manuscripts starts to indicate a significant functional narrowing of the Instr, Loc, and Gen, ultimately leading to the functional mergers described in the previous section. Of course, the relative chronology of the influx of L2 speakers and functional narrowing is very important. L2 speakers cannot have triggered changes that were already in progress when they started using the language, although they still could have accelerated the process. The potential effects of L2 speakers and the historical contact situations mentioned here are discussed further in section 3.2 below.

3.1.2.7. Analytism

Languages and the structures they contain can be broadly classified as synthetic, i.e., generally marked with inflections, or analytic, i.e., generally marked with free morphemes, in the case of explicit analytism, and word order, in the case of isolating, or implicit, analytism (Lindstedt 2014:170). One hypothesized cause of case loss is the general development from synthetic to

analytic structures. Both types of analytism have been attested as compensation for lost inflections in IE languages, but more often this entails the substitution of periphrastic expressions for inflections, resulting in explicit analytism. Only the least marked cases such as Nom and Acc are regularly compensated by word order alone. With respect to functional mergers without the total loss of the inflections involved, the expression of the marked function is much more likely to be analytically compensated than that of the unmarked function, for which the inherited inflection usually suffices (see Meiser 1992:198-199). These changes may in turn be motivated by certain types of contact.[115] For example, certain OE case uses were replaced with prepositional constructions in EME. Even in OE, there are examples of Dat marking assigned by a verb being reinforced by a preposition that takes Dat, e.g., *to* (see, for example, Van Gelderen 2000:214). Later, after a period of competition between complements in which the Dat marking is reinforced by a preposition and those in which it is used alone, i.e., a period of optional compensation, to use Meiser's (1992:197) terminology, some verbs that once assigned the Dat came to be used with prepositional constructions instead.

According to another view, as the system of morphological case marking broke down for other reasons, the greater explicitness of prepositional constructions over those relying on the moribund case system led to an increase in the former, e.g., to mark the recipient (Zehentner 2017:19). If this account is accurate, analytic tendencies cannot be the primary motivation for case loss, but they may have accelerated the total loss of morphological case. In fact, this motivation corresponds to Kulikov's (2013) Formal Markedness constraint. More generally, like the total or partial loss of a case's functions to another case, as described in the previous two sections, respectively, the displacement of a case's functions by analytical constructions first passes through a stage of variation, during which Kulikov's parameters may determine the outcome. If they can be shown to do so, that would add further support against a general analytic tendency; instead, analytic constructions displace case marking only when that is the optimal outcome based on the interaction of the relevant parameters.

Zehentner's (2017) proposal is supported by studies of sentence interpretation from a synchronic, child-language acquisition perspective. In a study of sentence interpretation by Hungarian preschool children, MacWhinney et al. (1985) considered how the cues of case marking, word order, animacy, and intonation were used to determine the grammatical roles of

[115] See section 3.2 below for further discussion of how contact may motivate analytism.

NPs. Their competition model assumes that each NP encountered in the processing of a sentence is a candidate for a particular grammatical role, e.g., subject. As more of the sentence is parsed, the strengths of the different candidates are updated based on cue validity. Cue validity can be defined as the product of cue availability and cue reliability (see MacWhinney 1978, MacWhinney et al. 1984). MacWhinney et al. found that animacy is an important cue for the youngest Hungarian children but declines in strength up to age 6. This follows from its lower availability, since subject and object can both be animate, and lower reliability, since the combination of an inanimate subject and animate object is possible if relatively rare. Word order is a weak cue at all ages, as expected given the essentially free word order of the major constituents. Adults and 6-year-olds, but not younger children, use stress as a cue when more reliable cues are not available. Finally, Hungarian children rely heavily on case marking from a young age and almost exclusively by age 6. This result is consistent with the full reliability and availability of case marking, namely *-t* for Acc in opposition to a zero-ending for Nom (1985:185-195). However, the younger children are not as reliant on case marking as Turkish children in a similar study by Slobin and Bever (1982), despite similar validity.[116]

In two additional experiments, MacWhinney et al. found that the second person singular possessive marker *-d* is sometimes interpreted as Acc *-t*, raising the issue of phonological detectability as an additional factor in cue strength: even if a cue is available and reliable, a child is less likely to rely on it if it is more difficult to detect. They ascribe the later acquisition of the Acc in Hungarian than in Turkish to lower detectability in Hungarian (1985:195-207). Overall, these results support the competition model of sentence interpretation, which predicts that case marking will be given less weight than other cues if it becomes unreliable, less available, or potentially even less detectable. It may be acquired later and later until it is not acquired as a productive form at all. This is one mechanism by which the loss of some case distinctions due to sound change could increase analytic tendencies and lead to the breakdown of the case system. Notably, this mechanism does not require sound change to neutralize all of the lost case distinctions directly.

As another example in the development of English, the preposition *of* began to be used as an alternative to the Gen: it is attested in partitive and possessive constructions no later than EME, and possibly even in LOE (Allen 2008:74). However, potential early examples may also

[116] Slobin and Bever (1982) designed sentences with definite objects, so the Acc suffix appeared consistently.

be interpreted as an original use of this preposition to indicate origin, which lends more credence to the theory that the use of *de* to indicate possession in Norman French, as in the OF example *la renomee de cel saint homme* 'the fame of that holy man', influenced or at least reinforced this development (Mitchell 1985: §1202-1203, Rickard 2003:59). This kind of French influence on EME morphology would suggest that contact with French may have played a bigger role in English case loss than is currently accepted. As expected, in a corpus linguistics study that I conducted of four EME texts, I also found examples of prepositional phrases being used with verbs that previously assigned the Dat or Gen to their sole or primary objects; some of these verbs were not attested with prepositional complements in OE, e.g., *missan* 'miss', while others were already attested with both bare and prepositional complements, e.g., *reccan* 'care for'. Both of these verbs assigned the Gen in OE but are found with *of*-phrases in Brut and/or The Owl and the Nightingale, in addition to bare Acc and Dat objects. This study is discussed further in section 5.2 below.

Standardization may work to preserve synthetic constructions that face competition from analytic alternatives. In a corpus study of the Gen and alternative constructions in High German and Dutch, Scott (2014:311-324) has found that the Gen was becoming more restricted in its use but experienced a resurgence at the expense of other constructions in the 15th and 16th centuries, around the time when standardization was taking place. In contrast, English and Swedish had already lost morphological case completely when standardization occurred. These facts form the basis for Scott's argument that standardization is a factor for language usage, even in everyday speech, and can therefore disrupt morphosyntactic change. With respect to High German and Dutch, the functions of the Gen were losing productivity in the same order as in English and Swedish, i.e., adverbial and adjectival uses were weakened earliest, adnominal uses last. However, case loss was delayed long enough, perhaps due to less intense contact in the Middle Ages, that case marking was still available for inclusion in the early norms developed by printers and in prescriptive grammars, leading to its resurgence, particularly in formal registers but also in informal language. In addition to everyday usage, exposure to normative language through education forms part of the input for language users, although it is not clear if this information has the same status in a user's mental grammar given that it is not acquired as a young child. Nevertheless, Dutch eventually lost all productive case marking, including the adnominal Gen. This use survived the longest but came to be limited to formal writing by the 19th century and

was further restricted to a few fragments with the spelling reform of 1947, which removed it from the standard language (ibid., 157, 206). In contrast to CSHG, Modern German dialects almost all lack a productive Gen case (Behaghel 1928:47). Most still have some morphological case distinctions, however (Scott 2014:209). These facts suggest that standardization can contribute to the rise and fall of particular constructions, both synthetic and analytic, but that is just one of several factors. Scott fails to account for the different outcomes in Dutch and High German, however.

Several analytical means for marking arguments have developed as alternatives to case inflections in the Balkan Sprachbund languages. In MB, the replacement of the Gen-Dat by the prepositional construction with *na* 'on' followed by an Acc was perhaps the most significant analytical development, as it overlapped with the merger of the Gen and Dat. Duridanov considers how and when *na* was grammaticalized in this function and why another preposition did not replace the Gen-Dat instead (1956:226-238). In other languages, the replacement of the Dat with a prepositional construction seems to have begun with reinforcement by a preposition already used with the Dat, e.g., *to* in English. Likewise, the first step in the loss of the Instr in MB was the earlier instrument-accompaniment merger, which can be viewed as the reinforcement of bare Instr uses with the preposition *sъ* 'with' (see ibid., 197-198). Therefore, the preposition *kъ* 'to', which took the Dat and appears to have increased in use at the expense of the bare Dat between early OCS manuscripts and later OCS and BChS manuscripts, may have been expected to continue this trend and fully replace the Gen-Dat (ibid., 227). Duridanov questions whether *kъ*, which was used to indicate goal, could even be extended to the more core Dat function of recipient (1956:227). On the other hand, the choice of *na*, a preposition not normally used with the Dat, as the replacement for the Dat also has parallels in other languages. For example, *ad* 'to' followed by the Acc competed with the Dat in VL and came to replace it in the Western Romance languages. The earlier goal-location merger meant that *na* with Acc could already be used in a broader array of functions. At least some of these functions overlapped with the Dat and Gen complements of certain verbs. While it likely began with the overlap between the goal functions of the Acc and Dat, the *na*-construction could then be extended to other functions of the Gen-Dat, including the core functions of recipient and possessor (ibid., 227, 232-233).

As for the timing of this development, Duridanov, among others, interprets the use of the prepositional *na*-construction in BChS manuscripts where Dat would be expected as evidence for an already-developed analytic alternative to the Dat in 13th-century spoken MB (1956:229-230). However, Lunt does not find credible evidence in these manuscripts for the use of *na* with recipients, the core function of the Dat (1965:306-307). Rather, its increased use occurred with mental state verbs, e.g., *radovati sę* 'to rejoice in/at something', *blagovoliti* 'to content oneself with', and *nadějęti sę* 'to trust in', which canonically take Dat objects in OCS. Thus, the total loss of the Gen-Dat form in favor of the *na*-construction likely did not occur until later than proposed by Duridanov. This is supported by Gyllin's finding of competition between *na*-constructions and the Gen-Dat in the damascenes, with the Gen-Dat attested longest in the IO function (1991:78-79). In fact, this pattern of functional narrowing is consistent with developments in other languages, and with the results of Vakareliyska's (1990) study of Wernicke's aphasics, which is discussed further in section 5.1 below. The core functions of cases are usually the last to be lost, at which point a total functional merger occurs, or they are retained but the case system carries less functional weight. In a confirmed instance of contact-induced change, Luraghi and Krstić (2018) find that Molisean Croatian (MC) has lost all but the IO function of the Dat, as described in section 3.2.4 below. Thus, the similar patterns in the development of Bulgarian and Macedonian, as well as the loss of syntactic productivity in Germanic languages described in the previous section, where the core function of a case may remain productive much longer, suggest that L2 speakers played a role, as in MC.

Another way to mark arguments analytically is clitic doubling of objects. Many IE languages have clitic doubling, especially Romance languages like French and Spanish (De Boel 2008:102). This is only possible if there are two parallel sets of long-form and clitic pronouns (Wahlström 2015:118, citing Guentchéva 2004:25). All South and West Slavic languages have clitic pronouns, but only those Slavic varieties within the Balkan Sprachbund have clitic doubling. In CSM, as well as the non-Slavic Balkan Sprachbund language Albanian, this construction is obligatory for IOs and other constructions with the Dat clitics. The level of grammaticalization of DO doubling varies among Balkan Sprachbund varieties (see Wahlström 2015:118-119). Friedman (2008:58-60) argues that the most grammaticalized use of clitic doubling in the Balkan Sprachbund, and therefore its center of innovation, is in Western Macedonia, the point of contact between western Macedonian, Central Gheg Albanian, and

Northern Aromanian. Clitic doubling is more restricted farther from this area: to the south it is inhibited by the conservative influence of Greek, in Eastern Bulgarian dialects it apparently only marks topicalization, and in Romanian it is limited by conditions such as partitivity and humanness.

In studies of grammaticality judgements by native speakers of CSB, Vakareliyska (1994, 2002) has shown that clitic doubling has also developed an argument-marking function. First, DO doubling is more acceptable when the object comes before the subject, where the Acc clitic serves to disambiguate the grammatical roles. Second, it is becoming more acceptable to omit *na*, the normal marker of IOs and other dative functions, when a Dat clitic is used. Given this tendency towards complementary distribution with the dative marker *na*, Dat clitics also mark this grammatical relation. Wahlström argues that clitic doubling in varieties of Bulgarian and Macedonian is most likely due to contact with the other Balkan Sprachbund languages. Although clitic pronouns arose independently, since they are well-attested in Romance languages as well as OCS and other Slavic languages, clitic doubling constructions are not found in Slavic languages outside the Balkan Sprachbund. Despite some differences in semantic criteria among the Balkan Sprachbund languages that reflect different degrees of grammaticalization, all of these criteria relate to information structure, which suggests a common origin (2015:128-129). Even in languages such as CSM and CSB, in which clitic doubling can disambiguate grammatical relations, information structure still plays an important role. For example, only [+specific] DOs are doubled in CSM (Cyxun 1968:112). Likewise, DO doubling in CSB still concerns topicality, particularly when the grammatical relations are already clear, e.g., when a clitic doubles a long-form pronoun, already formally marked as non-Nom. As Vakareliyska argues, however, long-form pronouns are already discourse-level Topics, even when not sentence-initial, i.e., in Topic position. When they are emphasized by a clitic, they are not being topicalized, therefore, but marked as a Focus (2002:3).[117]

A third type of analytical marking, mentioned above, is the use of prepositional constructions to mark DOs with human referents. Romanian uses the preposition *pe* followed by a noun in the Nom-Acc form. The Ohrid-Struga dialect of Macedonian has developed a construction with the preposition *na*, similar to a construction with the preposition *pi* in the local

[117] The primary Focus is usually sentence-final, so unless sentence-final, a doubled long-form pronoun is the secondary Focus.

Aromanian dialect. While *na* is also used to mark IOs in Macedonian dialects, both Macedonian *na* and Aromanian *pi* are used in locative constructions. This parallelism suggests that this additional function of a locative preposition, along with its specific semantic criterion, spread to the Ohrid-Struga dialect from the Aromanian dialect via structural borrowing (Wahlström 2015:127-128). The development of this construction in Romanian varieties parallels the construction with *a* in Spanish, discussed in sections 2.4.1.2 above and 3.1.2.4 above. Both were probably responses to at least some degree of case loss rather than primary causes.

Along with the borrowing of the Gen form to mark the animate masculine Acc in modern Slavic languages, the analytical strategies of clitic doubling, and prepositional marking of human DOs are examples of differential object marking (DOM), i.e., marking some objects different from others, despite there being no difference in their syntactic function. These differences are based on the referentiality of the object NP, its inherent reference, including animacy and humanness, its discourse relevance, such as definiteness and topicality, or its word-class, e.g., pronouns vs. nouns. As mentioned above, the specific criteria vary by marking strategy and by language within the Balkan Sprachbund. The question, then, is how much of the similarity among these languages can be attributed to contact, given that DOM is not typologically rare (Wahlström 2015:122-123, 129). Wählström relies on a quantitative analysis of 435 languages by Bickel et al. (2014:9-16) that tested whether there is a universal correlation between an argument's rank on the referential scale and its probability of being overtly marked, either with case or analytical constructions. Specifically, the researchers expected that objects high on the referential scale and subjects low on the scale were both more likely to be overtly marked. They found that 63% of the 149 splits in marking among different objects occur in only five language families, which include IE and fell into two proposed linguistic macro-areas, Eurasia and Sahul. Therefore, they concluded that the effect of the referential scale no longer appeared universal when potential genetic and areal connections among languages are excluded (ibid., 17-18). Based on these findings, Wahlström argues that referential scales can spread in micro- as well as macro-areas. The evidence from Bickel et al. (2014) suggests that the spread of referential scales was involved in the rise of clitic doubling and prepositional DOM in the Balkan Sprachbund languages. In contrast, there has not been a shared innovation in the Balkan Sprachbund to restore a distinctive Acc for objects high in referentiality. It is true that in CSB, at least, non-reduplicated long-form pronouns can be used instead of the unmarked clitic to add emphasis to

an object, making it higher in referentiality (see Vakareliyska 1994:125). This is simply the use of inherited forms, however, and does nothing to restore a distinctive Acc to nouns. Wahlström concludes that the increased chance of DOM in definite forms is most likely a by-product of grammaticalized definiteness rather than the spread of the referential scale, and that any distinctive Acc forms remaining have avoided neutralizing sound change (2015:126-129).

These developments across the Germanic, Romance, and Balkan Sprachbund languages certainly represent an increase in analytism, but it is not clear if this increase was primarily a cause or an effect of case loss. One possibility is that the initial reduction in the distinctiveness of case endings resulted from sound change.[118] When these developments occurred, some analytic structures may have already been used—whether developed internally or acquired through contact—but only as a less frequent and more syntactically constrained alternative to case marking. Compared to a reduced case system with a sufficient level of syncretism, these analytic structures would have become more preferable and expanded in frequency and acceptable contexts, eventually leading to the complete loss of the remaining case endings as they fell into disuse. Moscoso del Prado's (2014) Granger-causality analysis, which indicated that syntactic changes were more a result of morphological changes, might be extended to other languages to determine whether their increase in analytism was more of a cause or effect of case loss.

However, not all of the developments in these languages have been from synthetic to more analytic marking. New synthetic structures may have arisen due to the Germanic stress pattern: first the weak (regular) preterite from unstressed forms of 'do' and later the Scandinavian mediopassive from reflexive pronouns (primarily *sik*) suffixed to the verb, as well as the suffixed definite article (Norde 1997a:34, citing Van Haeringen 1962:48).[119] In addition, Germanic languages, including those without morphological case, often express genitival relations through compounds. In Dutch, for example, the concept of 'patriotism' is expressed with the compound *vaderlandsliefde*, while CL, a highly inflected language, requires two

[118] The specific effects of sound changes on case marking in the languages under consideration are described in chapter II above. The potential for a more general effect of prosodic change on case endings in these languages is discussed in section 3.1.1.1 above.

[119] Ringe argues that weak preterite forms derive from univerbation of a past participle and the imperfect **ded-* 'did' due to haplology: PGmc **salbōd(a d)(ed)ē* '(s)he anointed > **salbōdē* (cf. Gothic *salboda*) (2006:166-168). The Scandinavian mediopassive can be observed, for example, in CSS verb forms ending in *-s*, e.g., *kallas* '(s)he is called' < ON *kalla-sk* (Norde 1997a:32, citing Van Haeringen 1962:48). See section 3.1.2.3 above for discussion on the grammaticalization of the definite article.

separate words, e.g., *amor patriae*, meaning that CL could be considered more analytic than Dutch in this particular aspect (Norde 1997a:32, citing Van Haeringen 1962:57). Compounds are also one of the main ways to express genitival relations in CSHG. However, new synthetic case marking was unlikely to arise in the way that the postpositive definite article and verbal structures described above did because it would have required prepositions, the primary analytic compensation for case marking, to be reanalyzed as prefixes. Since nominal inflection has used almost exclusively suffixes, as well as some stem-inflection such as umlaut, adding new prefixes, as explained by Meiser (1992:205-206) and Wurzel (1984:87), would not be tolerated according to the principle of system adequacy.

Another argument against the importance of increasing analytism in case loss is that the verb systems were not simplified in the same way. CSB and CSM in particular have preserved highly synthetic verb systems despite undergoing case loss. It is also true that the Romance verb systems have remained more intact than their nominal systems, but they have still undergone many changes, most of which involve an increase in analytism. The Latin synthetic future, perfect, and passive systems were all replaced with analytic constructions, although the future construction was subsequently grammaticalized into new inflectional forms, which indicates that this is a highly relevant category for Romance speakers. Thus, analytism has increased in the verb systems of all Romance languages. According to theories presented in section 3.2.3 below, this could be partially due to the influence of L2 speakers. More extreme developments have occurred in some languages. In many phonetic contexts in spoken French, verbs have lost most person/number endings. Instead, clitic subject pronouns are required, in some dialects even when there is an overt subject noun. The loss of endings can be largely attributed to sound change, as with nouns, but replacement by analytic alternatives rather than repairing the inflectional system might indicate the effect of L2 speakers.

Germanic languages only preserve the IE tense distinction between past and non-past. PGmc lost most of the complex IE tense and aspect system. English lost almost all subject-verb agreement and the continental Scandinavian languages lost all subject-verb agreement. These simplifications could be the result of the same factors as case loss, including sound change and/or L2 speakers. The complex tense and aspect system found in ModE is the result of later innovations, primarily expressed through explicit analytism, i.e., auxiliaries. The languages that retain morphological case also retain subject-verb agreement to a comparable degree. In the

present active indicative, for example, Icelandic verbs regularly have five distinctive personal endings, CSHG verbs regularly have four, and Faroese verbs have three to four. The only exception is Dutch, which retains three distinctive personal endings in the present active indicative but has lost all case marking. Then again, Dutch lost morphological case relatively late compared to English or the standard continental Scandinavian varieties, and CSDu still retains three grammatical genders. Just as analytic constructions such as prepositions and clitics came to signal some of the same grammatical relations previously expressed by case inflections, most of the IE tense and aspect categories, or at least categories resembling them, were only retained through analytic constructions. For Germanic as well, then, it cannot be said that nominal inflection was simplified but verbal inflection was not. In conclusion, these outcomes for verbs do not necessarily rule out the effects of L2 speakers (or mutual bilingualism, perhaps) since they involve the same tendencies for explicit analytism, if to a lesser degree than in the nominal system. It is possible that the tendency for more extreme analytic developments with respect to nouns than verbs has its origins, at least in part, in PIE; at this stage, there was already a relatively high level of case syncretism in nominal inflection, but syncretism within verb paradigms was quite rare (see Meiser 1992:193). At the same time, the mechanism is clearly not as extreme as true creolization for any of the languages under consideration.

3.2. External Motivations

External motivations have been proposed to explain case loss in the Germanic, Romance, and Balkan Sprachbund languages, especially to account for the vastly different outcomes in different languages, despite starting with the same case system. These motivations are almost all related to the degree to which contact with other varieties occurred, from relatively complete isolation to extremely intense contact with multiple other varieties, and the sociolinguistic nature of these contact situations. Historical and other non-linguistic evidence often informs arguments about the nature of a particular contact situation, which in turn indicates what aspects of case loss and other linguistic developments can be attributed to that contact. However, the observed changes that are assumed to be the result of contact are often cited as evidence for the nature of that contact. Thus, there is a danger of circularity, as with most arguments concerning the motivations for case loss. This section evaluates various proposals about the nature and results of contact experienced by these languages, focusing on how case and gender reduction can be

explained. Specific contact situations and outcomes under consideration include dialect contact, sprachbunds, and structural convergence. The final subsection considers how developments in modern Slavic varieties in contact situations relate to similar developments in the languages under investigation, especially Bulgarian and Macedonian.

Within North Germanic, contact seems to have separated the more innovative continental varieties that have lost morphological case from the more conservative insular varieties that retain morphological case. For the development of continental Scandinavian, the most significant contact was with MLG. Jensen (1983:670) attributes the reduction from four to two cases, earliest in MDan, in part to the effects of the language mixture of the craftsmen and tradesmen who had moved into Scandinavian towns. In Sweden, this influence began with German settlement of cities starting in 1251, particularly in Stockholm, and peaked with the accession of a German king to the Swedish throne in 1364 (Norde 1997b:390-391). It is not clear what proportion of the population was German during this period, but by the 15th century the German population had grown high enough in influence, if not numbers, to control half the seats on the city council (Friman & Söderström 2008:28). This contact definitely resulted in a high number of loanwords and may have also influenced loss of inflection through the confusion of endings by L2 speakers. Notable differences included the use of the masculine Nom *-er* ending in MSw, e.g., *dagher* 'day', (cf. CSS *dag*) but not in MLG, e.g., *dach* (cf. CSHG *Tag*); and the use of postpositive definite articles in MSw as opposed to preposed articles in MLG (Norde 1997a:40-41). Under the right conditions, the influx of loanwords may also be a factor in case and gender loss in and of itself, as discussed in section 3.1.2.6 above. The main argument for the importance of this contact is that the decline of native inflectional systems was most advanced where contact with MLG was most substantial. Peripheral varieties like Icelandic and rural Swedish dialects did not change as much. Within Norwegian dialects, the retention of the definite Dat in inland dialects can be explained by their general tendency to preserve older systems, whereas coastal areas, with more outside contact, lose the Dat because they lose formal distinctions in general (Knudsen 1967:12).

However, the development of Faroese shows that isolation alone is not enough to motivate conservatism: despite being almost as isolated as Icelandic, Faroese is less conservative (Norde 1997a:35). The conservatism of Icelandic may have resulted from additional factors. Icelanders lived on isolated farms, meaning that dialectal differences could not develop (Norde

1997a:35, citing Torp 1982:124-125). This situation contrasts with the fishing villages and lack of a standard written language on the Faroe Islands (Haugen & Markey 1972:99). A homogenous monolingual community like the one in Iceland can maintain complexity in a way that is not possible when speakers must command multiple varieties (Werner 1984:220). Lithuanian is another conservative IE language in terms of both phonology and morphology (Balode & Holvoet 2001:43). As with Icelandic, most native Lithuanian speakers were not exposed to dialectal variation because they did not have much contact with each other on an individual level during much of the language's development, which suggests that lack of contact may be an important factor in conservativism. On the other hand, Faroese was also probably not as isolated from outside influence as Icelandic: Danish has clearly influenced the vocabulary of spoken Faroese (Barnes & Weyhe 1994:190). In contrast, Icelandic has few unassimilated loanwords from any language (Thráinsson 1994:188). There is also evidence that Danish has influenced Faroese structurally, as discussed in the following subsection.

The case systems of all the West Germanic languages have probably been affected by contact to some extent, but the intense nature of the contact experienced by English is often cited as the distinguishing factor from High German, which maintains morphological case and grammatical gender, and to a lesser extent from Dutch, which only maintains the latter. In the development of English, contact with at least three linguistic varieties most likely played a role: British Celtic, ON, and Norman French. Contact with British Celtic started in the 4th century and became widespread in the 5th. Contact with Danish and Norwegian invaders and settlers occurred primarily in the 9th and 10th centuries, with evidence from hundreds of Norse place names. Finally, contact with Norman French took place after 1066. Due to the high number of loanwords from ON and especially Norman French, these languages are traditionally assumed to have also influenced the structure of English more than British Celtic. However, Norman French speakers were never a very large percentage of the population, while there is reason to believe that speakers or former speakers of British Celtic constituted a high percentage (Trudgill 2010:8-9, 12). Furthermore, the prestige of Norman French made it a desirable source for loans, and the eventual shift of French speakers to English resulted in lexical transfer (Skaffari 2017:192). However, structural changes, e.g., case loss, are not as typical when a shift from a superstrate like Norman French occurs (Thomason 2010:37).

The role of contact in Western Romance case loss is unclear. Frankish, a West Germanic language, is generally considered a superstrate for French, Occitan, Catalan, and Italian. Sala (2013:203) finds this description questionable, however, because the Franks were dominant politically, but their language was less prestigious than Latin and had far fewer native speakers. Frankish nobles shifted to Romance, but even if their learning was imperfect, the much larger group of native speakers would not be motivated to imitate their speech. Thus, the historical and sociolinguistic context is problematic for the view that Frankish had structural influence on northern Gallo-Romance (see Wartburg 1950).

Sala distinguishes between incontrovertible and alleged effects of this contact. Frankish clearly had an effect on the French lexicon, providing six to seven hundred lexical loans. To a much lesser degree, functional vocabulary was also borrowed, including a few adverbs and derivational affixes, perhaps due to their expressiveness. Phonological effects such as the introduction of word-initial /h/ are clearly identifiable because this new phoneme occurred in Frankish loans, still survives in some dialects, and has residual phonotactic effects in standard French. The deletion of unstressed vowels and diphthongization of certain vowels in stressed open syllables in northern France may be the result of Frankish influence, specifically the heavy stress with exaggerated lengthening of open stressed syllables. This is less clear, however, since vowel reduction began in VL and similar outcomes are found in northern Italian dialects. These similarities could be due to Longobardic influence, but the details of this language and the contact situation are unknown, so the strongest claim that can reasonably be made is that the Germanic pronunciation patterns might have supported existing trends in Romance speech (2013:204-207). While OF still retained a two-case system, the set of endings was very limited due to the loss of unstressed vowels. Thus, a strong stress could have contributed to the opaque syncretism patterns that led to the breakdown of that system.

On the other hand, Sala (2013:208) mentions that Hilty (1968) attributes the longer retention of case distinctions in OF than other Western Romance languages to Frankish support. Sala (2013:208) considers two pieces of evidence in favor of this hypothesis. First, the two-case case system appears to have survived longest in northern and northeastern France, where Frankish influence would have been strongest. On the other hand, the same system also survived in southern France and possibly even early medieval Italo-Romance, outside the area normally assumed for strong Frankish influence. Second, Frankish proper names often adopted the OF

declension pattern with alternating stress distinguishing the Nom from the Acc, e.g., *Hues* ~ *Huón*, which is similar to the weak declension followed by these names in Frankish, e.g., *Hugo* ~ *Hugun*. The latter pattern did not involve stress alternations, however, and the use of the *-ón* suffix on Frankish names could have been an extension from native Romance words.

Case loss in all of the Germanic, Romance, and Balkan Sprachbund languages under investigation has been demonstrated to be motivated at least in part by phonological and morphological factors. At the same time, contact appears to have played a role for some of these languages. One theory as to how these internal and external factors interact was proposed by Jakobson: structural elements only spread when they correspond to the borrowing language's evolutionary tendencies (1938). Thus, contact might trigger or accelerate developments that would occur anyway (Weinreich 1953). While authors such as Cassano (1976) have argued against this theory, it can be seen in arguments such as Kulikov's (2013) that an existing pattern of case variation can be pushed towards a particular outcome by contact. For example, if multiple Balkan Sprachbund languages had variation between the Gen and Dat for possessors and recipients, once this developed into a functional merger in one language, it would be likely to trigger the same outcome in the other languages.

3.2.1. Dialect Contact

Contact between genetically related languages is considered dialect contact. Thomason (2001:77) argues that inflections spread faster between related languages than the contrary. The contact between MSc and MLG can be described as dialect contact (Jahr 1994:32-33), a phenomenon under which O'Neil (1978:269) observes that inflectional differences are rapidly neutralized. Jespersen (1935:75-76) attributes the more advanced loss of inflection in The Danelaw to contact between OE and OSc. However, there was already some loss of inflection in the Northumbrian dialect of OE, before contact with Scandinavian became significant (Thomason & Kaufman 1988:280, 303). Instead, this most likely resulted from internal changes and/or earlier contact with British Celtic.

One way that dialect contact can manifest is through semicommunication. Braunmüller (1989) demonstrates that the contact between MSc and MLG meets all the conditions for semicommunication: communication between speakers of many genetically related varieties, a preference not to use a lingua franca, and tolerance of other varieties. Semicommunication

among Scandinavians was already commonplace, so MLG could be included as well. The phonological and morphological similarities between MLG and MSc show that this was possible. For example, the *a*-stem paradigms in MLG and MSw were very similar. When contact for business began, MLG already had reduced inflection (Braunmüller 1995:63). To a lesser extent, MSw also had reduced inflection (Wessén 1968:136ff.). The major difference was the presence in MSw of the *-er* ending that distinguished masculine Nom from Acc, while MLG used the same form for both. Thus, this *-er* ending may have been interpreted by Low Germans as a derivational agentive suffix or as part of the stem (Braunmüller 1995:51). While this suggests that dialect contact may have led to the loss of *-er* in MSw, this ending was already lost in some OSw texts, before contact began, and survived in later texts, even after the influence of MLG had ended (Wessén 1968:137-141). The suffixed definite article may have been another source of confusion for Low Germans, but they may not have perceived the suffix and could just ignore it if they did (Braunmüller 1995:49). Although it was most likely possible for speakers to overcome the inflectional differences between MLG and MSw, there is no conclusive proof of mutual intelligibility (Norde 1997a:41). Thus, it is unclear what the nature of the contact between these languages was.

Another instance of dialect contact may explain the beginnings of certain morphological simplifications that surfaced in ME. As mentioned in section 3.1.2.4 above, there is evidence outside of case marking that dialect mixture occurred in the early development of OE. The use of different cases with the same preposition without any difference in meaning may be an example of case variation resulting from dialect mixture (Mitchell 1985: §1177(2), §1222). A result of dialect contact, such dialect mixture leads to regularization but not loss of morphological categories (Trudgill 2010:7). Thus, the contact and mixture of the different dialects spoken by the initial Germanic settlers of England can only explain certain simplifications, such as the reduction in nominal declension classes, but not the loss of morphological case or gender.

Dialect contact may lead to koineization, the stable result of the mixing of linguistic subsystems, which is used as a lingua franca, has a mixture of features from the contributing varieties, and is simplified compared to those varieties (Siegel 1985:363). The prototypical example of koineization is KG, the form of Greek that spread throughout the Macedonian Empire. KG has its origins in Great Attic, the result of mixing between the Ionic and Attic dialects. The convergence of these dialects began during the 5th century BCE due to Athenian

power over Ionians of the Maritime League (Adrados 2005:175-177). The resulting Koine experienced changes to its phonology, morphology, syntax, and lexicon, many of which can be characterized as simplification. While the loss of the dual and simplifications to the verbal system appear during the Hellenistic period (4th-2nd century BCE), most of the changes to nominal declension, including a fair amount of paradigm leveling, are only attested in the Roman period (2nd century BCE - 4th century CE) at the earliest. The Acc began to replace certain uses of the other cases, while the Gen tended to survive mainly in adnominal uses. The Dat did not disappear until later, during the Byzantine period, but the confusion with the Acc on objects of prepositions occurred much earlier, as did a general increase in the use of prepositional constructions (ibid., 193-196). As with dialect contact among Germanic varieties, including the possible koineization in the transition to EME discussed below, the dialect contact and resulting koineization of Greek accounts for certain simplifications, but not the complete loss of any case marking, for which other motivations are also required.

Dawson (2003:46) argues that the contact between OE and ON from c. 800-1000 CE represents koineization. OE and ON probably meet the criterion of mutual intelligibility required to be considered subsystems of one linguistic system: they share a similar lexicon and structure. Further evidence, though indirect, can be found in an OE manuscript, The Battle of Maldon, which depicts communication between Anglo-Saxons and Vikings. While poetic accounts should not be assumed to portray historical events factually, this poem does not mention any foreign language used by the attackers, thereby assuming mutual intelligibility (Curzan 2003:52). Simplification clearly occurred in Northern dialects of ME, and these dialects may exhibit a transformation to a more analytic system. This includes the loss of grammatical gender, case marking on nouns, and some inflections on verbs (Dawson 2003:49). This situation shares many characteristics with that of KG, so it seems reasonable to ascribe at least part of these changes to the koineization process, as Dawson does.

Andreose and Renzi (2013:332-333) uses syntactic and morphological features, some conservative, others innovative, to group the Romance languages. The results indicate partially similar groupings to earlier studies based on phonetics and lexicon: the most conservative area includes Romanian, Sardinian, and some central and southern Italian dialects, while the most innovative area is French, sometimes along with Occitan and northern Italian dialects. Although Romanian also has many innovations at every linguistic level that are unique among Romance

languages, it was not subject to the shared innovations of the other Romance languages, including the loss of case distinctions, particularly the Gen-Dat. This suggests that these innovations diffused among the Western Romance languages. In other words, language contact may have played a role in the total loss of morphological case in all Western Romance varieties, as has been proposed for varieties of continental Scandinavian and EME that lost case distinctions without undergoing the regular sound changes that would neutralize case distinctions and/or intense contact with outside groups. At the same time, the abundance of original innovations in Romanian supports the argument that contact with other Balkan Sprachbund languages played an important role in the development of its nominal morphosyntax.

Any argument for case loss due to dialect contact must assume either that one of the participating dialects had already lost case for some other reason or that the contact was intense enough to produce simplification without modeling on any existing dialect. There is less evidence for intense dialect contact of this kind than for the role of internal motivations and contact with more distantly related languages in case loss, so my study proposes that dialect contact accelerated and spread case loss without being a primary motivation.

3.2.2. Sprachbunds

The Balkan Sprachbund languages are the prototypical example of a Sprachbund, or linguistic area. These languages have influenced each other in phonology, vocabulary, and grammar. Sala (2013:213-214) argues that contact between a Pre-Romanian substrate and a South Slavic superstrate in areas both north and south of the Danube may have begun as early as the 7th century and had certainly occurred by the 11th century, based on loanwords in Romanian with phonetic characteristics of Bulgarian dialects before that date. It is not clear how long this contact lasted, but there has been permanent contact between the Romanian and Bulgarian peoples at least since the 11th or 12th century, after which continual Bulgarian migrations to north of the Danube took place. South of the Danube, there was permanent contact between the Aromanians and Bulgarians. Speakers of Byzantine Greek and the ancestor of Albanian were already in the Balkans when the Slavs arrived in the mid-6th century (Hupchick & Cox 2001, Map 8). However, the nature of their contact with South Slavic and Eastern Romance speakers during this time is not completely certain. At the very least, the appearance of vernacular Greek words in BChS gospel manuscripts, e.g., *alektorъ* ‘rooster’ in the 14th-century Curzon Gospel

(Mark 14.68, 72; Vakareliyska 2008), suggests that some BChS scribes, at least, may have been exposed to Greek speakers, not just Greek texts. Another interpretation is that the scribes may have already acquired Greek words, becoming part of their dialects during childhood. Thus, the beginnings of the Balkan Sprachbund, or at the very least, prolonged, intense contact between various pairs of Balkan languages, can be dated to the 11th or 12th century as a *terminus ad quem*. Contact among the IE Balkan languages remained at least as intense under Ottoman rule (late 14th to mid-19th centuries), in part because the population was divided into religious groups (millets) rather than geographical or ethnic units, so contact among members of the Christian majority occurred regardless of their native language (L1) (Lindstedt 2000:238).

Lindstedt (1998:91-95) argues that the Balkan grammatical features are characterized by explicit analytism and are not necessarily structural borrowings from particular source languages, but instead are new constructions that arose through the contact situation itself. He considers ten well-known features from the literature. Six of these are characterized by explicit analytism: a tendency for analytical constructions in place of case marking, the postposed definite article, the analytical comparison of adjectives, the *want*-future, finite complementation, and clitic object copying. Three other, non-analytic shared features are the Voc, aspect oppositions in the past tense, and evidentials, but none of these directly relates to syntactic structure, and the first two are more of a retention than an innovation, in contrast to the other eight features. Finally, the recipient-possessor merger can be expressed synthetically or analytically. CSB and CSM have all ten features, so they are closest to the ideal Balkan linguistic type. As mentioned in 3.1.2.7 above, clitic doubling is more widespread and grammaticalized in CSM than CSB, so the former is considered the most central language, which corresponds to its position at the geographic epicenter of the Balkan Sprachbund near Lake Ohrid, where CSM is in contact with Albanian, Greek, and Aromanian (Friedman 2008:58-60). These languages, along with CSRm, have six to seven fully manifested Balkan Sprachbund features, as well as two to three others in a more rudimentary form. On the other hand, only two of the features, the Voc and the *want*-future, are clearly evident in BCMS, while another two, aspect oppositions in the past tense and finite complementation, occur in a less developed form. Given the lack of other features in BCMS, it cannot be considered a full member of the Balkan Sprachbund.

These facts have several implications for case loss in Bulgarian and Macedonian. The patterns of case loss in these two languages align with three of Lindstedt's (1998) Balkan

grammatical features: a tendency for analytical constructions in place of case marking, the retention of the Voc, and syncretism of recipient and possessive expressions. The retention of more case distinctions in other Balkan Sprachbund languages means that the analytical tendency is less extreme for them. It is possible that Bulgarian and Macedonian experienced more case loss than the other Balkan Sprachbund languages due to a greater effect of sound changes. As argued by Wahlström (2015) and Lindstedt (1998, 2000, 2014, 2018), however, Sprachbunds are characterized by mutual influence between different pairs of varieties, with some directionality determined by sociolinguistic prestige. As the center of the Balkan Sprachbund, Macedonian and Bulgarian have been involved in the highest number of contact pairs, including languages of various prestige levels. They have influenced these other languages, especially those with equal and lower prestige, but they have also been influenced significantly, especially by those with equal and higher prestige.[120]

In response to earlier claims of areal convergence without a diachronic component in the analysis, scholars of areal linguistics, including Joseph (2008), now demand the inclusion of historical evidence (Wahlström 2015:93). For example, Thomason (2008b:49) lists five steps that must be taken to establish a claim of change due to contact:

(a) structural interference is unlikely to have occurred in just one instance, so the languages should be considered as a whole.
(b) the source language should be identified and the contact shown to be intense.
(c) the shared structural features should be identified in the source and receiving languages.
(d) it should be demonstrated that these features were not present in the receiving language before contact.
(e) it should be demonstrated that these features were present in the source language before contact.

The last three of these steps require historical evidence. Wahlström (2015:131-132) applies these steps to the Gen-Dat merger, as shown below.

The functional mergers of Gen-Dat that are characteristic of the Balkan Sprachbund languages may have spread through contact. Wahlström observes that these mergers generally do not involve the explicit structures that are typical of shared features but can still spread because

[120] Structural convergence based on mutual reinforcement and the prestige scale in the Balkan Sprachbund are discussed in more detail in the following section.

the Gen-Dat merger only involves the idea that a case form can express an additional function (2015:132). The generalization of the Loc to express both goal and location in some Balkan Turkish varieties is evidence that the goal-location merger can spread by contact (see Friedman 2006:35-38).[121] At the same time, this merger has occurred in many other languages in Europe, so a contact explanation is not necessary (Wahlström 2015:115). Similarly, Nomachi and Heine (2011) argue that the instrument-accompaniment merger in Bulgarian and Macedonian was based on contact with languages that already had this merger.

As the most notable feature of nominal morphosyntax among the Balkan Sprachbund languages, the recipient-possessor merger may have arisen in some languages due to contact. In CSB and CSM, recipients and possessors are analytically marked. However, when MB is considered instead as the precursor to both modern languages, the merged Gen-Dat case can be considered the most uniform aspect of the Balkan Sprachbund case systems, although it is not clear to what degree this feature is a retention of independent internal developments or an innovation (Wahlström 2015:131). Retentions can be the result of contact, but this is harder to prove than with innovations (Lindstedt 2002:305). Wahlström applies Thomason's (2008b:49) steps to this merger. First, there is a great deal of other structural interference among the Balkan Sprachbund languages. Second, the source language cannot be identified with certainty but extended mutual multilingualism suggests intense contact. Third, it is generally accepted that the proposed languages in contact all mark recipients in a way similar to possessors, although a counterargument is discussed below. Larger problems arise with the fourth and fifth steps, since it is hard to determine the status of recipient and possessor marking in most Balkan Sprachbund languages before contact. With no attested ancestor, this is the most difficult for Albanian. For Bulgarian and Macedonian, on the other hand, there is at least evidence from the earliest OCS texts which suggests that this merger cannot have occurred until at least the 10th century (2015:131-132). In addition, only singular *i*-stems had Gen/Dat syncretism in OCS, and the only

[121] Under the Ottoman rule of the late 14th to mid-19th centuries, there were native Turkish-speaking local government administrators, and contact with Turkish is demonstrated by the many Turkish loanwords, particularly for household items, in the Balkan Sprachbund languages. For example, using a loanword typology database for Romanian, in which 42% of the words are loans, Schulte (2009) found that 3% of the Romanian lexicon is borrowed from Turkish, 10% of which are household items, representing the largest lexical field. Likewise, CSM and CSB have a high number of Turkish loanwords; Falaleev (2016) has found that native CSM speakers underestimate the number, showing that they have been well-integrated phonologically or morphologically.

additional Gen/Dat syncretism expected due to sound change is in singular *jā*-stems with unstressed endings in MB dialects with vowel raising, as described in section 2.5.2.1 above.

Despite the uncertainty, several languages have been proposed as the source of the recipient-possessor merger. Reichenkron (1962:12) proposes a Romanian origin based on the existing Gen/Dat syncretism in CL *ā*-stems. This may have spread to the other declensions as the case system broke down in VL. Moreover, semantic overlap between Dat and Gen existed in CL, similar to the situation in Greek discussed in section 3.1.2.5 above: both were used to express possession. It must be noted, however, that an extension of the Dat from inalienable to adnominal possession would still have been required for this case to subsume the Gen in a functional merger. VL subscriptions in Dalmatia and the Danubian region appear to support early Gen/Dat syncretism in declension classes other than *ā*-stems. For example, *filius regi* 'the son of the king', with the consonant stem Dat singular form *-i* is attested in addition to *filius regis*, with the corresponding Gen form *-is* (see Catasso 2011:74). This particular example, however, can be explained by the regular loss of *-s* in Pre-Romanian, as described in section 2.4.2 above.

Catasso argues on syntactic grounds that the recipient-possessor merger is not a phenomenon of the Balkan Sprachbund. One of his main arguments is that the merger is not complete in Romanian and Albanian, because these languages require linking articles with possessive uses of the Gen-Dat but not with recipients (2011). This and other differences in syntax, however, are better attributed to more general differences between nominal and verbal syntax, given that the possessor is adnominal while the recipient is adverbial (see Wahlström 2015:152-153). With respect to Albanian, Spencer argues that the Gen is not a distinct morphological case from the Dat, but that the linking article is part of a possessum-agreement construction, in which a construction used for attributive modification is extended to possession. Therefore, linking articles, which also occur with other modifiers of nouns such as adjectives, should not be considered case marking (2007:247-248).

Catasso also raises the issue of differences in how the merged recipient-possessor is expressed and used across the Balkan Sprachbund languages (2011). The most obvious difference is that it is synthetic in some of the languages but analytic in others, as mentioned above. Another is the use of linking articles in only some of these languages. However, Wahlström argues convincingly that the latter arises due to differences in NP word order: linking articles only occur with right-branching modifiers, as a way to maintain the connection between

the modifier and its head noun. In contrast to Romanian and Albanian, which also have less agreement marking on adjectives, CSB and CSM have a strong left-branching preference for adjectives. However, it appears that in earlier stages of their development, as attested by OCS and BChS, adjectives typically followed the noun. However, it is unclear if the syntactic position of the adjectives represents OCS and BChS since the texts were translations from Greek. Both positions are common for adjectives in Greek, but notably, a second definite article must be used before an adjective in Greek when it follows the head noun. More generally, Wahlström argues that total identity should not be expected to result from convergence due to contact, since the borrowed feature must be integrated into a different existing morphosyntactic system in each language (2015:121, 153).

In Greek, the Acc and Gen competed in assuming the functions of the Dat during the first millennium, as mentioned in section 2.5.3 above. Likewise, analytical marking of recipients and possessors in Romanian appears to have decreased from the earliest 16th-century texts, as discussed in section 3.1.2.5 above. In both instances, it must have taken the Gen-Dat a long time to become the established marker of both recipient and possessor. Wahlström considers contact in the form of mutual bilingualism to have contributed to the eventual selection of the Gen-Dat over the competing structures. However, he states a problem with this explanation: the Northern Greek dialects in which the Acc fully replaced the Dat were geographically closest to the other Balkan Sprachbund languages, but if this process was complete before the contact began, the Gen-Dat would no longer be an available option and no competition process, in which contact might exert influence, would occur (2015:131-132). Therefore, competition between morphological cases does not seem to explain the situation completely, but that structural convergence does as the morphosyntax of languages in a Sprachbund becomes more similar.

3.2.3. Structural Convergence

One possible result of contact among multiple languages is structural convergence. Lindstedt, among others, considers the Balkan Sprachbund to be defined by a number of structural convergences. In terms of linguistic typology, these convergences are characterized by explicit analytism, as discussed in section 3.1.2.7 above. While these defining features of the Balkans are typologically common in the world and often even in other parts of Europe, their occurrence all together in the Balkan Sprachbund is notable (2014:169-170). As Friedman (1999:521) argues,

while Turkish was the state language during the Ottoman period (late 14th to mid-19th centuries) and is the most likely source for the grammaticalization of verbal distinctions in evidentiality, it had too many typological differences from the Indo-European Balkan Sprachbund to have a level of influence that would give rise to the observed convergence of features. Likewise, no single Balkan Sprachbund language can be the source of all convergent features. Instead, these languages, which were already somewhat similar, seem to have converged based on mutual reinforcement (Lindstedt 2014:171-172). It should be noted that these changes primarily involved featural convergence rather than material copying (ibid., 178). This can be contrasted with the direct borrowing of functional morphemes from ON to ME: derivational affixes, prepositions (*till* and *from*, as opposed to their OE cognates *to* and *fro*), and the third person plural pronouns *they*, *them*, and *their* (Dawson 2003:44-45). However, a few examples of material copying can also be found in the Balkans. The feminine Voc ending *-o* in Romanian was borrowed from South Slavic, while the masculine singular Voc ending *-e* probably continues the Latin Voc with support from the Slavic masculine Voc of the same form (Dimitrescu et al. 1978, 209-212).[122] The Romanian Voc form *-le*, which incorporates the definite article, was borrowed back into some Bulgarian and Macedonian dialects, e.g., *Božele!* 'God!' *Sestrole!* 'Sister!' (see Popović 1960:208). Both featural convergence and direct borrowing of functional morphemes indicate more intense contact than that associated with borrowing only of content words.

Lindstedt argues that Balkan features are not about simplification or complexification per se but are instances of explicit analytic marking that improves intertranslatability; he accounts for this tendency with three diachronic factors. The absence of the level of simplification in the Balkan Sprachbund that would be expected for creolization is clear from the complex verbal systems, as mentioned in section 3.1.2.7 above. On the other hand, an increased redundancy can be observed as a result of Balkan features (2019). In opposition to Hinrichs (2004), Dahl (2004:5-17) and Trudgill (2011:62) argue that redundancy actually represents an increase in complexity. According to Trudgill (2002:710-711), the loss of the infinitive in Balkan Sprachbund languages and its replacement by an embedded subjunctive clause increased both complexity and redundancy. In his example from Greek:

[122] CL did not have a distinctive Voc marker for the feminine (or neuter). The only distinctive Voc forms were *-e* or *-ī*, both for masculine singular *ō*-stems, depending on the preceding sound.

(8) θél-o na γráp-s-o
want-PRS.SG1 COMP write-PFV-PRS.SG1
'I want to write'

The first-person singular is marked twice: once in each clause. However, if the subjects of the matrix and embedded clauses are different, the Balkan pattern is simpler and more symmetrical, and arguably not redundant (Lindstedt 2019:72). In addition, because of this loss there is one fewer inflectional category. Many changes, including case reduction, the recipient-possessor merger, and the goal-location merger, decrease morphological complexity but are compensated by an increase in syntactic complexity through their substitution by enclitic articles, clitic doubling, and other analytic argument marking strategies (ibid., 73). Therefore, the Balkan Sprachbund does not fit the result of either of Trudgill's sociolinguistic types of contact: simplification due to adult L2 learning, especially short-term, or complexification due to long-term childhood bilingualism (2011:40-43). The Balkans also did not fit the sociolinguistic context: there was long-term contact in the Balkans, especially among adult men, but no evidence for a high degree of childhood bilingualism outside of certain ethnically mixed urban centers. The lack of a single language that combined state and religious authority for the Balkan Christian majority during the Ottoman period also played a role in distinguishing the situation of the Balkans from other types of contact (Lindstedt 2019:81-82).

The first external factor that increases analytism is borrowing. Analytic constructions have a higher probability of being borrowed than inflections (see, for example, Dahl 2004:127-128). This means that analytic constructions will gradually become more common when a language experiences grammatical borrowing. The source language does not need to be especially analytic, as long as it has some analytic structures. In addition, constructions may be borrowed in a way that makes them more analytic (Lindstedt 2019:74). Finally, evidence from language change suggests that analytic structures are not lost due to obsolescence any more than synthetic structures, so analytism should only increase due to borrowing (ibid., 76).

The Balkan Sprachbund had the social context necessary for large-scale borrowing, and Balkan features are assumed to have arisen through this borrowing: even those that developed through mutual reinforcement from more than one source were most likely copied multidirectionally (Lindstedt 2019:73-74). Support for this assumption comes from the fact that

the language varieties with the most Balkan features are those spoken in the areas with intense contact between multiple languages (Asenova 2002:17). An exception to this regularity may occur when children become bilingual, since they may not have the same preference as adult L2 learners for analytic structures (Trudgill 2011:40-43). However, as mentioned above, Lindstedt does not find evidence for enough childhood bilingualism in the Balkans to lead to increases in complexity due to borrowing (2019:81). Thus, analytism is expected to increase because of this borrowing.

Convergence of different structures is another factor in the increase of analytism. The greater the differences between a specific subsystem in multiple languages, the more analytic the result will be, because more change is required to reconcile the differences. This can be observed through a comparison of the Balkan nominal and verbal systems. The verbal systems, which already made similar distinctions, remained more complex, while the case systems were not similar enough until they were reduced to three distinctive forms: Nom, Acc, and Gen-Dat (Lindstedt 2019:77). Once the case systems had converged, further analytism may have even been resisted, e.g., in Romanian. In older Romanian texts, the reflexes of the Latin prepositions *dē* 'from, about' and *ad* 'to' with Acc objects, are used to indicate possessors and recipients, as in the Western Romance languages.[123] However, in CSRm several of these contexts only allow the bare Gen-Dat. This departure from the Romance pattern is likely a consequence of the influence of other Balkan Sprachbund languages: although it actually decreases analytism, it follows the Balkan tendency of using the same inflected case marking for recipients and adnominal possessors (Wahlström 2015:107). The fact that Romanian moved away from the analytic development of the Western Romance languages and came to prefer the combined Gen-Dat case to express these grammatical roles suggests that language contact can also work to prevent case loss (ibid., 110). Thus, this factor accounts for some of the Balkan analytism, but not for developments once convergence has been achieved, e.g., the loss of almost all case distinctions in Bulgarian and Macedonian (Lindstedt 2019:77).

The third factor that increases analytism is the existence of both L2 speakers and bilingual L1 speakers. A number of theories have been proposed as to what type of speakers were responsible for the convergent features of the Balkan Sprachbund. Lindstedt finds problems

[123] See sections 3.1.2.4 and 3.1.2.5 above for discussion on the analytic use of these prepositions in CL and Western Romance.

with theories that either the L1 speakers were entirely responsible due to influence from L2 or that L2 speakers were entirely responsible due to influence from L1 (2019:78). A third theory is that L1 speakers were trying to use structures that were easiest to understand by L2 speakers (Lindstedt 2019:78, citing Civ'jan 1965:14ff., 183ff.). However, this does not explain why speech in monolingual situations would change. Of the Balkan Sprachbund languages, Bulgarian and Macedonian were affected most by the analytic tendency. CSM is the standard language with the highest number of Balkan features, which, as Lindstedt points out, can be partly explained by its location at the geographical center of the Balkan Sprachbund. This does not explain the high analytism of less central varieties of Bulgarian and Macedonian, but their place in the middle of the prestige scale during the Ottoman period does (2019:77). The most prestigious Balkan language, Greek, and least prestigious, Romani, have fewer Balkan features than Bulgarian, Macedonian, and the other languages near the middle of the scale (Lindstedt 2000:232-234).[124] These languages in the middle have L2 speakers as well as L1 speakers who also regularly use other languages (Lindstedt 2019:79). This suggests more of a need for intertranslatability and hence explicit analytism. Thus, the existence of two groups of bilingual speakers using languages in the middle of the prestige scale may explain why these languages become more analytic than those with only one bilingual group (ibid., 79).

Overall, this proposal is relatively convincing, but it still relies on a few unproven assumptions. To work as proposed, a large enough proportion of the population must have been interacting with speakers of other languages. Men were more likely to do so because they were responsible for trade and for representing the family in state and religious matters. In addition, men participated in seasonal work migrations, while this was rare for women (see Lindstedt 2019:82, citing Hristov 2008:3). This does not rule out female bilingualism but makes it more difficult to confirm that enough speakers were involved to influence the language even among L1 speakers. Finally, the initial stages of case loss in Bulgarian and Macedonian (as well as Romanian) can be ascribed to phonological changes. It seems more likely that contact allowed for the spread of analytic alternatives to the already reduced case systems, making the complete abandonment of case marking possible in Bulgarian and Macedonian. Evidence from modern

[124] Lindstedt does not specifically address the presence of Balkan grammatical features in Turkish but notes that despite its status as the language of the state, Turkish was not prestigious among Christians (2000:238-239).

Slavic languages in contact situations supports this type of role for contact in case loss, as discussed in the following section.

The convergence of gender systems has also been attested among languages in close contact. For example, the dialects of the Indic languages Marathi and Urdu that are spoken in the Indian village of Kupwar have borrowed the semantic assignment system of Kannada, a Dravidian language also spoken in the village. Marathi varieties generally have the three IE genders, with semantic assignment only for rational referents, i.e., humans and gods, as in many IE languages. In the local dialect, however, just as in Kannada, the masculine gender is limited to male rational referents, the feminine is limited to female rational referents, and the neuter is used for all remaining non-rational referents. Urdu varieties also tend to have semantic assignment only for rational referents, but the neuter has been lost. As a result, the local dialect has adapted the semantic system of Kannada so that the masculine is the default category, with the feminine limited to female rationals (see Gumperz & Wilson 1971:155-156). These changes involved the borrowing of abstract assignment criteria rather than the material copying of gender forms. In Urdu, this required some modification to fit the existing morphology; semantic assignment was borrowed but not the number of categories. A similar adaptation may have occurred in Romanian.

Similarly, contact may be responsible for the retention of the neuter in Romanian as a gender category that experienced masculine agreement in the singular and feminine in the plural. According to Rosetti (1986:351), the neuter was lost in VL but reappeared during the Common Romanian period, which corresponds to the Pre-Romanian stage in section 2.4.2 above, as a result of the need to distinguish animates from inanimates. If this occurred, it would most likely have been due to contact with the other Balkan Sprachbund languages, which tend to maintain the neuter and treat animates and inanimates differently in various ways, i.e., the DOM strategies discussed in section 3.1.2.7 above. However, if the neuter truly disappeared and formerly neuter nouns took on another gender, speakers borrowing the neuter category would have no way to know which nouns ought to be reassigned to neuter, and it would be expected to consist of inanimates, regardless of their original CL gender. In fact, Romanian nouns have mostly retained their CL gender. Some inanimate masculine nouns have become neuter over time (see Baerman

et al. 2005:86, citing Windisch 1973:117, 196).[125] However, many inanimate masculines remain, not to mention almost all inanimate feminines.[126] This evidence suggests that the neuter was not restored by contact but maintained by mutual reinforcement. As discussed in section 3.1.2.1 above, ambigeneric agreement was probably the best option for the survival of the neuter category after the merger of the masculine and neuter in the singular. If only internal motivations are considered, a similar system could have developed just as easily in the Western Romance languages. In fact, as discussed in section 2.4.1.1 above, CSI maintains a similar but much smaller class of ambigeneric nouns, a development which began in VL. Given its shared development with Italian, the neuter in Romanian cannot be attributed to Balkan Sprachbund influence alone, but mutual reinforcement by these languages is the best explanation for why the neuter has survived as a productive class in CSRm but only as an unproductive one in CSI and not at all in most Western Romance languages. Without the same level of contact with other languages that retained the neuter, the most straightforward development in the latter group was to extend the merger of the masculine and neuter in the singular to the plural.

Coming from a slightly different perspective from Lindstedt, Myers-Scotton (2006:271) defines convergence as speech by bilinguals with surface forms from one language but part of the abstract structure borrowed from another. Myers-Scotton (2002:172) considers convergence to be largely unidirectional, generally from the language with higher prestige, but with the possibility of mutual influence. Trudgill (1989:234) points out that Danish has higher language contact than Faroese. This entails that Danish lost its case inflections faster than Faroese and replaced them with analytic alternatives. Based on interviews, grammatical judgement tests, and a corpus study, Petersen (2011:101-105) found that Faroese has converged with Danish structurally. While Faroese has been relatively isolated for much of its history, it has experienced intense language contact with Danish in the last two centuries. All Faroese speakers also speak Danish fluently, while very few speakers of Danish, the prestige language, speak Faroese. In addition to a strong influence on the lexicon, this stable asymmetrical bilingualism has resulted in convergence of a more synthetic language (Faroese) with a more analytic language (Danish).

[125] The Romanian neuter has also expanded by absorbing most loanwords denoting inanimate objects that come from a language without grammatical gender (see Schulte 2009:14).

[126] Inanimate feminine *a*-stems would have been unlikely to shift to the neuter regardless, since they were distinct from the major masculine/neuter classes in the unmarked singular, but gender confusion among consonant stem and *i*-stem nouns could easily have occurred if the neuter was reintroduced after disappearing.

Specifically, Petersen found that analytic comparatives are displacing native synthetic comparatives, and that prepositional constructions have become more acceptable as an alternative to the Dat for marking recipients in Faroese, in contrast to Icelandic (ibid., 106-108). Jónsson (2009) claims that the Dat is gradually being lost in Faroese. If so, it is because of the availability of alternative analytic structures due to contact, not because of sound change. I propose that other developments in Faroese, such as the decline of the Gen in speech, have been, at least in part, the result of a similar process.

Structural convergence may have also played an important role in the earlier case loss of Danish and the other continental Scandinavian languages. Scandinavia had the necessary social context, with sustained intense contact between different varieties due to trade and political entities including speakers of multiple varieties. This was particularly true in Norway, where the union with Denmark from the 14th to 19th centuries resulted in the use of Danish as a prestige language alongside Norwegian (Askedal 1994:219). For part of this period, from 1397 to 1523, Denmark, Sweden, and Norway were all united in the Kalmar Union, and from 1814 to 1905 Norway was affiliated with Sweden (Besters-Dilger & Braunmüller 2014:407). The structural changes observed during this time period are also suggestive of convergence. Varieties that had experienced some degree of case loss due to phonological erosion and/or simplification due to contact with MLG probably mutually reinforced these changes and spread them to dialects that did not have them yet. The analytic replacements to morphological case would have been favored when structural convergence occurred. Dano-Norwegian developed during this time in an instance of koineization, but even Norwegian dialects without such a heavy Danish element ended up much closer to Swedish dialects than to the other languages of the Western branch, Faroese and Icelandic. Morphological case was lost in NNw but all three genders were retained. The longer survival of the Dat case marking in isolated inland varieties of Norwegian and Swedish, even up to the present, adds further support for the important role of contact in case loss.

3.2.4. Contact-induced Change in Modern Slavic Languages

Luraghi and Krstić (2018) consider the use of morphological cases in Molisean Croatian (MC). This highly endangered variety is spoken in Southern Italy by about a thousand speakers whose ancestors from Dalmatia settled there in the 15th to 16th century. They have been in contact with

Italo-Romance speakers for over five centuries but were relatively isolated until the second half of the 20th century, when sociolinguistic conditions began to change due to improved means of communications. These developments have accelerated over the last few decades, and all of the active speakers among Luraghi and Krstić's informants were over 35 years old (2018). However, Marra (1999) studied semi-speakers, i.e., younger speakers who learned MC as children but considered themselves native Italian speakers. In addition to the informants, Luraghi and Krstić include textual evidence over the last 150 years to make diachronic comparisons. As used by native speakers, MC distinguishes five cases morphologically: Nom, Acc, Gen, Dat, and Instr. The Loc has formally merged with the Dat, as in BCMS, but some of its prepositional functions have been assumed by the Acc in a goal-location merger. The neuter category has been lost, but some forms are still used in impersonal constructions (2018:67, 76). This is similar to the situation in Romance languages such as Spanish, as discussed in section 2.4.1.2 above. In addition, gender distinctions have been neutralized in the plural, and declension classes have been reduced to two, corresponding one-to-one with gender. The animacy distinction among masculine nouns survives, but not systematically. MC has two numbers, but there are vestiges of the dual in older texts (ibid., 66-67).

The use of morphological cases in MC has undergone some significant changes compared to BCMS. Besides its usual uses as subject and nominal predicate, the Nom is also used as the complement of the preposition *di* 'at, to', which was grammaticalized from the adverb *gdi* 'where', originally in constructions with 'be'. When the verb was no longer required, the Nom was left as an object of a preposition, which allowed for its further extension to prepositions such as *do* 'of, from' (Luraghi & Krstić 2018:71-73, 79). Marra (2012) finds a potential source in a similar construction in Molisean Italian. However, a similar construction occurs in South American Spanish, so this may be a common grammaticalization pathway. Verbal arguments still retain the expected bare cases: Nom for subject, Acc for DO, and Dat for IO. The Dat has lost productivity in other uses. Prepositions that previously took the Dat (including the original Loc prepositions) now usually take the Acc or are replaced by other prepositions that already take the Acc, e.g., *niz* 'downwards' or *po* 'by, around' followed by the Acc instead of the original *po* plus Dat. Other cases have been reinforced by prepositions in most uses. The adnominal Gen is usually reinforced by the preposition *do*, although some speakers use the Nom with this preposition instead. The Gen is also used after other prepositions, and the bare

Gen is still consistently used in partitive constructions, possibly because of the rigid structure with a quantity word before the Gen. The preposition *s* 'with' regularly reinforces the Instr, even when denoting an instrument (Luraghi & Krstić 2018:78-83, 93).

Although the Acquaviva dialect allows for the centralization or dropping of final vowels, cases are still morphologically productive across the dialects of MC, since the same forms have been applied to Italian loanwords (Marra 2009:99). Luraghi and Krstić (2018:87-88, 93-94) argue that cases have undergone semantic bleaching despite morphological persistence. With reinforcement by prepositions, including the instrument-accompaniment merger, and the loss of case oppositions after prepositions, including the goal-location merger, cases carry less meaning of their own. Even in BCMS, the distinction between the Acc and Dat with the prepositions *na* 'on' and *u* 'in' is somewhat redundant because the choice of verb also indicates whether a goal or location is expected. When marking grammatical relations, cases are predictable based on verbal valency—presumably combined with word order, although Luraghi and Krstić do not mention this.

In contrast to native speakers, for whom the realization of cases has been stable over the last 150 years, semi-speakers tend to overextend the Nom and Acc in prepositional phrases. This often manifests as a zero ending but also occurs for feminine nouns and masculine nouns without a zero ending in the Nom. This involves the more frequent use of the Nom in the adnominal possession construction than by native speakers, as well as the use of the Acc or a zero ending in place of the Instr. The latter replacement tends to be associated with the use of the Italian loanword *ko*, *kon* 'with' instead of *s* (Marra 1999:148-149, 152-156). Luraghi and Krstić attribute the failure of semi-speakers to acquire cases, even when consistently realized by native speakers, to their semantic redundancy (2018:95). This failure can also be explained in terms of MacWhinney et al.'s (1985) competition model: i.e., other cues have become more valid than case marking. Due to their circumstances, semi-speakers may not acquire cues which they do not need for sentence interpretation, while normal L1 speakers would have deprioritized them but still have acquired them eventually.

There are a number of striking similarities between the changes to the case systems in the MC of semi-speakers and the changes in Bulgarian and Macedonian described above. All three underwent the instrument-accompaniment merger and the goal-location merger, and the same prepositions were involved. They also adopted an analytic adnominal possession construction,

although this was preceded by a recipient-possessor merger in Bulgarian and Macedonian, so *na* 'on' was used in this construction in addition to *ot* 'of, from', cognate to MC *do* (Luraghi & Krstić 2018:89). The order of case loss is quite similar as well, with peripheral cases lost first. Meyer has proposed the same failure to acquire cases due to semantic redundancy for Bulgarian, with L2 speakers in the same role as the semi-speakers of MC (1920). These similarities suggest that some of the same factors were involved.

The developments in MC have similarities to non-Slavic Balkan Sprachbund languages as well. Greek has also undergone both the instrument-accompaniment merger and the goal-location merger, and some Gen uses have been reinforced by *apó* 'from' (see Luraghi 2003:80). Unlike Greek and other Balkan Sprachbund languages, however, MC has not undergone a recipient-possessor merger and has retained the Instr much longer; even semi-speakers still use it to some extent, especially the Instr plural (Luraghi & Krstić 2018:96). Luraghi (2004) connects the loss of the Instr and later Dat in Greek to low token frequency but also a high degree of allomorphy. In contrast, the same Instr forms are used regardless of declension class in MC: *-om* in the singular and *-ami* in the plural. The latter form is also syncretic with the Dat, increasing its frequency (Luraghi & Krstić 2018:96). These differences in allomorphy may explain the different outcomes regarding the Instr.

Contact has clearly had an influence on the MC case system, but it is less clear how. Some developments are commonly seen as the result of Romance influence, particularly the reinforcement of the Instr and adnominal Gen with prepositions (see Marra 2012). Nomachi and Heine attribute the former, i.e., the instrument-accompaniment merger, to the influence of Romance and Germanic languages (2011). In contrast, Danylenko (2015) argues that Slavic languages in contact situations are prone to this development regardless of the contact language, simply due to the analytic tendency engendered by the contact situation itself. Likewise, in contrast to *de* 'of, from' in Spanish and CSF, both CSI and Molisean Italian have separate prepositions *di* (*de* for Molisean) 'of' and *da* 'from', so they do not fit as the direct target for the extension of MC *do* 'of, from' to the adnominal possession construction. In addition, Breu (2008) attributes the goal-location merger to Italian influence, since Italian varieties, including CSI, use the same prepositions for both, with no formal distinction due to the lack of morphological case. As noted above, however, the case alternation between goal and location with the prepositions *na* 'on' and *u* 'in' is not particularly meaningful even in BCMS, and

Luraghi (1989) argues that this is a common development as languages become more analytic, based on her conclusion that the same redundant case variation occurred in VL. In VL, however, the distinction between Acc and Abl singular had been almost completely neutralized by regular sound change, so the goal-location merger may have been phonologically motivated at first.

Luraghi and Krstić conclude that it remains unclear how contact brought about these developments that increase analyticity: they may be common to Slavic languages in contact situations generally, or just to those in contact with Romance and Germanic varieties; they may even be shared more generally by languages in Europe experiencing contact, or just by those in the Balkan area (2018:92). Since MC is not part of the Balkan Sprachbund, and since even BCMS has remained largely free of Balkan features, I propose that the similarities between developments in MC and the Balkan Sprachbund languages are more likely the result of similar starting configurations and types of contact than contact with the same language families. Specifically, MC, Bulgarian, and Macedonian appear to have undergone two stages of case loss. In the initial stage, functional mergers and reinforcement by prepositions increase the redundancy of case marking, but the core cases remain intact. In the second stage, the case system falls apart completely. Both steps appear to be caused, or at least accelerated, by contact, but the first involves stable bilingualism, while the second involves L2 speakers. Once the semantic contribution of the cases has been weakened, L2 speakers, or semi-speakers in the case of MC, fail to acquire the case forms. If the number of L2 speakers is high enough, this could affect native speech as well, particularly in the later generations exposed to L2 speech as children. For MC, language death is suggested by the fact that even the children of native speakers are learning MC more or less as an L2, leading to rapid changes that contrast with at least 150 years of relative stability (see Luraghi and Krstić 2018:96, citing Marra 1999:206).

For Bulgarian and Macedonian, on the other hand, the L2 speakers must have been shifting from some other Balkan Sprachbund language during the Ottoman period, and the children of native speakers were still acquiring the language as an L1, so the changes may have been more gradual, and language death did not result, but the preceding stage of almost complete case loss did. Further support for a more gradual process comes from the different rates at which Bulgarian and Macedonian dialects lost the last few case distinctions (except the Voc). Perhaps the rate of change depended on the proportion of L2 speakers to native speakers.

Other studies of Slavic languages in emigration show many similarities to the developments in MC. Sussex (1993:1016) discusses several studies that have consistently found reductions in nominal inflection, and attributes this to contact with West Germanic and Scandinavian languages, since these languages have simpler case systems than the Slavic languages under investigation.[127] In a longitudinal study of children of BCMS-speaking immigrants to Sweden, Stankovski et al. (1983:24) find that most have an incomplete case system at age seven. The degree of case loss varies, but several implicational universals hold, e.g., a child who uses the Gen will also use the Nom and Acc. Another finding is that the Loc is often replaced by the Acc. Other studies have found similar patterns of case reduction in the speech of other Slavic émigré communities, e.g., Preston and Turner (1984) for American Polish, and Gasiński (1986) for American Croatian. The latter study also considers specific case usage; notably, it finds that American-born Polish speakers use the Acc for the DO of a negated verb more frequently than Polish-born speakers who are also living in the US. In Standard Polish, the Gen is obligatory for this function. These studies add further support to the hypothesis that Slavic languages in certain sociolinguistic conditions tend to undergo some of the same patterns of case loss found in the history of Bulgarian and Macedonian, e.g., the replacement of the Loc with the Acc. Of course, as with MC, the linguistic communities under investigation are more likely in the process of shifting entirely to speaking English or another Germanic language, in contrast to the mutual bilingualism that is assumed to have occurred in the Balkan Sprachbund. Nevertheless, this is an indication that contact with a variety of different languages can lead to case loss in Slavic languages that have otherwise resisted these developments.

3.3. Conclusion

This chapter has discussed motivations for the loss of case and gender in the Germanic, Romance, and Balkan Sprachbund languages. Phonological factors indicate that prosodic change is the primary motivation for case loss in the Germanic, Romance, and Hellenic languages. Even if the degree to which case loss occurred in these languages cannot be fully explained by prosodic change, particularly the development of dynamic stress, it highlights the important role of vowel reduction, which often accompanies this development. The Slavic languages have stress-induced vowel reduction and case loss. These do in fact appear to be correlated: most

[127] CSB and CSM are not considered, because they had already lost almost all case distinctions.

Slavic languages have neither, and CSB has both. The frequency of nouns with final stress may explain why CSR and the other East Slavic languages have vowel reduction without case loss, but it is unclear for CSM, which needs another factor to explain why it has case loss despite a weak stress without vowel reduction.

Morphosyntactic and semantic factors include several proposals concerning the relationships between the loss of case distinctions and the grammaticalization of definite articles from demonstratives. There is typological evidence that these developments are correlated, perhaps due to functional overlap, but the direction of causation is unclear, and both may be the result of other factors. At the same time, the retention of more case distinctions on definite NPs than bare nouns in certain languages suggests that articles can also preserve case marking that would otherwise have been lost due to sound change.

The remaining morphosyntactic and semantic motivations involve several related aspects and stages of developments that often accompany case loss: case variation, functional mergers, functional narrowing, and analytic constructions replacing synthetic ones. Phonological change and paradigmatic pressure are among proposed sources of case variation. Thus, the parameters for determining the outcome of case variation can be interpreted as an explanation for developments that occur after the case system has been disrupted. Similarly, syntactic and semantic factors were the primary motivation for case loss instead of phonological and morphological factors. The cases can undergo functional mergers due to syntactic or semantic overlap.

Standardization has been proposed as a factor in preventing analytic constructions from totally supplanting synthetic ones. Scott (2014) argues that if standardization occurs before a case has been completely lost, as with the Gen in High German, it can be preserved alongside analytic alternatives, in contrast to languages like Swedish and English, in which standardization occurred around the same time period, but morphological case marking had already been completely lost.

External motivations involving different contact situations have also been discussed in this chapter. Various developments in nominal morphosyntax have been attributed to dialect contact, including increased case variation. While koineization may account for some early simplification in Greek and, more tentatively, in Germanic languages such as English, it cannot account for case loss on its own.

Alternatively, developments in both CSB and CSM could have occurred as a result of structural convergence with other Balkan Sprachbund languages. The grammatical features of the Balkan Sprachbund are characterized by explicit analytic marking, and that this is the expected result of structural convergence among multiple languages in contact. It has been proposed that three aspects of the contact situation in the Balkan Sprachbund increased analytism: structural borrowing, since analytic constructions are more likely to be borrowed and constructions can also become more analytic in the process; convergence of structures that originally differed significantly, e.g., the nominal systems of the Balkan Sprachbund languages, in contrast to their verbal systems, which were already more similar and remained much more complex; and high numbers of both L2 speakers and bilingual L1 speakers in the converging languages. Although certain of the aspects still lack sufficient historical evidence, this proposal is plausible as an explanation for analytic features in the Balkan Sprachbund languages generally, as well as for the higher degree of analytism in Bulgarian and Macedonian specifically. Based on more reliable evidence from modern Slavic languages that have undergone confirmed contact with non-Slavic languages, I offer a slightly different proposal concerning case loss in Bulgarian and Macedonian: that is, during an initial stage of stable bilingualism, case marking, likely already weakened by sound change, became more redundant due to functional mergers and reinforcement by prepositions; in a second stage, L2 speakers failed to acquire the relatively redundant case forms.

Mutual reinforcement among the Balkan Sprachbund languages may have preserved certain conservative features as well, although this is, by its nature, more difficult to confirm than their shared innovations. Notably, CSRm has retained some case marking on nouns and a productive ambigeneric class derived from the neuter, both in contrast to the Western Romance languages. On the other hand, all the Balkan Sprachbund languages have retained the neuter in some form, and all but CSB and CSM have retained a reduced case system.

A different type of structural convergence has occurred in Faroese due to asymmetrical bilingualism, with influence mostly only in one direction, i.e., from the prestige language, Danish. One of the developments is the increased acceptability of prepositional constructions to mark recipients instead of the Dat. The Dat may in fact be gradually disappearing; there is no sound change that can account for this, so if the Dat is eventually lost, it would confirm the role of analytic tendencies due to contact. A similar process was at least partially responsible for the

loss of a productive Gen in spoken Faroese, as well as for some of the case loss in the continental Scandinavian languages that cannot be explained by sound change and number profiling alone, such as the loss of the Dat in the standard varieties, in contrast to its survival in more isolated inland varieties of Norwegian and Swedish.

Several proposals that attempt to integrate multiple internal and external motivations for case loss have also been considered in this chapter. It has been proposed that case loss in Bulgarian and Macedonian resulted from a combination of mutual bilingualism, which caused functional mergers and analytic constructions to spread among the Balkan Sprachbund languages by structural convergence, and a high number of L2 speakers, who favored analytic alternatives over other ways of dealing with syncretism that arose due to sound change. In other Slavic languages such as BCMS, syncretism was instead eliminated by the adoption of more distinctive forms from unproductive declension classes, which can be seen as the profiling of number and case at the expense of declension. Earlier, working under the imposed Soviet philosophy of dialectical materialism, Duridanov (1956) drew attention to these different paths and also argued for the role of contact in determining which one a language would take. He also mentioned a third option, shifting case marking from nouns to articles, as in other Balkan Sprachbund languages as well as High German.

A usage-based constructional approach attributes case loss in the Germanic languages to functional narrowing, with lexical borrowing as an additional factor. Under this approach, the productivity of a construction depends on its type frequency and semantic coherence. For example, an argument structure containing a particular combination of case markings is generally more productive than a competing construction if it is used with more verbs, but a lower-frequency construction can still attract new verbs if they are similar in meaning to the verbs that already use it. Thus, the functional narrowing of a case marking occurs when one or more of the constructions containing it loses productivity to a construction containing a different case marking. A high number of loanwords from a particular language is a reliable indication that there has been sustained contact with that language, but not necessarily that any structural changes have occurred as a result. However, an influx of new verbs can lead to or accelerate functional narrowing by increasing the type frequency of constructions that are already highly productive, making the lower-frequency constructions even less productive in comparison. The Germanic languages that lost all morphological case distinctions are also those in which periods

of massive lexical borrowing occurred, and these periods coincide with the breakdown of their case systems. Incidentally, lexical borrowing has also been proposed as a factor in prosodic change since loanwords often differ in stress pattern from native words. Thus, an influx of loanwords from a language with different stress patterns can lead to a reanalysis of foot structure, occurred in ME.

In the following chapter, I present my quantitative analysis of case, gender, and number syncretism across different stages of a selection of Germanic, Romance, and Balkan Sprachbund languages.

CHAPTER IV.
INTERACTION OF CASE, NUMBER, AND GENDER CATEGORIES

Chapter III above considers many potential factors in the loss of morphological case and grammatical gender in the Germanic, Romance, and Balkan Sprachbund languages. The goal of the analysis in this chapter is to determine the extent to which several of these factors can account for case and gender loss in these languages. Broadly speaking, these are phonological change, discussed in section 3.1.1 above, and morphological change, discussed in sections 3.1.2.1 above and 3.1.2.2 above. My general hypothesis is that a combination of sound changes that directly neutralize distinctions, and analogical changes that profile a more relevant category, can account for the degree of case and gender loss in the Germanic, Romance, and Balkan Sprachbund languages.

The first part of the analysis measures the effects of these two types of change on number syncretism and, in turn, on how number syncretism relates to case loss. Specifically, it tests the hypothesis that decreases in number syncretism due to analogical processes are correlated with earlier or concurrent increases in number syncretism due to sound change, and that larger changes to number syncretism in both directions are correlated with more advanced case loss.

The second part measures the association of several factors concerning gender: gender-declension relationship, gender syncretism in the core cases (i.e., Nom and Acc) of NP-internal agreement targets, number of genders, and case loss. It tests several specific hypotheses: first, that a reduction in the number of genders occurs when gender and declension are totally dissociated but not when they are partially dissociated or totally associated; second, that case loss is correlated with the development of a more transparent gender-declension relationship; third, that the loss of all gender distinctions in the plural is more likely to occur when gender and declension are partially or totally dissociated; and finally, that more transparent gender-declension relationships are associated with higher gender syncretism.

In order to test the hypotheses involving number syncretism and gender syncretism, I created novel ways to measure the level of number syncretism in the nominal inflection system of a language and the level of gender syncretism in the inflection of agreement targets. In general terms, these indices, i.e., the number syncretism index and gender syncretism index, respectively, are measures of the proportion of forms in the system that are involved in syncretism.

The methodology followed for the analysis is discussed in section 4.1. The number syncretism analysis is presented in section 4.2. The analysis of gender is presented in section 4.3. This starts with a description of the gender-declension relationship in each stage, followed by an analysis of how these relationships are associated with the number of genders in a language and case loss. Next, gender syncretism on agreement targets and gender-declension relationships are analyzed together for individual languages, and then gender syncretism on agreement targets is analyzed in terms of groups based on gender-declension relationship in the modern stage. A more general discussion follows in section 4.4.

4.1. Methodology

To collect the data for the analysis, I examined the noun and agreement target paradigms across multiple stages of relevant Germanic, Romance, and Balkan Sprachbund languages, as well as BCMS, a South Slavic language whose standard variety is not part of the Balkan Sprachbund. Case marking on pronouns was not considered as part of my analysis. The stages of Germanic languages used were OIc, Icelandic, ONw, MNw, NNw, written Faroese, spoken Faroese, OSw, MSw, CSS, ODan, MDan, CSDan, OHG, MHG, CSHG, OS, MLG, ModLG, ODu, MDu, CSDu, OE, LOE, EME, and LME. The stages of Romance languages used were CL, WVL, EVL, Pre-Italian, CSI, ES, OF, MF, CSF, Pre-Romanian, and CSRm. The stages of Balkan Sprachbund languages used were LPS, PMB (eastern), MB (eastern), CSB, PMB (western), MB (western), CSM, MyG, AG, KG, ModGr, PA, and CSA. The stages of South Slavic languages used were MSrb and BCMS, in addition to those in the Balkan Sprachbund. The analysis is focused on historical case and gender loss, so the earliest attested stage was used as a starting point for most languages. Thus, some proto-languages discussed in chapter II were not included in the analysis, namely PIE and PGmc. However, two were included: Proto-Albanian, since no stage is attested before CSA, and LPS, for which evidence from OCS is used. In addition, reconstructed stages were sometimes interpolated between attested stages. These are generally the stages containing the prefix *Pre-*. In the number syncretism analysis, only one stage was included after case distinctions on nouns had been completely lost and number syncretism eliminated. Thus, LME, ES and MF were included in this part of the analysis, but not ModE, ModSp, or CSF. Similarly, modern stages were not included in the gender analysis if there was no change from the previous

stage. Thus, ModE and ModSp were excluded from this part as well, but CSF was included because gender syncretism has changed from MF.

4.1.1. Number Syncretism Measure

For the number syncretism analysis, I created tables showing the changes in noun declension between each pair of successive stages in the languages investigated. The total number of declension classes in each stage was counted based on the columns in these tables, with some differences noted below. These tables are provided with their respective languages in chapter II above; in section 4.2, descriptions of how these tables relate to the number of classes used in the analysis are included with the discussions of how number syncretism changed in each language. Tables were not created for pairs in which the second stage underwent total case loss; likewise, for stages with total case loss, the number of classes used in the analysis are not discussed in section 4.2 because the number syncretism is always zero. However, a table was created for the development of ModLG from MLG, during which all case distinctions were lost on nouns but not on NP-internal agreement targets, and for the development of ES from WVL, for comparison with the stages of the other Romance languages during the same time period, which had not yet undergone total case loss like ES.

I measured two types of number syncretism separately: number syncretism due to sound changes and number syncretism due to analogical processes. These measures are explained below. I determined the number syncretism that occurred synchronically within each declension class. Number syncretism refers to case/number forms that are identical to at least one form of the other number. Thus, the minimum requirement for number syncretism is a pair of syncretic case/number forms such as Nom singular/plural or Gen singular/Nom plural. A declension class with one such pair is considered to have two forms involved in number syncretism. Each additional form that is identical to the pair adds one form to the total. If a sound change resulted in number syncretism for a large proportion of nouns in a class, but still left a large proportion without number syncretism, these were divided into subclasses, and the number of forms involved were averaged across the subclasses. The number of forms involved in number syncretism from each declension class were added together to calculate the total for each stage under investigation. Next, the total number of syncretic forms was divided by the total number of forms in the declension system, i.e., by the product of the number of declension classes, cases,

and number categories,[128] to yield a number syncretism index, which I propose as an original way to compare the degree of number syncretism across stages and languages.

The following step involved a diachronic calculation of the effects of analogy. When an analogical process contributed to a change in number syncretism between stages, its effect was measured in terms of the change in number of syncretic forms. Usually, the process reduced the number of forms and pairs involved, but sometimes it increased them. The total effect of analogy on observed changes in number syncretism was calculated by summing up all of the increases (expressed with positive numbers) and decreases (expressed with negative numbers). If one or more analogical processes were optional during that stage, a range in the effect of analogy would result; this range was carried over further calculations. In other instances, an analogical process did not directly contribute to an observable change in number syncretism between stages. If the analogy had not applied, however, number syncretism would have increased. The number of forms that were thus prevented from participating in number syncretism were calculated separately. These values were used in two ways. First, they were added to the total number of syncretic forms in the later stage of each language pair. These totals represent the predicted degree of number syncretism if no analogical processes had applied. Second, they were added to the totals for the effect of analogy on observed changes. The result was a measure of the total effect of analogy on both observed and predicted number syncretism for each category. These two series of values were then used in further calculations.

Several steps were involved in measuring the decrease in forms involved in number syncretism due to analogy between two stages of a language. First, a second index of number syncretism was calculated by dividing the total number of syncretic forms including predicted forms by the total number of forms in the declension system, which was the same as in the earlier stage unless a case had been lost due to sound change alone; when this occurred, the new number of cases is used in calculating the total number of forms in the declension system. The decrease due to analogy was calculated by subtracting the first, synchronic index of number syncretism from the second, predicted index of number syncretism without analogy. In addition, the synchronic index of number syncretism was divided by the predicted index of number syncretism without analogy. The result of this calculation is the proportion of number syncretism

[128] Two number categories were assumed for all stages. Some early stages had dual forms on nouns, but these were not included in the analysis.

remaining after analogical processes applied. Subtracting this value from 1 and multiplying by 100% yielded the percent decrease in syncretic forms due to analogical processes. The increase due to sound change was calculated by subtracting the synchronic index of the earlier stage from the predicted index without analogy in the later stage. The percent increase due to sound change was then calculated by subtracting 1 from this value and multiplying by 100. The results of these various original calculations for each stage of the Germanic, Romance, and Balkan Sprachbund languages are included in Table 49, Table 50, and Table 51 in sections 4.2.1, 4.2.2, and 4.2.3, respectively; the results are also discussed by individual language in those sections.

4.1.2. Gender Types Criteria

Two factors in the gender analysis are relatively straightforward. The number of genders in each stage of the languages under investigation was gathered from grammatical descriptions of the languages. The main distinction concerning case for this part of the analysis was whether a language had lost all case distinctions by the modern stage. The other factors required more analysis.

The relationship between gender and declension in each stage was determined according to the categorization system used by Kürschner and Nübling (2011), as introduced in section 3.1.2.1 above. They describe four types of relationships. A Type 1 relationship is the total association of gender and declension, meaning there is exactly one declension for each gender. A Type 2 system has a partial association between gender and declension, i.e., each declension class is clearly associated with one gender, but at least one gender is associated with more than one declension class. A Type 3 relation is a partial dissociation, meaning that at least one declension class is associated with two genders, in addition to at least one gender having multiple associated classes. A Type 4 system has total dissociation of gender and declension, i.e., the choice of declension class has no connection to gender and is instead conditioned by more transparent features such as semantics or phonological form. However, Kürschner and Nübling (2011) do not explicitly state their criteria for a clear association between a particular class and a single gender, or for which classes they consider.

Therefore, I set the following criteria. First, minor classes were generally excluded, and a particular gender was only considered to be associated with a class if more than a handful of nouns with that gender belonged to that class; except where noted, the classes considered

generally correspond to those used in the number syncretism analysis, as described in sections 4.2.1, 4.2.2, and 4.2.3. For modern stages not described there, the major classes were determined based on grammatical descriptions. For the modern Germanic languages, particularly those without morphological case, I usually followed Kürschner and Nübling (2011), who generally indicate in some way which classes they consider major and minor, if not the reasoning behind those decisions.[129] Second, when there were two or more classes associated with different genders that generally corresponded, e.g., masculine and neuter *o*-stems, only one form had to differ between each pair of classes for them to be considered different classes, meaning there was a clear association of each with a single gender. Third, only indefinite forms were considered. Like preposed definite articles and adjectives, postpositive definite articles are treated as NP-agreement targets, and therefore analyzed with those and not as part of noun declension. These criteria are applied to each stage of the Germanic, Romance, and Balkan Sprachbund languages in sections 4.3.14.3.2, and 4.3.3, respectively. In this part only, PGmc was included to confirm that all of the languages under investigation can be traced back to a Type 3 system, even though most of the earliest attested Germanic languages are Type 2 according to the criteria.

For the analysis in sections 4.3.4 and 4.3.8, I divided languages within each genetic and/or areal group into subgroups based on the gender-declension type of their modern stage. Section 4.3.4 contains an analysis of how these subgroups are associated with the number of genders and case loss. The final factor in the gender analysis is gender syncretism on the core cases of NP-internal agreement targets. Gender is a morphological category for agreement targets but a lexical category for nouns, so gender distinctions can only be said to exist if they are marked on agreement targets. Thus, gender syncretism on agreement targets is a direct measure of how clearly gender categories are distinguished. This investigation focused on NP-internal agreement targets, as opposed to pronouns and predicate adjectives. Gender distinctions on pronouns are clearly not sufficient for the retention of grammatical gender on nouns, since many languages, including English, CSDu, and CSS, have more gender distinctions on pronouns than nouns. At the same time, these distinctions no longer represent grammatical gender, which is a

[129] Kürschner and Nübling (2011) categorize the gender-declension relationship for the following languages that are also included in my study: CSS (Type 2); Icelandic, Faroese, and CSHG (Type 3); CSDan and CSDu (Type 4). They also categorize the East Friesland dialect of Low German as Type 4, while my study considers Low German more broadly. They do not explicitly categorize earlier stages of any Germanic language, although they imply that all Germanic languages were once Type 3. They do not consider languages outside the Germanic family.

further indication that the survival of grammatical gender depends on distinctions within the NP (see De Vos & De Vogelaer 2011:249). NP-internal agreement targets include determiners and adjectives. Determiners include demonstratives, preposed articles, and postposed articles.

For the Germanic languages, this investigation only considers strong adjectives because, except in the earliest stages, weak adjectives almost exclusively occur together with more informative determiners. They only contribute additional disambiguating information in a few exceptional instances: for example, the ambiguity between plural and feminine singular for the CSHG article *die* is resolved by the distinction between *-en* and *-e* on a following weak adjective. This example concerns number syncretism, however, for which nouns themselves are expected to play a greater disambiguating role; the very small contribution of weak adjectives was not worth investigating. On the other hand, adjectives in Albanian are usually used with linking articles, so these are considered together as one set of agreement targets, in addition to the postpositive definite articles. In the following sections, the term *agreement targets* is used more specifically to refer only to demonstratives/definite articles and adjectives within the NP.

4.1.3. Gender Syncretism Measure

The gender syncretism index was calculated in a way somewhat analogous to the number syncretism index. However, the investigation of gender syncretism in my study is limited to comparisons within the same case and number. Unlike with number syncretism, gender syncretism between different cases was not measured because languages are often limited in the number of unique nominal inflections, so this number would be high even in languages with robust gender systems. In addition, morphological case can often be determined based on syntactic context, which could also help disambiguate gender syncretism between different cases. However, when there is gender syncretism within the same case, speakers must rely on differences in declension class, agreement targets, or their knowledge of the lexicon to determine gender. Therefore, this narrower type of gender syncretism is guaranteed to cause ambiguity when it occurs on agreement targets and gender cannot be determined from the declension of the noun. The analysis focuses on the core cases, Nom and Acc, because these occur with the highest frequency. Therefore, when speakers need to disambiguate the gender of a noun based on agreement targets, they will be in these cases most often.

Given two cases, Nom and Acc, and two numbers, singular and plural, there were four forms for each agreement target paradigm. If most agreement targets followed the same pattern of gender syncretism (even if they did not use the same markers to express these distinctions), further calculations were undertaken using only one paradigm. If determiners and adjectives followed different patterns, or if there were multiple major classes of adjectives,[130] each of these was counted as a separate paradigm. If there was minor variation among different determiners, the definite article or its precursor/equivalent, i.e., the distal demonstrative, was used for the determiner paradigm. For each of the four case/number forms in each paradigm, the gender syncretism was classified as follows: 0 if each form was distinctive, 2 if two forms were syncritic, 3 if three did, and 4 if four did. The values of the four forms across all of the relevant paradigms were then added together. Most of the stages have three gender categories. For these, the sum of the syncretism values above was divided by the total number of forms that could potentially participate in gender syncretism, i.e., by the product of the number of relevant agreement target paradigms, four, for the case/number forms, and three, for the gender categories, to yield the gender syncretism index. For stages with a different number of genders, two gender syncretism indices were calculated: one out of three genders to reflect the syncretism in terms of the inherited IE gender system, and another out of the number of remaining genders to reflect the syncretism that speakers were required to deal with as synchronic users of the system. The gender syncretism indices for each stage of the Germanic, Romance, and Balkan Sprachbund languages are included in Table 52, Table 75, and Table 87 in sections 4.3.5, 4.3.6, and 4.3.7, respectively; the results are also discussed by language in those sections. In section 4.3.8, the results are analyzed in terms of the subgroups based on gender-declension type described above.

4.2. Number Syncretism Analysis

My analysis measures the effects of two factors on number syncretism: sound change and analogy. The association of case loss with changes in number syncretism due to these factors is

[130] As mentioned above, weak adjectives are excluded from my analysis, and only the major strong adjective paradigm is considered for Germanic languages. However, most of the other languages under investigation have multiple adjective paradigms.

also analyzed. The analysis was conducted with regard to individual Germanic, Romance, and Balkan Sprachbund languages, as presented in sections 4.2.1, 4.2.2, and 4.3.3, respectively.

4.2.1. Analysis of Number Syncretism in Germanic Languages

This section provides the number of declension classes used for each Germanic language stage in the analysis. The results of the number syncretism analysis for each Germanic language are then discussed. Other factors are referenced in instances where the effects of analogical processes appear unexpected given the effects of sound change.

The results of the number syncretism analysis for the Germanic languages are summarized in Table 49, below. These results were calculated according to the methodology described above. The following paragraphs walk through examples of the calculations for stages of Icelandic and High German, noting how they correspond to the values in the table. The first column of the table shows the stages under investigation, with darker lines separating distinct languages. The second shows the number of declension classes. The number of morphological cases at that stage are provided in the third column. Along with the total number of case/number forms involved in number syncretism in all declension classes, the values in the second and third columns are used to calculate the number syncretism index in the fourth column. For example, OIc had 10 declension classes, 4 cases, and 2 numbers.[131] These values are multiplied together to yield the total number of forms in the declension system, i.e., 80 for OIc. OIc had 12 forms involved in number syncretism; this value is not included in the table but can be determined from the tables and discussion in chapter II. Therefore, 12 is divided by 80 to yield the number syncretism index of 0.15. The last three columns involve changes in the number syncretism from the previous stage. Since there is no stage earlier than OIc in this part of the analysis, these columns are not applicable for OIc.

Moving on to Icelandic, the number of declensions has increased to 11 and the number of forms involved in number syncretism has increased to 14, but there are still 4 cases and 2 numbers, so the new number syncretism index is 14 divided by 88, i.e., 0.16. Since there was no increase in number syncretism due to sound change, there is a zero in the seventh column, and potential syncretism index without analogy in Icelandic is the same as the number syncretism

[131] In most instances, the number syncretism index decreased due to analogy, so positive values are used for a decrease in this column (but not in any other column).

index in OIc, i.e., 0.15, as shown in the fifth column. Since the number syncretism index increased by 0.1 due to analogy, -0.01 is listed in the sixth column, along with -7%.[132] This percentage was calculated by dividing the number syncretism index in the fourth column by the potential syncretism index without analogy in the fifth column (i.e., 0.16/0.15), then subtracting that value (i.e., 1.07) from 1 (i.e., -0.07) and multiplying by 100% to convert from a decimal to a percent, i.e., -7%.

The calculations are slightly more complicated when sound change raises the number syncretism index and when there is optional analogy. The development of MHG from OHG provides an example of both. As with OIc and the other earliest stages, only one number syncretism index is calculated for OHG. It had 12 declension classes, 5 cases, and 2 numbers, i.e., 120 total forms in the declension system; 26 of these forms were involved in number syncretism. Thus, the index is 26 divided by 120, i.e., 0.22.

In MHG, 12 declension classes are still assumed, but the number of cases has decreased to 4, and the number of forms involved in number syncretism has increased to a range of 55 to 60 depending on whether optional analogical processes apply. Thus, the number syncretism index is a range from 0.57, i.e., 55 divided by 96, to 0.63, i.e., 60 divided by 96. When the effects of analogy are ignored, i.e., when only sound changes are assumed to have occurred, the number of forms involved in number syncretism is 60. Since sound change alone accounts for the loss of the Instr, the number of cases would still have decreased to 4 without analogy. Thus, the number syncretism index without analogy is 0.63. This is the same as the upper range of the number syncretism index with analogy because all of the analogical processes in MHG were optional, but this is not true for most of the stages. The decrease due to analogy is calculated by subtracting the syncretism index in the fourth column from the index without analogy in the fifth column, yielding a range of 0 to 0.06. The percent decrease is calculated by instead dividing the fourth column from the fifth column (i.e., 0.57/0.63 and 0.63/0.63), then subtracting those values (i.e., 0.9 and 1) from 1 (i.e., 0.1 and 0) and multiplying by 100%, i.e., 0%-10%. The increase due to sound change in the seventh column is calculated by subtracting the syncretism index in the fourth column for the previous stage instead, i.e., the 0.22 value for OHG, from the index without analogy in the fifth column, yielding 0.41. The percent increase is calculated by instead

[132] Two number categories were assumed for all stages of the languages under investigation, so it is not included in the table.

dividing 0.63 by 0.22 (i.e., 2.86), subtracting 1 from that value (i.e., 1.86), and multiplying by 100%, i.e., 186%.

Table 49. Number Syncretism in Germanic Languages

Stage	# of Declensions	Case	Number			
		# of Cases	Syncretism Index	(Potential) Syncretism Index Without Analogy	Decrease due to Analogy	Increase due to Sound Change
North Germanic						
OIc	10	4	0.15			
Ic	11	4	0.16	0.15	-0.01 (-7%)	0
ONw	10	4	0.15			
Fa (W)	11	4	0.18	0.15	-0.03 (-20%)	0
Fa (S)	11	4	0	0.18	0.18 (100%)	0
OSw	9	4	0.22			
MSw	9	4	0.06-0.43	0.43	0-0.37 (0-86%)	0.21 (95%)
CSS	6	0	0	0.25	0.06-0.43 (100%)	0
ODan	9	4	0.22			
MDan	9	2	0-0.28	0.6	0.32-0.6 (53-100%)	0.38 (172%)
CSDan	2–3	0	0	0.28	0.28 (100%)	0
MNw	10	4	0.05-0.5	0.36	-0.14 – -0.31 (-39% – -86%)	0.21 (140%)
NNw	3-4	0	0	0.67	0.05-0.5 (100%)	0
West Germanic						
OHG	12	5	0.22			
MHG	12	4	0.57–0.63	0.63	0–0.06 (0%–10%)	0.41 (186%)
CSHG	6	4	0.27	0.63	0.36–0.4 (57–63%)	0
OS	13	5	0.27			
MLG	12	4	0.54–0.66	0.67	0.01–0.13 (1–19%)	0.4 (148%)
ModLG	9	2	0	0.61	0.47–0.61 (100%)	-0.19–0.07 (-29%–13%)
ODu	9	4	0.38			
MDu	9	4	0.48–0.65	0.67	0.02–0.19 (3–28%)	0.29 (76%)
CSDu	2–3	0	0	0.65	0.48–0.65 (100%)	0
OE	12	4	0.29			
LOE	12	4	0.34–0.35	0.35	0–0.01 (0–3%)	0.06 (21%)
EME	3	4	0.39–0.56	0.83	0.27–0.44 (33–53%)	0.48–0.49 (137–144%)
LME	1	0	0	0.56	0.39–0.56 (100%)	0

4.2.1.1. Icelandic and Faroese

The 11 declension classes in Table 3 correspond to those used for Icelandic in the quantitative analysis. The masculine strong class with *-ar* in both the Gen singular and Nom plural is considered a declension class in Icelandic, where it gained productivity, but not OIc, for which only 10 classes are counted.

OIc already had a low degree of number syncretism, and Icelandic has not undergone any sound changes that disrupted nominal inflection. Therefore, it is not particularly surprising that no leveling that profiles number has applied. The 7% increase in the number syncretism index, from 0.15 in OIc to 0.16 in Icelandic, is mainly due to a reorganization of the strong declensions, specifically the addition of the class described in the previous paragraph. The *-ar* form already occurred in both the Gen singular and Nom plural of masculine nouns in OIc, just almost never together, so this extension does not seem unusual. This is considered an analogical change, since the *-ar* Gen probably spread from the other masculine class with it, and this combination already occurred in a feminine class as well. However, it did not profile number.

ONw is considered equivalent to OIc, with the same 10 declension classes. The 11 declension classes in Table 4 correspond to those used for Faroese (written and spoken) in the analysis. As in Icelandic, a new masculine strong class has been formed with *-ar* in both the Gen singular and Nom plural.

Like Icelandic, Faroese has not undergone any sound changes that affected nominal inflection. The small 20% increase in the number syncretism index, from 0.15 in ONw to 0.18 in written Faroese, is partially due to the addition of the class described in the preceding paragraph. The remainder of the increase is the result of analogical processes that leveled differences across declension classes associated with different genders. Number profiling played a minor role, if any, in these developments. In spoken Faroese, however, number syncretism was completely eliminated with the loss of a productive Gen and another analogical change that leveled differences across declension classes associated with different genders. Even if number profiling was not the primary motivation for these developments, they achieved an almost unique outcome among the languages under investigation: the total elimination of number syncretism without the loss of all case distinctions on nouns.[133] The lack of disruptive sound changes may have enabled this simplification of the nominal inflection system with minimal loss of inflectional categories. Without the same connection to a written tradition as Icelandic, spoken Faroese may have been more prone to morphological developments that increase transparency, making the system easier to acquire and process. The lack of standardization of the spoken language may have played a role as well.

[133] A similar outcome has occurred in Greek, but the total elimination of number syncretism is optional.

4.2.1.2. Swedish, Danish, and Norwegian

The 9 declension classes in Table 6 correspond to those used for OSw and MSw in the quantitative analysis. The 9 declension classes in Table 7, which are also the same classes as in Table 6, correspond to those used for ODan and MDan in the analysis. As mentioned in the previous section, the same 10 declension classes as in OIc are used for ONw in the analysis. These are also used for MNw and included in Table 8.

Sound change had a profound effect on nominal inflection in the development of Swedish, Norwegian, and especially Danish, in contrast to Icelandic and Faroese. Analogical processes profiling number also had a greater effect. Number syncretism was already higher in OSw/ODan, with an index of 0.22, than in OIc/ONw, at 0.15. The number syncretism index increased 95% to 0.43 in MSw as a result of sound change, but analogical processes optionally lowered the index 86% to 0.06. Sound changes had an even greater effect in MDan; these would have raised the number syncretism index 172% to 0.6. Analogical processes also had more of an effect than in MSw, lowering the index at least 53% to 0.28; with the right combination of optional processes, all number syncretism was eliminated. On average, analogical processes countered most or all of the increase in number syncretism due to sound change, but case syncretism increased even more than it would have with sound change alone, including the leveling of all remaining distinctions between Nom, Acc, and Dat in MDan.

Although it started with a lower index in ONw, Norwegian nominal inflection developed much more similarly to Swedish than to Icelandic or even Faroese; these developments were still not as extreme as in Danish, however. In MNw, the number syncretism index would have increased 140% to 0.36. The effects of analogical processes varied more than in MSw or MDan, ranging from a 39% increase up to 0.5 to an 86% decrease down to 0.05. The Nom and Acc had merged as a result of the analogical processes that usually applied by the end of the MNw period. In the development of CSDan, CSS, and NNw, any remaining number syncretism was eliminated, along with any case distinctions. Thus, the results for Swedish, Danish, and Norwegian support the hypothesis: large increases in number syncretism due to sound change were countered by large decreases due to analogy, ultimately leading to total case loss. In other words, number was profiled at the expense of case after the system was disrupted by sound change. However, this does not explain why Norwegian underwent neutralizing sound changes resembling those in Swedish while Icelandic and Faroese did not.

4.2.1.3. High German

The 12 declension classes used for both OHG and MHG in the quantitative analysis correspond to the 11 in Table 10, along with an additional distinction between masculine *a*-/*wa*-stems and *i*-stems. The 6 declension classes in Table 11 correspond to those used for CSHG. All of the masculine and neuter strong classes have merged into one singular paradigm with several options for the plural marker that largely correlate with gender.

Sound change had an even greater effect on nominal inflection in the development of High German than in Swedish and Danish. However, analogical processes had less of an effect, and this effect was delayed. The number syncretism index in OHG was 0.22, the same as in OSw/ODan. In MHG, it increased 186% to 0.63 due to sound change; analogical leveling optionally lowered the index 10% to 0.57. Due to paradigm reorganization alone, the index would have ranged from 0.54 to 0.63 in the development of CSHG, a decrease of up to 5% from the previous range. Additional analogical processes lowered the index to 0.23 in CSHG, a decrease of 57-63%. An ongoing analogical change, i.e., the extension of the strong masculine/neuter pattern of only distinguishing the Gen in the singular to weak nouns, optionally lowers the index to 0.2, a 13% decrease from 0.23, for a total decrease of 63-68% from MHG. After all of these developments, the number syncretism index in CSHG is almost the same as it was in OHG, but case syncretism is much higher.

Thus, number profiling has occurred in response to sound change, but it was delayed, and some case distinctions remain. More robust case marking has been retained on articles and other agreement targets. These had more salient pronominal forms that have remained distinctive after undergoing the same sound changes as nouns. In addition, more number syncretism appears to have been tolerated on agreement targets than on nouns: all of the plural forms are syncretic with at least one singular form, e.g., Nom/Acc plural and feminine singular, as shown in Table 66 below. Since these are used along with nouns, number marking on the noun itself may have been sufficient. In other words, number was not profiled on agreement targets as it was on nouns. In turn, this support from agreement targets may have prevented the leveling of the last few case distinctions on nouns. Standardization may also account for the retention of the Gen as a morphological case, in contrast to Low German and other German dialects.

4.2.1.4. Low German

The 12 declension classes used for MLG in the quantitative analysis correspond to the 9 in Table 13, along with additional distinctions in 3 of these classes: between masculine *a*-stems and long *i*-stems, between masculine *ja*-stems and shorts *i*-stems, and between long and short neuter *a*-stems. In OS, an additional distinction is assumed between feminine *(j)ō*-stems and short *i*-stems. The 9 declension classes used for ModLG in the quantitative analysis correspond to the 6 in Table 14, along with additional distinctions involving the choice of plural form in the 3 strong classes.

Low German developed similarly to High German until the modern stage, when there were additional sound changes and analogical processes. The number syncretism index in OS was 0.27, slightly higher than in OHG. In MLG, it would have increased 148% to 0.67 due to sound change alone, but analogical processes lowered the index at least to 0.66 and optionally to 0.54, a decrease of 1-19%. Depending on which of these processes applied in MLG, the number syncretism index varied from 0.47 to 0.61 after sound changes in the development of Low German, ranging from a 29% decrease to a 13% increase. By ModLG, however, all number syncretism had been eliminated by analogical leveling and the spread of distinctive plural forms across declension classes, along with all case distinctions for nouns. The lack of standardization in Low German, in contrast to High German, may have played a role in some of these developments. As in High German, agreement targets did not lose case marking to the same extent as nouns. In ModLG, however, there was just one additional distinction—between Nom and Acc in the masculine singular—and this was apparently not enough to support this distinction on nouns, where it was lost with the leveling of the weak masculine singular oblique form in the development of ModLG from MLG.

4.2.1.5. Dutch

The 9 declension classes used for both ODu and MDu in the quantitative analysis correspond to the 6 in Table 16, along with additional distinctions in 3 of these classes: between masculine *a*-stems and *i*-stems, between long and short neuter *a*-stems, and between feminine *ō*-stems and weak nouns.

The development of Dutch has similarities with the other continental West Germanic languages, but also with Swedish and Danish. The number syncretism index in ODu was 0.38,

the highest of the early Germanic languages. In MDu, it would have increased 76% to 0.67 due to sound change alone, but analogical processes lowered the index at least to 0.65 and optionally to 0.48, a decrease of 3-28%. Thus, analogy tended to have a slightly stronger effect on number syncretism in MDu than in MHG or MLG, but still much less than in MSw or MDan. In the development of CSDu, all number syncretism was eliminated, along with all case distinctions. This was the same outcome as in CSS and CSDan, adding further support to the number profiling hypothesis. In fact, it was also the same outcome as ModLG with respect to noun paradigms; the key difference from Low German is that Dutch has also lost all case marking on articles and other agreement targets. Although standardization occurred in Dutch in a way it did not in Low German, it was later than in High German, so it may have been too late to have an effect on these developments.

4.2.1.6. English

The 12 declension classes used for both OE and LOE in the quantitative analysis correspond to the 11 in Table 18, along with an additional distinction between feminine strong paradigm with *-a* in the Nom/Acc plural and those with *-e*. The 3 declension classes in Table 19 correspond to those used for EME in the quantitative analysis. Not all OE declension classes directly contributed forms to EME classes. Instead, their members adopted the endings of another class.

Of all the Germanic languages in the investigation, sound change had the greatest effect on nominal inflection in English. Analogical processes profiling number also had a profound effect, largely accompanying sound change as in Swedish and Danish, as opposed to the more delayed reaction in the continental West Germanic languages, particularly High and Low German. The number syncretism index in OE was 0.29, between OS and ODu. By LOE, however, it had increased 21% to 0.35 due to sound change; analogical leveling optionally lowered the index 3% to 0.34. In EME, it would have increased 137-144% to 0.83 due to sound change alone. This is both the highest potential level of number syncretism at any stage of a Germanic language and the largest increase in the index itself, at 0.48-0.49, even though it is not the largest percent increase. However, analogical processes lowered the index at least to 0.56 and optionally to 0.39, a decrease of 33-53%. In the development of LME, all number syncretism was eliminated, along with all case distinctions. These developments can still be attributed to

number profiling, but this does not really explain why they occurred earlier than in other Germanic languages.

4.2.2. Analysis of Number Syncretism in Romance Languages

This section continues the number syncretism analysis with regard to individual languages. As with the Germanic languages in the previous section, the number of declension classes used for each Romance language stage in the analysis are provided. Then the results of the analysis for each Romance language are discussed, with references to other factors where unexpected results occur. These results are summarized in Table 50.

Table 50. Number Syncretism in Romance Languages

Stage	# of Declensions	Case	Number			
		# of Cases	Syncretism Index	(Potential) Syncretism Index Without Analogy	Decrease due to Analogy	Increase due to Sound Change
CL	7	6	0.08			
WVL	7	2	0.18	0.27	0.09 (33%)	0.19 (238%)
PI	6	2	0.25	0.25	0	0.07 (39%)
CSI	3	0	0	0.25	0.25 (100%)	0
ES	6	0	0	0.21	0.21 (100%)	0
OF	9	2	0.58	0.44–0.31	-0.14 (-32%)	0.26 (144%)
MF	2	0	0	0.58	0.58 (100%)	0
CSF	2	0	0	0	0	0
EVL	7	6	0.27	0.27	0	0.19 (238%)
PR	7	3	0.36–0.43	0.38	-0.05–0.02 (-13%–5%)	0.11 (41%)
CSRm	7	3	0.19	0.36–0.43	0.17–0.24 (47%–56%)	0.19 (0%)

4.2.2.1. Western Romance: Italian, Spanish, and French

The 7 declension classes used for both CL and VL in the quantitative analysis correspond to the 5 in Table 22, along with 2 additional distinctions in one of these classes: between masculine/feminine consonant-stems and *i*-stems, and between masculine/feminine consonant-stems with a zero ending in the Nom singular and those with *-s*. The 6 declension classes used for Pre-Italian and ES in the quantitative analysis correspond to the 4 in Table 24 and Table 25, respectively, along with the same 2 additional distinctions in the masculine/feminine consonant-stems and *i*-stems. The 9 declension classes used for OF in the quantitative analysis correspond to the 5 in Table 26, which include separate masculine and feminine classes for the consonant-

stems and *i*-stems to capture their divergent development, along with the same 2 additional distinctions in each of these as in consonant-stems and *i*-stems in other Romance languages.

Sound change and analogical processes had a profound effect on nominal inflection in the Western Romance languages, with many similarities to the Germanic languages, particularly those that lost morphological case. Number syncretism was quite low in CL, with an index of 0.08. The number syncretism index increased to 0.18 in WVL but would have increased 238% to 0.27 if analogical processes had not applied. These reduced the index by 33%, countering about half of the increase in number syncretism. However, case syncretism increased even more than it would have with sound change alone, and of the 6 morphological cases in CL, only Nom and Acc remained. These developments resemble those in the transition from LOE to EME and from OSw/ODan to MSw and MDan, but with a few notable differences: the changes to the number syncretism index in WVL were less extreme than these Germanic languages in absolute terms; in terms of percent change, the increase due to sound change was larger due to a lower initial index, but the decrease due to analogy was lower; finally, more case distinctions were lost in the process, but there were more to begin with.

Following their common development in WVL, the Western Romance languages under investigation all developed differently in terms of number syncretism. By their modern stages, however, they have lost all case distinctions but no longer have any number syncretism. As reconstructed, the number syncretism index increased 39% to 0.25 in Pre-Italian as the result of sound change. Analogical processes did not directly affect number syncretism until the development of CSI, when all number syncretism was eliminated, along with all case distinctions. In the development of ES, all number syncretism was eliminated, along with all case distinctions, but the index would have increased 17% to 0.21 due to the disappearance of the neuter consonant-stems and *i*-stems if analogical leveling had not applied in the remaining declension classes. Part of the difference between Italian and Spanish may have to do with when they were first attested as vernacular languages, which has affected how they are divided into stages. In OF, the number syncretism index would have increased 144% to 0.44 due to sound change alone, but analogical processes actually increased it an additional 32% to 0.58. These analogical processes mostly reduced the differences between declension classes of the same gender; like the processes that increased number syncretism in Icelandic and Faroese, they were

generally not motivated by earlier increases in number syncretism due to sound change.[134] All number syncretism was eliminated in the development of MF, along with all case distinctions. In addition to being the only Romance language in which analogical processes temporarily increased number syncretism, French also experienced the largest increase in number syncretism of the Western Romance languages, both in absolute terms and percentage. The single increase of 0.26 from WVL to OF is comparable to the early increases attested in the early stages of all the Germanic languages except Icelandic and Faroese, particularly the 0.25 increase from OSw to MSw and the 0.29 increase from ODu to MDu. When this is combined with the 0.19 increase due to sound change in WVL, the overall increase of 0.45 in French is about as high as the 0.41-0.47 increase in High German. The overall increase of 0.26 in Italian is nearly identical to the 0.25 increase in Swedish, while the overall increase of 0.22 in Spanish is not much lower.

Thus, similar increases in number syncretism due to sound change in these Romance and Germanic languages were generally followed by analogical processes with similar effects in terms of both number syncretism and case loss, with the exception of the temporary increase in number syncretism in OF. Overall, these results indicate that the number profiling hypothesis does not apply only to the Germanic languages.

4.2.2.2. Eastern Romance: Romanian

The same 7 declension classes are used for both EVL and Pre-Romanian in the quantitative analysis as for CL and WVL. These also correspond to the 5 in Table 27 and Table 28, along with the same 2 additional distinctions in one of these classes: between masculine/feminine consonant-stems and *i*-stems, and between masculine/feminine consonant-stems with a zero ending in the Nom singular and those with *-s*. The 7 declension classes used for Romanian in the quantitative analysis correspond to the 7 in Table 30, which include separate masculine and feminine classes for the *ā*-stems as well as the consonant-stems and *i*-stems to capture their divergent development. The 2 additional distinctions in consonant-stems and *i*-stems are no longer relevant in Romanian, so they are not included.

[134] See section 2.4.1.3 above for a discussion of these analogical processes. It should be noted that even in stages of languages where analogical processes resulted in a net decrease in number syncretism, some analogical processes had other motivations and/or increased number syncretism slightly. However, these types of analogy did not play as significant a role as in OF.

In the development of Romanian, sound change had a similar effect on nominal inflection as in the Western Romance languages, but analogical processes had a smaller, delayed effect. In the development of EVL from CL, which had an index of 0.08, number syncretism increased 238% to 0.27, as it would have in WVL if, as is assumed for EVL, analogical processes had not applied. In Pre-Romanian, the number syncretism index would have increased 41% to 0.38 due to sound change. When analogical processes—some optional—are considered as well, the index varied from 0.36 to 0.43, ranging from a 13% decrease to a 5% increase. These processes included the leveling of the last distinctive Abl form to the Acc and the last distinctive Gen forms to the Dat, reducing the number of morphological cases from 6 to 4. In the development of CSRm, analogical processes lowered the number syncretism index to 0.19, a decrease of 47-56%. Case syncretism increased, however, and the last distinctions between the Nom and Acc were leveled, leaving 3 morphological cases for nouns. The trajectory of the number syncretism index from CL to CSRm resembles the one from OHG to CSHG: an early increase due to sound change accompanied by some analogy, but without much decrease in number syncretism until more significant analogy in the development of the modern stage. The changes to the number syncretism index in Romanian were less extreme in absolute terms but actually larger in terms of percent change due to a lower initial index. In addition, case distinctions were not lost in the development of High German as they were in Romanian, but the latter was initially attested with more cases.[135] As in High German, however, morphological case was retained in Romanian despite overall increases in number syncretism in the same range as languages that lost case. In fact, the overall increase in Romanian was 0.3, higher than in Italian and Spanish. One feature shared by CSHG and CSRm is the retention of robust case marking on articles despite limited distinctions on bare nouns; this is a likely factor in distinguishing their outcome from closely related languages that underwent similar sound changes. Unlike in CSHG, however, there is no number syncretism in the CSRm postpositive article paradigm, so the additional salience of these forms compared to nouns may have been sufficient for the retention of case distinctions; it is not clear that number is a less relevant category for agreement targets than nouns in CSRm, as it appears to be in CSHG.

[135] All of the languages under investigation began with the same PIE case system, but experienced different degrees of case loss before they were first attested.

4.2.3. Analysis of Number Syncretism in Balkan Sprachbund languages

This is the final section of the number syncretism analysis with regard to individual languages. As in the previous sections, this one provides the number of declension classes used for each Balkan Sprachbund language in the analysis. A discussion of the results for each Balkan Sprachbund language follows. A summary of these results in presented in Table 51.

Table 51. Number Syncretism in Balkan Sprachbund Languages

Stage	# of Declensions	Case	Number			
		# of Cases	Syncretism Index	(Potential) Syncretism Index Without Analogy	Decrease due to Analogy	Increase due to Sound Change
LPS	10	7	0.32			
PMB (Eastern)	10	7	0.48–0.5	0.52	0.02-0.04 (4–8%)	0.2 (63%)
MB (Eastern)	9	4	0.18	0.50	0.3-0.32 (62–64%)	0
CSB	4	2	0	0.18	0.18 (100%)	0
PMB (Western)	10	7	0.42	0.44	0.02 (5%)	0.12 (28%)
MB (Western)	9	4	0.13	42	0.29 (69%)	0
CSM	4	2	0	0.13	0.13 (100%)	0
MSrb	10	6	0.21–0.54	0.47	-0.07 – 0.26 (-14% – 55%)	0.15 (47%)
BCMS	8	6	0.17–0.18	0.2–0.52	0.02–0.35 (10–67%)	0
MyG	8	6	0			
AG	8	5	0.03	0.04	0.01 (25%)	0.04 (400%)
KG	8	5	0.15	0.15	0 (0%)	0.12 (400%)
ModGr	8	4	0-0.06	0.18	0.12–0.18 (67–100%)	0.03 (20%)
PA	2	5	0.25			
CSA	2	5	0.14	0.31	0.17 (55%)	0.06 (24%)

4.2.3.1. Eastern South Slavic: Bulgarian and Macedonian

The 10 declension classes in Table 33 and Table 34 correspond to those used for LPS in the quantitative analysis, as well as eastern and western dialects of PMB. The same 10 declension classes occur in Table 35 and Table 36 correspond to those used for eastern and western dialects of MB.

In the development of Bulgarian and Macedonian, sound change and analogical processes affected nominal inflection in ways similar to many of the Germanic and Western Romance languages. Number syncretism was already quite high in LPS, with an index of 0.32. This was higher than in most of the early Germanic languages. In eastern dialects of PMB, number

syncretism would have increased 63% to 0.52 due to sound change alone, but analogical processes lowered the index at least to 0.5 and optionally to 0.48, a decrease of 4-8%. Sound change had slightly less of an effect in western dialects; the number syncretism index increased to 0.42 but would have increased 28% to 0.44 without analogy. Thus, there was a 5% decrease due to analogy. However, the majority of analogical processes, primarily consisting of the mergers of the Instr and Loc with the Acc and of the Gen with the Dat, are assumed to have taken place in the development of MB from PMB. In eastern dialects of MB, these lowered the index to 0.18, a decrease of 62-64%. The index decreased 69% to 0.13 in the western dialects. In the development of CSB and CSM, all number syncretism was eliminated, along with all case distinctions. The trajectories of the number syncretism index from LPS to CSB and CSM have some similarities with those from OSw/ODan to CSS and CSDan. When the developments in PMB and MB are considered together, they resemble those in MSw and MDan, with analogical processes countering the increase due to sound change. In both major dialects of MB, the effects of these processes were even greater, bringing the index below its starting point in LPS, while MSw and MDan had indices with a range that included the starting point in OSw/ODan. The extra decrease in MB was only made possible by large increases in case syncretism, including the reduction in morphological cases from 7 to 4. The resulting case system does not appear to have been sustainable, leading to the loss of all nominal case distinctions in CSB and CSM, as in CSS and CSDan. These parallels add further support to the validity of the number profiling hypothesis outside of the Germanic languages.

4.2.3.2. Western South Slavic: Bosnian-Croatian-Montenegrin-Serbian

The 10 declension classes in Table 38 correspond to those used for both LPS and MSrb in the quantitative analysis. The 8 declension classes in Table 39 correspond to those used for BCMS.

Sound change had a similar effect on nominal inflection in the development of BCMS as in Bulgarian and Macedonian, but analogical processes had a smaller effect, and many were optional at first. In the development of MSrb from LPS, which had an index of 0.32, number syncretism would have increased 47% to 0.47 due to sound change alone. Depending on which analogical processes applied, including variation among many forms in the plural oblique cases, the number syncretism index varied from 0.21 to 0.54 in MSrb, ranging from a 55% decrease to a 14% increase. Due to paradigm reorganization alone, the index would have ranged from 0.2 to

0.52 in the development of BCMS, a decrease of 4-5% from the previous range. Additional analogical processes, including the leveling of the Loc to the Dat, or very rarely the Instr, lowered the index at least to 0.18 and optionally to 0.17, a decrease of 10-67%. The increase in number syncretism due to sound change in MSrb falls between the increases in western and eastern dialects of PMB. In terms of analogical processes, however, BCMS has diverged from Bulgarian and Macedonian, and its overall trajectory more closely resembles High German and Romanian. Differences include higher initial number syncretism in LPS than in OHG and especially CL, along with less of an increase due to sound change. When the very early changes from CL to EVL are excluded, however, developments in Romanian appear much more similar to those in BCMS. In addition, more of a decrease due to analogy was possible in MSrb than in MHG or Pre-Romanian, but on average, number syncretism was about as high in MSrb as in Pre-Romanian. Most notably, analogical processes lowered number syncretism to a similar index in all three languages. In all three, this was accompanied by large increases in case syncretism but with minimal loss of morphological cases. In BCMS, however, this path of development does not appear to have the same connection with agreement targets; although some forms of agreement targets are more salient than the corresponding forms on nouns, both have essentially the same syncretism pattern, and there is no definite article.

4.2.3.3. Greek

The 8 declension classes in Table 43 correspond to those used for all stages of Greek. Up to KG, there was no distinction between masculine and feminine consonant-stems, so they are combined in Table 42 but still considered separate classes for consistency. In MyG and AG, *i*-stems and *u*-stems followed the same basic pattern as consonant-stems of the same gender, so they are combined as well in Table 40. Of the *i*-stems and *u*-stems, however, only feminine *i*-stems remained productive in KG and ModGr; these are the third class counted for the column of classes from masculine and feminine consonant-stems, *i*-stems, and *u*-stems in Table 41.

Greek has experienced some increase in number syncretism due to sound change during each stage of its development, as well as decreases due to analogical processes during two stages. However, these changes were still much smaller less in absolute terms than most Germanic and Romance languages. As reconstructed, there was no number syncretism in MyG nominal inflection. The number syncretism index increased to 0.03 in AG, but would have been

0.04 without analogy, specifically the merger of the Instr and Dat, which accounts for the 25% decrease. However, number syncretism was very low either way, and it most likely did not motivate this merger, as described in section 2.5.3 above. The number syncretism index grew to 0.15 in KG as the result of sound change, technically a 400% increase, but still low overall. Analogical processes did not directly affect number syncretism in KG, but they lowered the index at least to 0.06 in ModGr and optionally eliminated all number syncretism, a decrease of 67-100%. However, the index would have increased 20% to 0.18 without analogy. The trajectory of the number syncretism index from KG to ModGr closely resembles the one from ONw to spoken Faroese. One case was lost as part of these developments in both languages: Gen in Faroese, Dat in Greek. Unlike the Gen in ONw and written Faroese, however, the Dat in KG was not involved in much number syncretism; rather, it was often syncretic with other cases of the same number, particularly in the singular. In Greek, other analogical processes such as the retention of *-n* only in the Gen plural played a larger role in reducing number syncretism. These parallels suggest that another aspect of the number profiling hypothesis holds outside of the Germanic languages: smaller increases in number syncretism due to sound change are countered by less extreme analogical processes, so number can be profiled with minimal loss of other categories.

4.2.3.4. Albanian

The 2 declension classes in Table 45 correspond to those used for both Proto-Albanian and CSA in the quantitative analysis.

In the development of Albanian, sound change had a smaller effect on nominal inflection than in all of the languages under investigation except Icelandic and Faroese, but analogical processes still lowered number syncretism significantly. As reconstructed, Proto-Albanian had moderately high number syncretism, with an index of 0.25, but the actual index may have been lower: if the unreconstructable forms, i.e., the Voc singular forms of both declension classes and the *ā*-stem Gen-Dat and Abl plural forms, were not involved in any number syncretism, the index would have been 0.2, and additional unreconstructable declension classes such as *i*-stems may have changed the index as well. The number syncretism index decreased to 0.14 in CSA but would have increased 24% to 0.31 if analogical processes had not applied. Thus, there was a 55% decrease due to analogy. Case syncretism increased, however, and the last distinctive

indefinite Acc form was leveled to the Nom. The trajectory of the number syncretism index from Proto-Albanian to CSA has some similarities with the one from EVL to CSRm. As analyzed, the latter took place across three stages, making it possible to observe a delay before most of the decrease due to analogy, similar to High German, but this was not possible in the former, which took place across only two stages. After an early increase from CL, however, EVL had only slightly more number syncretism than Proto-Albanian. In both, analogical processes more than countered a subsequent increase due to sound change, and only the distinction between the Nom and Acc was lost in the process. Both languages have also retained some case distinctions on definite nouns but not indefinite nouns. However, it is less likely in Albanian than in Romanian that case marking on articles was an important factor in the retention of morphological case, since sound changes were less disruptive in the former, and bare nouns still make a lot more distinctions.

4.3. Gender Analysis

This section considers various factors that relate to changes in gender. On the noun side, gender-declension relationships are considered. Their association with the number of genders and case loss are then analyzed. The focus then shifts to gender syncretism on agreement targets, but this is still analyzed with reference to the gender declension relationships, first in individual languages and then in groups.

4.3.1. Gender-Declension Relationship Types in Germanic Languages

In this section, the criteria established in section 4.1 above are applied to each stage of the Germanic languages included in the gender analysis in order to determine their gender-declension type.

4.3.1.1. Proto-Germanic

In PGmc, not all declension classes were clearly associated with a single gender. As shown in Table 2, there was no distinction between masculine and feminine *i*-stems, *u*-stems, or consonant-stems. The corresponding neuter classes could be distinguished based on at least the Nom/Acc singular, and most likely the Nom/Voc/Acc plural, although it cannot always be reconstructed. Similarly, neuter *a*-stems could be distinguished from masculine *a*-stems in the

Nom/Voc singular and the Nom/Voc/Acc plural. Therefore, PGmc had a Type 3 declension system.

4.3.1.2. Icelandic and Faroese

In OIc, each major declension class was clearly associated with one gender. As shown in Table 3, there were both masculine and feminine classes for *a*-stems and *i*-stems, but these were consistently distinguished in the Nom singular, where masculine had *-r* but feminine had a zero ending, and the Acc plural, which was the same as the Nom plural for feminine but lacked the *-r* of this form for the masculine. Neuter strong nouns were distinguished by a zero ending in the Nom/Acc plural. Among the weak classes, the neuter was the same as the feminine in the Nom singular and Gen plural, and it was the same as the masculine in the other singular cases, but taken together, each gender still had a different declension pattern in the singular, and all three genders were distinctive in the Nom and Acc plural. Therefore, OIc had a Type 2 declension system. The situation has remained essentially the same in Icelandic, so it is treated as a Type 2 language, not Type 3, as proposed—without justification—by Kürschner and Nübling (2011:378).

ONw is considered equivalent to OIc. Each declension class is still associated with a single gender in Faroese, although not quite as clearly. As shown in Table 4, the masculine Acc plural is no longer distinctive; it was leveled to the Nom plural, matching the pattern in the feminine. However, masculine strong nouns are still distinguished by *-ur* in the Nom singular. Additional analogical changes in spoken Faroese eliminated the distinction between neuter and feminine weak nouns in the Nom/Acc plural, but these can still be distinguished by the singular declension pattern. Therefore, both written and spoken Faroese are treated as Type 2 for my analysis. Kürschner and Nübling (2011:378) instead ascribe a Type 3 system to Faroese, as with Icelandic.

4.3.1.3. Swedish, Danish, and Norwegian

Each major declension class was clearly associated with one gender in OSw. As shown in Table 6, all three genders were distinctive in the Acc plural, and masculine strong nouns were distinguished by *-er* in the Nom singular. These distinctions were the same as in OIc/ONw, as were the distinctions among weak nouns. In addition, a zero ending in the Dat singular was more

reliably associated with feminine strong nouns than in OIc/ONw. Therefore, OSw also had a Type 2 declension system. In MSw, the masculine-feminine distinction in the Acc plural was usually neutralized by sound change, but the neuter was still distinctive. The strong masculine and neuter Dat singular forms were increasingly leveled to a zero ending during this period, so this was no longer associated with feminine strong nouns. The strong masculine Nom singular form was sometimes leveled to a zero ending as well, and the Gen singular form *-s* spread to all classes, at least as an option. As a result, masculine and feminine strong nouns could no longer be reliably distinguished. However, the distinctions among weak nouns remained. Therefore, a Type 3 system developed in MSw. It should be noted that this was only a Type 3 system as long as the masculine and feminine remained distinctive categories on agreement targets; once they merged into the common gender, the same declension classes were once again only associated with a single gender each. As a result of total case loss, declension is only indicated by plural markers in CSS. Four of the six plural markers are limited to either the neuter or the common gender. The other two markers, a zero ending and *-er*, are still primarily associated with the neuter and common gender, respectively. Most of the exceptional common nouns with a zero plural and neuter nouns with the *-er* plural have formal and/or semantic features that clearly indicate their gender (see Andersson 1994:279). Despite these exceptions, Kürschner and Nübling argue that CSS has a Type 2 system because a noun's gender can almost always be predicted based on its declension (2011:370-371, 378). It could be argued that CSS actually has a Type 3 system, but my analysis follows Kürschner and Nübling.

ODan is considered equivalent to OSw. Not all declension classes were still clearly associated with a single gender in MDan. As shown in Table 7, sound change neutralized the masculine-feminine distinction in the Acc plural, as in MSw. More extreme vowel reduction in MDan also neutralized all distinctions between masculine and feminine weak nouns. The Gen singular form *-s* became an option in all classes, as in MSw, and the strong masculine Nom singular and strong masculine/neuter Dat singular forms were leveled to a zero ending. Thus, all distinctions between masculine and feminine classes were lost on nouns, paralleling the loss of the masculine-feminine distinction on agreement targets. Along with the spread of the *-æ(r)* plural to some neuter strong nouns, these could no longer be reliably distinguished from other strong nouns, i.e., those now belonging to the common gender resulting from the masculine-feminine merger. However, neuter weak nouns remained distinctive in the Nom/Acc plural.

Thus, the declension system in MDan had shifted to Type 3. In CSDan, gender has been totally dissociated from declension, which is now indicated only by plural markers as a result of total case loss. More neuter nouns have adopted plural forms previously associated with the common gender. Meanwhile, some common-gender nouns have a zero plural, which had been strongly associated with the neuter. Instead, declension classes are conditioned by prosodic-phonological and semantic features. For example, nouns ending in *-r* all take the *-e* plural form, as do animate monosyllabic nouns ending in a consonant. Nouns with the derivational suffixes *-ing* and *-ling* that refer to humans generally take *-e*, while others take the *-er* plural form (see Kürschner & Nübling 2011:372-373). This behavior contrasts with more conservative declension systems such as CSHG, in which a particular derivational suffix almost always has the same declension (and gender). Thus, CSDan has a Type 4 system because a noun's gender cannot be reliably predicted based on its declension (ibid., 378).

As mentioned in the previous section, each major declension class in ONw was clearly associated with a single gender. The sound changes and optional analogical processes that applied in MSw also applied in MNw. There were several additional processes that sometimes occurred. Masculine strong nouns sometimes became more distinctive by adopting the regular development of the Nom singular (*-e*) in the Acc as well. On the other hand, weak nouns could no longer be reliably distinguished in the singular due to the interchangeable use of the Nom and Acc/Dat/Gen forms and the extension of the masculine Nom form to the feminine paradigm as well. However, masculine weak nouns were still distinctive in the Nom/Acc plural. Neuter strong nouns were also distinctive in these forms, as in MSw. Thus, a Type 3 system had developed in MNw, with even more overlap among genders than in MSw. Although total case loss resulted in plural markers being the only indication of declension in NNw as well as CSS, NNw developed quite differently from CSS in other respects. Masculine and feminine have not merged, and a different plural form is primarily associated with each gender: *-ar* for masculine, *-er* for feminine, and a zero plural for neuter (see Enger 2004:55-56). Therefore, Kürschner and Nübling classify NNw declension as a Type 1 system (2011:374). However, Askedal lists *-er*, *-r*, and a zero plural as other possibilities for masculine, *-ar* and *-r* as other possibilities for feminine, and *-o* as another possibility for neuter (1994:230). The multiple options for each gender and overlap between masculine and feminine would indicate a Type 3 system, but it is possible that most of these forms are rare enough to justify a Type 1 classification. At the very least, it appears that

some dialects, such as those of Rana and Røros, have an almost consistent one-to-one relationship between gender and declension (see Enger 2004:60). Despite the exceptions, my analysis follows Kürschner and Nübling in treating NNw as Type 1.

4.3.1.4. High German

In OHG, each declension class had a clear association with one gender. As shown in Table 10, all strong masculine classes had a corresponding neuter or feminine class, but each could be distinguished by at least one form. Corresponding masculine and neuter classes differed in the Nom/Acc plural. For example, masculine *a*-stems, *wa*-stems, and *ja*-stems had *-a* for this form, while neuter *a*-stems had a zero ending, and neuter *ja*-stems had *-i*. On the other hand, masculine and feminine *i*-stems differed in the Gen, Dat, and Instr singular. Other feminine strong classes had no corresponding classes of another gender. Among weak nouns, all three genders were distinctive in the Acc singular and Nom/Acc plural. Therefore, OHG had a Type 2 declension system. In MHG, however, vowel reduction and deletion neutralized many of the gender distinctions described above. Masculine and feminine *i*-stems still differed in at least the Gen singular, but the Nom/Acc plural was *-e* for both masculine and neuter *ja*-stems, and a zero ending for masculine *a*-stems, *wa*-stems, and *i*-stems with stem-final nasals and liquids, as well as neuter *a*-stems. Masculine and feminine weak nouns could no longer be distinguished, and the neuter was only distinctive in the Acc singular. Thus, MHG had developed a Type 3 declension system. As shown in Table 11, the merger of the strong and weak feminine classes has resulted in a distinctive feminine declension in CSHG, easily distinguishable from masculine and neuter strong nouns in both numbers, as well as masculine weak nouns in the singular. Although the neuter weak declension has been reduced to a few irregular nouns, the overlap between masculine and neuter strong nouns has only grown stronger. Therefore, it is clear that CSHG retains a Type 3 declension system, as proposed by Kürschner and Nübling (2011:378).

4.3.1.5. Low German

In OS, some declension classes were not clearly associated with a single gender. As shown in Table 13, there was a corresponding neuter or feminine class for all strong masculine classes. Most could be distinguished by at least one form. Corresponding masculine and neuter classes usually differed in the Nom/Acc plural. For example, masculine *a*-stems had *-os* or *-a* for this

form, while neuter *a*-stems had *-u* or a zero ending. However, neuter *ja*-stems were not distinguishable from masculine short *i*-stems, even in the Nom/Acc plural. Masculine and feminine short *i*-stems differed in at least the Gen singular, but long *i*-stems did not have this distinction due to the extension of the *-es* form to the feminine paradigm. Among weak nouns, all three genders were distinctive only in the Acc singular. Therefore, OS had a Type 3 declension system. In MLG, vowel reduction led to further dissociation of gender and declension. In addition to the classes that were not distinguishable in OS, masculine *a*-stems and neuter short *a*-stems became indistinguishable from masculine and feminine long *i*-stems. There were no longer any gender distinctions on weak nouns except in the Acc singular, in which the masculine had *-en*, the neuter had *-e*, and the feminine could have either form. Gender and declension were not totally dissociated in MLG, so it retained a Type 3 system. In ModLG, gender has been almost entirely dissociated from declension, which is now indicated only by plural markers since all case distinctions have been lost on nouns. As shown in Table 14, the loss of *e* in the final syllable and analogical processes have resulted in a zero ending for all nouns in the singular, while the *-en* and *-s* plural forms have spread to strong nouns of all genders. A zero plural with an accompanying stem change such as umlaut or final consonant deletion is mostly limited to masculine and neuter nouns, but this is not productive and also occurs on a few feminine nouns. Regardless, Kürschner and Nübling state that each declension in a Type 3 system is associated with one or two genders (2011:378); in ModLG, the two major plural forms are used with all three genders. Therefore, the declension system in ModLG is better classified as Type 4 than Type 3.[136]

4.3.1.6. Dutch

In ODu, each declension class was associated with one gender, but not as clearly as in most of the other early Germanic languages. As shown in Table 16, both strong masculine classes had a corresponding neuter or feminine class, but both could be distinguished by at least one form. Masculine and neuter *a*-stems differed in the Nom/Acc plural, while masculine and feminine *i*-stems differed in the Gen singular. The weak declension was the same for all three genders in the

[136] Kürschner and Nübling also ascribe a Type 4 system to Low German, but they only consider the dialect of East Friesland, in which the masculine and feminine have merged as a common gender, and plural forms are more clearly conditioned by prosody than in other Low German dialects (2011:369).

plural, and two genders were syncretic for every singular form, but taken together, each gender still had a different declension pattern in the singular. Therefore, ODu had a Type 2 declension system. In MDu, however, vowel reduction and analogical processes neutralized most of the gender distinctions described above. Masculine and neuter strong nouns still had some distinctive Nom/Acc plural forms, but *-e* was possible for both. No reliable gender distinctions remained on weak nouns. However, feminine *i*-stems could still be distinguished from masculine and neuter strong nouns based on the Gen singular. Thus, MDu had developed a Type 3 declension system. In CSDu, along with the merger of masculine and feminine as the common gender, gender and declension have been totally dissociated; due to total case loss, only plural markers indicate the latter. These markers are conditioned by prosody, with a trochaic foot as the preferred output: monosyllabic nouns generally take *-e(n)*, while polysyllabic nouns generally take *-s* (see Kürschner & Nübling 2011:375-376, De Schutter 1994:458). The former derives from the weak plural, which was already spreading to strong feminine and neuter nouns in MDu. The latter most likely derives from one of the masculine *a*-stem forms; it may have spread under the influence of French loanwords, which had this plural and were largely polysyllabic (see De Schutter 1994:459). Thus, CSDu has a Type 4 system, as Kürschner and Nübling propose (2011:378).

4.3.1.7. English

In OE, each major declension class was clearly associated with one gender. As shown in Table 18, a neuter or feminine class corresponded to each strong masculine class, but each could be distinguished by the Nom/Acc plural. Feminine strong nouns were also characterized by *-e* in the Gen singular, in contrast to *-es* for masculine and neuter strong nouns. On the other hand, the weak declension was the same for all three genders in the plural, and at least two genders were syncretic for every singular form, but taken together, each gender still had a different declension pattern in the singular. Therefore, OE had a Type 2 declension system. Due to the merger of unstressed back vowels in LOE, neuter strong nouns could no longer be reliably distinguished from feminine strong nouns based on the Nom/Acc plural. However, they could still be distinguished based on the Gen singular, and weak declension retained the same distinctions. Thus, a Type 2 system persisted in LOE. In EME, however, gender was almost entirely dissociated from declension. As shown in Table 19, sound changes neutralized all gender

distinctions on weak nouns. During this period, almost all of the strong declension classes lost productivity as their members adopted the masculine *a*-stem forms or, in some dialects, the weak forms.[137] A zero ending in the Nom/Acc plural remained an option for some neuter strong nouns. Based on the existence of this distinctive neuter class, it could be argued that EME had a Type 3 system. However, this class was also losing productivity, and the system also had characteristics of Type 4. In particular, the two major classes were used with all three genders, as opposed to a Type 3 system, in which each declension is associated with only one or two genders (see Kürschner & Nübling 2011:378). Thus, Type 4 is a better categorization for EME. In LME, grammatical gender had been lost entirely, and only the declension derived from masculine *a*-stems remained productive. Kürschner and Nübling did not account for the total loss of gender in their typology, but total dissociation is implied by total loss, so Type 4 remains the best categorization.

4.3.2. Gender-Declension Relationship Types in Romance Languages

The gender-declension type of each stage of the Romance languages included in the gender analysis is determined in this section using the criteria from section 4.1 above.

4.3.2.1. Western Romance: Italian, Spanish, and French

In CL, not all declension classes had a clear association with a single gender. As shown in Table 22, there was no distinction between feminine and masculine *ā*-stems, nor between masculine and feminine consonant-stems and *i*-stems. Masculine *ā*-stems were a relatively small class, but consonant- and *i*-stems of both genders were common. Neuter nouns of all classes ended with a distinctive *-a* in the Nom/Voc/Acc plural. In addition, neuter *o*-stems could be distinguished from masculine *o*-stems based on the Nom/Voc singular, while neuter consonant- and *i*-stems were distinctive in the Acc singular. Therefore, CL had a Type 3 declension system. Sound changes and analogical processes in VL, including the absorption of several minor classes by other classes, did not affect the distinctions described above, so a Type 3 system persisted.

In Pre-Italian, several changes to declension occurred, but these did not change its relationship to gender. Many neuter nouns were reanalyzed as masculine *o*-stems beginning in WVL. Some neuter consonant- and *i*-stems were reinterpreted as masculine or feminine nouns of

[137] This process is described in more detail in section 3.1.2.2 above.

the corresponding class, so the class of neuter consonant- and *i*-stems was basically eliminated. Sound change neutralized the distinction between the remaining neuter *o*-stems and their masculine counterparts in the Nom singular, but they remained distinctive in the plural. As shown in Table 24, both feminine classes remained indistinguishable from their masculine counterparts. Thus, Pre-Italian retained a Type 3 system. Analogical changes in the development of CSI did not change the relationship between gender and declension. However, the number of nouns still following the paradigm inherited from neuter *o*-stems has decreased to the point where it is no longer part of the productive declension system.

In ES, similar changes to declension occurred as in Pre-Italian and CSI; they likewise did not change the relationship between gender and declension. One difference was that the loss of the neuter classes was already complete in ES. Since both feminine classes remained indistinguishable from the corresponding masculine nouns, as shown in Table 25, a Type 3 system persisted in ES. No changes to declension or gender have occurred in the development of ModSp.

In OF, changes to declension resulted in a clear association of each declension class with one gender. As in ES, the neuter category had been lost, although a few became indeclinable masculine nouns. As shown in Table 26, masculine consonant- and *i*-stems came to be distinguished from their feminine counterparts in the Nom plural, as the former began to converge with masculine *o*-stems. Similarly, the few remaining masculine *ā*-stems retained an *-e* in all forms but adopted the *o*-stem syncretism pattern. Thus, OF had developed a Type 2 declension system. With the loss of the remaining case distinction in favor of the Acc forms, MF was left with a single productive declension class for each gender. Except for a few indeclinable and irregular nouns, there was a total association between gender and declension, i.e., a Type 1 system. CSF retains the same system.

4.3.2.2. Eastern Romance: Romanian

In Pre-Romanian, a number of sound changes and analogical process applied to nominal inflection, as shown in Table 28. However, these did not change the number of productive declensions or genders, nor their relationship. Thus, a Type 3 system persisted in Pre-Romanian. In CSRm, changes to declension, particularly analogical processes, have brought about a clear association of each declension class with one gender. As shown in Table 30, the Gen-Dat

singular form of masculine *ā*-stems, consonant-stems, and *i*-stems was leveled to the Acc form by analogy with the syncretism pattern for masculine *o*-stems. Masculine *ā*-stems also adopted the plural form *-i* that the other masculine classes had inherited by regular sound change. In contrast, the Gen-Dat singular form of feminine *ā*-stems, consonant-stems, and *i*-stems have remained distinct from the other singular forms. Despite differences in the specifics, these developments produced the same result as in OF: one syncretism pattern shared by all masculine classes and another by both feminine classes. CSRm has also retained distinctive neuter classes; these lack all case distinctions like the masculine classes but are distinguished from them by their plural forms. Thus, CSRm has developed a Type 2 system.

4.3.3. Gender-Declension Relationship Types in Balkan Sprachbund Languages

This section applies the criteria in section 4.1 above to each stage of the Balkan Sprachbund languages included in the gender analysis to ascertain their gender-declension type.

4.3.3.1. Eastern South Slavic: Bulgarian and Macedonian

In LPS, not all declension classes were clearly associated with one gender. As shown in Table 32, there was no distinction between feminine and masculine *(j)ā*-stems, as in CL. There were consonant-stems of all genders, as well as masculine and neuter *(j)o*-stems, and masculine and feminine *i*-stems, but these were all consistently distinguished in the Nom/Voc plural. In addition, neuter nouns of both classes had distinctive Acc forms in the singular and plural, and neuter *(j)o*-stems also had distinctive Nom/Voc singular forms. Finally, feminine consonant- and *i*-stems had a distinctive Instr singular form. Therefore, LPS had a Type 3 declension system.

In PMB, the relationship between gender and declension did not change, despite several developments in the latter. Consonant-stems and masculine *i*-stems lost productivity, reducing the number of classes that had significant overlap among the genders. As shown in Table 33 and Table 34, sound changes and other analogical processes neutralized distinctions among different classes in a few case forms but did not result in any full mergers. However, some number of masculine *(j)ā*-stems remained, and these continued to be declined the same as feminine *(j)ā*-stems, so a Type 3 system persisted in PMB. In PMB (and other early Slavic languages), an additional distinction in animacy arose in the Acc singular of masculine *(j)o*-stems. Analogical processes did not fundamentally change the relationship between gender and declension in MB,

although declension classes were distinguished by fewer forms due to the loss of case distinctions, as shown in Table 35 and Table 37. Gender and declension became slightly more dissociated with the almost complete loss of case distinctions in CSB and CSM. Neuter *(j)o*-stems are distinctive in the singular and plural forms, but indefinite feminine *i*-stems can no longer be reliably distinguished from masculine *(j)o*-stems, and the *(j)ā*-stem paradigm is still productive for both of these genders. Thus, CSB and CSM retain a Type 3 system.

4.3.3.2. Western South Slavic: Bosnian-Croatian-Montenegrin-Serbian

MSrb underwent some of the same changes to declension as PMB and MB, likewise without a substantive change in its relationship to gender. Consonant-stems and masculine *i*-stems were absorbed by other paradigms, which reduced the number of classes with limited distinctions among the genders. As shown in Table 38, sound changes and other analogical processes neutralized distinctions among classes of different genders in a few case forms, but the only mergers were among classes of the same gender. However, masculine *(j)ā*-stems continued to follow the same paradigm as feminine *(j)ā*-stems, so a Type 3 system persisted in MSrb. As in PMB, masculine *(j)o*-stems gained an additional animacy distinction in the Acc singular. The relationship between gender and declension did not fundamentally change due to analogical processes in BCMS, although increased case syncretism slightly reduced the number of forms that distinguished declension classes, as shown in Table 39. Thus, BCMS retains a Type 3 system.

4.3.3.3. Greek

In MyG, not all declension classes had a clear association with one gender. As shown in Table 41, there was no distinction between masculine and feminine consonant-stems, *i*-stems, and *u*-stems. This was similar to the situation in CL; in contrast, however, masculine *ā*-stems could be distinguished from feminine *ā*-stems in the Nom and Gen singular. Neuter nouns of all classes had a distinctive *-a* in the Nom/Voc/Acc plural, as in CL and LPS. In addition, neuter *o*-stems were distinct from masculine *o*-stems in the Nom/Voc singular, while neuter consonant- and *u*-stems were distinctive in the Acc singular. Therefore, MyG had a Type 3 declension system. Sound changes and analogical processes in AG, including the merger of the Instr with the Dat, did not have much effect on the distinctions described above. The Voc singular of feminine *ā*-

stems was leveled to the Nom, further distinguishing this class from masculine *ā*-stems. Thus, a Type 3 system persisted in AG. In KG, *u*-stems and masculine *i*-stems lost productivity, leaving only consonant-stems with no distinction between masculine and feminine. Due to the loss of vowel length distinctions, the masculine *ā*-stem Voc singular was no longer reliably distinct from the feminine *ā*-stem form, as shown in Table 42. However, these classes could still be distinguished based on the Nom and Gen singular. Thus, KG retained a Type 3 system.
In ModGr, changes to declension have resulted in a clear association of each declension class with one gender. As shown in Table 43, masculine and feminine consonant-stems merged with *ā*-stems; in the process, they adopted the distinctive *ā*-stem forms in the Nom and Gen singular. In addition, feminine *i*-stems almost entirely merged with one of the feminine *ā*-stem paradigms in the singular due to a combination of a vowel merger and analogical changes. As a result of these developments, ModGr has developed one singular declension pattern shared by all masculine classes and another by all feminine classes. The former can be generalized as a vowel followed by -*s* in the Nom and a vowel in the Voc/Acc/Gen, the latter as a vowel in the Nom/Voc/Acc and a vowel followed by -*s* in the Gen. Neuter classes follow have the same Nom/Voc/Acc syncretism as feminine classes but can have a Gen singular form with or without an -*s* after the vowel. However, they are clearly distinguished from all masculine and feminine classes by their Nom/Voc/Acc plural form. Thus, ModGr has developed a Type 2 system with a strong resemblance to the one in CSRm. Both profile the masculine-feminine distinction using the syncretism pattern in the singular. The loss of the Dat did not affect the relationship between gender and declension in ModGr because it had not been useful in distinguishing classes anyway.

4.3.3.4. Albanian

In Proto-Albanian, some declension classes lacked a clear association with a single gender. The full declension system is unreconstructable, since only a few classes have survived or even left traces. However, there is evidence for masculine and feminine *i*-stems, and these likely shared the same paradigm, as in MyG and CL, but it is also possible they differed slightly as in LPS. Nevertheless, the Proto-Albanian declension system is probably best classified as Type 3. On the other hand, gender became more closely associated with declension as the number of classes was reduced during this period: all nouns that became *o*-stems were reanalyzed as masculine or

neuter, and all nouns that became *ā*-stems as feminine. Due to a combination of sound changes and analogical processes in the development of CSA, the two *ā*-stem plural paradigms are the same as two of the four *o*-stem plural paradigms. Feminine nouns can still be distinguished based on the *ā*-stem singular forms, but ambigeneric nouns, which have *o*-stem singular forms and ambiguous plural forms, and neuter nouns, which are only used in the singular with *o*-stem forms, cannot be distinguished from each other or from masculine nouns. Thus, CSA has retained a Type 3 system when all four gender categories are considered, but there are no longer any declension classes containing both masculine and feminine nouns, the two most frequent genders.

4.3.4. Analysis of Gender-Declension Relationship by Number of Genders and Case Loss

As described in the preceding three sections, all of the languages under investigation have a Type 2 or 3 declension system in their earliest attested stage, but all four types are attested in the modern stages. Although the earliest attested stages of all the non-Germanic languages under investigation were Type 3, except for OS among the Germanic languages. The rest were Type 2, but PGmc was very likely Type 3, so a Type 3 system is the assumed starting point for all languages investigated. Of the Germanic languages that have retained a Type 2 system after developing one early, i.e., Icelandic, Faroese, and Swedish, the first two have retained three genders, but in Swedish the masculine and feminine have merged into the common gender, in opposition to the neuter. Norwegian is the only Germanic language that developed a Type 1 system, and it has retained three genders. Only High German redeveloped and retained a Type 3 system, and it has also retained three genders. The Germanic languages that have become Type 4 differ in the number of genders retained: Low German has retained three genders, Danish and Dutch have retained the same two genders as Swedish, and English has lost all grammatical gender.

Like the Germanic languages, the Romance languages have developed several different relationships between gender and declension, even though all descend from the same CL system. In the Romance languages that have remained Type 3, i.e., Italian and Spanish, only the masculine and feminine remain productive. Romanian retains ambigeneric category, which has developed a Type 2 declension system with three genders. French has developed a Type 1 system with two genders. Of the Balkan Sprachbund languages, Bulgarian, Macedonian, and

Albanian have remained Type 3, while Romanian and Greek have developed Type 2 declension systems. All retain three genders except Albanian, which has developed a fourth, as described in section 2.5.4 above. BCMS has also remained Type 3 and retained three genders.

As discussed in section 3.1.2.1 above, Kürschner and Nübling describe a strong association between number of genders and gender-declension type. Among the contemporary Germanic varieties they consider, those with Type 1 and Type 3 declension systems all have three genders and those with a Type 4 system all have two genders. They argue these associations are systematic: a Type 3 system retains complexity in genders and declensions, a Type 1 system maintains the complexity of three genders through the reduction of declensions to one for each gender, and a Type 4 system has enough simplification in both genders and declensions to allow for their dissociation. A Type 2 system can have two or three genders and some simplification in declension (1987:377-381). The hypothesis for this part of the analysis is that these associations hold in Romance and Balkan Sprachbund languages as well as the Germanic languages.

In fact, these associations generally hold for the contemporary Germanic languages considered in the present investigation, except that ModLG retains three genders in a Type 4 system, at least in some dialects. They also hold for the Balkan Sprachbund languages: all have three (or four) genders in a Type 2 or 3 system. However, these associations do not hold up when extended to the Romance languages under investigation. CSI, ES, and CSF have Type 1 or 3 systems with only two genders; only CSRm, with a Type 2 system, has three genders. Thus, all three hypothesized associations fail to hold for at least one language when extended to all of the languages in my analysis: the loss of a gender category did not accompany the total dissociation of gender and declension in ModLG, while it did accompany partial dissociation in CSI and ES, and total association in CSF.

Kürschner and Nübling also consider some effects of case loss on gender-declension type. Fewer morphological cases mean fewer forms to distinguish declension classes, and total case loss leaves only the plural markers in this function. Thus, case loss is usually accompanied by a reduction in the number of declension classes and, therefore, a shift away from a Type 2 or 3 system (1987:367, 370-371). This tendency is taken as the hypothesis for this part of the analysis. The Germanic languages under investigation follow this tendency for the most part: the

three languages that retain case marking on nouns have Type 2 or 3 systems, and CSS is the only Type 2 language with total case loss.[138]

The Romance languages do not follow this tendency, however. CSRm retains case marking and has a Type 2 system, but those with a Type 3 system have undergone total case loss. The number of declension classes has been reduced in most Romance languages, but not in Romanian. Throughout their development, however, all of the major classes have been distinguishable based on existing forms, even after these were limited to one singular and one plural form as a result of total case loss. In combination with the class of consonant- and *i*-stems making no distinction between masculine and feminine since CL, this meant that case loss did not directly lead to a shift away from a Type 3 system. The shift to a Type 2 system in OF and CSRm only occurred when masculine consonant- and *i*-stems underwent analogical changes to bring their case syncretism patterns closer to those of masculine *o*-stems. In fact, these changes were only possible due to the retention of case distinctions in these languages. Thus, Type 2 systems are a simplification from Type 3 systems in the sense that gender and declension are more transparently linked, but this simplification does not appear to be caused by case loss and may even be inversely correlated with total case loss. However, this shift may be motivated by some of the same factors that motivate case loss. Similarly, many of the earliest attested Germanic languages had developed Type 2 systems, including the most conservative in terms of case and gender, adding further support to the idea that this transparency can help preserve these categories.

The Balkan Sprachbund languages also lack the association of case loss with a change from a Type 2 or 3 system. Two of three Type 3 languages have experienced total case loss, but neither Type 2 language has. Similar to the Western Romance languages, a reduction in declension classes accompanied case loss in Bulgarian and Macedonian, including the loss of all distinctions between feminine *i*-stems and masculine (*j*)*o*-stems, except the Voc singular. This development reinforced the overlap of genders in a single declension class, which has also continued in (*j*)*ā*-stems. Thus, both of the hypotheses directly based on Kürschner and Nübling's findings for their selection of Germanic languages fail when extended to all of the Germanic, Romance, and Balkan Sprachbund languages considered in my present analysis.

[138] As mentioned in section 4.2.2 above, Low German has only retained case marking on articles, which do not help distinguish declension class.

4.3.5. Analysis of Gender Syncretism on Agreement Targets and Gender-Declension Relationship in Germanic Languages

This section and the following two sections introduce the gender syncretism index into the analysis. For each language, the relevant paradigms for the calculation of the index are presented. These are followed by a discussion of how the index has changed over time and how these changes relate to other developments concerning gender, particularly the number of genders and the gender-declension relationship. Germanic languages are considered in this section.

The results of the gender syncretism analysis for the Germanic languages were calculated according to the methodology described in section 4.1 above and are summarized in Table 52. Examples of the calculations for stages of Swedish are provided here, along with their corresponding values in the table. The stages under investigation are listed in the first column of the table, with darker lines separating distinct languages. The second and third columns show the number of declension classes and the number of morphological cases at that stage, respectively; these values do not directly factor into the calculation of the gender syncretism index but are provided for reference, since they relate to other factors in the gender analysis. The number of genders and the gender-declension type is listed in the fourth and fifth columns, respectively. In contrast to the number syncretism analysis, the same four core forms are used for all stages in the calculation of the gender syncretism indices in the sixth column: Nom and Acc, singular and plural. The gender syncretism in these forms is determined for each of the major agreement target paradigms.

In OSw, for example, there was no gender syncretism within any of the four case/number forms in the paradigms for determiners, including the postpositive definite article, as well as strong adjectives, so the gender syncretism is simply 0. In MSw, there was variation due to ongoing sound changes and analogical processes. At a minimum, the masculine and feminine shared the same Acc plural form in both of the major paradigms, for a total of four forms involved in gender syncretism. In addition, there was sometimes masculine-feminine syncretism in the Nom and Acc singular of both paradigms, adding 8 forms, i.e., 2 genders in 2 cases in 2 paradigms, for twelve total forms involved. The number of paradigms (i.e., 2) is multiplied by 4, for the forms under consideration, and 3, for the gender categories, to yield the total number of forms that could potentially be involved in gender syncretism, i.e., 24. The gender syncretism

index is the proportion of forms that could be involved in gender syncretism that actually are, i.e., 4 divided by 24, or 0.17, at a minimum, and 12 divided by 24, or 0.5, at a maximum. These two values are listed in the sixth column, but their average, 0.33, is used for the analysis. Since MSw still has three gender categories, no other calculations are needed.

CSS only has two gender categories, however, so two gender syncretism indices are calculated. For consistency, gender syncretism is still considered in terms of the four core case forms, even though the Nom and Acc have merged; the form for each number is simply counted twice. There is masculine-feminine syncretism throughout the definite article and strong adjectives, reflecting their merger as the common gender. The neuter is still distinctive in both the singular and plural of the definite article. Assuming three genders, this paradigm has 8 forms involved in gender syncretism, i.e., 2 genders in 2 cases and 2 numbers. The neuter is also distinctive in the singular of strong adjectives, but all three genders are syncretic in the plural, so this paradigm has 10 syncretic forms, i.e., 2 genders in 2 cases in the singular plus 3 genders in 2 cases in the plural. Thus, a total of 18 forms are involved in gender syncretism; this value is divided by 24, as for MSw, to yield the index out of three genders, i.e., 0.75. The second index is calculated out of the two remaining genders, so only syncretism between common and neuter is considered. Thus, only the plural of strong adjectives has gender syncretism, with 4 total forms for the 2 genders in 2 cases. The total number of forms that could be involved in gender syncretism is also lower; the number of paradigms and the number of case/number forms remain the same at 2 and 4, respectively, but they are only multiplied by 2 for the genders, yielding 16. Thus, the index out of the two remaining genders is 4 divided by 16, i.e., 0.25.

Table 52. Gender Syncretism on Agreement Targets in Germanic Languages

Stage	# of Declensions	Case	Gender		
		# of Cases	# of Genders	Type	Syncretism Index
North Germanic					
OIc	10	4	3	2	0
Ic	11	4	3	2	0
ONw	10	4	3	2	0
Fa (W)	11	4	3	2	0.17
Fa (S)	11	4	3	2	0.08
OSw	9	4	3	2	0
MSw	9	4	3	2	0.17–0.5 (avg. 0.33)
CSS	6	0	2	2	0.75 (out of 3 genders) 0.25 (out of 2 genders)
ODan	9	4	3	2	0
MDan	9	2	2	3	0.67 (out of 3 genders) 0 (out of 2 genders)
CSDan	2–3	0	2	4	0.83 (out of 3 genders) 0.5 (out of 2 genders)
MNw	10	4	3	3	0.33
NNw	1	0	3	1	0.58
West Germanic					
OHG	12	4	3	2	0
MHG	12	4	3	3	0.33
CSHG	6	4	3	3	0.5
OS	13	4	3	3	0.63
MLG	12	4	3	3	0.71
ModLG	9	2	3	4	0.5–0.83 (avg. 0.67)
ODu	9	4	3	2	??
MDu	9	4	3	3	0.79
CSDu	2–3	0	2	4	0.92 (out of 3 genders) 0.75 (out of 2 genders)
OE	12	4	3	2	0.33
LOE	12	4	3	2	0.58
EME	3	4	3	4	0.75
LME	1	0	0	4	1

4.3.5.1. Icelandic and Faroese

Table 53. Old Icelandic/Old Norwegian Agreement Targets

	Definite Articles						Strong Adjectives					
	Sg.			Pl.			Sg.			Pl.		
	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.
Nom	-(i)nn	-(i)n	-(i)t	-nir	-nar	-(i)n	-r	-Ø	-t	-ir	-ar	-Ø
Acc	-(i)nn	-(i)na	-(i)t	-na	-nar	-(i)n	-an	-a	-t	-a	-ar	-Ø

(Faarlund 2004:39)

Table 54. Icelandic Agreement Targets

	Definite Articles						Strong Adjectives					
	Sg.			Pl.			Sg.			Pl.		
	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.
Nom	-(i)nn	-(i)n	-(i)ð	-nir	-nar	-(i)n	-ur	-Ø	-t	-ir	-ar	-Ø
Acc	-(i)nn	-(i)na	-(i)ð	-na	-nar	-(i)n	-an	-a	-t	-a	-ar	-Ø

(Thráinsson 1994:155)

Table 55. Written Faroese Agreement Targets

	Definite Articles						Strong Adjectives					
	Sg.			Pl.			Sg.			Pl.		
	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.
Nom	-(i)n	-(i)n	-(i)ð	-nir	-nar	-(i)ni	-ur	-Ø	-t	-ir	-ar	-Ø
Acc	-(i)n	-(i)na	-(i)ð	-nar	-nar	-(i)ni	-an	-a	-t	-ar	-ar	-Ø

(Barnes & Weyhe 1994:201)

Table 56. Spoken Faroese Agreement Targets

	Definite Articles						Strong Adjectives					
	Sg.			Pl.			Sg.			Pl.		
	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.
Nom	-(i)n	-(i)n	-(i)ð	-nir	-nar	-(i)ni	-ur	-Ø	-t	-ir	-ar	-Ø
Acc	-(i)n	-(i)na	-(i)ð	-nir	-nar	-(i)ni	-an	-a	-t	-ir	-ar	-Ø

(Barnes & Weyhe 1994:201)

In OIc, all three genders were distinguished on the Nom and Acc forms of NP-internal agreement targets, and this has continued in Icelandic. In terms of gender, this makes it the most conservative language under investigation. ONw started out the same as OIc but an analogical change to the masculine Acc plural of agreement targets in written Faroese neutralized its distinction from the feminine, raising the gender syncretism index to 0.17. Icelandic and Faroese have experienced little to no change in gender syncretism on agreement targets, so it is not surprising that they have also retained three genders and the same Type 2 relationship between gender and declension.

4.3.5.2. Swedish, Danish, and Norwegian

Table 57. Old Swedish/Old Danish Agreement Targets

	Definite Articles						Strong Adjectives					
	Sg.			Pl.			Sg.			Pl.		
	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.
Nom	-(i)nn	-(i)n	-(i)t	-(i)nir	-(i)nar	-(i)n	-er	-Ø	-t	-ir	-ar	-Ø
Acc	-(i)nn	-(i)na	-(i)t	-(i)na	-(i)nar	-(i)n	-an	-a	-t	-a	-ar	-Ø

(Delsing 2002:930, 933)

Table 58. Middle Swedish Agreement Targets

	Definite Articles					
	Sg.			Pl.		
	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.
Nom	-(e)n(n)	-(e)n	-(e)t	-(e)ne	-(e)næ	-(e)n
Acc	-(e)n(n)	-(e)næ	-(e)t	-(e)næ, -(e)ne	-(e)næ	-(e)n
	Strong Adjectives					
	Sg.			Pl.		
	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.
Nom	-er, -Ø	-Ø	-t	-e	-æ	-Ø
Acc	-æn, -Ø	-æ, -Ø	-t	-æ, -e	-æ	-Ø

(Mørck 2005:1140-1141)

Table 59. Contemporary Standard Swedish Agreement Targets

Definite Articles				Strong Adjectives		
Sg.		Pl.		Sg.		Pl.
Common	Neut.	Common	Neut.	Common	Neut.	All Genders
-(e)n	-(e)t	-na	-en, -a	-Ø	-t	-a

(Andersson 1994:280, 281)

As in OIc and ONw, there was no gender syncretism on the core cases of NP-internal agreement targets in OSw. The gender syncretism index increased at least to 0.17 in MSw with the neutralization of the masculine-feminine distinction in the Acc plural. This was the same development as in written Faroese but due to sound change rather than analogy. During the MSw period, the distinction between masculine and feminine in the Nom singular of the postpositive definite article began to be neutralized by sound change. This distinction was also lost with increasing frequency on strong adjectives due to the leveling of the masculine form to a zero ending, already present in the feminine form. The masculine and feminine Acc singular forms of strong adjectives were often leveled to a zero form as well. Likewise, the feminine Acc singular form of the definite article was frequently leveled to the Nom singular form. If all of these changes applied, the gender syncretism index increased to 0.5, but there was still a masculine-

feminine distinction in the Nom plural. For MSw, my analysis assumes an index of 0.33, the average of 0.17 and 0.5. With the loss of the Nom-Acc distinction in CSS, it appears the forms of agreement targets that survived were those without a masculine-feminine distinction, i.e., the Nom singular and Acc plural. In the singular, this followed from the completion of the optional leveling processes in MSw. In the plural, the masculine Nom form may have been leveled to the feminine Nom form, as in the singular, or to the masculine Acc form. These changes resulted in the complete merger of the masculine and feminine as the common gender. In addition, strong adjectives lost all gender distinctions in the plural when the zero ending of the neuter plural was leveled to the common form *-a*, possibly to avoid syncretism with the zero ending of the common singular form. Other determiners also lost this distinction, but sound change was more likely a factor. Only the postpositive definite article retained a gender distinction in the plural, despite syncretism with the common singular form. As a result of these changes, the gender syncretism index increased to 0.75. However, the index recalculated out of the two remaining genders only increased to 0.25.

As described above, sound change accounts for an increase in masculine-feminine syncretism. Case loss seems to have played an important role, but if different forms had been retained as the Nom-Acc distinction was lost, masculine and feminine could have remained distinctive in both numbers. Morphosyntactic factors unrelated to gender may have determined which case forms survived. A preference for a zero form in unmarked categories like the singular is well-attested cross-linguistically, as discussed in section 3.1.2.1 above, and would account for the completion of the merger in the singular. It is also possible that agreement target forms without the masculine-feminine distinction survived because this distinction had already been lost on nouns, in part due to a similar leveling process in the masculine Nom singular as on strong adjectives. Case loss left only plural markers to indicate declension, but these were no longer useful in distinguishing masculine and feminine after sound changes in MSw. Thus, the loss of distinctive singular forms on nouns and the resulting Type 3 system may have been a major factor. The merger of the masculine and feminine allowed the declension system to reacquire the stronger association with gender it had had in OSw. It is certainly possible for agreement targets to make gender distinctions despite the lack of reliable differences in declension, as in Type 4 languages, but most of these only retain two genders, so there was usually a simplification of agreement targets along with declension. Overall, it seems more likely

but by no means certain that the total loss of the masculine-feminine distinction on nouns played a larger role in the merger of these genders than the partial loss of this distinction on agreement targets due to sound change.

Table 60. Middle Danish Agreement Targets

Definite Articles				Strong Adjectives			
Sg.		Pl.		Sg.		Pl.	
Common	Neut.	Common	Neut.	Common	Neut.	Common	Neut.
-(æ)n	-(æ)t	-(æ)næ	-(æ)n	-Ø	-t	-æ	-Ø

(Mørck 2005:1140-1141)

Table 61. Contemporary Standard Danish Agreement Targets

Definite Articles			Strong Adjectives		
Sg.		Pl.	Sg.		Pl.
Common	Neut.	All Genders	Common	Neut.	All Genders
-et	-en	-ne	-Ø	-t	-e

(Haberland 1994:330)

Although ODan and OSw were the same, the masculine and feminine completely merged as the common gender in MDan, raising the gender syncretism index to 0.67. However, there was no gender syncretism when only the two surviving genders are considered. The earlier merger than in Swedish can be attributed to more extreme vowel reduction, which neutralized the masculine-feminine distinction in the Nom and Acc plural of all agreement targets, and the completion of processes that were still in progress in MSw. These included the analogical leveling of several singular forms as the Nom-Acc distinction was lost. In CSDan, remaining distinctive neuter forms were leveled to the common form on all plural agreement targets, including the postpositive definite article. Thus, the index increased to 0.83, even higher than in CSS. However, the gender syncretism index only increased to 0.5 when recalculated out of the two surviving genders.

The higher index in CSDan than in CSS, whether out of two or three genders, reflects the total loss of gender distinctions in the plural of agreement targets. In turn, this was likely connected with the dissociation of gender and declension, which was only indicated in the plural. This dissociation was already beginning in MDan with the tendency to avoid zero plurals on nouns, which had distinguished the neuter (see Kürschner & Nübling 2011:372). Thus, the

dissociation of gender and declension more likely motivated the loss of gender distinctions in the plural than vice versa, but the directionality is not as clear for the masculine-feminine merger.

Table 62. Middle Norwegian Agreement Targets

	Definite Articles						Strong Adjectives					
	Sg.			Pl.			Sg.			Pl.		
	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.	Masc.	Fem.	Ne.
Nom	-(e)n(n)	-(e)n	-(e)t	-(e)ne	-(e)næ	-(e)n	-er, -e, -Ø	-Ø	-t	-e	-æ	-Ø
Acc	-(e)n(n)	-(e)næ	-(e)t	-(e)næ, -(e)na	-(e)næ	-(e)n	-æn, -e, -Ø	-æ, -Ø	-t	-æ, -e	-æ	-Ø

(Mørck 2005:1140-1141)

Table 63. Neo-Norwegian Agreement Targets

Definite Articles						Strong Adjectives		
Sg.			Pl.			Sg.		Pl.
Masc.	Fem.	Neut.	Masc.	Fem.	Neut.	Common	Neut.	All Genders
-(e)n	-a	-(e)t	-(a)ne	-(e)ne	-a	-Ø	-t	-e

(Askedal 1994:230-231)

As mentioned in the previous section, there was no gender syncretism on the core cases of agreement targets in ONw. As in MSw, sound change neutralized the distinction between masculine and feminine in the Acc plural in MNw, raising the gender syncretism index at least to 0.17. Likewise, additional neutralizations sometimes resulted from other sound changes and analogical processes during this period; these were the same as those described for MSw above, except that the Nom and Acc masculine singular forms of strong adjectives was sometimes *-e*, the regular outcome of the Nom singular. In contrast to leveling to the zero form, which was also an option, this preserved the masculine-feminine distinction. If all of the processes that neutralized gender distinctions applied, however, the index still increased to 0.5 as in MSw, leaving only the Nom plural with a masculine-feminine distinction. Therefore, this analysis also assumes an index of 0.33 for MNw. Despite only the minor difference in gender marking on agreement targets between MSw and MNw described above, NNw has a very different outcome from CSS in that it retains three genders. Nevertheless, many of the same changes applied and the gender syncretism increased to 0.58. As in CSS, the forms that remained after the loss of the Nom-Acc distinction lacked a masculine-feminine distinction. This distinction did not survive on most strong adjectives, where more distinctive forms were possible in MNw than in MSw. Instead, a new distinctive feminine singular form *-a* developed on determiners and a few strong

adjectives; like determiners, these adjectives have a masculine singular form with *-n*, e.g., *eigen* 'own'. It is unclear why this *-n* was lost from the feminine form, which had been apparently syncretic with the masculine form by the end of the MNw period. The loss of *-n* does not seem to have been a regular sound change, but it is possible that some dialects underwent the loss of word-final unstressed /n/ and the shortening of word-final unstressed /n:/ to [n], in contrast to the merger of both phonemes as [n]. If this change occurred optionally or only in some dialects, it may have been morphologized. The change from *-e* to *-a* could have been influenced by the ONw strong adjective form *-a* or the definite form *-ena*. It is also possible that the development was entirely analogical, but other determiners have feminine singular forms without *-n* or *-a*, e.g., the feminine indefinite article *ei* vs. masculine *ein* 'a(n)'. Whether this change was partly or fully analogical, it may have arisen as a way to restore a distinction in the unmarked singular that had become limited to the marked plural, before the distinction in the plural was lost. However, this unstable situation could have been resolved by the total merger of the masculine and feminine as in CSS, so perhaps the distinctive masculine and feminine strong adjective forms remained in use longer in Norwegian than Swedish, even though they were eventually lost for the majority of strong adjectives in both. The development of a Type 1 system in NNw is likely connected to the retention of three genders; the total association of gender and declension enhances the memorability of both (see Kürschner & Nübling 2011:374). With one fewer gender categories to learn, the same need for a reduction in declensions was not present in the development of CSS, so these appear to be a kind of tradeoff. Although gender and declension are linked in NNw, they are not marked in the same way. Declension is marked entirely in the plural, while there are more reliable gender distinctions on the singular of agreement targets. Though less extreme, this reflects the same tendencies as the total loss of gender distinctions in the plural alongside a reliance on plural forms to mark declension in other Germanic languages such as CSHG. As discussed in section 3.1.2.1 above, Kürschner and Nübling (ibid., 357-360) consider these developments a form of number profiling.

4.3.5.3. High German

Table 64. Old High German Agreement Targets

	Demonstratives					
	Sg.			Pl.		
	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.
Nom	dër	diu	daz̨	dē, dea, dia, die	deo, dio	diu, (dei)
Acc	dën	dea, dia (die)	daz̨	dē, dea, dia, die	deo, dio	diu, (dei)
	Strong Adjectives					
	Sg.			Pl.		
	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.
Nom	-ēr	-iu	-az̨	-e	-o	-iu
Acc	-an	-a	-az̨	-e	-o	-iu

(Wright 1906:55)

Table 65. Middle High German Agreement Targets

	Definite Articles					Strong Adjectives				
	Sg.			Pl.		Sg.			Pl.	
	Masc.	Fem.	Neut.	Masc./Fem.	Neut.	Masc.	Fem.	Neut.	Masc./Fem.	Neut.
Nom	dër	diu	daz̨	die	diu	-er	-iu	-ez̨	-e	-iu
Acc	dën	die	daz̨	die	diu	-en	-a	-ez̨	-e	-iu

(Wright 1917, §55, §68)

Table 66. Contemporary Standard German Agreement Targets

	Definite Articles				Strong Adjectives			
	Sg.			Pl.	Sg.			Pl.
	Masc.	Fem.	Neut.	All Genders	Masc.	Fem.	Neut.	All Genders
Nom	der	die	das	die	-er	-e	-es	-e
Acc	den	die	das	die	-en	-e	-es	-e

(Eisenberg 1994:366)

Like the early North Germanic languages, OHG lacked gender syncretism on the core cases of agreement targets. In MHG, vowel reduction neutralized the masculine-feminine distinction in the Nom/Acc plural of all agreement targets, raising the gender syncretism index to 0.33. The neuter was still usually distinctive in the plural in MHG, but all gender distinctions in the plural were lost by CSHG, raising the index to 0.5. It is possible that the neuter Nom/Acc plural form -*iu* lost its distinctiveness due to vowel reduction, as the masculine and feminine forms did earlier, since the same neutralization occurred in the feminine singular: Nom -*iu* came to be the same as Acc -(*i*)*e*. For the definite article in particular, confusion of the full forms *diu* and *die*, which is already attested in MHG, may have resulted from the use of the same reduced forms for

both: *de* before a consonant, and *d'* before a vowel (see Wright 1917, §68). These distinctions may have been subsequently leveled on other agreement targets by analogy with the article. However, there were also motivations for the analogical leveling of these forms, if sound change alone did not neutralize them. The Nom and Acc were never distinguished in feminine noun paradigms, so this may have been an important distinction for agreement targets. In contrast, this distinction occurred in both masculine singular agreement targets and weak nouns. Given other developments in High German that had strengthened the distinction between feminine and the other genders, distinguishing the neuter but not the feminine in the plural would have made less sense than in the continental Scandinavian languages, where the neuter was more strongly distinguished.

Unlike Icelandic and Faroese, the other Germanic languages that have not lost all case distinctions on nouns, High German developed a Type 3 declension system. This occurred in MHG, when vowel reduction neutralized many distinctions in nominal inflection, including among classes. In contrast, vowel reduction has not occurred in Icelandic and Faroese, and it had minimal effects on nominal inflection in Swedish, the other Type 2 language. Thus, vowel reduction may explain the different outcomes. In High German, it is also at least partially, and possibly fully, responsible for the loss of all gender distinctions in the plural of agreement targets, as discussed in the previous paragraph. All of the other Germanic languages that experienced such a loss developed a Type 4 declension system. In contrast to those languages, however, CSHG retains case distinctions, and certain forms are still limited to classes associated with at most two genders, e.g., Gen singular *-(e)s* for masculine and neuter strong nouns. Thus, High German has undergone some of the developments associated with Type 2 Germanic languages and others associated with Type 4 Germanic languages, so it follows that it has developed an intermediate type of gender-declension relation.

4.3.5.4. Low German

Table 67. Old Saxon Agreement Targets

	Demonstratives					
	Sg.			Pl.		
	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.
Nom	thê, thie (se)	thiu	that	thê, thea, thie, tha	thê, tha	thiu, thia, thea, thie, the
Acc	thena, thana, thane, then, than	thia, thea, thie, the, tha, thi	that	thê, thea, thie, tha	thê, tha	thiu, thia, thea, thie, the

	Strong Adjectives					
	Sg.			Pl.		
	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.
Nom	-Ø	-Ø	-Ø	-e, -a	-a	-Ø, -a
Acc	-an, -en	-a	-Ø	-e, -a	-a	-Ø, -a

(Gallée 1891:73, 84)

Table 68. Middle Low German Agreement Targets

	Definite Articles				Strong Adjectives			
	Sg.			Pl.	Sg.			Pl.
	Masc.	Fem.	Neut.	All Genders	Masc.	Fem.	Neut.	All Genders
Nom	dê, di(e)	dû̂, dê, di(e)	dat	dê, di(e)	-(er)	-(e)	-Ø	-e
Acc	den(e)	dû̂, dê, di(e)	dat	dê, di(e)	-en	-e	-Ø	-e

(Lasch 1914:204, 2018)

Table 69. Modern Low German Agreement Targets

	Definite Articles				Adjectives				
	Sg.			Pl.	Sg.			Pl.	
	Masc.	Fem.	Neut.	All Genders	Masc.	Fem.	Neut.	Masc./Fem.	Neut.
Nom	de	de	dat	de	-e(n), -Ø	-e, -Ø	-(e)t, -Ø	-en, -Ø	-Ø
Acc	den	de	dat	de	-en, -Ø	-e, -Ø	-(e)t, -Ø	-en, -Ø	-Ø

OS had the highest gender syncretism index of the Germanic languages investigated, at 0.63. Agreement targets did not reliably distinguish gender in the plural, and strong adjectives did not in the Nom singular either. In MLG, the feminine Nom singular form of determiners was often leveled to the masculine form, raising the index to 0.71. Strong adjectives are often uninflected in ModLG. With only the singular distinctions on determiners inherited from MLG, the index would have increased to 0.83. When strong adjectives are inflected, however, all three genders can be distinguished in the singular forms, due to the use of the Acc forms in the Nom as well, and the neuter is distinguished in the plural by remaining uninflected. Thus, the lack of gender distinctions in the plural would be partially undone, and the index would have decreased to 0.5.

For ModLG, this analysis assumes a gender syncretism index of 0.67, the average of 0.5 and 0.83.

As in Danish, the high index in ModLG was associated with the development of a Type 4 declension system. However, Low German still retains three genders even though they are only consistently distinguished in the Acc singular. This may be possible because ModLG also retains case distinctions on agreement targets but not nouns. Another difference from Danish is that the loss of gender distinctions in the plural of agreement targets clearly occurred before the dissociation of gender and declension, and a distinctive neuter plural has even been partially restored. Thus, the early loss of gender distinctions in the plural of agreement targets may have contributed to the spread of plural markers to nouns of different genders, a process that served to profile number, and therefore a Type 4 system. However, the fact that nouns of all genders and declensions have come to share the same zero form throughout the singular was probably at least as important, and this was the result of sound change and number profiling rather than the influence of agreement targets, which still distinguish gender in the singular.

4.3.5.5. Dutch

Table 70. Middle Dutch Agreement Targets

	Definite Articles				Adjectives		
	Sg.			Pl.	Sg.		Pl.
	Masc.	Fem.	Neut.	All Genders	Masc.	Fem./Neut.	All Genders
Nom	die	die	dat	die	-e	-e	-e
Acc	dien	die	dat	die	-en	-e	-e

(van der Wal & Quak 1994:76)

Table 71. Contemporary Standard Dutch Agreement Targets

Definite Articles			Adjective
Sg.		Pl.	Sg./Pl.
Common	Neut.	All Genders	All Genders
de	het	de	-e

(De Schutter 1994:462-464)

There is not enough evidence to determine gender syncretism on the core cases of agreement targets in ODu, but it was likely similar to OS due to the genetic and geographic proximity of these languages, and since the situation in MDu was similar to MLG: both had no reliable gender distinctions in the Nom singular of adjectives, or the plural of any agreement target and

masculine-feminine syncretism in the Nom singular of determiners. However, MDu also had feminine-neuter syncretism in the Acc singular of adjectives, for which weak forms were used regardless of syntactic context. Thus, the gender syncretism index was 0.79, slightly higher than in MLG. In CSDu, the loss of final *-n* neutralized the last remaining masculine-feminine distinction in the Acc singular of all agreement targets. As a result, adjectives no longer reliably distinguish gender, but there is still a distinction between neuter and the common gender in the singular of determiners, and the gender syncretism index increased to 0.92. However, the index only increased to 0.75 when recalculated out of the two remaining genders.

Thus, CSDu has lost more gender distinctions on agreement targets than ModLG, but only slightly more than when strong adjectives are uninflected in ModLG. As discussed in Section 2.3.3.3 above, dialectal evidence suggests that the completion of the masculine-feminine merger was the result of *-n* loss, not the loss of the Nom-Acc distinction. As in ModLG, the loss of gender distinctions in the plural of agreement targets may have been a factor in the dissociation of gender and declension, while the reverse can be ruled out based on chronology.

4.3.5.6. English

Table 72. Old English Agreement Targets

	Demonstratives				Strong Adjectives					
	Sg.			Pl.	Sg.			Pl.		
	Masc.	Fem.	Neut.	All Genders	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.
Nom	se	sēo	þæt	þā	-Ø	-u	-Ø	-e	-a	-u
Acc	þone	þā	þæt	þā	-ne	-e	-Ø	-e	-a	-u

(Marsden 2010:374)

Table 73. Late Old English Agreement Targets

	Demonstratives				Strong Adjectives			
	Sg.			Pl.	Sg.			Pl.
	Masc.	Fem.	Neut.	All Genders	Masc.	Fem.	Neut.	All Genders
Nom	se, seo, þe	seo	þet, þæt	þa	-Ø	-u	-Ø	-e
Acc	þone, þæne	þa	þet, þæt	þa	-ne	-e	-Ø	-e

Table 74. Early Middle English Agreement Targets

	Definite Articles				Strong Adjectives			
	Sg.			Pl.	Sg.			Pl.
	Masc.	Fem.	Neut.	All Genders	Masc.	Fem.	Neut.	All Genders
Nom	se, þe	þa	þet, þæt	þa	-Ø	-e, -Ø	-Ø	-e
Acc	þone	þa	þet, þæt	þa	-ne, -Ø	-e, -Ø, -ne	-Ø, -ne	-e

OE had a lower gender syncretism index, at 0.33, than the other early West Germanic languages except OHG. In the plural, strong adjectives distinguished all three genders, but there were no gender distinctions on determiners. In addition, the masculine and neuter both had a zero form in the Nom singular of strong adjectives. In LOE, all gender distinctions were neutralized in the plural of strong adjectives as well, raising the index to 0.58. This was likely connected to vowel reduction, although the analysis in section 2.3.3.4 above still assumes a front-back distinction for unstressed vowels, meaning feminine Nom/Acc plural *-a* and neuter *-u* could not be distinguished, but masculine *-e* was still distinctive. Thus, analogy with determiners may have also played a role in the spread of *-e* to the feminine and neuter, or vowel reduction was actually more advanced than assumed. In EME, the singular of strong adjectives joined the plural in no longer reliably distinguishing gender. This resulted from the extension of the zero form from the masculine and neuter Nom singular to the feminine, as well as the extension of the masculine Acc form *-ne* or the zero form, already in the Nom and the neuter Acc, to other genders in the Acc. As a result, the gender syncretism index increased to 0.75. By LME, all gender distinctions on determiners and adjectives had been lost, along with the loss of grammatical gender.

As in Low German and Dutch, the loss of gender distinctions in the plural of agreement targets occurred before the spread of plural markers to nouns of different genders, and the former may have enabled the latter. The causality is much less clear for the spread of singular case forms such as Gen *-(e)s* to nouns of different genders and the loss of gender distinctions in the singular of agreement targets, which also involved the spread of forms to other genders. In fact, the spread of singular case forms on both nouns and agreement targets involved the profiling of case over gender during the period when morphological case was in decline but not yet totally lost. Thus, gender loss may have been a side effect of case loss, but it is not clear why English did not retain at least one gender distinction in the singular in the way that other Type 4 languages did. English even still has the forms to mark this distinction, *the* and *that*, but instead this distinction has been repurposed. The equivalents of these forms, as well as other determiners with an analogous distinction, are the only NP-internal gender markers remaining in CSDu, so they would seem to be sufficient. However, English did not retain the formal distinction on other determiners in the same way; *this* derives from the neuter form, which was already commonly

used for all three genders in EME texts.[139] In addition, the indefinite article already had the same Nom form *ān* for all three genders in OE, but the indefinite article similarly lacks gender distinctions in CSDu (see De Schutter 1994:463). Nevertheless, Heltveit (1958) views the lack of a distinctive neuter form on the indefinite article as a weak point in OE gender marking, in contrast to the consistent use of *-t* for the neuter singular across most determiners as well as strong adjectives.

4.3.6. Analysis of Gender Syncretism on Agreement Targets and Gender-Declension Relationship in Romance Languages

This section continues the gender analysis with regard to individual languages. As with the Germanic languages in the previous section, the paradigms used in calculating my original gender syncretism index are presented for each Romance language, starting with their shared development from CL to VL. Changes in the index between stages are then discussed, and connections to other gender developments are noted, with a focus on the number of genders and the gender-declension relationship.

Table 75. Gender Syncretism on Agreement Targets in Romance Languages

Stage	# of Declensions	Case	Gender		
		# of Cases	# of Genders	Type	Syncretism Index
CL	7	6	3	3	0.28
WVL	7	2	3	3	0.36
PI	6	2	3	3	0.42
CSI	3	0	2	3	0.78 (out of 3 genders) 0.33 (out of 2 genders)
ES	6	0	2	3	0.78 (out of 3 genders) 0.33 (out of 2 genders)
OF	9	2	2	2	0.75 (out of 3 genders) 0.25 (out of 2 genders)
MF	2	0	2	1	0.72 (out of 3 genders) 0.17 (out of 2 genders)
CSF	2	0	2	1	0.83 (out of 3 genders) 0.5 (out of 2 genders)
EVL	7	6	3	3	0.36
PR	7	3	3	3	0.36
CSRm	7	3	3	2	0.78

[139] See Appendix A for descriptions of how demonstrative forms were used in a selection of the EME texts analyzed in chapter V.

4.3.6.1. Western Romance: Italian, Spanish, and French

Table 76. Classical Latin Agreement Targets

	Demonstratives						Adjectives < o-/ā-stems					
	Sg.			Pl.			Sg.			Pl.		
	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.
Nom	ille	illa	illud	illī	illae	illa	-us	-a	-um	-ī	-ae	-a
Acc	illum	illam	illud	illōs	illās	illa	-um	-am	-um	-ōs	-ās	-a

	Adjectives < i-stems				
	Sg.			Pl.	
	Masc.	Fem	Neut.	Masc./Fem.	Neut.
Nom	-is	-is	-e	-ēs	-ia
Acc	-em	-em	-e	-ēs	-ia

(Bennett 1913:34, 38)

There was already some gender syncretism on the Nom and Acc forms of agreement targets in CL, with an index of 0.28. In the most common class of adjectives, which used *ā*-stem forms for feminine and *o*-stem forms for masculine and neuter, there was masculine/neuter syncretism in the Acc singular, as with *o*-stem nouns. Similar forms were used on determiners, but the neuter had a distinctive Nom/Acc singular form, so it was not syncretic with the masculine in either case. In the less common *i*-stem class of adjectives, there was usually no distinction between masculine and feminine, as with *i*-stem nouns, but the neuter was almost always distinctive in the core cases. With the loss of *-m* in VL, the neuter was no longer distinctive in the Acc singular of *i*-stem adjectives. The distinctive neuter Nom/Acc form on determiners was leveled to the more common *o*-stem form, spreading masculine/neuter syncretism to the Acc singular of determiners. These changes raised the gender syncretism index to 0.36.

Table 77. Western Vulgar Latin/ Eastern Vulgar Latin Agreement Targets

	Definite Articles						Adjectives < o-/ā-stems					
	Sg.			Pl.			Sg.			Pl.		
	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.
Nom	elle, elli	ella	ellu	elli	elle	ella	-us	-a	-u, -o	-i	-ae	-a
Acc	ellu, ello	ella	ellu	ellos	ellas	ella	-u, -o	-a	-u, -o	-os	-as	-a

	Adjectives < i-stems			
	Sg.		Pl.	
	Masc./Fem.	Neut.	Masc./Fem.	Neut.
Nom	-es	-e	-es	-ia
Acc	-e	-e	-es	-ia

(Grandgent 1907:157, 164)

Table 78. Pre-Italian Agreement Targets

	Definite Articles						Adjectives < o-/ā-stems					
	Sg.			Pl.			Sg.			Pl.		
	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.
Nom	li	la	lo	li	le	la	-o	-a	-o	-i	-e	-a
Acc	lo	la	lo	lo	le	la	-o	-a	-o	-o	-e	-a

	Adjectives < i-stems				
	Sg.			Pl.	
	Masc.	Fem.	Neut.	Masc./Fem.	Neut.
Nom	-i	-i	-e	-i	-ia
Acc	-e	-e	-e	-i	-ia

(Alkire & Rosen 2010:206)

Table 79. Italian Agreement Targets

Definite Articles				Adjectives < o-/ā-stems			
Sg.		Pl.		Sg.		Pl.	
Masc./Neut.	Fem.	Masc.	Fem./Neut.	Masc./Neut	Fem.	Masc.	Fem./Neut.
il, lo (l')	la (l')	i, gli	le	-o	-a	-i	-e

Adjectives < i-stems	
Sg.	Pl.
All Genders	All Genders
-e	-i

(Alkire & Rosen 2010:205, 189)

The gender syncretism index increased from 0.36 in WVL to 0.42 in Pre-Italian. The increase resulted from the loss of *-s*, which neutralized the distinction between masculine and neuter in the Nom singular of *o*-/*ā*-stem adjectives, as with *o*-stem nouns. In CSI, the last remaining case distinction between Nom and Acc was lost in agreement targets as well as nouns. The Nom singular forms that had remained distinctive were leveled to the Acc as part of this process, leaving all singular neuter forms syncretic with the masculine, along with the feminine for *i*-stem adjectives. In addition, the neuter plural forms fell out of use, and feminine plural forms came to be used with the small group of neuter nouns that had not been reanalyzed as masculine or feminine. With masculine agreement in the singular and feminine agreement in the plural, these nouns have become ambigeneric. Since the masculine and feminine were already syncretic for *i*-stem adjectives, these no longer have any gender distinctions. As a result of these developments, the gender syncretism index increased to 0.78, but it actually decreased to 0.33 when limited to the two genders that remained productive.

Despite total case loss and the loss of a productive neuter, CSI has retained a Type 3 declension system rather than develop a more transparent Type 2 or Type 4 system like the Germanic languages that experienced total case loss on nouns. A number of different

developments may explain this difference in outcome. One is that agreement targets in Romance languages tend to match noun declension much more closely than in Germanic languages. This parallelism makes both systems easier to learn and remember than in Germanic languages where inflection on determiners, adjectives, and nouns often varied significantly. In addition, the most common noun and adjective classes in CSI have overt endings in the singular (*-o* for masculine, *-a* for feminine) as well as the plural. This also makes gender more memorable and avoids a major consequence of total case loss in Germanic languages: the reliance on just one form, the plural, to indicate declension. The divergence between these two families can be attributed largely to the different sound changes they underwent. Those in VL and Pre-Italian neutralized many case distinctions, but in terms of gender, they only affected the distinction between masculine and neuter in the singular, leading to the loss of productivity of the latter. In contrast, Germanic languages tend to have more extreme vowel reduction and a higher prevalence of zero endings in the singular.

Table 80. Early Spanish Agreement Targets

Definite Articles				Adjectives < o-/ā-stems				Adjectives < i-stems	
Sg.		Pl.		Sg.		Pl.		Sg.	Pl.
Masc.	Fem.	Masc.	Fem.	Masc.	Fem.	Masc.	Fem.	All Genders	All Genders
ele	ela	elos	elas	-o	-a	-os	-as	-(e)	-es

(Alkire & Rosen 2010:190, 206)

Table 81. Modern Spanish Agreement Targets

Definite Articles				Adjectives < o-/ā-stems				Adjectives < i-stems	
Sg.		Pl.		Sg.		Pl.		Sg.	Pl.
Masc.	Fem.	Masc.	Fem.	Masc.	Fem.	Masc.	Fem.	All Genders	All Genders
el	la (el)	los	las	-o	-a	-os	-as	-e	-es

(Alkire & Rosen 2010:191, 205)

In ES, the gender syncretism index increased to 0.78, the same as CSI. As part of the loss of the last surviving case distinction, the remaining distinctive Nom forms were leveled to the Acc for adjectives, as with nouns. This made singular neuter forms of adjectives syncretic with the masculine, as was already true for the Acc in WVL. It was the Nom singular forms of determiners that survived, however, so the neuter forms remained distinctive and were repurposed for use with nonspecific antecedents (see Pharies 2007:115-117). Since these forms were no longer used in agreement with nouns, singular determiners are also considered to have

masculine-neuter syncretism in my analysis. Similarly, the neuter plural forms of all agreement targets fell out of use. In contrast to CSI, no class of ambigeneric nouns developed, so masculine-neuter syncretism is assumed for all agreement targets. Like CSI, however, the gender syncretism index instead decreased to 0.33 when recalculated out of the two remaining genders. Gender syncretism remains the same in ModSp.

As in CSI, the retention of a Type 3 declension system in ModSp despite the loss of the neuter and all case distinctions can be attributed to the overt marking of declension in the singular of most nouns with endings that correspond to those on agreement targets and the lack of sound changes that would have neutralized these. However, one major difference was that final *-s* was lost in Italian but not Spanish, and this may account for the different treatments of the neuter. Due to *-s* loss in Italian, the *o*-stem neuter plural form *-a* was not only the same as the *ā*-stem feminine singular, it had the same basic form as the other plural forms: a single vowel. This likely made it easier for speakers to think of *-a* as a plural form and for the feminine plural forms of agreement targets to be associated with it, in accordance with the principle of system adequacy. In Spanish, on the other hand, the neuter plural was very different from the masculine and feminine *-s* plurals, so it instead merged entirely with the masculine, for the declension of both nouns and agreement targets. Ultimately, however, the ambigeneric class has lost productivity in CSI, so both languages have the same basic outcome. The difference between the gender syncretism index calculated out of two and three genders indicates the processing advantage gained by the complete loss of the neuter as a productive category, and this was the same whether or not the neuter first passed through an ambigeneric stage.

Table 82. Old French Agreement Targets

	Definite Articles				Adjectives < o-/ā-stems				Adjectives < i-stems			
	Sg.		Pl.		Sg.		Pl.		Sg.		Pl.	
	Masc.	Fem.	Masc.	Fem.	Masc.	Fem.	Masc.	Fem.	Masc.	Fem.	Masc.	Fem.
Nom	li, l’	la	li	les	-s	-e	-Ø	-es	-s	-Ø	-Ø	-s
Acc	le, lo, l’	la, l’	les	les	-Ø	-e	-s	-es	-Ø	-Ø	-s	-s

(Rickard 2003:50, 51)

Table 83. Middle French Agreement Targets

Definite Articles			Adjectives			
Sg.		Pl.	Sg.		Pl.	
Masc.	Fem.	All Genders	Masc.	Fem.	Masc.	Fem.
le, l’	la, l’	les	-Ø	-e	-s	-es

(Alkire & Rosen 2010:191-192)

Table 84. Contemporary Standard French Agreement Targets

Definite Articles			Adjectives			
Sg.		Pl.	Sg.		Pl.	
Masc.	Fem.	All Genders	Masc.	Fem.	Masc.	Fem.
le (l')	la (l')	les	-Ø	-e	-s	-es

(Alkire & Rosen 2010:191, 205)

The gender syncretism index increased to 0.75 in OF. As in ES, the neuter category was lost, and its forms fell out of use on agreement targets. They are assumed to be syncretic with the masculine forms in my analysis since there was already some overlap in CL, and most, but not all, neuter nouns became masculine. An additional increase resulted from the neutralization of the masculine-feminine distinction in the Acc plural of the definite article. However, this was more than offset by the new masculine-feminine distinction that arose in the Nom singular and plural forms of *i*-stem adjectives with the redistribution of the *-s* and zero endings by analogy with the pattern in *o*-/*ā*-stem adjectives, similar to the developments in the corresponding noun classes. When recalculated out of the two surviving genders, however, the gender syncretism index decreased to 0.25. With the loss of the last surviving case distinction in MF, Nom forms were leveled to the Acc if they were not already syncretic for all agreement targets, as with nouns. However, feminine *i*-stem adjectives adopted the distinctive *ā*-stem forms, leaving only plural determiners without any gender distinction. Thus, the gender syncretism index decreased to 0.72, or to 0.17 out of two genders. In CSF, sound change neutralized the gender distinction on adjectives ending in an oral vowel. The definite article also lacked a gender distinction before nouns beginning with a vowel. When these agreement targets are treated as one group, along with a group each for the more frequent adjectives and determiners that do not have these properties, the gender syncretism index has increased to 0.83, or to 0.5 out of two genders.

Like the other Western Romance languages, French has lost the neuter category. However, it has also developed a more transparent declension system, possibly in response to more extreme sound changes. These resulted in a declension system in OF with similarities to those in Germanic languages, particularly the prevalence of zero endings. Different syncretism patterns came to be associated with masculine and feminine, likely as a way to compensate for the lack of overt markers. This was no longer an option with the total loss of case in MF, however, and *-e* developed an even stronger association with the feminine, based on its presence in *ā*-stem forms. It is not clear why French developed a total association between gender and

declension as opposed to a total dissociation as in CSDan and CSDu, two of the three Germanic languages under investigation that retain two genders but no case distinctions. One difference that may have been a factor is that all gender distinctions were lost on the plural of agreement targets in CSDan and CSDu but not in French. As discussed above, this development may have enabled the dissociation of gender and declension in CSDu, but it more likely worked in the opposite direction in CSDan. However, other Germanic languages, including some Alsatian dialects of German, which are not included in my analysis, and NNw, which is, have developed a Type 1 system like French, even though they also lack gender distinctions in the plural of agreement targets. On the other hand, these languages retain three genders, and among the languages discussed by Kürschner and Nübling, they find that only those with three genders develop Type 1 systems, and only those with two genders develop Type 4 systems (2011:377-381).[140] However, my analysis has found exceptions to both of these tendencies, including French, so the reason for its development of a Type 1 system is still unclear.

4.3.6.2. Eastern Romance: Romanian

Table 85. Pre-Romanian Agreement Targets

	Definite Articles						Adjectives < o-/ā-stems					
	Sg.			Pl.			Sg.			Pl.		
	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.
Nom	(el)le	(e)a	(el)lu	(el)l'i	(el)le	(e)a	-ui	-ă	-u	-i	-e	-ă
Acc	(el)lu	(e)a	(el)lu	(el)lui	(el)le	(e)a	-u	-ă	-u	-ui	-e	-ă

	Adjectives < i-stems					
	Sg.			Pl.		
	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.
Nom	-i	-i	-e	-i	-i	-ie
Acc	-e	-e	-e	-i	-i	-ie

(Alkire & Rosen 2010:281)

Table 86. Contemporary Standard Romanian Agreement Targets

Definite Articles					Adjectives < o-/ā-stems				Adjectives < i-stems	
Sg.			Pl.		Sg.		Pl.		Sg.	Pl.
Masc.	Fem.	Neut.	Masc.	Fem./ Neut.	Masc.	Fem.	Masc.	Fem.	All Genders	All Genders
-(u)l, -le	-a	-(u)l, -le	-i	-le	-∅	-ă	-i	-e	-e	-i

(Alkire & Rosen 2010:189, 283)

[140] Kürschner and Nübling's (2011) findings are discussed in more detail in section 3.1.2.1 above.

In Pre-Romanian, the gender syncretism index remained 0.36, as in EVL. In CSRm, however, it increased to 0.78. Nom singular forms that had stayed distinctive were leveled to the Acc as this distinction was lost for nouns and agreement targets. As a result, the neuter is no longer distinct from the masculine in any singular form, and all gender distinctions were lost in the singular of *i*-stem adjectives. These developments were very similar to those in CSI, and they were also accompanied by neuter plural forms of agreement targets falling out of use in favor of the feminine forms. Thus, ambigeneric agreement has developed for neuter nouns, except with *i*-stem adjectives, which make no gender distinctions. In contrast to CSI, these have remained productive, and CSRm is still considered to have three genders in my analysis.

As in CSI, masculine and feminine plural forms lacking *-s* in the same way as the neuter plural form probably contributed to the development of ambigeneric agreement. In contrast to CSI and ModSp, however, CSRm has developed a Type 2 declension system but also retained a productive ambigeneric class. Thus, it can be considered to have undergone more simplification than the Type 3 Romance languages in terms of the transparency of the gender-declension relationship, but less in terms of gender categories. This tradeoff has parallels to the divergent developments in Swedish and Norwegian discussed in section 4.3.5.2 above. It is possible that the simplification in Italian and Spanish occurred as it did because they did not have the morphological material needed to develop a clear association of each declension class with one gender. As in OF, this was accomplished in Romanian by the association of one syncretism pattern with masculine classes and another with feminine classes, meaning it was only possible because there were still case distinctions on nouns.

4.3.7. Analysis of Gender Syncretism on Agreement Targets and Gender-Declension Relationship in Balkan Sprachbund languages

This section of the gender analysis focuses on individual languages. As in the previous sections, this one presents the relevant paradigms for calculating my original gender syncretism index in each Balkan Sprachbund language, then discusses changes in the index over time, noting how they relate to other developments, especially in number of genders and gender-declension relationship.

Table 87. Gender Syncretism on Agreement Targets in Balkan Sprachbund Languages

Stage	# of Declensions	Case	Gender		
		# of Cases	# of Genders	Type	Syncretism Index
LPS	10	7	3	3	0.17
PMB (Eastern)	10	7	3	3	0.39
MB (Eastern)	9	4	3	3	0.39
CSB	4	2	3	3	0.5
PMB (Western)	10	7	3	3	0.31
MB (Western)	9	4	3	3	0.25
CSM	4	2	3	3	0.5
MSrb	10	6	3	3	0.17
BCMS	8	6	3	3	0.17
MyG	8	6	3	3	??
AG	8	5	3	3	0.36
KG	8	5	3	3	0.36
ModGr	8	4	3	2	0.17
PA	2	5	3	3	??
CSA	2	5	3	4	0.73

4.3.7.1. Eastern South Slavic: Bulgarian and Macedonian

Table 88. Late Proto-Slavic Agreement Targets

	Demonstratives						Hard Long-Form Adjectives					
	Sg.			Pl.			Sg.			Pl.		
	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.
Nom	tъ	ta	to	ti	ty	ta	-ъjь	-aja	-oje	-iji	-yję	-aja
Acc	tъ, togo	tǫ	to	ty	ty	ta	-ъjь, -ajego	-ǫjǫ	-oje	-yję	-yję	-aja
	Soft Long-Form Adjectives						Hard Short-Form Adjectives					
	Sg.			Pl.			Sg.			Pl.		
	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.
Nom	-ьjь	-aja	-eje	-iji	-ęję	-aja	-ъ	-a	-o	-i	-y	-a
Acc	-ьjь, -ajego	-ǫjǫ	-eje	-ęję	-ęję	-aja	-ъ, -a	-ǫ	-o	-y	-y	-a
	Soft Short-Form Adjectives											
	Sg.			Pl.								
	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.						
Nom	-ь	-a	-e	-i	-ę	-a						
Acc	-ь, -a	-ǫ	-e	-ę	-ę	-a						

(Huntley 1993:144-145, 147)

Table 89. Pre-Middle Bulgarian (Eastern) Agreement Targets

	Demonstratives					Hard Short-Form Adjectives				
	Sg.			Pl.		Sg.			Pl.	
	Masc.	Fem.	Neut.	Masc./Fem.	Neut.	Masc.	Fem.	Neut.	Masc./Fem.	Neut.
Nom	-t	ta	to	ti	ta	-Ø	-a, -ŭ	-o, -u	-i	-a, -ŭ
Acc	-t, togu	tŭ	to	ti	ta	-Ø, -a, -ŭ	-ŭ	-o, -u	-i	-a, -ŭ

	Soft Short-Form Adjectives[141]					
	Sg.			Pl.		
	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.
Nom	-Ø	-a, -ŭ	-e, -i	-i	-e, -i	-a, -ŭ
Acc	-Ø, -a, -ŭ	-ŭ	-e, -i	-e, -i	-e, -i	-a, -ŭ

Table 90. Pre-Middle Bulgarian (Western) Agreement Targets

	Demonstratives					Hard Short-Form Adjectives				
	Sg.			Pl.		Sg.			Pl.	
	Masc.	Fem.	Neut.	Masc/Fem.	Neut.	Masc.	Fem.	Neut.	Masc./Fem.	Neut.
Nom	-t	ta	to	ti	ta	-Ø	-a	-o	-i	-a
Acc	-t, togu	tŭ	to	ti	ta	-Ø, -a	-a	-o	-i	-a

	Soft Short-Form Adjectives					
	Sg.			Pl.		
	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.
Nom	-Ø	-a	-e	-i	-e	-a
Acc	-Ø, -a	-a	-e	-e	-e	-a

Table 91. Middle Bulgarian (Eastern) Agreement Targets

	Definite Articles					Hard Short-Form Adjectives				
	Sg.			Pl.		Sg.			Pl.	
	Masc.	Fem.	Neut.	Masc./Fem.	Neut.	Masc.	Fem.	Neut.	Masc. /Fem.	Neut.
Nom	-t	ta	to	ti	ta	-Ø	-a, -ŭ	-o, -u	-i	-a, -ŭ
Acc	-t, togu	tŭ	to	ti	ta	-Ø, -a, -ŭ	-ŭ	-o, -u	-i	-a, -ŭ

	Soft Short-Form Adjectives					
	Sg.			Pl.		
	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.
Nom	-Ø	-a, -ŭ	-e, -i	-i	-e, -i	-a, -ŭ
Acc	-Ø, -a, -ŭ	-ŭ	-e, -i	-i	-e, -i	-a, -ŭ

[141] Due to vowel contraction, hard and soft long-form adjectives developed the same endings as the soft short-form adjectives, with the exception of the animate masculine Acc singular form *-ego*.

Table 92. Middle Bulgarian (Western) Agreement Targets

	Definite Articles					Hard Short-Form Adjectives				
	Sg.			Pl.		Sg.			Pl.	
	Masc.	Fem.	Neut.	Masc./Fem.	Neut.	Masc.	Fem.	Neut.	Masc./Fem.	Neut.
Nom	-t	ta	to	ti	ta	-Ø	-a	-o	-i	-a
Acc	-t, togu	ta	to	ti	ta	-Ø, -a	-a	-o	-i	-a

	Soft Short-Form Adjectives					
	Sg.			Pl.		
	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.
Nom	-Ø	-a	-e	-i	-e	-a
Acc	-Ø, -a	-a	-e	-i	-e	-a

Table 93. Contemporary Standard Bulgarian Agreement Targets

Definite Articles						Hard Adjectives			
Sg.			Pl.			Sg.			Pl.
Masc.	Fem.	Neut.	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.	All Genders
-ŭt, -a, -ta	-ta	-to	-ta, -te	-te	-ta	-Ø, -i	-a	-o	-i

Soft Adjectives			
Sg.			Pl.
Masc.	Fem.	Neut.	All Genders
-i	-a	-e	-i

(Scatton 1993:202, 207)

Table 94. Contemporary Standard Macedonian Agreement Targets

Definite Articles						Hard Adjectives			
Sg.			Pl.			Sg.			Pl.
Masc.	Fem.	Neut.	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.	All Genders
-ot	-ta	-to	-ta, -te	-te	-ta	-Ø, -i	-a	-o	-i

Soft Adjectives			
Sg.			Pl.
Masc.	Fem.	Neut.	All Genders
-i	-a	-e	-i

(Friedman 1993:264, 267)

LPS already had a small amount of gender syncretism in the core cases of agreement targets, with an index of 0.17. Specifically, there was masculine/feminine syncretism in the Acc plural. In PMB, the merger of **i* and **y* neutralized the distinction between masculine and feminine in the Nom plural for hard-stem adjectives. In eastern dialects of PMB, this distinction was no longer reliable for soft-stem adjectives as well, due to vowel raising in unstressed syllables. In addition, the adoption of the Gen singular form for the Acc singular with animate masculine referents meant that this form could no longer be reliably distinguished from the feminine Acc singular for short-form adjectives. In both CSB and CSM, however, agreement targets no longer reliably distinguish gender in the plural. The plural form *-i*, already used for the masculine and

hard-stem feminine, has also replaced the neuter and soft-stem feminine forms. In the eastern dialects of PMB that developed into CSB, the latter had already merged with the masculine forms when unstressed, as endings are for most adjectives (see Scatton 1993:207). In the western dialects that developed into CSM, however, sound change cannot account for the replacement of the soft-stem feminine form, just as it cannot account for the replacement of the neuter forms in CSB or CSM. Multiple postpositive definite article forms are used in the plural when they are attached directly to the head noun, but they are phonologically conditioned in this position, so they do not reliably indicate gender. Thus, the gender syncretism index has increased to 0.5 in both CSB and CSM.

Like CSI and ModSp, CSB and CSM have retained a Type 3 declension system despite near-total case loss and an increase in gender syncretism. As in those Romance languages, this may be explained by the persistence of the same overt endings in the singular for the most common feminine and neuter noun classes and the corresponding agreement targets; a zero ending is associated with masculine nouns and agreement targets, but also occurs on feminine *i*-stems, so the system is not quite as transparent. On the other hand, CSB and CSM have lost all gender distinctions in the plural of agreement targets and fewer noun classes are distinctive in the plural than the singular. However, these types of developments cannot lead to the dissociation of gender and declension if there are still singular forms to distinguish the declensions, as seen in CSHG. In addition, CSB and CSM retain three genders like CSHG, and this seems to make the development of a Type 4 system less likely, though not impossible as in ModLG. It is even clearer why CSB and CSM have not developed a stronger association between gender and declension: they could not make use of syncretism patterns to distinguish gender as in OF and CSRm, and there are too many classes to develop a Type 1 system.

4.3.7.2. Western South Slavic: Bosnian-Croatian-Montenegrin-Serbian

Table 95. Middle/Modern Bosnian-Croatian-Montenegrin-Serbian Agreement Targets

	Demonstratives						Long Adjectives					
	Sg.			Pl.			Sg.			Pl.		
	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.
Nom	tāj	tā	tō	tī	tē	tā	-ī	-ā	-ō	-ī	-ē	-ā
Acc	tāj, tog(a)	tū	tō	tē	tē	tā	-ī, -ōg(a)	-ū	-ō	-ē	-ē	-ā

	Short Adjectives					
	Sg.			Pl.		
	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.
Nom	-∅	-a	-o	-i	-e	-a
Acc	-∅, -ōg(a)	-u	-o	-e	-e	-a

(Browne 1993:325, 328)

In MSrb, the gender syncretism index did not increase from its value of 0.17 in LPS. The hard- and soft-stem paradigms merged as they did with nouns, i.e., the soft-stem forms survived in the masculine Acc plural and feminine Nom/Acc plural. Therefore, unlike in PMB, these forms remained distinct from the masculine Nom plural despite undergoing the same merger of **i* and **y*. No additional sound changes or analogical processes affected gender distinctions in the core cases of agreement targets, so BCMS also retains an index of 0.17. Given the lack of changes to gender syncretism, it is not surprising that BCMS has retained three genders and the same Type 3 declension system as LPS.

4.3.7.3. Greek

Table 96. Ancient Greek Agreement Targets

	Definite Articles						Adjectives < o-/ā-stems					
	Sg.			Pl.			Sg.			Pl.		
	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.
Nom	ho	hē	to	hoi	hai	ta	-os	-ā, -ē	-on	-oi	-ai	-a
Acc	ton	tēn	to	tous	tās	ta	-on	-ān, -ēn	-on	-ous	-ās	-a

	Compound Adjectives < o-/ā-stems				Adjectives < u-/ā-stems					
	Sg.		Pl.		Sg.			Pl.		
	Masc./Fem.	Neut.	Masc./Fem.	Neut.	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.
Nom	-os	-on	-oi	-a	-us	-eiā	-u	-eis	-eiai	-ea
Acc	-on	-on	-ous	-a	-un	-eiān	-u	-eis	-eiās	-ea

(Sihler 1995:348-350)

Table 97. Koine Greek Agreement Targets

	Definite Articles						Adjectives < o-/ā-stems					
	Sg.			Pl.			Sg.			Pl.		
	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.
Nom	o	e	to	oi, ü	e	ta	-os	-a, -e	-on	-oi, -ü	-ε	-a
Acc	ton	ten	to	tus	tas	ta	-on	-an, -en	-on	-us	-as	-a
	Compound Adjectives < o-/ā-stems						Adjectives < u-/ā-stems					
	Sg.			Pl.			Sg.			Pl.		
	Masc./Fem.		Neut.	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.
Nom	-os		-on	-oi, -ü	-oi, -ü	-a	-üs	-ia	-ü	-is	-iɛ	-ɛa
Acc	-on		-on	-us	-ous	-a	-un	-ian	-ü	-is	-ias	-ɛa

Table 98. Modern Greek Agreement Targets

	Definite Articles						Adjectives < o-/ā-stems (including compound)					
	Sg.			Pl.			Sg.			Pl.		
	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.
Nom	o	i	to	i	i	ta	-os	-a, -i	-o	-i	-es	-a
Acc	to(n)	ti(n)	to	tus	tis	ta	-o	-a, -i	-o	-us	-es	-a
	Adjectives < u-/ā-stems											
	Sg.			Pl.								
	Masc.	Fem.	Neut.	Masc.	Fem.	Neut.						
Nom	-is	-ia	-i	-is	-ies	-ea						
Acc	-i	-ia	-i	-is	-ies	-ea						

Agreement targets are not well-attested in MyG, but there was already some gender syncretism in their core cases in AG, with an index of 0.36. As in CL and LPS, the most common class of adjectives used *ā*-stem forms for feminine and *o*-stem forms for masculine and neuter. These had masculine/neuter syncretism in the Acc singular, as with *o*-stem nouns. The forms used on determiners were similar to those on adjectives, but the neuter was not syncretic with the masculine in the Acc singular because it had a distinctive Nom/Acc singular form. Another common class of adjectives used *ā*-stem forms for feminine and *u*-stem forms for masculine and neuter; like determiners, these had no gender syncretism in the core cases. Compound adjectives used the masculine *o*-stem forms for the feminine as well; along with the much rarer consonant-stem adjectives, which also have the same forms for masculine and feminine, these account for most of the gender syncretism in AG. The gender syncretism index did not change in KG, but it decreased to 0.17 in ModGr. The feminine Nom plural form of the definite article was no longer distinct from the masculine form, primarily due to sound change. However, this increase in gender syncretism was more than offset by compound adjectives adopting *ā*-stem forms for the feminine by analogy with other adjectives. The loss of *-n* was an important sound change in the development of nominal inflection, but it was blocked where it would have increased gender

syncretism, i.e., in the Acc singular of the definite article and other determiners. It was not blocked for the most common adjective classes, which already had masculine/neuter syncretism in the Acc singular.

ModGr has developed a Type 2 declension system through the association of one singular declension pattern with masculine classes and another with feminine classes, as in CSRm. However, the neuter remained distinctive on both nouns and agreement targets, so gender syncretism did not increase as it did with the development of the ambigeneric class in CSRm. Thus, while it is clear how a Type 2 declension developed, it is not entirely clear why this additional transparency was necessary. However, this is one of several ways in which Greek has undergone simplification, despite retaining three morphological cases and three genders. These must be explained by factors beyond those measured directly in my analysis.

4.3.7.4. Albanian

Table 99. Contemporary Standard Albanian Agreement Targets

	Definite Articles				Linking Articles				Adjectives (most common pattern)		
	Sg.			Pl.	Sg.			Pl.	Sg.	Pl.	
	Masc.	Fem.	Neut.	All Genders	Masc.	Fem.	Neut.	All Genders	All Genders	Masc.	Fem.
Nom	-i, -u	-a	-(i)t	-t(ë)	i	e	të (e)	të	-∅	-∅	-a
Acc	-in, -un, -në	-n, -së	-(i)t	-t(ë)	të (e)	të (e)	të (e)	të	-∅	-∅	-a

(Newmark 1982:160-163, 181, 183)

Agreement targets in Proto-Albanian cannot be reconstructed with enough certainty to determine the level of gender syncretism. In CSA, the distribution of gender marking on agreement targets is quite complex. This is largely due to the existence of four gender categories; one of these, the neuter, is only used in the singular due to the semantics of the remaining neuter nouns. The postpositive definite article has distinctive masculine, feminine, and neuter forms in the Nom and Acc singular but no gender distinctions in the plural. Adjectives are used with linking articles. In the Nom singular, there are distinctive masculine, feminine, and neuter forms of the linking article, but generally no difference in the form of the adjective itself. In the Nom/Acc plural, on the other hand, most adjectives have a feminine form that is distinct from the form used for the masculine plural, which is the same as the singular form used for all genders. In the Acc singular,

there are no gender distinctions on the linking article or the adjective. Finally, ambigeneric nouns have masculine agreement in the singular and feminine agreement in the plural. When calculated out of four genders in the singular and three in the plural, the gender syncretism index for the forms described above is 0.73.

With four genders, CSA has a unique agreement system among the languages investigated in my analysis. Of the Type 3 languages, it resembles CSHG most closely. Neither has lost all case distinctions on nouns, which has allowed at least one major distinction in declension to survive: one class for feminine and another for the other genders. This would not be possible based on the plural forms alone, since in both languages, all plural forms that are available for feminine nouns are also used in classes associated with at least one other gender. One difference from CSHG, as well CSB and CSM, is that some agreement targets still distinguish gender in the plural in CSA. This is critical for the retention of an ambigeneric gender; it would be indistinguishable from the masculine without a distinction between masculine and feminine agreement in the plural. Given that the postpositive definite article lacks gender distinctions in the plural, the ambigeneric category appears to be under greater threat than in CSRm, which has a gender distinction in the plural of more agreement targets, although still not all. As discussed in section 2.5.4 above, the ambigeneric pattern may have developed in Albanian due to Romanian influence. The ambigeneric is still in the process of displacing the neuter in CSA, and this may provide some insight into how this occurred in the development of Romanian. There are differences, however. In Romanian, the neuter first lost all distinctions with the masculine in the singular, i.e., for both noun classes and agreement targets. In CSA, neuter nouns are only used in the singular, and they are only distinguished from the masculine by agreement, not in terms of noun declension.

4.3.8. Analysis of Gender Syncretism on Agreement Targets by Gender-Declension Relationship

In this section, changes in gender syncretism are analyzed in terms of the groups based on gender-declension type that have been established in the preceding sections. I find that Type 3 gender-declension systems are not associated with the lowest gender syncretism index, while Type 1 and 4 systems are associated with the highest index. Type 2 systems actually have the lowest index on average.

As expected, gender syncretism on agreement targets increased the most in Type 1 and 4 languages. However, there was much less differentiation between Type 2 and 3 languages, as well as the Type 1 language in Germanic. Gender syncretism increased the same amount in MHG, the Type 3 Germanic language, as MNw, the Type 1 language, and MSw, a Type 2 language. This increase was more than in the other Type 2 Germanic languages, since masculine and feminine could no longer be distinguished in the plural in MHG but still could in the Nom plural in the Type 2 languages, including MSw, as well as the Acc plural in Icelandic. Gender syncretism did not increase in Type 2 Romance and Balkan Sprachbund languages, but it still remained slightly higher than in Type 3 Balkan Sprachbund languages. For the Type 3 Romance languages, the large increase can be partially attributed to the loss of the neuter category, which was complete in ES and in progress in Pre-Italian. As calculated for this part of the analysis, the lowest possible index with only two genders is 0.33. This explains why these Type 3 languages are closer to OF and MF, which were in the process of developing a Type 1 declension system and had already undergone a merger of the neuter with the masculine. Although the merger of the masculine and feminine was in progress in some of the Type 4 Germanic languages, more of the large increases among these languages can be attributed to the general loss of gender distinctions in the plural.

My investigation of the languages found that it was not until the development of the modern stage that they actually developed relationships between gender and declension other than Type 2 or 3, with the exception of EME and MF, mentioned above. The Type 4 languages continue to have the highest index among the Germanic groups, followed by the Type 1 language. Type 2 languages continue to have the lowest index, but there is significant divergence within the Type 2 languages: masculine and feminine have completely merged in CSS, but Icelandic and spoken Faroese have no gender syncretism on the core cases of agreement targets.

Within both the modern Romance and Balkan Sprachbund languages, gender syncretism on agreement targets does not appear to be correlated with gender-declension type. The apparent correlation within the Germanic languages can mostly be attributed to the number of genders and whether gender distinctions have been lost in the plural, both of which factor heavily into the gender syncretism index as calculated for this part of the analysis. Type 1, 2, and 3 languages have three genders, with the exception of CSS, while Type 4 languages have two genders or no gender, with the exception of ModLG. No Type 1 or 2 languages have lost all gender distinctions

in the plural, but all Type 3 and 4 languages except ModLG. In contrast, no Romance language has retained three genders or lost all gender distinctions in the plural, so these developments are not correlated with gender-declension type. All Balkan Sprachbund languages have retained at least three genders, so number of genders does not vary by gender-declension type. However, the Type 3 languages have lost most or all gender distinctions in the plural, while the Type 2 languages have not, which explains the lower average for Type 2. Overall, if gender syncretism on agreement targets is correlated with gender-declension type, it is Type 2 languages that have the least gender syncretism, not Type 3 languages. This is somewhat unexpected, since Type 3 is arguably the most conservative and least transparent relationship between gender and declension.

When calculated out of the three genders that all of the languages under investigation had in their earliest attested stages, the gender syncretism index strongly reflects the number of genders that are still distinctive in each language, along with the loss of gender distinctions that are limited to particular forms of agreement targets. This provides a reliable indication of when mergers occurred but not why they occurred; the gender syncretism index increases when genders merge, so it would be circular to argue that they merged because of an increase in the index. However, the gender syncretism index can also be calculated out of the number of genders remaining in a language, in the same way the number syncretism index is calculated out of the number of case-number forms remaining. This reflects the synchronic state of gender agreement in each stage of a language and, therefore, how difficult it might be for learners to acquire and speakers to use. Thus, it is a better way to measure correlation between gender syncretism on agreement targets and gender-declension type.

In the stages following those when the languages under investigation were first attested, most still had three gender categories, but the masculine-neuter merger had already occurred in some of the Romance languages. When calculated out of the remaining genders, the average gender syncretism index only increased from 0.36 in WVL to 0.38 in the Type 3 Romance languages, i.e., Pre-Italian, which still had three genders, and ES, which only had two. This is 0.22 lower than the average index of 0.6 when calculated out of three genders, but it is still 0.02 higher than in WVL. When calculated out of the two remaining genders, the index decreased to 0.25 in OF, then further to 0.17 in MF. These are 0.5 and 0.55 lower than the respective indices calculated out of three genders, and the index in OF is 0.11 lower than in WVL. Pre-Romanian still had three genders, so its index was 0.36 regardless. Thus, when controlling for the number

of genders, the stronger the eventual association between gender and declension, i.e., the lower the Type in the modern stage, the lower the gender syncretism index.

In the modern stage, CSI has also lost a productive neuter category. As a result, the average index out of the two remaining genders decreased by another 0.05 to 0.33 in the modern stage of the Type 3 Romance languages, i.e., CSI and ModSp. This is 0.45 lower than the average index of 0.78 calculated out of three genders. In CSRm, on the other hand, the index increased by 0.42 to 0.78 regardless. When calculated out of three genders, these three languages all have the same index, since all retain an adjective class without gender distinctions, and none retain distinctive neuter forms. However, the difference in their indices when calculated out the remaining genders reflects the simplification achieved by losing a gender entirely, as in CSI and ModSp, instead of retaining a gender, as in CSRm. As mentioned above, there may have been a trade-off between more transparent gender marking on agreement targets, as in CSI and ModSp, and a more transparent gender-declension relationship, as in CSRm. In contrast, both systems became more transparent in French, perhaps a necessity due to the more extreme sound change it experienced. The gender syncretism index out of the two remaining genders increased by 0.33 to 0.5 in CSF. This is 0.28 lower than the index of 0.83 calculated out of three genders. Although CSF has the highest index of the modern Romance languages when calculated out of three genders, it is lower than CSRm when calculated out of two. Gender syncretism on agreement targets and gender-declension type are related in a somewhat more expected way in the modern Romance languages. As in the previous stages, Type 1 has a lower index than Type 2. This suggests that once languages move away from Type 3, a stronger association between gender and declension is still linked with a lower gender syncretism index.

Of the Germanic languages, only MDan experienced gender loss before the modern stage. When calculated out of the remaining genders, the average syncretism index for the Type 4 languages increased to 0.56 instead of 0.73, i.e., 0.17 less. More Germanic languages lost gender categories in the modern stage. In this stage of the Type 2 Germanic languages, i.e., Icelandic and spoken Faroese, which retain three genders, as well as CSS, which has two, the average gender syncretism index out of the remaining genders decreased from 0.19 in the previous stage to 0.11. This is 0.17 lower than the average index of 0.28 calculated out of three genders and 0.08 lower than in the previous stage. The average gender syncretism index out of the remaining genders increased from 0.56 to 0.64 in the modern stage of the Type 4 Germanic

languages, i.e., CSDan and CSDu, which have two genders, as well as ModLG, which retains three. LME is not included in this average because no gender categories remained, so this form of the index cannot be calculated. This average is 0.21 lower than the average index of 0.85 when calculated out of three genders, and it is 0.09 lower than in the previous stage calculated out of three genders, but 0.08 higher than the previous stage calculated out of the remaining genders. NNw and CSHG both retain three genders, so either way, they have an index of 0.58 and 0.5, respectively. As in the earlier stage of Romance languages, modern Germanic languages with a stronger gender-declension association have a lower gender syncretism index out of the remaining genders. When calculated this way, the index of 0.25 for CSS is much less of an outlier among the Type 2 Germanic languages than the index of 0.75 out of three genders.

However, the low index for Type 2 Germanic languages contrasts with the modern Romance languages. The higher index for CSHG, the Type 3 language, can be explained by its loss of gender distinctions in the plural, which also partially explains the higher average index in Type 3 Balkan Sprachbund languages. The relatively high average index for Type 4 Germanic languages, even when calculated out of the remaining genders, indicates that agreement targets are not very reliable indicators of gender. This reflects the impact of gender distinctions being lost in the plural, separate from the loss of gender categories, and provides further evidence of the link between the former and the dissociation of gender and declension. In general, Type 4 languages are the most likely to lose gender distinctions in the plural, and this also occurs in some Type 3 languages, but no Type 1 or 2 languages in my analysis have lost all gender distinctions in the plural. Thus, the hypothesized correlation of gender loss in the plural with a partial or total dissociation of gender and declension holds, but the hypothesized association of more conservative gender-declension relationships with a lower gender syncretism index only holds if a Type 2 relationship is considered at least as conservative as Type 3, which would go against commonly held assumptions. In fact, these two hypotheses are mutually exclusive to some extent because the prevalence of gender loss in the plural in Type 3 languages means that they tend to have a high gender syncretism index.

4.4. Discussion

Morphological changes usually profile more semantically relevant categories, i.e., make them more salient and more memorable, often at the expense of less relevant categories (Kürschner

and Nübling 2011). At least for IE languages, the most relevant morphological category in nominal inflection is number, followed by case, and finally declension class. Gender is a lexical category for nouns that is often closely linked to declension class, but it is only a morphological category on agreement targets. In the development of Germanic and other IE languages, the original class markers have often fused with case and number markers. Sound change has clearly played a role in these combinations, but they also tend to occur in accordance with relevance, as shown in Figure 1. In other words, declension markers were reinterpreted as case or number markers, perhaps due to the reduction or loss of the latter due to sound change. Later, case markers were also reinterpreted as number markers.

Figure 1. The relevance-driven path of covert declension markers (adapted from Kürschner & Nübling 2011:361)

Class markers have often disappeared from the unmarked values of their new host categories. This tendency is illustrated in Figure 2. It accounts for the increasing frequency of zero marking in the singular paradigms of Germanic and other IE languages, particularly the Nom singular, which is unmarked for both case and number. The use of marking on the plural makes the number category more salient as opposed to the unmarked singular. Likewise, the marking on the Gen makes this morphological case observable.

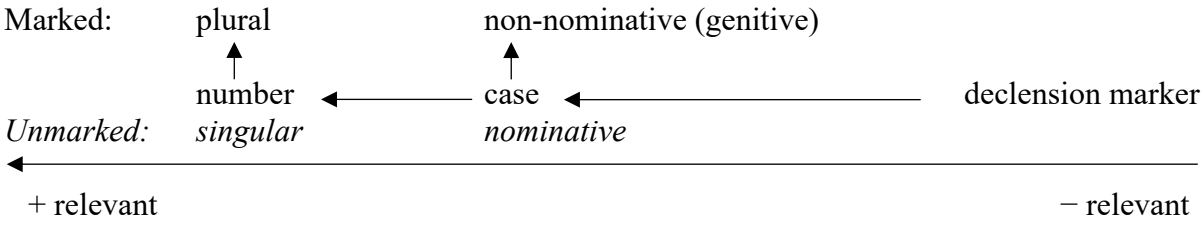


Figure 2. The disappearance of declension markers from unmarked category values, and their persistence on marked ones (adapted from Kürschner & Nübling 2011:362)

My general hypothesis for my study is an extension of Kürschner and Nübling's proposal concerning gender and declension and findings for a selection of Germanic languages. Overall, my analysis supports this hypothesis, namely that case and gender loss in the Germanic, Romance, and Balkan Sprachbund languages results from a combination of sound change that neutralizes morphological distinctions in the lost categories and analogical change that profiles more relevant categories whose distinctions have also been neutralized by sound change. The analysis of number syncretism in nominal inflection in section 4.2 above supports the idea that number was usually profiled at the expense of case. It shows that sound change very often increased syncretism among case forms of different numbers, i.e., number syncretism, in addition to case forms of the same number, i.e., case syncretism. It also indicates that analogical changes to nominal inflection between two stages of a language almost always had an overall effect of reducing the level of number syncretism.

4.4.1. The Effects of Phonological Changes on Case, Number, and Gender

In the analysis above, I find that larger increases in number syncretism due to sound change were usually followed by larger decreases due to analogical change. More precisely, the reduction in number syncretism by analogical processes seemed to depend on the level of number syncretism after sound change. This accounts for the variation in number syncretism across the earliest stages of the languages under investigation, which were first attested at different times and had already diverged drastically in terms of nominal inflection. For example, in some Germanic languages, there was an increase of number syncretism due to sound change in MSw, MDan, MNw, MHG, MLG, MDu, and LOE/EME. However, as the languages develop, the number syncretism was resolved by analogical processes. The same process occurred in some Romance and Balkan Sprachbund languages under investigation, i.e., WVL, EVL, PI, OF, Pre-Romanian, PMB (Eastern), PMB (Western), MSrb, AG, and KG. Thus, analogical processes that reduce number syncretism are not a response to sound change per se but to high levels of number syncretism, which often but not always result from sound change. One of the major differences across languages was how quickly analogical processes countered increases in number syncretism. In some languages, i.e., MSw, MDan, PMB (Eastern), PMB (Western), ModGr, and CSA, analogical changes occurred more or less concurrently with sound changes, resulting in relatively little change in number syncretism from one stage to the next, despite the potential for

a large increase due to sound change if analogical processes had not occurred. In other languages, i.e., MNw, MHG, MLG, OF, EVL, Pre-Romanian, MSrb, and AG, analogical processes were delayed until a later stage of the language, well after the sound changes had occurred. Some of this variation may be connected to how languages are divided into stages, a process that necessarily has arbitrary aspects.

Morphological blocking of a sound change can allow a form to remain distinctive in both case and number. For instance, in the development of MSw, the morphological blocking of *-r* in the Nom singular for masculine *i*-stems/*u*-stems prevented number syncretism in the Nom/Acc plural to occur. In the same vein, in ModGr, the morphological blocking of *-n* in the Gen plural removed this form's long-standing syncretism with the Acc singular for masculine *o*-stems, as well as the Nom/Voc/Acc for neuter *o*-stems, but the Gen plural has also remained distinct from other plural forms.

A form's stem can be leveled with no change to the ending itself. This is how umlaut was limited to the plural in many Germanic languages: singular stems that had acquired umlaut by regular sound change were leveled to the singular stem without umlaut at some point after umlaut was no longer an active phonological process. However, this sometimes still resulted in the loss of case distinctions because the ending had already been lost and umlaut was its only distinguishing feature.

Number profiling explains some loss of gender categories, but other factors are clearly important as well. Sometimes sound change neutralizes gender distinctions on agreement targets, just as it neutralizes case distinctions. Vowel reduction in particular appears to be a critical factor in the loss of gender distinctions, since they are so often conveyed by vowel differences alone. However, vowel reduction can neutralize distinctions in one morphological category but not another, depending on the vowels used to express them. For example, the Romance languages and Bulgarian both underwent processes that neutralized the distinctions between *e* and *i*, as well as *o* and *u*, at least in certain contexts. These were important in distinguishing case and number in these languages, e.g., LPS masculine *jo*-stem Nom/Voc plural *-i* vs. Acc plural *-ę* (later *-e*) and neuter *o*-stem Nom/Voc/Acc singular *-o* vs. Dat singular *-u*. However, they were not particularly important in distinguishing gender in the core cases, at least in the singular. As can be seen in Table 76, no agreement target paradigm in CL had forms for two genders that differed only in the presence of *e/ē* vs. *i* or *o/ō* vs. *u*, the vowels that merged in VL. Several gender distinctions

were neutralized due to other sound changes in VL, such as *-m* loss, but not vowel reduction itself. Table 88 shows that masculine and feminine were distinguished by the contrast between *i* and *ę* in the Nom plural of soft-stem adjectives in LPS, but no distinctions in the singular relied on this contrast. In both eastern and western dialects of PMB, the merger of *i* and *y* neutralized this same distinction for hard-stem adjectives, and the denasalization of *ǫ* neutralized the distinction between the feminine and the animate masculine in the Acc singular of short-form adjectives. However, both of these processes applied to vowels regardless of stress, so they are not considered vowel reduction. Even considering vowel mergers more broadly, they had a limited effect on gender distinctions in the singular. Therefore, instances of vowel reduction such as these are not considered significant for the purposes of this discussion about gender.

Vowel reduction also neutralized the distinction between *e* and *i* in Swedish and Norwegian. As in the Romance languages and Bulgarian, this distinction was more important in distinguishing case and number than gender. Table 57 reveals that the number distinction in the Nom relied on the contrast between *e* and *i* for strong adjectives in OSw, and this was also true for masculine *i*-stem nouns. However, no gender distinction relied on this vowel contrast. The situation is slightly more opaque in Norwegian because the epenthetic *e* had not yet been inserted in ONw, as in Table 53, but the loss of the same number distinction as in MSw that resulted from its insertion followed by the optional loss of *-r* in MNw can be seen in Table 62; as in Swedish, there is no potential for the loss of gender distinctions from this vowel neutralization. Thus, the languages under investigation in which there has been some vowel reduction, but not vowel reduction that is significant by this definition, include Swedish, Norwegian, Italian, Spanish, Romanian, and Bulgarian.[142]

4.4.2. The Effects of Analogical Processes on Case, Number, and Gender

As discussed in section 4.2 above, the exceptions to the tendency that analogy reduces number syncretism fall into a few categories with different explanations. In the insular languages of North Germanic, Icelandic and written Faroese, there was not much number syncretism anyway, and the changes involved reorganization of declension classes in ways that made the system

[142] See chapter II for further discussion of sound changes in these and other languages, including those considered to constitute vowel reduction. See section 3.1.1.2 for a more detailed discussion of vowel reduction and its connection to the loss of distinctions in nominal inflection.

more regular overall. In contrast, in MSrb and Pre-Romanian there were periods of transition with much variation in the forms used, resulting in a range of possible number syncretism that could be lower or higher than what it would have been with only sound change. In the following stages of both languages, i.e., BCMS and CSRm, analogical developments resulted in less number syncretism than in the preceding stage, i.e., LPS and EVL. Thus, the overall effect of analogical processes in these languages was still to reduce number syncretism.

Other languages went through similar periods of variation, but these periods were not all captured by the analysis; when they were, the range had at most the same level of number syncretism as without analogical changes. This reflects the fact that some analogical changes were optional at that stage, but the effect of optional processes that increased number syncretism was never greater than the effect of obligatory processes that caused a decrease. Only in MSrb and Pre-Romanian were the effects of optional processes enough to result in a potential temporary increase in number syncretism. The most notable exception occurred in OF; it appears to involve profiling gender marking and is addressed below when the focus of this discussion shifts from interactions with other categories involving case to those involving gender.

I also find a correlation between case loss and decrease in number syncretism as a result of analogy. This supports the idea that number marking was profiled at the expense of case. However, the nature of the analogical processes a language and its speakers had available makes case loss a more or less necessary consequence of number profiling. The most common analogical process found in the investigation is a form of leveling, i.e., one form changing to match another in the same paradigm. This necessarily involves the loss of the distinction between the two forms involved; in order to profile number marking by eliminating number syncretism, a form must be leveled to the form of a different case with the same number, hence the reduction in case distinctions.

Another analogical process that can profile both number and case marking is the extension of a form from one declension class to another. The classes involved often already share some number of forms; in addition to setting up the context for the analogy to occur, this can create a feedback loop that leads to a full merger of two declension classes. As discussed in 3.1.2.2 above, when a partial or full merger of classes occurs, the forms of the larger class are more often retained. However, forms of the smaller class are more likely to survive if they are more salient, thereby maintaining or restoring distinctions that would have been lost if the more

frequent forms had been adopted. Thus, this is a process by which number and/or case are profiled at the expense of declension class.

As mentioned above, analogical processes in the development of OF from WVL actually increased number syncretism. This appears to be an exception where gender marking was profiled through declension at the expense of number. The major classes associated with each gender became more similar to other classes of the same gender. In the pattern that most masculine nouns shared after these changes, both singular forms were syncretic with the plural form of the other case but distinct from each other. Thus, the changes could be considered to profile case as well as gender, but certainly not number. The loss of the number distinction in the small group of neuter *o*-stems that became indeclinable masculine nouns accounts for the largest increase in number syncretism in the analysis. However, it is not clear this number syncretism was significant in the development of nominal inflection more broadly because these nouns were only used with a plural meaning, so there was no confusion. Nevertheless, there is still reason to believe that gender behaved as a more relevant category in the development of French than in other languages. In MF, the productive adjective and noun classes converged on one class for each gender. In the process, all number syncretism and all case distinctions were eliminated. Thus, number and gender were profiled at the expense of case and declension class.

In the development of ModGr, the analogical processes that profiled gender marking were not as closely connected to those that profiled number marking as in CSRm. In OF, the masculine pattern was characterized by all forms participating in number syncretism, and in CSRm, the feminine pattern involves syncretism between the plural and the Gen-Dat singular, but the masculine and feminine patterns in ModGr only concern syncretism among the singular cases: Voc/Acc/Gen for masculine vs. Nom/Voc/Acc for feminine. For *ā*-stems that retained the vowel *a* in singular forms, the other form in each pattern, i.e., Nom for masculine and Gen for feminine, had become syncretic with the Acc plural form *-as* in KG, but this syncretism was eliminated for both genders by the adoption of consonant-stem forms in the plural. While this merger of *ā*-stems and consonant-stems in the plural profiled number, their merger in the singular profiled gender, as consonant-stems adopted *ā*-stem forms, which already followed the patterns described above. Meanwhile, feminine singular *i*-stems merged with the feminine *ā*-stems that had developed the vowel *i* in singular forms. This merger partially resulted from sound change, but without analogy the *i*-stems would not have fit the syncretism pattern for

feminine or masculine, since neither the Nom nor the Gen was syncretic with the Voc/Acc. However, the Nom singular form *-is* was leveled to the Voc/Acc, bringing *i*-stems in line with other feminine classes, while also eliminating the syncretism among the Nom singular and Nom/Voc/Acc plural. Thus, gender and number were profiled by the same process, but another optional process essentially reversed this profiling of number. Namely, the *i*-stem Gen singular form sometimes changed from *-eos* to *-is*, the *ā*-stem form, making it syncretic with the Nom/Voc/Acc plural. Although *i*-stems already fit the general feminine syncretism pattern without this optional change, *-eos* is an exception among singular feminine forms in not having the same vowel as the Nom/Voc/Acc form. Thus, its replacement by *-is* can be seen as part of the overall simplification of nominal inflection in Greek. On the other hand, the optionality of this change, without which there is no number syncretism in ModGr nominal inflection, can be seen as a testament to the high relevance of number; alternatively, resistance to this change can be attributed to the influence of Katharevousa and earlier conservative forms of written Greek, as Horrocks does (2010:288, 462-463). It is noteworthy, however, that the retention or restoration of *-eos* and of the final *-n* in Gen plural forms both serve to profile number. It is possible that such learned forms were more likely to be accepted into everyday usage if they marked a relevant distinction. The retention of *-n* in the masculine Acc singular form of determiners and clitic pronouns, particularly when confusion with the neuter would otherwise be likely, adds further support to this idea. Under this assumption, the acceptance of these forms into the spoken as well as written standard, but not other archaizing features of Katharevousa such as the Dat, would indicate that number and gender are more relevant categories than case in ModGr.

Number and gender marking also appear to have been profiled together in other languages. The relationship between gender and declension became more transparent in the development of CSRm and ModGr, as in OF. Unlike in OF, however, this was accomplished by analogical processes that also reduced number syncretism, or at least did not increase it. In the development of both OF and CSRm, masculine nouns in classes that also contained many feminine nouns underwent analogical changes that differentiated them from those feminine nouns by adopting the syncretism pattern of masculine *o*-stems, the largest masculine class. In OF, the masculine pattern happened to be the one with more number syncretism, so the overall number syncretism in the system increased as a result. In CSRm, however, the masculine pattern has no number syncretism but also no case distinctions, in contrast to the pattern retained by

feminine nouns, so number syncretism decreased overall instead. The different levels of number syncretism in masculine *o*-stems cannot be attributed to sound change alone, since the Nom singular and Acc plural had become syncretic due to sound change in EVL but were leveled to the Acc singular and Nom plural, respectively, in the development of CSRm. A similar leveling eventually happened in MF but may have been resisted in OF because it would have resulted in a merger of *o*-stems with consonant-stems, meaning that no declension pattern would be associated with only the masculine. There was no risk of such a merger in Romanian, for at least two reasons. First, final vowel deletion without a trace only applied to *-u*, in contrast to OF, where it applied to all vowels except *-a*, so *o*-stems and consonant-stems still differed in the Acc singular, as well as the Nom/Voc singular once this was leveled to the Acc in both classes. Second, only *o*-stems had syncretism between the Acc and Gen-Dat singular, until other masculine nouns adopted this syncretism as well. Thus, less extreme sound changes and the retention of the Gen-Dat may have both allowed number and gender to be profiled together in the development of CSRm, while only one could be profiled in OF. Case and gender marking appear to have been particularly relevant in OF, but by MF, number had become more relevant again, suggesting that the situation in OF was not sustainable.

Total case loss also does not specifically have an effect on gender marking as a category because some languages lost case marking entirely but still kept gender marking. For example, CSI, ES, and MF still have gender marking despite the loss of morphological case. Their case loss has a significant effect on the number of gender markings because it reduces the number of forms that can carry gender distinctions. The gender of the reduced forms does not specifically change whether the forms are informative regarding gender.

4.5. Conclusion

This chapter has focused on two of the broad factors proposed as motivations for case and gender loss in the Germanic, Romance, and Balkan Sprachbund languages: phonological change and morphological change. The results by and large support the general hypothesis that the loss of morphological case and grammatical gender in these languages can be explained by the combined effects of sound changes that directly neutralize morphological distinctions in these categories and analogical changes profiling a more relevant category. The analysis also refutes some of specific hypotheses, particularly those based on the findings of Kürschner and Nübling

(2011), which were limited to the Germanic languages. In addition to adopting their typology of gender-declension relationships, I created original measures of number syncretism in noun paradigms and gender syncretism in agreement target paradigms. The methods for calculating these indices and other measures derived from them are described in section 4.1.

The analysis using the number syncretism index has found that increases in number syncretism as a result of sound change are generally correlated with concurrent or subsequent decreases in number syncretism as a result of analogical change, and that case loss is more likely when the effects of both types of change on number syncretism are larger, as hypothesized. Several languages were exceptional with respect to one or more of these associations, however, most notably High and Low German, but also Bulgarian, Macedonian, and French. Factors beyond the focus of my analysis provide potential explanations for these exceptions, including the role of articles, developments related to functional mergers, different contact situations, and standardization.

The gender analysis has tested the association of several factors, including the gender syncretism index and gender-declension relationship. The first part of the gender analysis has found that a gender category can still be lost in languages that retain a Type 3 system or develop a Type 1 system, and that the loss of a gender category is not necessary for the development of a Type 4 system. my study has not found an association between total case loss and a shift from a Type 2 or 3 system. Thus, the two hypotheses that follow most directly from Kürschner and Nübling's (2011) findings fail to hold when extended beyond the Germanic languages. A third hypothesis based less directly on their findings also failed to hold: lower gender syncretism was not entirely correlated with more conservative gender-declension relationships. However, the discrepancy can be explained by the correlation of gender loss in the plural, which greatly increases gender syncretism, with a dissociation of gender and declension, including the partial dissociation found in a Type 3 system, which is considered the most conservative. Despite the failure of Kürschner and Nübling's (2011) findings to extend well beyond the Germanic languages, the results still support their proposal about the important role of relevance in analogical change, including changes resulting in the loss of case and gender.

The following chapter considers two studies on case assignment by verbs, which provide more insight into the nature of functional narrowing.

CHAPTER V.
STUDIES ON FUNCTIONAL NORROWING

Chapter III discussed various motivations behind the loss of case and gender, with one of the internal factors being discussed is functional narrowing in section 3.1.2.6 above. This chapter considers that factor in more detail. Two studies on functional narrowing, are included in this chapter. The first is Vakareliyska (1990), a study of patients with a form of Wernicke's aphasia (specifically acoustico-amnestic aphasia), but the insights gained from it may have applications for the functional narrowing in unimpaired L1 and L2 speech. The second is my own corpus study of Old English (OE) and Early Middle English (EME) texts. The purpose of that study was to build on previous research into case variation in functional narrowing and the subsequent functional narrowing of the Gen and Dat in favor of the Acc during the transition from OE to Middle English (ME), a critical period in the development of English. Competition with prepositional constructions is also considered in that study.

Barðdal (2009) provides a theoretical basis for explaining the results of the second study, which reveal the gradual loss of productivity of the Gen and Dat in EME as they were displaced by the Acc and prepositional constructions. Together, the two studies provide insight into how and why case functions lose productivity faster than other case functions, including potential connections to language contact. Their connection to each other and to arguments made by Barðdal (2009) and others are discussed at the end of this chapter, together with a *post hoc* analysis of the changes to the case use among verbs that are semantically similar in CSR and OE.

5.1. The Functional Narrowing by Wernicke's Aphasia Patients

Vakareliyska aims to account for the choice of Acc or Dat as the primary object of a verb based on the latter's semantic features, with a focus on Balto-Slavic languages. In her model, the Dat canonically marks a non-active participant with ego-involvement, who therefore attracts the anticipatory interest of the speaker (1990:150-151). Vakareliyska later uses the term *cognitive engagement* to describe this feature of the Dat (1994). Vakareliyska notes that animacy is a necessary condition for *cognitive engagement*, except in metaphorical constructions, but is not a sufficient condition in itself. The Acc is the unmarked case for the non-active participant, but is still marked relative to the Nom. The theory that word meaning is composed of semantic features

is supported by studies with Wernicke's aphasics, who have incomplete lexical comprehension (1990:154-162, 210). For example, Vakareliyska notes that Rinnert and Whitaker (1973) demonstrated that patients had difficulty distinguishing between nouns whose referents shared certain features, and attributed this to these patients' lack of, or lack of access to, the distinctive features of these words. Another experiment found the most semantic impairment when patients had to use non-dominant properties to classify objects (Koemeda-Lutz et al. 1987). Based on these studies, Vakareliyska suggests that Wernicke's aphasics are more likely to retain access to dominant or typical semantic features of a word than to peripheral features. For verbs, at least, she proposes that the former comprise the core meaning and tend to be universal, while the latter tend to be language-specific and sometimes dependent on metaphor. Therefore, patients with mild to moderate Wernicke's aphasia may be able to understand the core meaning of a verb but still make errors in case marking if it depends on language-specific features (1990:162-168, 210-212). Thus, Vakareliyska's study concerns the loss of morphological case marking, but due to brain injury rather than in the historical development of languages.

Vakareliyska tested her proposal by comparing the case errors made by Wernicke's aphasics with different verbs. The participants were mostly native Russian speakers, but a few Bulgarian and Latvian speakers were also tested. All the patients were tested on-site in their home countries. Participants were asked to choose between an Acc and Dat pronoun in various verbal constructions. Each participant was tested once on each construction, so the number of errors for a particular construction was equivalent to the number of participants who made an error. Relatively few participants made errors in the use of Dat for IOs or the use of Acc for DOs of verbs involving physical activity, e.g., CSR *udarit′* 'hit', and verbs of perception, e.g., CSR *videt′* 'see'. For the latter category of verbs, this demonstrates that having an experiencer instead of agent as subject does not cause confusion. For the primary object of verbs that did not involve physical activity, however, there were approximately twice as many errors in the choice of the correct case, which was either Dat or Acc depending on the verb. For example, CSR *pomogat′* 'help' requires the Dat and CSR *sprosit′* 'ask, inquire of' requires the Acc. The test included a group of doublets in which two verbs have nearly the same meaning but different case constructions, e.g., CSR *dosaždat′* 'annoy' with the Dat and *bespokoit′* 'disturb' with the Acc. This group of verbs elicited a number of errors approximately as high as the verbs without semantic doublets in the test. Impersonal expressions with a Dat or Acc experiencer, e.g., CSR

X-Dat *povezlo* ‘X was in luck’ and X-Acc *vzorvalo* ‘X blew up’, also elicited a high number of errors (1990:77-118).

In a more detailed discussion of the results, Vakareliyska accounts for the higher number of errors with some verbs than others, both within and across the categories tested. In general, the verbs that elicited the most errors have language-specific features that cause them to be associated with a different case than expected from the core meaning. Other verbs had a lower but still significant number of errors. These tend to have a mix of features associated with each case; in other words, *cognitive engagement* may be less of a core part of the meaning than with verbs that are clearly expected to assign the Dat, but a total lack of *cognitive engagement* is not predicted, as it is with verbs that are clearly associated with the Acc. An evenly balanced number of features associated with each case may also explain case variation in the speech of unimpaired speakers. For example, unimpaired Latvian informants reported that *sist* ‘hit’ can be used with the Acc or Dat without any noticeable difference in meaning (1990:213-261). Thus, in addition to overlapping case functions, paradigmatic pressure, and sound change, case variation may result from ambiguity or contradictions among the semantic features of a verb.

Vakareliyska compares the loss of access to more peripheral features than core features by Wernicke’s aphasics to the position of an L2 speaker, who understands the core meaning, presumably based on the equivalent in their L1, but may not grasp language-specific semantic features (1990:211). If this parallel holds, Vakareliyska’s findings could provide another specific mechanism connecting an influx of L2 speakers to functional narrowing of morphological cases.

5.2. Study of the Functional Narrowing of Case Use in OE/EME Texts

This section investigates the assignment of case marking by verbs in three OE and EME texts: Beowulf, Layamon’s Brut, and The Owl and the Nightingale (O&N). The EME texts are representative of the period when significant morphological case ambiguity and case loss first became observable. The motivations for OE case marking and changes in later texts will also be examined in this section, including the semantic associations between the case markings and the verbs that assign them. This is a close study of a small corpus, and while it serves as part of a much broader cross-linguistic study of case loss, the conclusions reached here are applicable specifically to English. The limited nature of this study necessitates conservative and/or tentative conclusions.

In the transition from OE to EME, during the 11th and 12th centuries, OE case distinctions were lost, and the functions of the cases that remained morphologically distinct became more limited. Uses that were marked by the Genitive (Gen) or Dative (Dat) marker in OE began to be marked by the Accusative (Acc) marker or prepositions instead. While some linguists claim that morphological case was lost entirely (Lightfoot 1999:117-125), others argue that at least the Gen did not lose its status as a morphological case until later in the Middle English (ME) period (Allen 2002:57). This disagreement hinges on whether the possessive marker (originally *-es*, later -'*s*) still marked the morphological Gen or whether it instead marked simply the grammatical role of Possessor.

Another way to indicate the semantics associated with a verb is with a preposition. Even in OE, there are examples of Dat marking assigned by a verb being reinforced by a preposition that takes Dat (see, for example, Van Gelderen 2000:214). Later, after a period of competition between unreinforced Dat objects and complements in which the Dat marking is reinforced by a preposition, some verbs that once assigned a morphological case marking other than Acc came to be used *with* prepositions instead.

There are also some descriptions of the semantic classes of verbs that assign the various case markings in OE. Lass (1994) provides some semantic classification of these verbs and the connections these classes have to the meanings of the case assigned. He supports his classification by relying on Kuryłowicz (1964). However, his classification is not comprehensive in terms of the individual verbs or classes of verbs under investigation. Despite these shortcomings, the present study uses Lass (1994) as a starting point for grouping the OE verbs under investigation into semantic classes. Some semantic classes of Gen-assigning verbs that are not categorized by Lass but categorized by Bungenstab (1933) are used to give a complete semantic classification of the verbs under study. Clearly, by EME, the classification of these semantic classes does not hold because of the breakdown of the case system.

There is a limited number of studies that investigates the functional narrowing of the OE/EME case assignment by verbs. For example, Van Gelderen (2000) and Allen (2008) have limited examples of changes in case assignment in EME texts. My investigation aims to add new contributions to the literature by tracking the case assignment by verbs over time through using quantitative and statistical measures. Thus, my work helps to fill a gap in the field of OE and EME studies.

My study also is compared to Vakareliyska's study to show how loss of productivity of functions for aphasic speakers are similar in some way to the decline of case marking in later stages of English. Although having very different methodologies, these two studies generally contribute to answering issues about the early loss of productivity for functions relative to other case functions.

This corpus study aims to provide further insight into variation among the morphological cases and the functional narrowing of the Gen and Dat in favor of the Acc, and, in some instances, prepositional constructions, during a critical period in the development of English. The methodology is described in section 5.2.1. Connections between verb semantics and case assignment are considered in section 5.2.2. In section 5.2.3, the analysis of the functional narrowing of case use in the selected texts is presented, including a subsection on prepositional alternatives to case marking.

5.2.1. Methodology

The study employed the following methodology. After making a selection of verbs that could assign Dat and/or Gen in OE, I sorted these into semantic classes based on their meaning and use in OE, as described in the following section. For the main investigation, I used an online transcription version of each of the three texts in the corpus to search for objects of the specific verbs that had been selected. I then categorized the objects that had been found by their morphological case marking.

I referenced the glossaries of Marsden (2010) and Klaeber (2008) to determine which verbs could take objects in morphological cases other than Acc. For Marsden (2010), I found 36 verbs with Gen objects and 52 with Dat objects in this glossary. For Klaeber (2008), I found 25 verbs with Gen objects in addition to those in Marsden (2010) and 64 additional verbs with Dat objects. The combined list contained 61 verbs taking the Gen and 116 verbs taking the Dat. Verbs from this list were selected for further investigation if they could be confirmed to occur at least once in Layamon's Brut, the longest text investigated. Since the glossaries of Marsden (2010) and Klaeber (2008) both cover a limited subset of OE texts, I checked Mitchell (1985:455-464) for the full combination of case markings each selected verb could assign in

OE.[143] I also checked the Dictionary of Old English (DOE), an online source, for the verbs whose entries have been published, and generally found the same case assignments, although additional prepositions were sometimes listed as complements. The selected verbs are listed in Table 100 in columns based on these case-assignment categories.

Table 100. Verbs by Case-Assignment Category

+Gen	+Dat	+Gen/Dat	+Gen/Acc	+Dat/Acc	+Gen/Dat/Acc
missan 'miss'	*behōfian* 'be required of' (animate)	*bedǣlan* (*be)*, (*fram)*, (*of)* 'deprive of'	*ābīdan/onbīdan* 'await'	*dēman (be*), (*ofer)* 'decide, judge'	*beniman* 'deprive of'
reccan (be) 'care for'	*beorgan* 'protect'	*berēofan/berēafian (æt)*, (*fram)*, (*on)* 'deprive of'	(*ge)bīdan* 'await, experience'	*forlēosan* 'lose'	*brūcan* 'use, enjoy'
(be)weddian 'promise, wed'	*oƿyncan* 'displease' (animate)	*gilpan (for)* 'boast, rejoice'	*biddan* 'ask for' (inanimate)		*(ge)fandian* 'search out, experience'
	(*ge)unnan* 'grant' (animate)	*helpan* 'help'	*hēdan* 'heed'		*tilian* 'provide' (inanimate)
	þyncan 'seem'		(*ge)unnan* 'grant' (inanimate)		*tilian* 'provide for' (animate)
			wēnan 'expect, think'		(*ge)wealdan* 'control, possess'
			wundrian (æfter), *(be)*, *(fram)*, *(on)*, *(ymb)* 'admire'		*weorpan* 'throw'
					wilnian (to) 'desire'

Verbs that can appear with or without the *ge-* prefix were treated as a single verb for the purposes of this study. Verbs with other prefixes were generally treated separately. However, *ābīdan* and *onbīdan* 'await' were grouped together because they have the same meaning, and there are not enough data for *onbīdan* to constitute a separate verb. For similar reasons, *beweddian* 'promise, wed' was grouped with *weddian*; moreover, forms with the prefix are limited to the past participle and infinitive in the texts under investigation, in complementary distribution with unprefixed finite forms. On the other hand, each non-Nom argument of verbs with multiple objects was considered separately. Thus, *tilian* 'provide' (inanimate), 'provide for' (animate), was included twice: once for the thing provided and again for the person provided for;

[143] The verb *weorpan* 'throw' is not included in Mitchell (1985:455-464) or the DOE, so I continued to rely on the glossaries of Marsden (2010) and Klaeber (2008) for the cases assigned by this verb.

likewise, the thing granted and the person to whom it is granted were both included for (*ge)unnan* 'grant' (animate). I only included arguments for which the Gen or Dat was listed as an option in Mitchell (1985:455-464). Thus, only the thing requested was included for *biddan* 'ask for' (inanimate), not the person asked. Likewise, the object that is taken was included for *bedǣlan*, *beniman*, and *berēofan/berēafian* 'deprive of', but not the person deprived. For *ofþyncan* 'displease' (animate), only the displeased person was included. In OE, the displeasing thing could be Nom or Gen; the impersonal construction with the Gen occurs once in Beowulf but not in any of the later texts under investigation, so I excluded the inanimate argument from consideration. Similarly, there are two distinct argument structures for *behōfian* 'be required of' (animate): one with a Nom person-in-need and a Gen thing required, the other with a Dat person-in-need and, optionally, a Nom thing required. The former is attested once in Beowulf but not at all in the later texts, so it was excluded; only the Dat person-in-need was investigated. For consistency, verbs are cited throughout this study in their OE infinitive form, even when discussed in the context of EME. Some verbs also take prepositions in OE, these prepositions are provided with parentheses in Table 100.

I first investigated Beowulf as well as consulting the DOE entries and the list of OE verbs that Mitchell (1985:454-464) discussed for the assignment of cases/prepositions by verbs in OE. Beowulf's date of composition is strongly contested, with various scholars arguing for dates ranging from the middle of the 7th century, not long after missionaries had introduced writing to the English, to the beginning of the 11th, when the copying of the single surviving manuscript occurred. However, the first half of the 8th century seems most likely based on linguistic evidence (see Klaeber 2008:clxiii-clxxx). Data for Beowulf were obtained from an online transcription version of the full text of the poem at http://www.heorot.dk/beowulf-rede-text.html. Data for Layamon's Brut, likely written sometime in the early 13th century, was obtained from an online transcription version of the text based on the Cotton Caligula manuscript at http://quod.lib.umich.edu/c/cme/LayCal?rgn=main;view=fulltext. The final text investigated was The Owl and the Nightingale (O&N), written around 1200. Data for O&N were obtained from an online transcription version of the text based on the Cotton Caligula manuscript at https://quod.lib.umich.edu/c/cme/OwlC?rgn=main;view=fulltext.

Searching for objects of the verbs compiled from the glossaries above was accomplished in one of two ways, depending on the text. For Beowulf, the glossary in Klaeber (2008) lists the

line numbers where each verb occurs, grouped by the case of any object. Thus, it was simply a matter of going to these lines, copying them, and saving them for analysis. For the remaining texts, I entered the different possible spellings of the present, past, and past participle stems, as found in the Middle English Dictionary (MED) and accounting for orthographical tendencies of each text, into the browser search bar on each website listed above.

Many of the verbs searched were not found in the texts or else were not found with any complements. For those that were found, the first step was to confirm that the search result was, in fact, a verb and not a similarly spelled noun, adjective, or other lexical categories. Next the complement was identified. If no nominal object was found that instance of the verb was excluded from further analysis. Adverbial complements like *þer-to* 'to that' and clausal complements beginning with a complementizer such as *þat* 'that', *hu* 'how', and *ȝif* 'if' were also excluded. However, prepositional complements of the verb that contained a nominal were counted in a separate category.[144]

Each object found was categorized based on the case markings of which its form might be an instance, according to the procedure detailed in this paragraph. Any determiner, adjective, noun, or pronoun form present was considered in determining case. If two or more coordinated objects were present, they were assumed to have the same possible cases; thus, cases that were possible based on the form of one but not the other were eliminated. The possible cases for a given word were determined based on the forms listed in the paradigm to which it belonged, as compiled in the grammatical overview for that text.[145] However, many nouns do not inflect exactly like the example nouns listed in the grammatical overviews. Their gender in OE as well as the list of forms found in the MED were also taken into consideration in determining case. Context and translations were sometimes used to limit the possible case forms to either singular or plural, but translations are not always literal and sometimes a determination of number could not be made, and more morphological cases were possible. If the Acc, Dat, and Gen were all possible, then the case marking was categorized as ambiguous (Amb).

[144] The term *complement* will be used here to refer collectively to nouns, pronouns, and prepositional phrases that function as complements of the verb. This term excludes clausal and adverbial complements.

[145] See the Appendix A for grammatical overviews of each text except Beowulf whose grammar is consistent with standard descriptions of OE.

Once morphological case possibilities were assigned to each object, I counted the number of nouns and pronouns with each combination of case marking possibilities for each verb. The noun category includes substantive uses of adjectives such as *al* 'all' and *mare* 'more'. The pronoun category includes pronominal demonstratives, interrogatives, and inflected relatives, in addition to personal pronouns.

Statistical analysis using the chi-square test and Fisher exact test are obtained to determine if there is a statistical significance between the observed and expected data. This is performed by using Beowulf as expected data versus Brut and O&N as observed data. These statistical methods are utilized to show the statistical significance between two morphological case markings: the Gen and Dat.

5.2.2. Semantics of Verb Categories

This section attempts to explain why the verbs under investigation assigned a given case marking in OE. Lass (1994:234-238) serves as a starting point for the different semantic classes of OE verbs that can take the Gen and/or Dat; some Gen-assigning verbs not mentioned by Lass are categorized based on Bungenstab (1933), and additional insight is drawn from other sources. Comparisons to the semantic equivalents of these verbs in CSR are noted as well; these help confirm the connections between verb semantics and case assignment, since the Gen, Dat, Instr, and prepositional constructions are all productive in CSR, in addition to the Acc construction. Lass (1994:234-238) includes some semantic equivalents of these verbs in CL, in which the Gen, Dat, Abl (which absorbed the Instr), and prepositional constructions were also productive; I have noted several others as well.

5.2.2.1. Genitive

Lass lists two semantic classes of OE verbs that take Gen objects: verbs of perception, mental state, or desire; and verbs of deprivation or separation. He lists the following Gen verbs of perception, mental state, or desire: Gothic *gairnjan* 'long for', *lustōn* 'desire'; OE *gyrnan*, *wilnian* 'desire', *wundrian* 'admire', *hogian* 'intend'; OIc *kenna* 'know', *minnask* 'remember'; CL *cupīre* 'desire', *obliviscor* 'forget', *meminī* 'remember', *miror* 'admire'; AG *epithumeĩn* 'long for', *memnẽsthai* 'remember', *aisthánesthai* 'perceive' (1994:236-237). Of these, the Gen/Dat/Acc-assigning verb *wilnian* 'desire' and the Gen/Acc-assigning verb *wundrian* 'admire'

are investigated here.[146] Bungenstab categorizes more of the verbs investigated. In addition to these two verbs, which are listed as a verb of desire/expectation and a verb of emotion, respectively, the Gen-assigning verb *reccan* 'care (for)' and the Gen/Acc-assigning verb *hēdan* 'heed' are listed as verbs of perception/reflection (1933:27). The Gen/Acc-assigning verbs *ābīdan/onbīdan* 'await', (*ge)bīdan* 'await, experience', and *wēnan* 'expect, think' are listed as verbs of desire/expectation (ibid., 33).[147] Lass connects this class of verbs to the partitive function of the Gen: these verbs do not affect their objects completely. In fact, their objects are often completely unaffected by the action (1994:237). For example, the act of expecting has no influence on what is expected. When used in the sense of enjoyment, the Gen/Dat-assigning verb *gilpan* 'boast, rejoice' and the Gen/Dat/Acc-assigning verb *brūcan* 'use, enjoy' fit in this category as well.[148] The objects of these verbs are also unaffected by the action. For example, someone's enjoyment of something does not necessarily bring about any changes in the thing enjoyed.

The Gen-assigning verb *(be)weddian* 'promise, wed' can also be analyzed as involving a cognitive component.[149] If someone promises something, they then have a commitment to provide that thing, mainly in the mind of whoever they promised it to, but also in their own mind. The thing promised is not brought into existence or changed when the promise is made, but by an additional action performed later. Another verb involving mental activity more than mental state is *dēman* 'decide, judge', which this study categorized as taking Dat or Acc in OE but which does not fit semantically in any of the Dat classes. This verb has a similar relationship to its object as *(be)weddian* 'promise, wed', but the object only exists in the mind of the subject. Van Kemenade (1987:83) lists a number of these verbs as taking a Gen object that specifies what the mental state or experience is concerned with: *wilnian*, *gilpan*, *weddian*, *wundrian*, *reccan*.

[146] The CSR equivalents, *žaždat´* 'desire' and *ljubovat´sja* 'admire', assign the Gen and Instr, respectively (ORD, 2000; see also Jakobson 1936, noting separation or going away from as one of the meanings of the CSR case marking.).

[147] The semantic equivalents of these verbs in CSR, e.g., *ždat´* 'await, expect', takes Gen with indefinite or abstract objects—[-specific]—and Acc with definite or concrete objects—[+specific] (ORD, 2000).

[148] Bungenstab (1933:7, 38) lists *brūcan* 'use, enjoy' as a verb of use/experience and *gilpan* 'boast, rejoice' as a verb of emotion.

[149] The CSR equivalent, *obeščat´* 'promise', takes the Acc (ORD, 2000).

Although the connection is not as clear, he also includes *wealdan* in this class of verbs when it has a Gen object.

Lass lists the following verbs of deprivation or separation that assign Gen: Gothic *ga-hrainjan* 'cleanse', *ga-þarban* 'abstain'; OE *bedǣlan* 'deprive', *berȳpan* 'despoil'; OIc *sakna*, *missa* 'lose'; CL *leuāre* 'relieve', *dēsinere* 'cease from'; AG *apestéresthai* 'deprive' (1994:237). Of these, one verb and one cognate of a verb are investigated here: the Gen/Dat-assigning verb *bedǣlan* 'deprive' and the Gen-assigning verb *missan* 'miss', cognate to OIc *missa*. Bungenstab lists these as a verb of separation and a verb of lacking, respectively (1933:44, 51). Two synonyms of *bedǣlan* are investigated here: *berēofan/berēafian* 'deprive of' is also Gen/Dat-assigning verb of separation, but *beniman* 'deprive of' is attested as taking Acc in addition to Gen and Dat.[150]

Bungenstab also includes the verb *behōfian* 'require' as a verb of lacking (1933:51).[151] Lass connects this class of verbs to the use of Gen to express removal or separation. This Gen function is similar to the Abl of separation: the Gen can mark origin, as in *England's greatest poet*, and origin is equivalent to "extraction of X out of Y" (1994:233, 237). As mentioned in section 5.2.1 above, the argument structure for *behōfian* with the thing required in the Gen is not investigated here due to a lack of attestations in the later texts, but its occurrence in Beowulf is consistent with the semantics. The Gen/Dat/Acc-assigning verb *weorpan* 'throw' may also fit this category in that it causes separation between the agent and the object, or the Gen may be used to mark a partitive object in contrast to a non-partitive Acc object. However, this verb can also take an object marked for Dat to indicate instrumentality, and Mitchell (1985:562-563) considers the Gen object of this verb to be an instrumental Gen, although he does not adequately justify this categorization.

A number of other verbs that could take a Gen object do not fit either of Lass's categories: the Gen/Dat/Acc-assigning verb *(ge)fandian* 'search out, experience' and the Gen/Acc inanimate objects of the verbs *biddan* 'ask for' and (*ge)unnan* 'grant', as well as the

[150] As mentioned above, the Gen/Dat object for the *bedǣlan*, *berēofan/berēafian* as well as the Gen/Dat/Acc object for *beniman* that mean 'deprive of' is the thing that is taken. The equivalent verbs of deprivation take Gen objects for the thing taken in CSR, e.g., *lišat´* 'deprive, bereave of' (ORD, 2000).

[151] The CSR equivalent, *trebovat´* 'require', takes the Gen (ORD, 2000).

Gen/Dat/Acc inanimate object of *tilian* 'provide'.[152] Bungenstab assigns these to four different classes (1933:7, 14, 19, 24). These are not verbs of mental state, except perhaps *(ge)fandian*, which overlaps in meaning with (*ge)bīdan* 'await, experience'. However, they are similar in that they do not affect their object. In fact, the object of some of these verbs is not even required to have an existing referent. Therefore, the unaffectedness of the object is a better explanation for the situation in OE. If the object is affected, as it may be with verbs like *tilian* 'provide' and (*ge)unnan* 'grant', it is only partially affected, connecting these verbs to the partitive meaning of the Gen.

5.2.2.2. Dative

As a result of syncretism among IE Dat, Abl, Loc, and Instr case markings, the OE Dat is assigned to the objects of a number of semantic verb classes: impersonal verbs, verbs of separation, and what Lass describes as verbs of serving, confiding, or trusting (Lass 1994:237-238). For impersonal verbs, the Dat object functions semantically as an experiencer. Lass lists the following impersonal verbs with a Dat experiencer: OE *þyncan* 'seem', *mislimpan* 'fail', *gespōwan* 'succeed', CL *persuadēre* 'persuade', AG *deĩ* 'have need' (ibid., 238). Of these, only the Dat-assigning verb *þyncan* 'seem' is investigated here, but the animate Dat objects of *ofþyncan* 'displease' and *behōfian* 'be required of' are also experiencers. Lass traces this use to the possible origin of the IE Dat as an alternative form of the Loc used with personal nouns: like the Dat of possession, the Dat with impersonal verbs indicates location with respect to the experiencer (ibid., 237-8). However, evidence that the Dat experiencer has a Loc origin is lacking. Vakareliyska's proposal of *cognitive engagement* as the core semantic feature of the Dat with verbs provides a better account of the Dat experiencer, in that the experiencer is affected cognitively by the action but not physically (1994:15). This account works better because *cognitive engagement* is much more important than location in defining the semantic relationship between the action and the experiencer; it is also more economical, since *cognitive engagement* is already assumed for Dat objects of other verbs. Grimm (2011:22) demonstrates that

[152] Two CSR equivalents, *iskatʹ* 'search for', *prositʹ* 'ask for', take Gen with indefinite or abstract objects—[-specific]—and Acc with definite or concrete objects—[+specific]. However, *žalovatʹ* 'grant' consistently takes the Acc (ORD, 2000).

experiencers have the same properties as recipients, i.e., sentience and qualitative change, providing further support for the importance of cognitive involvement for both of these uses. Lass lists the following verbs of serving, confiding, or trusting as assigning Dat: OE *þegnian* 'serve', *betǣcan* 'entrust'; CL *seruīre* 'serve', *fidere*, *confidere* 'trust', *nitor* 'rely on' (1994:238).[153] Neither of the OE verbs listed is among those investigated here, but there are a few verbs from my study that clearly fit into this class: the Gen/Dat-assigning verb *helpan* 'help'[154] and the Dat-assigning verb *beorgan* 'protect'. In addition, the animate objects of *tilian* 'provide for', which could be Gen, Dat, or Acc in OE, and *(ge)unnan* 'grant', which could only be Dat, indicate someone who is served. Lass does not discuss how this class might be connected to one of the functions of the Dat or the morphological cases it merged with, but *cognitive engagement* accounts for many of these verbs that assign Dat: the person helped must be aware of the help and consider it to be help for the action to succeed. However, *beorgan* and certain other verbs of serving are exceptional in that they do not require *cognitive engagement*.[155] Van Kemenade links the use of Dat with these verbs to the Dat referent's role as an experiencer, including with *beorgan* and *gehelpan* 'help' (1987:80). However, as with experiencers of impersonal verbs, this amounts to the same thing as *cognitive engagement* and does not account for the exceptional verbs of serving. Grimm (2011:23) provides support for the requirement of *cognitive engagement* for most of these verbs but also fails to account for the exceptions, which have Dat objects that lack another expected property: they may be completely unaffected by the action. For these verbs, the peripheral status of the Dat—the case that absorbed the Loc, Abl, and Instr, all peripheral cases associated more strongly with adjuncts than arguments—may be the best explanation for the Dat: the action is not directed at the Dat referents, but elsewhere, for their benefit.

Lass lists the following verbs of separation as assigning Dat: OE *ætwindan* 'escape' and *linnan* 'cease' (1994:238). Neither of these verbs is found in the texts under investigation here. However, the fact that these verbs of separation assign Dat, while others assign Gen, indicates a

[153] The CSR verbs *služit´* 'serve' and *doverjat´* 'confide in, trust' both assign the Dat as well (ORD, 2000). Lass (1994) does not mention that *betǣcan* 'entrust' assigns Dat to the person entrusted, but Acc for the thing entrusted, making it similar to a canonical ditransitive verb like *ġiefan* 'give' (see Mitchell 1985:462).

[154] The CSR equivalent, *pomogat´* 'help', takes the Dat (ORD, 2000).

[155] The CSR equivalent, *zaščiščat´* 'protect', assigns the Acc (ORD, 2000).

semantic overlap. This may explain why both cases are possible with the synonyms *bedǣlan*, *beniman*, and *berēofan/berēafian*, all meaning 'deprive of'. The Dat/Acc-assigning verb *forlēosan* 'lose' also involves separation. Lass connects the use of Dat to objects that were originally Abl (ibid., 238). The only semantic commonality between separation of the referent, as originally indicated by the Abl, and other uses of the Dat is the peripheral status of the referent. Semantic overlap accounts for the use of both Dat and Gen with verbs of separation.

The Gen/Dat/Acc-assigning verb (*ge)wealdan* 'control, possess' does not fit any of these classes well but may take Dat some of the time because it originally took Instr. Another possibility is that *wealdan* has a Dat experiencer (Van Kemenade 1987:80) like some of the verbs discussed above. The Gen/Dat/Acc-assigning verbs *weorpan* 'throw' and *brūcan* 'use, enjoy', as well as the Gen/Dat-assigning verb *gilpan* 'boast, rejoice', most likely also assigned Instr originally, since the thing being thrown, used, or boasted about is a kind of instrument for the enjoyment or rejoicing of the subject referent. In the meaning of all four of these verbs, the agent causes the referent of the Dat to act as an instrument on behalf of the agent and under the agent's control.[156] The use of the Dat as well as Gen and Acc for the object of *(ge)fandian* when it means 'experience' and the inanimate object of *tilian* 'provide' could also derive from an instrumental function: these are the objects that the agent uses to have an experience or to provide for the animate Dat referent, respectively. The fact that the Dat is not an option when *(ge)fandian* means 'search out' supports this explanation, since the thing that is sought after cannot be considered an instrument used in the process of searching (see "fandian" 5, DOE, 2021).

5.2.3. Analysis

Tables of results and discussion for each text are given in subsections 5.2.3.1- 5.2.3.3. The columns of the tables represent the OE case-assignment categories for the verbs investigated in that text. The rows are the case markings of which the forms of the objects might be an instance.

[156] The semantic equivalents of most these verbs in CSR take the Instr: *vladetʹ* 'control, possess', *polʹzovatʹsja* 'use, enjoy', and *xvastatʹ* 'boast' (ORD, 2000). The CSR verb *švyrjatʹ* 'throw' assigns either the Instr or the Acc depending on whether the object has an incidental or a direct role in the sentence (Jakobson 1936/1984:79). This same distinction could also account for at least some of the variation in the case marking assigned by OE *weorpan* 'throw'. One CL equivalent, *potior* 'possess, be master of', can take Gen, Abl, or Acc; another two, *fruor* 'enjoy' and *gaudēre* 'rejoice', can take the Abl or Acc. The CL verb *iacere* 'throw' only assigns the Acc (Lewis & Short, 1879).

When two case markings are listed, it means that the form of the object is syncretic between these two morphological cases. The label "Prep" means that a prepositional complement is used with the verb instead. Section 5.2.3.4 provides a table summarizing the most relevant case-assignment results across all of the texts. Section 5.2.3.5 discusses prepositional complements in more detail.

5.2.3.1. Beowulf

Table 101. Results for Beowulf

Beowulf	Historical Case-Assignment Category						
Observed Object Case	**+Gen**	**+Dat**	**+Gen/Dat**	**+Gen/Acc**	**+Dat/Acc**	**+Gen/Dat/Acc**	**Total**
Gen	100% (2)		42% (5)	39% (12)		30% (10)	29% (29)
Dat		69% (11)	50% (6)		80% (4)	33% (11)	32% (32)
Acc				13% (4)	20% (1)	9% (3)	8% (8)
Gen/Dat				13% (4)			4% (4)
Gen/Acc				3% (1)		9% (3)	4% (4)
Dat/Acc		25% (4)					4% (4)
Amb		6% (1)	8% (1)	32% (10)		18% (6)	18% (18)
Prep							
Total	100% (2)	100% (16)	100% (12)	100% (31)	100% (5)	100% (33)	100% (99)

All objects of all case-assignment categories of verbs in Beowulf at least had the possibility of being marked by one of the expected morphological cases, as determined based on Mitchell (1985:455-464). This is not surprising since Beowulf is representative of OE poetry, which tended to be relatively conservative. The verbs that assign Gen, i.e., *missan* 'miss' and *reccan* 'care for', each have one unambiguously Gen noun object, as in *wǣpna* 'weapons' in the following example:[157]

[157] Here and elsewhere, nominal forms that are described as unambiguously one morphological case cannot be interpreted as any other case.

(9) *þæt sē ǣglǣca for his wonhȳd-um*
that the.MASC.NOM.SG assailant.MASC.NOM.SG for his recklessness-FEM.DAT.PL
wǣpn-a ne recc-eð
weapon-NEUT.GEN.PL NEG care.for-PRES.3SG
'that the assailant does not care for weapons in his recklessness' (Beowulf, ll. 433-434)

For the verbs that assign Dat, i.e., *beorgan* 'protect', *ofþyncan* 'displease', (*ge)unnan* 'grant' (animate), and *þyncan* 'seem', 69% (11 instances) of the complements are unambiguously Dat. Most of the other objects are syncretic Dat/Acc forms. In the following example, *ealdre* 'life' is the Dat object of *burgan*, from *beorgan*:

(10) *ac hȳ on holt bug-on ealdr-e burg-an*
but they on forest.NEUT.ACC.SG sink.PST-PL life-NEUT.DAT.SG protect.PST-PL
'but they sank to the forest [and] protected their life' (Beowulf, ll. 2598-9)

However, the ratio of case markings for verbs that assign more than one marking to their object was not as obvious before the investigation was conducted. For the verbs that assign either Gen or Dat, i.e., *bedǣlan* 'deprive of', *berēofan/berēafian* 'deprive of', *gilpan* 'boast, rejoice', and *helpan* 'help', 42% (five instances) of objects are unambiguously Gen, 50% (six instances) are unambiguously Dat, and the rest could be interpreted as Gen or Dat. Thus, these verbs favor Dat over Gen slightly, at least in Beowulf. In the following sentence, *mæges* 'kinsman' is the Gen object of *helpan*:

(11) *ofer mín gemet mæg-es help-an*
beyond my measure.NEUT.ACC.SG kinsman-MASC.GEN.SG help-INF
'to aid my kinsman beyond my measure' (Beowulf, l. 2879)

For the verbs that assign Gen or Acc, i.e., *ābīdan/onbīdan* 'await', (*ge)bīdan* 'await, experience', *biddan* 'ask for' (inanimate), *hēdan* 'heed', and *wēnan* 'expect, think', 52% (16 instances) of objects can be interpreted as Gen but not Acc, 13% (4 instances) are unambiguously Acc, and the rest could be interpreted as Gen or Acc. Thus, these verbs favor

Gen over Acc, but examples with the Acc are still found. In the following example, *līcsār* 'body-pain' is the Acc object of *gebād*, from *gebīdan*:

(12) *lic-sar* *gebād* *atol* *ǣglǣca*
body-pain.NEUT.ACC.SG experience.PST.3SG awful.MASC.NOM.SG assailant.MASC.NOM.SG
'the awful assailant experienced body-pain' (Beowulf, ll. 815-6)

For the verbs that assign either Dat or Acc, i.e., *dēman* 'decide, judge' and *forlēosan* 'lose', four of objects are unambiguously Dat and one is unambiguously Acc, indicating that these verbs favor Dat over Acc. In the following example, *dōme* 'glory' is the Dat object of *forlēas*, from *forlēosan*:

(13) *þǣr* *hē* *dōm-e* *forlēas*
there he glory-MASC.DAT.SG lose.PST.3SG
'there he lost his glory' (Beowulf, l. 1470)

For the verbs that assign Gen, Dat, or Acc, i.e., *beniman* 'deprive of', *brūcan* 'use, enjoy', *(ge)fandian* 'search out, experience', *tilian* 'provide' (inanimate), (*ge)wealdan* 'control, possess', *weorpan* 'throw', and *wilnian* 'desire', 30% (10 instances) of objects are unambiguously Gen, 33% (11 instances) are unambiguously Dat, 8% (8 instances) are unambiguously Acc. Thus, the Gen and Dat have similar levels of productivity with these verbs, while the Acc is less common. In the following sentence, *Gēatum* 'Geats' is the Dat object of *wealdan*:

(14) *lēt* *ðone* *bregostōl* *Bīowulf* *heald-an* *Gēat-um*
let.PST.3SG the.MASC.ACC.SG throne.MASC.ACC.SG Beowulf hold-INF Geat-MASC.DAT.SG

weald-an
rule-INF
'he let Beowulf hold the throne to rule the Geats' (Beowulf, l. 2389)

Overall, verbs that take either Acc or another case marking tend to assign the Acc less frequently in Beowulf. Factoring out ambiguous forms, the Dat is preferred over the Acc more strongly than the Gen is, and the Dat is also preferred slightly over the Gen when those are the only two choices. While these results may seem unexpected, given that the Acc is the most common case marking on objects generally, the Gen and/or Dat are better fits semantically for verbs that can assign one or both of these case markings as well as the Acc. Verbs that can assign multiple case markings in OE often have distinct meanings associated with each case marking, as can be seen in the entries of these verbs in Mitchell (1985:455-464), the DOE, and the glossaries of Marsden (2010) and Klaeber (2008). For example, Klaeber (2008) lists *(ge)bīdan* as taking Gen when it means 'await' and Acc when it means 'experience'. Thus, the preferences for certain case markings over others when both are attested with a specific verb may be a result of how often the verb is used with each meaning in Beowulf.

5.2.3.2. Layamon's Brut

Table 102. Results for Brut

Layamon's Brut	**Historical Case-Assignment Category**						
Observed Object Case	**+Gen**	**+Dat**	**+Gen/Dat**	**+Gen/Acc**	**+Dat/Acc**	**+Gen/Dat/Acc**	**Total**
Gen			1% (1)	3% (2)	5% (1)	2% (4)	2% (8)
Dat	7% (1)	64% (59)	25% (24)	14% (11)	9% (2)	3% (6)	21% (103)
Acc	53% (8)	3% (3)	5% (5)	26% (20)	55% (12)	52% (95)	29% (143)
Gen/Dat		3% (3)	2% (1)	6% (5)		1% (1)	2% (10)
Gen/Acc			5% (5)	5% (4)	5% (1)	9% (16)	5% (26)
Dat/Acc	7% (1)	26% (24)	37% (36)	9% (7)	9% (2)	13% (24)	19% (94)
Amb	7% (1)	2% (2)	13% (13)	35% (27)		14% (25)	14% (68)
Prep	27% (4)	1% (1)	12% (12)	3% (2)	18% (4)	6% (10)	7% (33)
Total	100% (15)	100% (92)	100% (97)	100% (78)	100% (22)	100% (181)	100% (485)

Table 103. The Use of Gen vs. Non-Gen and Dat vs. Non-Dative in Beowulf and Brut

Observed	**+ Gen**	**− Gen**	**Total**	**+ Dat**	**− Dat**	**Total**
Beowulf	29	70	99	32	67	99
Brut	8	477	485	103	382	485
Total	37	547	584	135	449	584
($p < .00001$)				(p .017106)		

In Brut, almost all the expected case markings are productive with all case-assignment categories of verbs. The one exception is the lack of unambiguous Gen objects of Gen-assigning verbs. More generally, there is a strong tendency away from Gen objects and a moderate one away from Dat objects. There are also more substitutions of Gen for Dat and vice versa among the objects of verbs that only assigned one of these in OE, with Dat replacing Gen much more often than vice versa.

According to the chi-square test, the shift away from the Gen from Beowulf to Brut is extremely statistically significant ($p < .00001$). Similarly, the shift away from Dat is also statistically significant ($p = .017106$), but it is lower than the shift from the Gen. Therefore, this is evidence that the shift away from the Gen is more than the Dat. Of all complements, 29% are unambiguously Acc, which can be interpreted as a tendency to assign this particular morphological case. Brut and O&N are the only texts investigated to use a prepositional complement with any of the verbs in question. Brut does this for every category of verb and with a higher percentage overall, with prepositional phrases accounting for 7% (33 instances) of all verbal complements. See subsection 5.2.3.5.1 below for a more detailed discussion of prepositional complements in Brut.

For the verbs that historically assign Gen, i.e., *missan* 'miss', *reccan* 'care for', and *(be)weddian* 'promise, wed', no complements in Brut are unambiguously Gen, and only one can be interpreted as Gen; 27% (4 instances) are prepositional. The rest are Dat (one), Dat/Acc (one) or Acc (eight). While Beowulf only assigns the Gen for this category, Brut shows a preference away from Gen and prominent use of Acc. In the following example, *Ælene* 'Alene' is the possible Gen/Dat/Acc object of *biwedded*, the past participle of *(be)weddian*:

(15) *& Custance haue-d-e Ælene biwedd-ed to quen-e*
and Constance have.AUX-PST-3SG Alene.ACC/DAT/GEN marry-PP to queen-DAT.SG
'and Constance had taken Alene as his wife and queen' (Brut, l. 5504)

As a female name, *Ælene* could be any case form, but it is most likely Acc like the other six objects of *(be)weddian* in Brut.

For the verbs that historically call for Dat, i.e., *behōfian* 'require', *beorgan* 'protect', *ofþyncan* 'displease' (animate), (*ge)unnan* 'grant' (animate), and *þyncan* 'seem', 95% (88

instances) of complements can be interpreted as Dat; only one occurrence is prepositional. The remaining 3% (3 instances) are Acc. Therefore, like Beowulf, this case assignment category still prefers the use the historical Dat in Brut. In the following example, *us* 'us' is one of the Dat/Acc pronoun objects of *bihoueð*, from *behōfian*:

(16) *& ȝet hit us bihou-eð*
and yet it.NOM us behoove-PRES.3SG
'and yet it behooves us' (Brut, l. 8400)

For the Gen/Dat-assigning verbs, i.e., *bedælan* 'deprive of', *berēofan/berēafian* 'deprive of', *gilpan* 'boast, rejoice', and *helpan* 'help', one of complements in Brut is unambiguously Gen, 25% (24 instances) are Dat, 5% (5 instances) are Acc, and the remaining 12% (12 instances) are prepositional. The use of Acc objects in Brut is an innovation in this case-assignment category. Nevertheless, Brut still uses the historical case assignments; the Gen and Dat objects are definitely still productive with these verbs, but the Dat is clearly dominant and there are almost as many unambiguously Acc objects as possible Gen objects. In the following example, the noun *liues* 'life/lives' is a possible Gen object of the past participle *bidæled*, from *bedælan*:

(17) *seouen þusend þer lei-en liu-es bidæl-ed*
seven thousand there lie.PST-PL life-ACC.PL/GEN deprive-PP
'seven thousand lay there deprived of life/their lives' (Brut, l. 8668)

If it is singular, *liues* must be Gen, but if plural, it could be interpreted as Acc or Gen.[158] The form *liue(n)* is expected for Dat plural, so this is not considered as an option. In the following example, however, *þan liue* 'the life' is the Dat object of *biræfued*, the past participle of *berēofan/berēafian*:

[158] In Brut and O&N, the masculine/neuter Gen singular noun ending *-es* was used much more consistently than distinct Gen plural endings. For more details, see the noun paradigms in Appendix A.

(18) *Þer wes moni riche Brut; biræfu-ed þan liue*
there was many noble Briton; bereave-PP the.DAT.SG life
'There was many a noble Briton bereft of his life' (Brut, l. 7626)

For the Gen/Acc-assigning verbs, i.e., *ābīdan/onbīdan* 'await', (*ge)bīdan* 'await, experience', *biddan* 'ask for' (inanimate), *hēdan* 'heed', (*ge)unnan* 'grant' (inanimate), *wēnan* 'expect, think', and *wundrian* 'admire', two of complements are explicitly Gen, 14% (11 instances) are Dat, and 26% (20 instances) are Acc, two are prepositional, and 35% (27 instances) are ambiguous. The key difference in Brut is the decreased use of Gen objects than in Beowulf and the overall preference for the Acc. This is expected as the use of the Gen declined considerably in Brut. In the following example, *him* 'him' is the Dat pronoun object of the infinitive *ibide*, from *gebīdan*:

(19) *no in nauer na-re burȝ-e ne dur-st-e he him ibid-e*
NEG in never NEG-FEM.DAT.SG city-DAT.SG NEG dare-PST-3SG he.NOM him.DAT await-INF
'he did not dare await him in any city ever' (Brut, l. 15414)

For comparison, one of the nouns that is ambiguous between Gen and Acc (but most likely Gen) is provided in example (5) in section 3.1.2.6 above, reproduced here for convenience:

(5) *þere læi þa uerde. þeos weder-es abid-en*
there lay the army. the.GEN.SG weather-GEN.SG wait-3PL
'there lay the army waiting for good weather' (Brut, II. 14093–4)

Therefore, the Gen is still productive with these verbs, but not to the same degree as Acc, and even Dat.

For the verbs that traditionally assign Dat or Acc in OE, i.e., *dēman* 'decide, judge' and *forlēosan* 'lose', 9% (two instances) of the complements in the data are unambiguously Dat, 55% (12 instances) are unambiguously Acc, and one is Gen, 18% (four instances) are prepositional. There is a clear preference for Acc over Dat, whereas in Beowulf, there is a preference for Dat over Acc. In the following example, *alle uolke* 'all folk' is the Dat object of *demeð*, from *dēman*:

(20) *þenne ure Drihte dem-eð all-e uolk-e*
when our Lord judge-PRES.3SG all-DAT.SG folk-DAT.SG
'when our Lord judges all people'. (Brut, l. 11507)

For the Gen/Dat/Acc-assigning verbs, i.e., *beniman* 'deprive of', *brūcan* 'use, enjoy', *(ge)fandian* 'search out, experience', *tilian* 'provide' (inanimate), (*ge)wealdan* 'control, possess', *weorpan* 'throw', and *wilnian* 'desire', in contrast to Beowulf, the complements of Gen (one) and Dat (3%, six instances) are significantly reduced and there is an increased use of the Acc (52%, 95 instances). Beowulf does not use prepositional complements for this case assignment category but Brut does by 6% (10 instances) of the complements. In the following example, *þan leoden* 'the peoples' is the object of *fondede*, from *fandian*:

(21) *he fer-d-e feor ȝeond moni lon[d][159] and fond-ed-e þan leod-en*
he go-PST-3SG far beyond many lands and seek-PST-3SG the.DAT.PL people-PL
'he went far beyond many lands and sought their peoples' (Brut, l. 3286)

Thus, the Acc is the most productive case with these verbs, followed by the Dat and prepositions.

While nouns and pronouns both occur in unambiguously Gen forms at about the same low frequency, pronouns are much more likely than nouns to be unambiguously Dat. Part of this is due to the fact that quite a few nouns decline with *-e* in the Acc singular, so there is syncretism with the Dat (and Gen, for feminines). In addition, for nouns that are expected to have distinct Acc and Dat forms, the forms without *-e* sometimes occur in clearly Dat contexts. If the Dat *-e* was becoming optional, then that is another source of Dat/Acc syncretism. First and second-person pronouns also have Dat/Acc syncretism, but the third-person pronouns are more common and have distinct Dat forms.

Allen notes that the originally-Dat *him* 'him' is sometimes used where Acc *hine* 'him' would be expected but finds only two examples of *hine* used where *him* would be expected. She attributes this to the scribes' attempts to copy faithfully a language they did not fully command (1995:198-199). Van Gelderen takes this to mean that the scribes distinguished the Acc and Dat

[159] Square brackets in an edited text enclose one or more hypothetical letters that the editor has determined to have been lost through physical damage.

semantically but were losing the distinction morphologically in the third-person pronouns (2000:212). This distinction was already being lost in the first and second-person pronouns in OE, where the specially marked Acc forms, *mec* 'me' and *þec* 'you', were less frequent in relation to the Dat forms, *me* 'me' and *þe* 'you', than in the third-person pronouns (ibid., 244). Even in Beowulf, the first and second-person Dat forms were used in environments where Acc was expected, but this did not occur for the third-person until the EME texts. Van Gelderen cites (using different terminology) the greater functional narrowing of Acc forms in favor of Dat forms on first and second-person pronouns in comparison to third-person pronouns as an explanation for this split between the first/second-person and third-person, both in terms of the merger of Acc and Dat and behavior with impersonal, passive, and ergative constructions (ibid., 245-246). If the distinction between Acc and Dat was also being lost in third-person pronouns, it suggests the Acc forms were undergoing functional narrowing for these paradigms as well. Analogy may have also played a role in ending the person split that existed in OE and EME.

The lower productivity of the Gen relative to the Dat and of both relative to the Acc has a number of potential explanations. For pronouns, substitutions of Dat for Gen can be attributed to the general replacement of the Acc with Dat forms, after the Acc had replaced the Gen with particular verbs. But there are still some substitutions of case markings in nouns, and substitutions of Gen for Dat in pronouns. It seems that the case markings assigned by certain verbs have changed from OE. There may also have been dialectal differences in OE that carried through to Brut. However, in the instances when case marking really did change between OE and Brut, a usage-based approach accounts for these changes in case-assignment categories due to differing type frequencies of constructions with semantic similarities. Regardless of the cause of unexpected case marking, the Dat and Gen do not appear to be merging completely in their use, since there are still significant differences in case distribution among different case-assignment categories of verbs.

5.2.3.3. The Owl and the Nightingale (O&N)

Table 104. Results for O&N

The Owl and the Nightingale	Historical Case-Assignment Category						
Observed Object Case	**+Gen**	**+Dat**	**+Gen/Dat**	**+Gen/Acc**	**+Dat/Acc**	**+Gen/Dat/Acc**	**Total**
Gen						12% (1)	2% (1)
Dat		13% (1)	33% (3)				7% (4)
Acc	50% (4)		11% (1)	40% (4)	50% (8)	75% (6)	39% (23)
Gen/Dat			33% (3)				5% (3)
Gen/Acc				10% (1)			2% (1)
Dat/Acc	25% (2)	75% (6)	11% (1)	10% (1)	25% (4)	12% (1)	25% (15)
Amb		13% (1)		40% (4)	25% (4)		15% (9)
Prep	25% (2)		11% (1)				5% (3)
Total	100% (8)	100% (8)	100% (9)	100% (10)	100% (16)	100% (8)	100% (58)

Table 105. The Use of Genitive vs. Non-Genitive and Dative vs. Non-Dative in Beowulf and O&N

Observed	**+ Gen**	**− Gen**	**Total**	**+ Dat**	**− Dat**	**Total**
Beowulf	29	70	99	32	67	99
O&N	1	58	59	4	55	59
Total	30	128	158	36	122	158
($p < .00001$)				(p .0002)		

For the most part, Gen and Dat are no longer productive in O&N, but these case markings have not been completely lost. Counting only the unambiguous assignment of Gen and Dat versus others in Beowulf and the O&N to yield the statistical significance by using Fisher exact test. The results show that, as in Brut, the shift away from the Gen from Beowulf to O&N is also extremely statistically significant according ($p < .00001$). Likewise, the shift away from Dat is also highly statistically significant (p=.0002), but it is lower than the shift from the Gen. This shift from Dat is higher in O&N than Brut. Therefore, the tendency away from Gen and Dat is another piece of evidence in both Brut and O&N. There is an increased use of the Acc (39%), which is even higher than Brut (29%), which indicate that O&N is advanced in breakdown of morphological case than Brut.

Three verbs under investigation in two case-assignment categories each have one prepositional complement, with prepositional phrases accounting for three of all verbal complements. For a more detailed discussion of prepositional complements in O&N, see section 5.2.3.5.2 below. For the verbs expected to assign the Gen, i.e., *missan* ‘miss’ and *reccan* ‘care

for', it is not even possible to interpret any complements as Gen; two are prepositional. The rest are Acc (four) or Dat/Acc (two). Thus, all the objects might be instances of Acc, expanding the use of Acc as in Brut. The following is an example for this category:

(22) *ne recche ich noȝt of winteres reue*
NEG care I nothing.ACC of winter deprivations
'I care nothing for the deprivations of winters' (O&N, l. 458)

For the one Dat-assigning verb found in O&N, i.e., *þyncan* 'seem', the forms of its experiencers could all be interpreted as Dat. However, none of the nouns and only one pronoun is unambiguously Dat, i.e., *heom* 'them':

(23) *forþan heom þuh-t-e þat heo had-d-e*
because them.DAT seem-PST-3SG that she have-PST-3SG
'because it seemed to them that she had' (O&N, l. 1661)

In this respect, seven out of eight of these experiencers could actually be Acc, and possibly all of them, since *heom* is a third-person pronoun, and these may be undergoing the same merger of the Dat and Acc as in the other texts.

For the verbs that historically assign Gen or Dat, i.e., *gilpan* 'boast, rejoice', and *helpan* 'help', none of the complements in the data can be interpreted as Gen but not Dat, 33% (three instances) are Dat, and 30% (three instances) are Gen/Dat; one is prepositional, and the remaining one is Acc. For the verb *helpan,* 'help', 77% (7 instances) of objects were clearly not Acc. As in Brut, this assignment category still uses the historical morphological cases with novel uses of Acc and prepositions. In the following sentence, *hom* 'him' is the Dat object of *helpe*, from *helpan*:

(24) *ich hom helpe*
I they.DAT.PL help
'I help them' (O&N, l. 484)

For the Gen/Acc-assigning verbs, *ābīdan/onbīdan* ‘await’, *biddan* ‘ask for’ (inanimate), *hēdan* ‘heed’, and *wēnan* ‘expect, think’, 50% (five instances) of complements can be interpreted as Acc but not Gen, and the other 50% (five instances) can be interpreted as either Gen or Acc. This indicates that all complements can be interpreted as Acc. Thus, like Brut, the Acc is the dominant utilization in this case assignment category. In the following example, *rem* ‘hue and cry’ is the Acc object of *abide*, from *ābīdan*:

(25) *Ȝef eni mon schal rem abide*
if any man shall cry.ACC.SG await
‘if anyone awaits a hue and cry’ (O&N, l. 1215)

For the verbs expected to assign Dat or Acc, i.e., *dēman* ‘decide, judge’ and *forlēosan* ‘lose’, 50% (8 instances) of the complements can be interpreted as either Dat or Acc, and the other 50% (8 instances) are unambiguously Acc. Thus, every complement for these two case-assignment categories could be interpreted as Acc. Like Brut, the Acc is prevalent in this category. In the following example, *his hou* is the Acc object of *uorlost,* from *forlēosan*:

(26) *an his hou neuer ne uorlos-t*
and its color.ACC.NEUT.SG never NEG lose-PRES.3SG
‘and never loses its color’ (O&N, l. 619)

For the Gen/Dat/Acc-assigning verbs, i.e., *beniman* ‘deprive of’, *(ge)fandian* ‘search out, experience’, (*ge)wealdan* ‘control, possess’, and *weorpan* ‘throw’, 75% (6 instances) of complements are unambiguously Acc; of the two remaining complements, one (12%) is a Dat/Acc noun, and the other is the Gen pronoun *þas* ‘that’, the only unambiguously Gen object found in this text, used with *fondi*, from *fandian*:

(27) *forþi hit þoh-t-e fondi [þ]as*
therefore it think-PST-3SG find that.GEN.SG
‘Therefore she set out to try it’ (O&N, l. 1443)

Like *heom/hom* ‘them’ in examples (23) and (24) above, most of the unambiguously Dat forms in O&N are third-person pronouns, so these alone are not sufficient evidence for the Dat remaining productive. However, there is one occurrence of the verb *helpan* ‘help’ with an unambiguously Dat non-pronominal object, *þan sunfulle* ‘the sinful [man]’:

(28) *an þan sunfulle ich helpe alswo*
and the.DAT.SG sinful I help also
‘and I also help the sinful [man]’ (O&N, l. 891)

This verb also occurs with two noun objects that could be interpreted as either Dat or Gen, i.e., *monne* ‘men’ in example (29) and *þare niȝtegale* ‘the Nightingale’ in example (30):

(29) *Ich helpe monne on eiþer halue*
I help man.GEN/DAT.PL on either half
‘I help men on either side’ (O&N, l. 887)

(30) *to helpe þare niȝtegale*
to help the.GEN/DAT.SG Nightingale
‘to help the Nightingale’ (O&N, l. 1719)

These three examples are unlikely to be affected by a process like the third-person pronoun Dat-Acc merger. Therefore, the Dat and possibly also Gen are still marginally productive, at least with *helpan*.

5.2.3.4. Case-Assignment Summary by Text

Table 106 is a summary of the most relevant case marking results by case-assignment category across all of the texts. The percentages are out of the total number of nominal and prepositional complements, even though the percentages for prepositional complements are not included in this table. Only the relevant case marking interpretations are included. For example, “unambiguous Gen” refers to objects that can only be interpreted as Gen. “Possibly Gen but not Dat” refers to objects that can be interpreted as Gen and possibly also as Acc, but that cannot be

interpreted as Dat. “Possibly both” refers to objects that can be interpreted as either of the expected cases and possibly also as the unexpected case, i.e., they include “ambiguous” forms.

Table 106. Case-Assignment Summary by Text

Case-Assignment Category	Object Case	Beowulf	Brut	O&N
+Gen	Unambiguous Gen	100% (2)	0% (0)	0% (0)
	Possible Gen	100% (2)	7% (1)	0% (0)
+Dat	Unambiguous Dat	50% (6)	64% (59)	13% (1)
	Possible Dat	58% (7)	96% (88)	100% (8)
+Gen/Dat	Unambiguous Gen	38% (5)	1% (1)	0% (0)
	Unambiguous Dat	54% (7)	25% (24)	30% (3)
	Possible Gen but not Dat	38% (5)	6% (6)	0% (0)
	Possible Dat but not Gen	54% (7)	62% (60)	44% (4)
	Possibly both	8% (1)	15% (14)	8% (1)
+Gen/Acc	Unambiguous Gen	39% (12)	3% (2)	0% (0)
	Unambiguous Acc	13% (4)	26% (20)	40% (4)
	Possible Gen but not Acc	52% (16)	9% (7)	0% (0)
	Possible Acc but not Gen	13% (4)	35% (27)	50% (5)
	Possibly both	35% (11)	40% (31)	50% (5)
+Dat/Acc	Unambiguous Dat	80% (4)	9% (2)	0% (0)
	Unambiguous Acc	20% (1)	55% (12)	50% (8)
	Possible Dat but not Acc	80% (4)	9% (2)	0% (0)
	Possible Acc but not Dat	20% (1)	59% (13)	50% (8)
	Possibly both	0% (0)	9% (2)	50% (8)
+Gen/Dat/Acc	Unambiguous Gen	30% (10)	2% (4)	12% (1)
	Unambiguous Dat	33% (11)	3% (6)	0% (0)
	Unambiguous Acc	9% (3)	52% (95)	75% (6)
	Possible Gen	57% (19)	41% (46)	12% (1)
	Possible Dat	51% (17)	31% (56)	12% (1)
	Possible Acc	36% (12)	88% (160)	75% (6)

Brut represents a stage in the decline of morphological case marking. All of the expected cases are still available for all case-assignment categories of verbs to assign, so case marking has not been lost, but there is a strong tendency away from Gen objects and a moderate one away from Dat objects. The O&N is further along in the decline of morphological case marking than Brut: the Gen and Dat are generally not available for verbs to assign, but these case markings have not been completely lost. In these two texts, prepositional complements serve as another alternative to the assignment of case marking based on verb semantics. Brut shows a higher frequency of

prepositional complements and a more even distribution across case-assignment categories, although there are still many verbs which never take prepositional complements.

5.2.3.5. Prepositional Complements

In OE, prepositional complements already alternated with the Gen and Dat objects of many verbs. Prepositions have narrower meanings than case markers, but their meanings often fall within the general meaning of a particular morphological case marking, making them suitable alternatives. No prepositional complements were found for the verbs under investigation in Beowulf*,* but competition between these two systems for expressing meaning can be observed in other OE texts. The results of my analysis suggest that the use of prepositional complements expanded at the expense of the Gen and Dat in EME, at least in some texts. As mentioned in sections 5.2.3.2 and 5.2.3.3, Brut and O&N are two such texts: both continue the use of prepositional complements observed in OE—often with different prepositions, however—and they also occasionally have a prepositional complement with verbs that are not known to have had this option in OE. Brut has prepositional complements for every category of verb, and prepositional phrases account for 7% (33 instances) of all verbal complements. In the following sentence, the prepositional phrase *at live* 'of life' is a complement of the verb form *birœueden*, the past tense of *berēofan/berēafian* 'deprive of':

(31) *heo mid cnif-en birœu-ed-en heom at liu-e*
they with knife-PL deprive-PST-PL them.DAT of life-DAT.SG
'they deprived them of life with knives' (Brut, l. 7679)

O&N uses prepositional complements for two case-assignment categories of verbs, accounting for 5% (3 instances) of all verbal complements. In the following sentence, the prepositional phrase *of þine songe* 'about your song' is a complement of the verb *ȝelpst*, from *gilpan* 'boast, rejoice':

(32) *Ac ȝet þu ȝelp-st of þine song-e*
but yet you.NOM.SG boast-PRES.2SG about your song-DAT.SG
'Furthermore, you boast about your song' (O&N, l. 971)

Certain patterns are apparent with respect to the verbs that take prepositional complements and the specific prepositions that are used. Even in these two texts, most verbs never take prepositional complements at all. Those that do are limited to one or two different prepositions. In Brut, seven different prepositions are used overall. The following sections address the use of prepositional complements in these two texts, including the semantic motivations for the choice of preposition with each verb. For each text, a table listing the number of times each preposition appears with each verb is provided. Verbs are grouped by case-assignment category.

5.2.3.5.1 Layamon's Brut

Table 107. Prepositional Complements in Brut

Case assignment Category	**Prepositions**							
+Gen	*after*	*at*	*for*	*of*	*on*	*over*	*to*	**Total**
reccan				3	1			4
+Dat								
behōfian							1	1
+Gen/Dat								
bedǣlan				2				2
beniman		2						2
berēofan/berēafian		4		3				7
gilpan			2	1				3
+Gen/Acc								
wundrian					2			2
+Dat/Acc								
dēman				4				4
+Gen/Dat/Acc								
(ge)fandian				1	1			2
tilian (inanimate)					1			1
(ge)wealdan						2	1	3
wilnian	1			1				2
Total	1	6	2	15	5	2	2	33

For the verbs historically assign the Gen, four of the complements are prepositional. Of the three verbs in this category, only *reccan* 'care for' has prepositional complements. Three of its four

prepositional complements contain the preposition *of* and one potential prepositional complement contains *on*. Prepositional complements already occurred with *reccan* in OE, usually with the preposition *be* (Mitchell 1985:461). With *reccan* and other verbs of mental state, *be* carried the meaning 'about, concerning' in OE ("be, big" I.D.1.c., DOE, 2021). In Brut, *of* is probably used similarly, meaning 'about, concerning' with a verb of mental attitude ("of" 23.a.(d), MED, 2021).[160] In the following example, the preposition *of* appears to be used in this meaning as the complement of *rohten*, the past tense of *reccan*:

(33) *ȝif heo grið soh-t-en; & of his freond-scipe roh-t-en*
if they peace seek-PST-PL and about his friendship care-PST-PL
'if they sought peace, and cared about his friendship' (Brut, l. 6889)

The preposition *on* was also used with *reccan* in OE to denote the person charged with something (Mitchell 1985:461). However, *reccan* does not appear to be used with the meaning 'charge' in Brut. In fact, the preposition *on* more likely introduces a locational adjunct in this sentence:

(34) *Þat we com-en to lond-e ne recche we on wulche leoden*
that we come-PL to land-DAT NEG care we on which nation
'That we come to land; we do not care on which nation' (Brut, l. 2297)

However, the meaning 'about, concerning' was also possible for *on* ("on" 21.a, MED, 2021), like *be* in OE and *of* in the other prepositional complements of this verb in Brut. In addition, *on* was also used alongside the Gen in OE with other verbs of mental attitude (Bungenstab 1933:58). Therefore, the use in example (33) could theoretically be a complement as well; this just seems unlikely given the context of the sentence. The preposition *of* indicates the focus of the mental attitude denoted by *reccan*, and *on* possibly does as well. Thus, these are reasonable replacements for a Gen object. For *of*, there may also be a parallel with the substitution of *of*-phrases for possessive and partitive Gen constructions, especially since *of* was not one of the

[160] The preposition *o* 'about' and a Loc object is required to complete the meaning of the CSR equivalent *zabotit'sja* 'care for' (ORD, 2000).

prepositions used with *reccan* in OE. However, since this replacement is not uniform for these Gen-assigning verbs, that alone is not enough of a justification.

For the verbs that historically call for Gen or Acc, only two of the complements are prepositional. Of the seven verbs in this category, only *wundrian* 'admire' has prepositional complements. In OE, prepositional complements with *æfter*, *be*, *fram*, *on*, and *ymb* occurred with this verb (Mitchell 1985:464). The preposition *on* is used for its two complements in Brut, and the MED translates this combination as 'wonder at' ("on" 21.b(h), MED, 2021). This reflects the same force that *on* seems to carry with other verbs of mental state. However, both examples present issues. In the first, *on* quite clearly introduces the complement of *wundreden*, from *wundrian*, but *on* itself takes a clause introduced by *wheonene* 'whence, from where' rather than a noun as its object:

(35) *þet al folc wundr-ed-en on wheonene com swa feir mon*
that all folk wonder-PST-PL about whence come.PST.3SG so fair man
'that everyone wondered about where so fair a man came from' (Brut, l. 3515)

Thus, this cannot be considered a replacement of the Gen, but it still complements the meaning of *wundrian* in the same way as other verbs of mental state. In the second example, *on* quite likely introduces a locational adjunct rather than a complement, because the clause introduced by *what* 'what' does not fit in the sentence except as the complement of *awundred*, the past tense of a prefixed form of *wundrian*, and that verb could only take one complement:

(36) *Þa wes he awundr-ed; on þissere wurlde-riche what weor-e*
then was he wonder-PP on this.FEM.DAT.SG world-kingdom what be.PST.SUBJ-SG

þis tacninge; þa he isæh þer on grund-e
this.NOM.SG omen REL he see.PST.3SG there on ground-DAT
'then he was astonished on this material world what this omen might be, that he saw there on the ground' (Brut, ll. 7972-3)

Thus, prepositions were at least sometimes used with the Gen/Acc-assigning verb *wundrian* to complement its meaning as with other verbs of mental state that assigned the Gen, but the prepositional alternatives were already being used in OE when the Gen was still productive.

For the verbs that traditionally assign Gen or Dat, 14 of the complements are prepositional. Four of the five verbs in this category have prepositional complements; only *helpan* 'help' does not have any. Of the 14 prepositional complements, six contain the preposition *at*, six *of*, and two *for*. The two complements with the preposition *for* and one of the complements with *of* are the only complements found with *gilpan* 'boast, rejoice' in Brut, making it the only verb examined there with no noun or pronoun objects. The preposition *for* was already used to complement verbs of rejoicing in OE with the meaning 'at, over, on account of' ("for" A.14.c., DOE, 2021), including *gilpan* (Mitchell 1985:458). In Brut, both prepositions *of* and *for* are probably used with the sense 'about, concerning' ("for" 14, "yelpen" 1(b), MED, 2021), as with other verbs of mental state. Despite slightly different glosses, this sense essentially continues that of *for* with *gilpan* in OE. In the following sentence, the preposition *for* appears to have this sense as the complement of *ȝelpen*, from *gilpan*:

(37) *ne mihten þer nan ȝelp-en; for oðere kunn-es scenche-n*
NEG might there none boast-INF about other kind-GEN.SG drink-PL
'might none there boast about drinks of another kind' (Brut, l. 11453)

The meaning of the construction does not seem to change in the sentence that has *of* instead:

(38) *Þis wes þat ilke bord; þat Brutt-es of ȝelp-eð*
this.NOM.SG was the.NEUT.NOM.SG same board REL Briton-PL about boast-PRES.PL
'This was the same board that Britons boast about' (Brut, l. 11455)

More generally, it is possible that *for* replaces more Dat senses and *of* more Gen senses, given the uses of these prepositions in later stages of English (see, for example, Lightfoot 1999:121). However, it seems that these prepositions are used without any difference in meaning in Brut, just as there was no apparent difference in meaning between the use of Gen and Dat with this verb in Beowulf.

The other prepositional complements occur with the three synonyms meaning ‘deprive of’. Two of these took prepositional complements in OE: *be*, *fram*, and *of* with *bedǣlan* (Mitchell 1985:456, “be-dǣlan” DOE, 2021) and *æt*, *fram*, and *on* with *berēafian* (Mitchell 1985:460, “be-rēafian” DOE, 2021). The preposition *æt* had the meaning ‘from’ with verbs of deprivation, including the verb *(ge)niman*, which is related to *beniman* (“æt” I.D.4., I.G.1.e., DOE, 2021); the other prepositions carried the same basic meaning. In Brut, the preposition *at* is used both times *beniman* has a prepositional complement and four times with *berēafian* as well. In all but one of these instances, the object of *at* is *live* ‘life’, as in the following sentence with *benimen*, from *beniman*:

(39) *þat he me n-ol-d-e vt driu-en. binim-en me æt þan liu-e*
that he me NEG-want-PST-3SG out drive-INF deprive-INF me at the.DAT.SG life-DAT.SG
‘that he didn’t want to drive me out, or deprive me of my life’(Brut, l. 4389)

The other has the object *þan hefde* ‘the head’, which is used to convey essentially the same idea, i.e., execution:

(40) *elc-ne Cristine mon; at þan hefd-e he binom*
each-MASC.ACC.SG Christian man of the.DAT.SG head-DAT.SG he deprive.PST
‘he deprived each Christian man of his head’ (Brut, l. 5436)

These instances of *at* likely continue the sense of *æt* with the same or related verbs in OE. However, Bungenstab argues that the preposition *at* appears with verbs of deprivation as a replacement for a Gen use that he connects to the “Instr of Respect” (1933:57). Thus, it is possible that *at* is used in this sense, i.e., ‘with respect to’, as the complement of these verbs in Brut, with the phrase as a whole translating as ‘deprive of life’ (“at” 8(d), MED, 2021) or more loosely, ‘put to death’ (“binimen” 2, MED, 2021). This assumes that the verb indicates a type of separation, while *at* specifies what is separated without indicating separation itself. If this is the case, the complements with *at* are replacing an Instr function inherited by the Dat, rather than the inherited Abl function of separation or the Gen of separation. The preposition *of* is used both times that *bedǣlan* has a prepositional complement, as in the following sentence:

(41) *þat þas Bruttene wes of cniht-e bidæl-ed*
that this.NOM.SG Briton was of knight-DAT.SG deprive-PP
'that this Briton was deprived of a knight' (Brut, l. 6048)

The preposition *of* is also used in the other three prepositional complements of *berēafian*, as in the following sentence:

(42) *Þus wes þas kineriche; of heora king-e birœu-ed*
thus was this.NOM.SG kingdom of their king-DAT.SG deprive-PST
'Thus this kingdom was deprived of their king' (Brut, l. 1447)

The preposition *of* in these sentences has the meaning 'from, of' denoting deprivation ("of" 4.a(c), MED, 2021). Thus, *of* more clearly replaces the separation function of the Gen and Dat than *at*, but the latter probably does as well.

Of the five verbs expected to assign Dat, only *behōfian* 'require' has one prepositional complement, with *to*:

(43) *mid wepn-en. and mid hors-en; swa bihou-eð to cniht-en*
with weapon-PL and with horse-PL, as behoove-PRES.3SG to knight-PL
'with weapons and with horses, as knights require' (Brut, l. 12676)

This verb was not used with prepositional complements in OE (Mitchell 1985:458). This suggests that *to* is used to reinforce the Dat with *behōfian* in Brut, a role it has served for various Dat functions, as mentioned in sections 3.1.2.7 above (see also "to" 32(f), MED, 2021).

For the verbs expected to assign either Dat or Acc, four of the complements are prepositional. As with Gen-assigning verbs, this category has relatively few complements overall. Of the two verbs in this category, only *dēman* 'decide, judge' has prepositional complements. All four prepositional complements contain the preposition *of*. In OE, this verb sometimes had complements with prepositions, including *be* 'regarding' and *ofer* 'over' ("dēman" I.A., DOE, 2021). However, it is not clear how common these were, since Mitchell (1985:456) does not indicate this possibility. In Brut, the preposition *of* could be indicating the

focus of a mental activity, meaning ‘about, concerning’ as with the verbs of mental state. Thus, *of* would be supplanting *be* in this meaning, as it seems to have done with *reccan*. This seems to be its meaning with *demen*, from *dēman*, in the following sentence:

(44) *Ælles ne cunne we dem-en; of Arður-es deð-en*
Else NEG can we deem-INF about Arthur-GEN.SG death-DAT.SG
‘Otherwise we cannot decide about Arthur’s death’ (Brut, l. 11508)

In sentences such as this one, *of* could also be indicating the source of perception, in the sense ‘as a result of’ (“of” 20(a,b), MED, 2021), with the phrase translating as ‘make an inference from’ (“demen” 10, MED, 2021).[161] Regardless, these prepositions provide further evidence that *dēman* behaves more like a verb that would be expected to take Gen than one that would be expected to take Dat.

For the verbs expected to assign Gen, Dat, or Acc, only eight of the complements are prepositional. Four of the seven verbs in this category have prepositional complements. Of the eight prepositional complements, the prepositions *of*, *on*, and *over* each occur twice, while *after* and *to* each occur once. Of these four verbs, only *wilnian* ‘desire’ took a prepositional complement in OE, with *to* (Mitchell 1985:457, 462-464). In Brut, *wilnian* is used once with *after*, probably in the sense of ‘for, in pursuit of’ with a verb of desire (“after” 6(d), MED, 2021; Bungenstab 1933:57):

(45) *heo wiln-ed-en after worre*
they desire-PST-PL for war
‘they wished for a war’ (Brut, l. 1314)

This use of *after* resembles the use of *to* in OE; both prepositions indicate the goal of the action (Bungenstab 1933:57). The other prepositional complement of *wilnian* contains *of*, probably meaning ‘about, concerning’ as with other verbs of mental attitude:

[161] With the CSR equivalent *sudit´* ‘judge’, the preposition *o* ‘about’ and a Loc object is used to express the topic of the judgements, i.e., *sudit´ o* means ‘make judgements about’ (ORD, 2000). This choice of preposition in CSR corresponds more with the *of* meaning ‘about, concerning’ than ‘as a result of’, suggesting the former meaning is used with *dēman*.

(46) *for heo al mid wrong-e; wiln-ed-en of ure lond-e*
because they all with wrong-DAT.SG desire-PST-PL about our land-DAT.SG
'because they all wrongly desired our land' (Brut, l. 13627)

The verb *(ge)fandian* 'search out, experience' is also used once with *of*, possibly with a partitive sense, i.e., 'some of' ("of" 14.(d), MED, 2021), or perhaps to indicate that the object is not affected by the action, since this is the assumed motivation for the use of Gen with *(ge)fandian*:

(47) *for to let-en fondi-en; of his main strong-e*
for to let-INF experience-INF of his unusual strong-DAT.SG
'to let his unusual strength be tested' (Brut, l. 930)

These uses of prepositions with Gen/Dat/Acc-assigning verbs make sense as a substitute for Gen objects. As with *reccan* 'care for' and *wundrian* 'admire', the meaning of *on* with *(ge)fandian*, as well as *tilian* 'provide', is less clear. In the sentence with *(ge)fandian*, it may be used to indicate that the object is unaffected, as the Gen and possibly *of* are used with this verb:

(48) *fondi-en on feht-e; ȝif he hine ouercumen mæhte*
search-INF out fight-DAT.SG if he him.ACC overcome might
'to search out a fight, so that he might overcome him' (Brut, l. 2760)

This interpretation is supported by the combination of *on* and *(a)waiten* with the meaning 'wait for' ("on" 22.(c), MED, 2021). This verb is a synonym of *ābīdan* 'await' borrowed from French ("awaiten", "waiten", MED, 2021). Like *(ge)fandian*, *ābīdan*, and a number of other verbs that assigned Gen in OE, its action does not affect the object. With *tilian*, a locational sense is more likely than with *(ge)fandian*:

(49) *Heo mak-ed-en tun-es; heo til-ed-en on eorð-en*
they make-PST-PL town-PL they plow-PST-PL on earth-ACC/GEN/DAT.SG
'They made towns; they plowed (on) the earth' (Brut, l. 971)

However, this phrase is treated as a complement because *eorðen*, the object of *on*, also appears once as an Acc object of *tilian* in Brut, and it is likely that both of these sentences have the same general sense:

(50) *þa cheorl-es heo uloȝ-en; þa til-ed-en þa eorð-en*
the churl-PL they drive.off.PST-PL REL plow-PST-PL the.ACC.SG earth-ACC.SG
'they drove off the churls that plowed the earth there' (Brut, l. 10458)

Two of the three prepositional complements of *(ge)wealdan* 'control, possess' contain the preposition *over* and one contains *to*. The preposition *over* brings attention to the status difference in relationships of authority or control ("over" 7(a), MED, 2021):

(51) *& ic hem ȝeu-e al þa winn-e; þe ich æm wald-inge ouer*
and I them.DAT give-1SG all the.NOM.PL possession-PL REL I am possess-PRESP over
'And I give them all the possessions that I own' (Brut, l. 1548)

This is more of a Dat sense, as it shows the lower status of the object. The preposition *to* may be a reinforcement of the Dat as with *behōfian* 'require', but it can also be used with expressions of ruling with the same sense as *over* ("to" 26, MED, 2021):

(52) *& he him-seolf wen-d-e. alse he wald-e to his lond-e*
and he him.DAT-self believe-PST-3SG as he control-PRES.SUBJ.SG to his land-DAT.SG
'And he himself believed as if he controls his land' (Brut, l. 6498)

This prepositional complement more likely replaces the Dat than Gen as well. Thus, some prepositional complements with Gen/Dat/Acc-assigning verbs appear to replace the Gen, others appear to replace the Dat, and a few are difficult to explain as replacements for either case.

Overall, *of* is the most common preposition in prepositional complements by quite a large margin, with fifteen occurrences. Of these, eleven could be replacing the Gen. Some clearly do so, while the five occurrences with verbs of separation could be replacing either the Dat or the Gen. Moreover, the other four occurrences are with *dēman* 'decide, judge', which is categorized

as a verb that takes the Dat/Acc but seems to fit better semantically as a verb of mental state assigning the Gen. Thus, all the occurrences of *of* could be considered replacements of Gen-like uses, even with verbs that do not themselves assign Gen. As mentioned above, these replacements are not all EME innovations; some verbs already had prepositional complements in OE, either with the same prepositions or with different prepositions. Among the verbs considered here, only *bedǣlan* 'deprive of' appears to have been used with *of* in both OE and Brut, with a separation meaning, i.e., 'from'. In addition to spreading to *berēafian* 'deprive of' in this use, *of* has supplanted *be* in indicating the focus of a mental attitude, as with *reccan* 'care for' and *dēman*. It is also used in this way with *gilpan* 'boast, rejoice' and *wilnian* 'desire', which took prepositional complements with *for* and *to*, respectively, in OE. Finally, *of* occurs once with *(ge)fandian*, which was not attested with prepositional complements in OE. Here it seems to be replacing a different Gen function, either partitive or unaffectedness.

The next most common preposition is *at*, which occurs six times. These instances are all with verbs of separation that assign Gen or Dat and are all followed by the same object (*live* 'life'), so they do not indicate much about general trends with *at*. However, some of these verbs were already used with *æt* meaning 'from' in OE, so its use is not an EME innovation. The preposition *on* is used five times, all with verbs expected to assign Gen, although one has a clause as its complement, and several may be locational adjuncts, particularly with *reccan* 'care for', *wundrian* 'admire', and *tilian* 'provide', verbs without any objects in Brut that could be interpreted as Gen. On the other hand, the Gen with these verbs may have been supplanted at least in part by prepositions including *on*, which may be used alongside *of* with *reccan* to indicate the focus of a mental attitude. Both of these prepositions are also used with *(ge)fandian* 'search out, experience', but likely not with the same sense as with *reccan*.

The preposition *over* occurs twice, both times to show lower status, a Dat sense, with *(ge)wealdan* 'control, possess', a verb of ruling. Its OE equivalent *ofer* already occurred with another verb of ruling, *dēman* 'decide, judge'. The preposition *to* is used twice, both reinforcing the Dat with verbs that did not have any prepositional complements in OE, i.e., *(ge)wealdan* 'control, possess', where *to* may have a similar sense as *over*, and *behōfian* 'require'. The preposition *for* also occurs twice, both times with *gilpan* 'boast, rejoice', a verb of mental state that was already used with this preposition in OE, in addition to the Gen and Dat. The use of *for* alongside *of* in Brut mirrors the two case options for this verb in OE. The preposition *after* only

occurs once, with *wilnian* ‘desire’, a verb which could take complements with Gen, Dat, Acc, or *to* in OE; *after* could be a replacement for the Gen and/or *to*.

Overall, these results indicate that prepositional complements in Brut often continue similar constructions in OE, either with the same preposition or a different preposition used with a similar meaning. However, there are also novel uses of prepositions, suggesting that prepositional constructions have played a role in the functional narrowing of the Gen and Dat.

5.2.3.5.2 The Owl and the Nightingale (O&N)

Table 108. Prepositional Complements in O&N

+Gen	*of*
missan	1
reccan	1
+Gen/Dat	
gilpan	1
Total	3

Two of the six verbs that traditionally assign Gen have one prepositional complement each, and both of the complements contain the preposition *of*. Neither verb was used with this preposition in OE: *reccan* ‘care for’ was used with other prepositions, as described in the previous section, while *missan* ‘miss’ did not take prepositional complements (Mitchell 1985:460). The preposition *of* is most likely used with *reccan* with the same sense as in Brut:

(53) *of non-e winter-e ich ne recche*
about no-DAT.SG winter-DAT.SG I NEG care
‘I do not care about winter’ (O&N, l. 533)

As used with *missan*, *of* probably denotes deprivation, but might also be specifying the area of failure, meaning ‘with regard to’, with the phrase translating ‘fail in’ or ‘fail to get’ (“of” 4.a(e), MED, 2021):

(54) *ȝif þe uox mis-t of al þis dwole*
if the.NOM.SG fox miss-PRES.3SG of all this deception
‘If the fox runs out of all this deception’ (O&N, l. 825)

In either sense, this *of* is clearly replacing the Gen ("missan" 3, MED, 2021).

Of the two verbs expected to assign Gen or Dat, only *gilpan* 'boast, rejoice' has one prepositional complement, with *of*, which is provided in (32) above. In fact, this is the only complement found with *gilpan* 'boast, rejoice' in O&N, making it the only verb examined in this text with no noun or pronoun objects, a status it also carries in Brut. Since *gilpan* already had complements with *for* in OE, its prepositional complements may have gained in productivity until they completely replaced the Gen and Dat. The preposition *of* is most likely used in O&N with the same sense as in Brut, more likely replacing the Gen than the Dat.

In O&N, all three prepositional complements have the preposition *of*. There are not enough data to draw any definite conclusions, but it seems that all instances of prepositional complements with verbs in this text replace the Gen, or at least Gen-like uses, i.e., complements with different prepositions that already alternated with the Gen in OE. This can be viewed as part of a general trend, which also includes *of* being used instead of the Gen in possessive and partitive constructions. However, the use of *of* with these verbs can also be motivated by the specific sense of the Gen required by each verb: two denote sources/areas of focus for mental verbs and one denotes deprivation. It could be that *of* began replacing specific uses of the Gen based on different senses of the preposition, and eventually these became common enough that *of* could be extended to any former use of the Gen, whether or not it fit with the specific meaning.

5.2.4. Summary

The above corpus study has investigated the case assignment by verbs in a selection of OE and EME texts. Analysis of the data indicates that the texts lie on a spectrum from Beowulf, with fully productive Gen and Dat, to Brut and O&N, where the Acc is the dominant productive form of case marking. The extreme functional narrowing of the case markings in the latter text may indicate the complete loss of morphological case marking and replacement by grammatical relations. Unlike in Beowulf, Brut and O&N have a particularly strong tendency away from Gen objects, followed by the Dat.

In Brut and O&N, prepositional complements serve as another alternative to the Gen and Dat. Brut shows a higher frequency of prepositional complements and also a more even

distribution across case-assignment categories, although there are still many verbs which never take prepositional complements. In both texts, one verb, *gilpan* 'boast, rejoice', is attested without any noun or pronoun objects, only prepositional complements. There are some tendencies in terms of which prepositions seem to replace the functions of which case markings, although more promising connections can be drawn between the meaning of the verb and the preposition.

5.3. Comparison and Conclusion

Two studies have been considered in the preceding sections. In the first, as noted in 5.1 above, Vakareliyska (1990) found that patients with mild to moderate Wernicke's aphasia made more errors in choice of case marking with certain categories of verbs than with others. Based on previous research that Wernicke's aphasics were more likely to retain access to a word's core semantic features, which tend to be universal, than to peripheral features, which tend to be language-specific, she argues that case marking that depends more on the latter type of features accounts for the higher error rate on some verbs. This has implications for case use by unimpaired L1 speakers as well as L2 speakers. For unimpaired L1 speakers, case variation may result from a verb's having conflicting case-associated semantic features, even if the conflicting cases themselves are relatively distinctive in their other functions. In a sense, this is just shifting the focus on semantic overlap from the cases to the verbs. If L2 speakers are similar to Wernicke's aphasics in that both groups tend to have better access to the universal, core meanings of verbs than the language-specific, peripheral meanings, as Vakareliyska suggests, then her findings could provide the connection between L2 speakers and functional narrowing. Specifically, even if the L2 speakers have an L1 with a similar case system and successfully acquire the core meanings of verbs and morphological cases, differences in the peripheral features between L1 and L2 might still cause them to make frequent errors in functions, as well as other functions in which one case is often displaced by another, i.e., after prepositions. In addition to the higher frequency of IO constructions for the Dat and adnominal constructions for the Gen, for example, the retention of these core case functions longer than the case assignment by verbs and prepositional uses can thus be explained by greater cross-linguistic variation in the latter.

In the second, the corpus study of OE and EME texts in 5.2, I found that the Gen and Dat lost productivity in a roughly chronological progression that supports Barðdal's (2009, 2008) usage-based constructional approach to functional narrowing that constructions with higher type frequencies tend to attract verbs from constructions with lower type frequencies over time, especially when there is semantic overlap. This accounts for the increase in Acc objects with verbs that previously assigned other cases. In both of the OE glossaries used in the investigation, the Gen had a lower type frequency than the Dat. The type frequency of both constructions was much lower than the Acc object construction. This also supports the proposal by Bybee (1985, 1995) that type frequency is a primary indicator of productivity. This can be observed via the decrease in Dat objects, on the one hand, and the possibly complete lack of Gen objects, on the other, is the difference in type frequency between Dat assigning and Gen-assigning verbs. Based on the glossary in Klaeber (2008), 88 verbs that can assign the Dat to their primary or sole object were used in Beowulf, but only 32 that could assign the Gen. The type frequency of Acc-assigning verbs was not calculated, but is assumed to be much higher than any other case-assignment construction. Assuming that the relative type frequencies in Beowulf are representative of the general situation in OE, the Gen object construction already had a much lower type frequency than the Dat object construction.

The results are consistent with this situation in OE: the Gen and Dat both lost productivity to the Acc as the texts became more innovative, but the process was more rapid for the Gen. Some replacement of the Gen by the Dat was also attested, supporting the idea that constructions with lower type frequency can still attract new members if semantic coherence among the existing members and semantic similarity to the new members are both sufficient. Prepositional constructions also competed with the Gen and Dat in two of the texts. Given the variety of prepositions attested in these constructions, they necessarily had a lower type frequency than the Acc, but they also seem to have been much more coherent, due to the narrower meaning of prepositions. Thus, it follows that prepositional constructions were able to attract certain verbs based on semantic similarity, even as all but one of these verbs were also used with the Acc in the same texts. In conclusion, both of Barðdal's (2008, 2009) factors of productivity can be observed to have an effect on constructions in the OE and EME texts investigated. The results are also consistent with Suttle and Goldberg's (2011) concept of coverage as the determining factor in productivity: the Acc achieved high coverage with a

variety of new verbs through high type frequency and high variability, i.e., low semantic coherence, i.e., the resemblance of a new instance of a construction to previous instances is a measure of its semantic similarity, while the prepositional constructions achieved high coverage through low variability but high similarity to the new verbs.

Besides providing insight into the progression of functional narrowing in English, the results of my corpus study may also provide some clues as to the origins and motivations behind this process in English and other Germanic languages, and potentially IE languages more broadly. Barðdal suggests that an influx of borrowed verbs accelerated the process because these new verbs were attracted to the high type frequency constructions, the Acc construction, further increasing its type frequency (2009:133-134). This may have played some role in the later texts, but further investigation would be needed to determine how many verbs of Norman French origin are actually attested in them. Already in Beowulf and other OE texts, however, the Acc was competing with the Gen and/or Dat for many of the verbs. Moreover, the semantic analysis of the verbs that were used with the Gen and/or Dat suggests that these constructions were not particularly coherent. It was difficult to determine why some verbs took a particular case in OE, when these constructions still appeared robust. It should be noted, however, that only a fraction of the verbs attested with these constructions in OE were selected for investigation. Thus, there are most likely gaps in their semantic coverage; further research could address these gaps and provide a more complete picture.

Section 5.2.2 above discusses the semantic classes of verbs that take Gen or Dat in OE. For instance, the semantics of verbs that traditionally take Gen objects in OE include verbs of perception, mental state, desire, and deprivation or separation. The semantic connections are no longer parallel between the semantics of the verb and the morphological case assigned by EME, as can been seen in Brut and O&N since these verbs can take any morphological case marking or prepositions. For example, the existence of both Gen and Dat semantic classes containing verbs of separation may explain some of the substitutions and uncertainty found in certain verbs' assignment of these two morphological cases, as does the use of both to express possession. In fact, one of three verbs of separation investigated, *beniman* 'deprive of', historically take Acc besides Gen and Dat, while its synonymous verbs *bedǣlan* and *berēofan/berēafian* take only Gen or Dat. This indicates that even in OE there is a semantic overlap between verbs but by EME this overlap expands to include other verbs. Given these semantic overlaps, differences in type

frequencies may explain why the Acc replaces both the Dat and Gen, while the Dat replaces the Gen more often than vice versa.

Contact may have contributed to functional narrowing in other ways, however. The mixture of West Germanic dialects that gave rise to OE may have contributed to greater case variation. If L2 speakers, like Wernicke's aphasics, have more difficulty choosing the correct case with verbs where that case differs from what is expected based on the core semantic features of verbs, then we would expect to see an earlier shift to the Acc for verbs with less clear semantic motivation for the case they assign in OE, e.g., Dat-assigning verbs that lack the core feature of *cognitive engagement*. I did not investigate this aspect specifically in the original corpus study, but a post hoc analysis found some correlation in the changes to case use among semantically similar verbs in CSR and OE.

Only one verb in my corpus study, *helpan* 'help', which assigned the Gen or Dat in OE, was the direct semantic equivalent of a verb in Vakareliyska's (1990) study, CSR *pomogatʹ* 'help'. The feature of *cognitive engagement* by the non-active participant is a core part of this verb's meaning, although not quite to the same extent as for verbs with both a DO and an IO. In Vakareliyska's CSR aphasia test, three out of 16 patients made case errors with this verb, which is below her cutoff for considering a verb to have predominantly Dat-associated features (ibid., 172-187). This was also one of the verbs in my corpus study that most clearly retained non-Acc case marking in EME texts. In Beowulf, *helpan* has two Gen noun objects, two Dat pronoun objects, and one ambiguous noun object. Unambiguously Dat and, to a lesser extent, Gen objects still occur in Brut, but a few are clearly Acc and more could be Acc. In O&N, all three pronoun objects are Dat or Gen. Even if these were a result of the ongoing Dat/Acc merger on third-person pronouns, three of the verb's five noun objects in the study were Dat or Gen, and only one was unambiguously Acc, so the Dat and possibly also Gen were still in use with this verb. Thus, the verbs meaning 'help' appear to have semantic features that inhibit the use of Acc marking on their objects, both in the EME texts and by Wernicke's aphasics speaking CSR in Vakareliyska's study.

Other verbs investigated in these two studies are not direct semantic equivalents but belong to the same semantic classes. The verbs meaning 'help' fit in Lass's class of verbs of serving, confiding, or trusting (1994:238), one of the classes most clearly associated with *cognitive engagement*. Other CSR verbs from Vakareliyska (1990) that fit into this class are

sovetovatʹ ‘advise’ and *veritʹ* ‘believe’. The verbs *soobščitʹ* ‘inform’ and *lʹstitʹ* ‘flatter’ are marginal fits: they require *cognitive engagement* but not cooperation. The first three of these verbs received case errors from four of 16 patients, right at the cutoff for predominantly Dat-associated features, meaning they had almost as many Acc-associated features as Dat ones. However, *lʹstitʹ* received six errors, suggesting an even closer proportion of features (ibid., 92-93, 175, 229-237).

As noted in section 5.2.2.2 above, the other OE verbs investigated that fit this class are the Dat-assigning verb *beorgan* ‘protect’, as well as (*ge)unnan* ‘grant’ and *tilian* ‘provide for’, with respect to their animate objects, which could be exclusively Dat and either Gen, Dat, or Acc in OE, respectively. The first of these verbs does not require *cognitive engagement*, while the other two are essentially IOs that were included in the investigation because the inanimate objects of these verbs were not canonical DOs but could be either Gen or Acc and either Gen, Dat, or Acc in OE, respectively. Therefore, Vakareliyska’s results predict that the Dat should be less durable with *beorgan* than on the animate objects of *tilian* and (*ge)unnan*. However, the texts did not contain many objects of these three verbs. In Beowulf, *beorgan* has one ambiguous and two Dat noun objects. Of the later texts, only Brut contains objects of this verb: one Dat/Acc noun and two Dat pronouns. These results are ambiguous; the Dat could still be productive with this verb, or these forms could reflect the Acc, given the ongoing Dat/Acc merger on third-person pronouns. The verb *(ge)unnan* has one Dat and three Dat/Acc pronoun objects in Beowulf, and one Dat and one Dat/Acc pronoun object in Brut. Thus, the results with these two verbs are also ambiguous, and the prediction based on Vakareliyska cannot be evaluated conclusively.

Impersonal verbs occurred in both studies, but without much similarity in meaning beyond the *cognitive engagement* associated with Dat experiencers. My corpus study only investigated Dat-assigning impersonal verbs, which Vakareliyska notes are more common in CSR, although she only investigated two Dat-assigning verbs against four Acc-assigning impersonal verbs (1990:99-100). Two of the Acc-assigning verbs elicited a high number of errors: nine with X-Acc *znobit* ‘X is chilled’ and seven with *tošnit* ‘X feels nauseous’. The other two, X-Acc *vzorvalo* ‘X blew up’ and X-Acc *tjanet* ‘X longs for’, had only one error each. Vakareliyska attributes the much higher error rate for the first two verbs to interference from semantically similar Dat-assigning constructions and lower than normal transitivity for an Acc-

assigning verb, in contrast to the second two, which are metaphorical extensions of clearly transitive verbs involving physical activity (ibid., 258-261). Three patients made errors with each of the Dat-assigning impersonal verbs: X-Dat *vletelo* 'X got into trouble', and X-Dat *povezlo* 'X was in luck' (ibid., 93).

As noted in section 5.2.2.2 above, the OE impersonal verbs included in my corpus study are *þyncan* 'seem', *ofþyncan* 'displease', and *behōfian* 'be required of'. All three took Dat experiencers in OE, so it is no surprise that all of their complements are unambiguously Dat in Beowulf. In Brut, three of the ten nouns and 53 of the 76 pronouns that occur with them are unambiguously Dat, while only two nouns and one pronoun are unambiguously Acc. An additional one noun and one pronoun are Gen/Dat, and the rest could be Dat or Acc. In O&N, one ambiguous noun, one Dat pronoun, and six Dat/Acc pronouns occur with *þyncan*. Thus, the use of Dat on experiencers of these impersonal verbs appears relatively durable in EME texts, just as Wernicke's aphasics speaking CSR made relatively few errors on Dat-assigning impersonal verbs. It is clear from the noun objects that cannot be Acc that the Dat was still productive with impersonal verbs in Brut, but the Dat and Dat/Acc pronoun forms in O&N do not provide the same level of certainty; even the Dat forms in O&N could reflect the Acc because they are third-person pronouns.

Vakareliyska investigated two groups of verbs with canonical DOs: verbs involving physical activity and verbs of perception. These verbs do not involve *cognitive engagement* and tended to elicit few case errors, with the exception of *narisovat′* 'draw' in the first group; Vakareliyska attributes the five errors with this verb of creation to an alternative interpretation with a Dat beneficiary that even some normal controls noted. Other verbs involving physical activity all received three or fewer errors: three with *udarit′* 'hit', two each with *spasti* 'save' and *ubit′* 'kill', and one each with *pocelovat′* 'kiss' and *napugat′* 'frighten'. Verbs of perception also elicited three or fewer errors: three each with *videt′* 'see' and *ljubit′* 'love', two each with *slyšat* 'hear' and *uvažat′* 'respect', and one each with *znat′* 'know' and *slušat′* 'listen to' (1990:81, 88-95, 222-224).

Verbs that exclusively assigned the Acc in OE were not included in my corpus study, but some comparisons can be made to verbs that involve physical activity or perception, even though they assigned cases other than the Acc to various degrees. The verb *weorpan* 'throw' involves physical activity; in addition to the Acc, it sometimes assigned Dat to focus on the incidental role

of the object (i.e., as an instrument) or Gen, for unclear reasons discussed in section 5.2.3.1 above. Since it does not have core Dat-associated semantic features like *cognitive engagement* and there is no clear motivation for assigning the Gen, this verb is expected to stop assigning the Dat and Gen earlier than the verbs with predominantly Dat-associated features discussed above, e.g., *helpan* 'help'. In Beowulf, one noun object in each case, i.e., Gen, Dat, and Acc, occurs with *weorpan*. In Brut, ten of its 18 noun objects and three of its four pronoun objects are unambiguously Acc, and the other noun objects could all be Acc as well; there is one Dat third-person pronoun object, but this could also reflect the Acc. Similarly, both pronoun objects and two of its three noun objects are Acc in O&N; the other is Dat/Acc. Thus, there is no conclusive evidence for non-Acc case assignment in any text investigated except Beowulf; instead, the data suggest that the Dat and Gen quickly lost productivity with *weorpan* during the EME period, as predicted.

A number of OE verbs under investigation involve perception. Some of these verbs were assigned to this class or to a closely related class in the previous literature, as noted in section 5.2.3.1 above: the Gen-assigning verb *reccan* 'care for', the Gen/Acc-assigning verbs *hēdan* 'heed', *wēnan* 'expect, think', and *wilnian* 'desire', and the Gen/Acc-assigning verb *wundrian* 'admire'. The first two most clearly involve perception as opposed to other mental activity and are the closest in meaning to any of the CSR verbs included by Vakareliyska (1990), namely *uvažat′* 'respect'. The verb *wundrian* is not so different in meaning from *ljubit′* 'love', but it might be less clearly associated with the Acc, since the CSR verb *ljubovat′sja* 'admire' (which is reflexive and thus cannot also have an Acc DO) assigns Instr (see ORD, 2000). Likewise, *wilnian* 'desire' corresponds to CSR *žaždat′* 'crave', which assigns Gen, and *wēnan*, at least in its meaning 'expect', corresponds to CSR verbs such as *ždat′* 'await, expect', which take Gen with [-specific] objects and Acc with [+specific] objects (ibid.). Therefore, this comparison will focus on *hēdan* and *reccan*, which are predicted, based on Vakareliyska's results, to shift relatively quickly to assigning only the Acc in EME texts. These each have one Gen noun object in Beowulf, and *hēdan* also has a Gen/Acc noun object. In Brut, *hēdan* has two noun objects, one Acc and one ambiguous, and two Acc pronoun objects, while *reccan* has two Acc noun objects and one Dat pronoun object, as well as four prepositional complements; the third-person Dat pronoun is particularly likely to reflect the Acc here, since Dat was not one of the cases assigned by *reccan* in OE. In O&N, *hēdan* has two Acc noun objects; *reccan* has three Acc noun objects,

one Acc pronoun object, and one prepositional complement. Thus, all of these two verbs' objects in the texts after Beowulf could be Acc, and only one could still be Gen. However, the prepositional complements with *reccan* suggest that its semantic features are not as heavily Acc-associated as Vakareliyska (1990:216-217) proposes for the CSR verbs of perception in her study.

Although the verbs meaning 'help' were the only direct semantic equivalents included in both studies, a comparison of verbs belonging to the same semantic classes has revealed some correlation between the developments in case use in EME texts and the case errors made by Wernicke's aphasics speaking CSR. In both studies, the Dat was slightly less likely to be replaced by the Acc for the object of the verbs meaning 'help' than other verbs of serving, confiding, or trusting. The Dat was also relatively durable for Dat-assigning impersonal verbs. On the other hand, non-Acc case marking with verbs of perception and *weorpan* 'throw', a verb involving physical action, lost productivity in EME texts more quickly, just as aphasics relatively consistently used the Acc with verbs in these classes. These parallels allow for the possibility, but by no means confirm, that a high number of L2 speakers could have contributed to the same types of changes to case use in the development of EME from OE that Wernicke's aphasics apply synchronically in their use of CSR. Even without the influence of L2 speakers in EME, semantic features inconsistent with the choice of case could have led to the gradual loss of productivity of the Gen and Dat in favor of the Acc. Some of the same inconsistencies almost certainly existed in other Germanic languages, however, but in the development of High German, for example, the Gen and Dat lost productivity much more slowly. Thus, a potential role for contact in accelerating the process of functional narrowing still remains.

This chapter has considered two studies on case use that generally support the accounts of functional narrowing discussed in section 3.1.2.6 above, including prepositional constructions, as discussed in section 3.1.2.7 above. The two studies suggest several potential connections to language contact, but by no means confirm that these played a role. Considered together, they have also revealed areas for future research.

CHAPTER VI.
CONCLUSION

The preceding chapters have examined the loss of morphological case and grammatical gender in the Germanic, Romance, and Balkan Sprachbund languages. To my knowledge, my study is the first to compare case and gender loss, and simplification, among 18 IE languages. Previous efforts were limited to a single branch or a few languages. Several approaches have been used to answer how and why the loss of case and gender categories has occurred to the extent it has in these languages. The discussion has generally divided the factors involved in case and gender loss into internal motivations, including phonological, morphosyntactic, and semantic, and external motivations, i.e., contact situations. This chapter synthesizes the findings and conclusions of the preceding chapters, observing these divisions among factors.

6.1. Phonological Factors

It is relatively clear that sound change has played a significant role in the loss of case and gender distinctions in the Germanic, Romance, and Balkan Sprachbund languages. In the historical development of the languages investigated, only Icelandic and Faroese have not lost any such distinctions as a direct result of regular sound change; these two languages have maintained almost all of the case and gender distinctions they inherited, except the Gen in spoken Faroese. The loss of this category cannot be attributed to sound change, even partially; sound change has been established as a factor in almost every other reduction in case and gender categories that has been examined. On the other hand, the loss of only one category, the Instr in High and Low German, can be fully attributed to regular sound change.

Many of the segmental sound changes that neutralized distinctions in nominal inflection may best be considered as consequences of prosodic change. Specifically, a shift to rhythmic dynamic stress has occurred in the development of all Germanic and Romance languages, as well as Greek and Albanian. Vowel reduction and deletion, the loss of phonemic vowel length, and the loss of final consonants all follow from this shift, particularly when weight-sensitive stress developed. Even some changes that do not appear to be regular may have been prosodically-motivated in particular phrasal contexts. Later some of these languages, e.g., the continental Scandinavian languages and French, moved away from a dynamic stress, but not before it

triggered many neutralizing segmental changes. Notably, Icelandic and Faroese have dynamic stress, but it has not become weight-sensitive. This conservative prosody explains why they did not undergo neutralizing sound changes and, by extension, significant case loss. More generally, the degree to which case loss occurred in the Germanic, Romance, and Hellenic languages cannot be fully explained by prosodic change, but the development of rhythmic dynamic stress could still be the primary motivation for the simplification of nominal inflection that has occurred in these languages.

The Slavic languages have diverged more in their prosodic development, so they provide a suitable test for this proposal. Three different prosodic outcomes can be observed in those investigated in depth in my study: BCMS has retained a pitch accent, CSB has developed a strong free dynamic stress that has triggered vowel reduction, and CSM has developed a weaker stress fixed on the antepenultimate syllable with no associated vowel reduction. More generally, stress-induced vowel reduction and total case loss on nouns appear to be correlated among the Slavic languages: most have neither and CSB has both. The frequency of nouns with final stress may explain why CSR and the other East Slavic languages have vowel reduction without case loss, but an additional motivation would be needed to explain why CSM has lost (almost) all case distinctions on nouns despite a weak stress without vowel reduction.

6.2. Morphosyntactic and Semantic Factors

To the extent that phonological factors cannot account for all reductions in case and gender categories, morphosyntactic and semantic motivations must have existed. In fact, they must have played a central role in these developments, when only the regular sound changes established in the literature are considered since these leave almost all categories distinctive in at least one paradigm.

6.2.1. The Merger of Declension Classes

My analysis in chapter II also reveals the significant role of merging declension classes over time in the loss of case and gender. Even without case and gender loss, declension classes tend to be simplified, including the number of forms they comprise. Major classes expanded over time, absorbing less productive, smaller classes, and further increasing as a default. CL and PS *o*-stems were the default for masculine and neuter nouns *ā*-stems for feminine nouns. For example, *u*-

stems merged with *o*-stems in MSrb. In PGmc the *a*-stems were much larger than other classes, with more masculine than neuters, and *ō*-stems were the largest feminine class. For instance, *i*-stems merged with *a*-stems in OHG.

Declension classes comprised of nouns largely of the identical gender are much more likely to merge than those that are not. Even without much formal similarity, shared gender can allow a minor class to be absorbed. For example, feminine *ē*-stems merged with the feminine *ā*-stems, and masculine *u*-stems merged with masculine *o*-stems by VL. In the transition from OE to ME, masculine *a*-stems expanded as the default, and by EME, this was the only productive class that remained, and grammatical gender had been lost, e.g., the merger of masculine *ja*-stems/*i*-stems with masculine *a*-stems. Forms of the smaller class are more likely to survive if they are more salient, e.g., the spread of -*ov*- in MSrb.

The loss of gender distinctions seems to be facilitated by similarity in forms in the languages under investigation. As with the merger of masculine and feminine into a common gender in MSw, MDan, and MDu. This formal similarity can be observable in masculine *a*-stems and feminine *ō*-stems in MSw, and MDan. Likewise, masculine *a*-stems forms are similar to feminine *i*-stems by MDu. A parallel process occurred in the development of VL from CL as with the loss of the neuter beginning in Vulgar Latin, merging with the masculine. By VL, the forms of masculine and neuter declension classes are syncretic except the Nom singular in the masculine and neuter *o*-stems.

6.2.2. Relevance, Case Syncretism, Number Syncretism, and Gender

For nouns in IE languages, number is the most relevant category, followed by case and then declension class. Analogical processes tend to profile more relevant categories, even if distinctions in less relevant categories are lost. In other words, syncretism among case forms within the same number (case syncretism) is preferable to syncretism among case forms of different numbers (number syncretism). This tendency often manifests as the leveling of a case form, usually in the unmarked singular, to another form of the same number. In the singular, this form is usually the least salient, as in the frequent leveling to an uninflected form. The least salient form is often the least marked form, i.e., the Nom, but if the Nom happens to be more salient than the Acc, not only is it less likely to be the target for leveling, but it is prone to leveling itself. When a form is leveled in the plural, an unmarked form is usually the target, but

this form tends to be relatively salient as a frequent marker of the marked plural. In the relatively rare instances of leveling among marked plural forms, the most salient form is preferred because it is marked in both case and number.

The morphological blocking of a sound change is another mechanism by which number syncretism is avoided while maintaining the salience of a marked form. For example, in the development of MSw, number syncretism in the Nom/Acc plural was avoided by the morphological blocking of *-r* in the Nom singular for masculine *i*-stems/*u*-stems. Similarly, in the development of ModGr from KG, the number syncretism in the Gen plural with the Acc singular for masculine *o*-stems as well as the Nom/Voc/Acc for neuter *o*-stems were eliminated by the morphological blocking of *-n* in the Gen plural. Relatedly, a phonological process such as umlaut may be morphologized, usually, so it only applies in the marked value of the most relevant category, e.g., the plural, and so it can apply in this category even when the phonological conditions for the change were not met. All of these processes are in accordance with the iconic principle, i.e., the tendency for more noticeable forms in more distinct categories. However, the loss of case distinctions in the plural is less often a direct response to number syncretism than the consequence of the same case being lost in the singular.

However, developments in some languages are exceptional in terms of these associations. The development of OF from WVL is exceptional in a fundamental way: number syncretism actually increased due to analogical processes. As mentioned above, analogical processes profiled gender rather than number, but number marking was profiled again in the development of MF and the associations held over the history of French as a whole. The effects of sound change and analogical processes on number syncretism in High and Low German were about as major as in the Germanic languages that lost all nominal case distinctions. However, both High and Low German retain case distinctions on articles and other agreement targets, and bare nouns in CSHG still preserve some of these distinctions. Differing paradigmatic pressures on agreement targets is a potential explanation for the outcome in these languages; however, this does not explain why case distinctions were not able to survive in a similar way in other Germanic languages since all have definite articles. The development of Low German is more similar to that of the Germanic languages with total case loss because it, too, lost all case distinctions on nouns, but case marking is not entirely lost on articles. In the development of Romanian, number syncretism increased as a result of sound change and decreased as a result of

analogical processes to a similar extent as in the Western Romance languages. As in High German, however, some case distinctions have been retained in Romanian on nouns and even more on articles. The development of the postpositive article in Romanian is the most obvious morphosyntactic difference from the Western Romance languages that might be connected to this (the postpositive article did not cause the association with the Sprachbund, it is a symptom of it). This feature is associated with the Balkan Sprachbund languages; mutual reinforcement among these languages may have also influenced Romanian in other ways.

Two other Balkan Sprachbund languages, Bulgarian and Macedonian, have experienced much the same increases in number syncretism due to sound change as another South Slavic language, BCMS; in the latter, however, analogical processes did not lower number syncretism as much, and much less case loss occurred. One or more additional factors must have led to the different outcomes in these closely related languages. The most obvious source of divergence is the contact that Bulgarian and Macedonian had with other Balkan Sprachbund languages because of their geographic location. On a morphological level, these developments manifested as the leveling of case distinctions beyond what would be expected in response to neutralizations resulting from sound change alone. In other Slavic languages such as BCMS, syncretism was instead eliminated by the adoption of more distinctive forms from unproductive declension classes, which can be seen as the profiling of number and case at the expense of declension. In fact, the adoption of more distinctive dual forms in the Dat/Loc/Instr plural in BCMS can be considered the profiling of case marking at the expense of number marking. This is not to say that the distinction between the dual and plural was lost for this reason; rather, the dual was already falling out of use, probably due to its low frequency. When a function merger occurs in this way, the forms of the less marked category usually survive unless some other principle applies. In BCMS, it was likely the iconic principle that favored the retention of the more salient dual forms in the highly marked Dat/Instr/Loc plural.

Gender marking on agreement targets is closely connected to declension class on nouns, but their relationship varies across languages. Gender marking on NP-internal agreement targets appears to be critical for the retention of gender categories, so it follows that the degree to which gender categories are lost is correlated with differences in these relationships among the Germanic, Romance, and Balkan Sprachbund languages. Contrary to what Kürschner and Nübling (2011) have hypothesized for a selection of Germanic languages, however, the total

dissociation of gender and noun declension can occur without the loss of a gender category. Total dissociation is often correlated with the semanticization of gender categories along more transparent lines, but only if there is also a reduction to two genders, e.g., CSDan, CSDu. In addition, gender loss can still occur in languages that retain a partial dissociation of these categories or develop a total association, e.g., MDan. Moreover, the most conservative gender-declension relationships are correlated with lower gender syncretism on the core cases of agreement targets, but systems with a partial association have the least gender syncretism, rather than a partial dissociation, which appears to be the most conservative gender-declension relationship, based on its occurrence in PGmc, CL, LPS, MyG, and likely Proto-Albanian. Finally, the loss of gender distinctions in the plural, e.g., CSDan MDu, CSHG, and CSS, are correlated with a partial or total dissociation of gender and declension; the high levels of gender syncretism associated with a lack of distinctions in the plural can then explain why more gender syncretism has occurred with a partial dissociation than with a partial association.

Case loss is also a significant factor in the loss of gender categories, as well as changes in the relationship between gender and declension since a decrease in case forms is also a decrease in forms that can be used for distinguishing genders on both agreement targets and nominal declension classes. In languages with total case loss on nouns and agreement targets but still have gender, such as CSS, CSDan, CSDu, CSF, CSI, and ModSp, only the singular and plural forms are available to distinguish gender and declension.

Plural allomorphy profiles number category, so it is much more likely to be retained, but singular forms are frequently uninflected for the reasons described above. Noun declensions can be distinguished based on the plural form alone, as in all of the Germanic languages that have undergone total case loss on nouns but still retain gender categories (usually two, but three in NNw). On agreement targets, however, number appears to be a less relevant category since agreement targets are used together with nouns, whose number marking may be sufficient for number disambiguation. Therefore, gender distinctions on agreement targets are not expected to survive in the plural unless they also survive in the singular. Despite its important role in gender loss, I have not found total case loss on nouns to be correlated with the development of a total association or dissociation of gender and declension; this phenomenon is another way in which these more transparent gender-declension relationships are not as closely connected to other

simplifying or transparency-increasing developments. This could be because sufficient transparency in one aspect averts the need for more transparency in another.

The interaction between phonological and morphological factors has the clearest role in gender loss, as with case loss. Declension is generally the least relevant category on nouns, so it is particularly prone to simplification. The same sound changes often neutralize gender distinctions on agreement targets as well as case and declension distinctions on nouns. However, in the core cases, some sound changes have a much greater effect on case distinctions than on gender distinctions.

6.2.3. Refunctionalization and Principles of Selection

My study shows that when the nominal system of a language is interrupted by accidental changes, the remaining forms are often repurposed to make new functions. The loss of distinctions among declension classes for nouns is preferable to both case and number syncretism. This tendency often manifests as the extension of a distinctive marker from one class to one or more classes in which a relevant distinction has been neutralized. Such an overstable marker is more likely to survive in a functional merger of cases or the merger of declension classes, even if it belonged to the more marked (i.e., less frequent) case or class; the latter type of merger can result from a feedback loop by which some overlap in forms between classes can lead to the further extension of forms. When a complete breakdown of case and/or gender marking occurs, overstable markers (and remaining distinctive forms more generally) are often refunctionalized, using principles of selection. For example, the *-s* Gen has become a clitic or phrasal affix in most of the Germanic languages that have lost all case distinctions on nouns. These processes follow the principles of distinctive strength and system adequacy; they can profile number and case together, in contrast to leveling, which usually profiles number at the expense of case. However, some developments in the Germanic, Romance, and Balkan Sprachbund languages appear to have been exceptional in terms of relevance. In the development of OF from WVL, number syncretism actually increased as gender marking was profiled through declension. In the development of CSRm and ModGr, number and gender marking were profiled together at the expense of case. Overall, however, relevance and related principles can account for the vast majority of observed analogical processes in these language groups.

6.2.4. Functional Mergers

Several instances of case loss can be attributed to functional mergers in the languages under investigation. Case variation may result in a functional merger. In addition, the semantic overlap between morphological cases triggers case mergers. In the early stages of the languages, a number of functional mergers occurred. In Greek and Slavic, the Abl merged with the Gen. The Loc merged with the Dat in MyG, BCMS, and Molisean Croatian. In Germanic, the Alb, Instr, and Loc merged with Dat. In Romance languages, the Instr and Loc merged with the Abl.

One of the characteristic features of the Balkan Sprachbund is the functional merger of the Gen and Dat in languages with case marking. This may have spread among the Balkan Sprachbund languages after developing initially in one, or it may have developed by process of structural convergence, the focus of section 3.2.3. This functional merger is not unique to these Balkan Sprachbund languages as it also occurred in other languages outside the Balkan. For example, some Germanic languages experienced the merger of the Gen with the Dat, as in EME. My investigation of the morphological case in OE and EME texts, presented in section 5.2, has provided a more detailed view of the decline of case marking in the period in a language's development leading up to the total loss of morphological case marking. It gives an idea of how case oppositions that are semantically motivated, at least in part, can devolve into (almost) free case variation and the functional narrowing of one case at the expense of another or a prepositional construction. The study shows that even in OE, there is a semantic overlap.

Other functional mergers found in the Balkan Sprachbund include the goal-location merger. This merger also occurred in Germanic languages as the merger of Instr with the Dat. Another functional merger that occurred in the Balkan Sprachbund is the instrument-accompaniment merger. All of these functional mergers occurred in the development of Bulgarian and Macedonian. As it has been argued, these Balkan Slavic languages developed these mergers due to contact. As a result of the lack of language contact in BCMS, it did not develop these mergers. On the other hand, it is possible that the functional mergers developed internally, especially the goal-location and instrument-accompaniment mergers, which are common in IE languages more generally. Other instances of case mergers can be explained by the usage-based approach, as discussed in the following subsection.

6.2.5. Functional Narrowing and Analytic Constructions

Several related aspects and stages of morphosyntactic development often accompany case loss: functional mergers, functional narrowing, and analytic constructions replacing synthetic ones. These developments, which likely all pass through a stage of variation, either among case markers or among case markers and analytic constructions, can often be attributed to phonological change and paradigmatic pressure, the motivations discussed above. Sometimes semantic and/or syntactic overlap provides a better explanation, however, a combination of factors must be assumed. The loss of productivity of the Gen in spoken Faroese through functional narrowing has no basis in sound change; although Faroese did earlier profile number in some classes, this is not a sufficient motivation since Icelandic retains the number syncretism involving the Gen. The loss of the Gen in other Germanic languages may have occurred in a similar way since its forms were quite distinctive, particularly in the singular and in the languages without *-s* plurals. This could also be true for WVL, although all Gen singular forms in masculine and feminine classes were syncretic with the Nom plural, so number profiling can also account for this loss.

My OE/EME study also shows that the Gen and Dat lost their productivity to the Acc object construction in an approximately chronological progression: that is, the Dat also sometimes replaced earlier Gen uses. In OE, the Dat and Gen had already occurred with far fewer verbs than the Acc, and the Gen had the lowest type frequency. Similar results can be gained from Vakareliyska's (1991) study on patients with Wernicke's aphasia, as these patients tend to assign the core meanings of verbs more than the peripheral cases. My findings generally support the idea proposed by Mańczak (1980) that lower-frequency forms are assumed to be more vulnerable to analogical change. Since the oblique cases are less frequent than the core cases, they are more susceptible to being replaced by the core cases. The results also support the usage-based approach proposed by Bybee (1985, 1995) and Barðdal (2009, 2008) that higher frequency forms attract lower frequency forms to be similar. The less frequent case marking was absorbed by the case marking that had a core grammatical role and was more frequent, especially when there was semantic overlap. The less frequent case marking was absorbed by the case marking that had a core grammatical role and was more frequent. For example, in spoken Faroese, the Gen merged with the Acc. The Instr and Loc merged with the Acc in Middle Bulgarian. In ModGr, the Dat, Instr, and Loc were subsumed by the Acc. The change of case-

assigning verbs in favor of core cases is common. For example, in Greek, most prepositions that assigned the Dat and Gen came to assign the Acc instead. In Romanian, verbs that assigned Gen now take an Acc or prepositional complement, not the Dat-Gen case marking. Similarly, OE verbs that historically assigned Gen or Dat came to take Acc instead by EME. These Dat and Gen constructions may have had more in common semantically with higher frequency Acc constructions than with the overlapping meanings of the Dat and Gen.

The outcomes of my investigation also support the study conducted by Harmon and Kapatsinski (2017) that when forms have adequately similar functions, a speaker may merely pick the form which is the most accessible in their lexicon, even if that form is not the best option, possibly causing the expansion of a form's use through recurrent extension. The broadening of forms tends to associate with increasing frequency while narrowing associates with decreased frequency. The corpus study also shows that the Gen and Dat were sometimes replaced by constructions with prepositions, which have narrower meanings than case markings. The use of prepositions instead of case marking supports the proposal by Plank (1979:619) that analytic means must exist before the category is lost. Presumably, similar developments occurred in many of the Germanic, Romance, and Balkan Sprachbund languages.

6.3. External Factors

The role of contact in the loss of case and gender distinctions is subject to considerable debate. The evidence for contact-induced change is often circumstantial at best. However, it is clear through evidence such as loan words that contact among speakers of particular languages occurred. There are several notable contact situations involving the languages under investigation in my study. The influence of Norman French on the ME lexicon is clear, but it there likely was not enough speakers of Norman French in Britain for their use of ME as an L2 to have an effect, and the loss of case and gender categories had already begun in LOE, before their arrival in 1066. Contact with ON speakers in the 9th and 10th centuries more likely had an influence on nominal morphosyntax in ME. There is less historical evidence of a shift of a high number of British Celtic speakers to (pre-)OE in the 4th and 5th centuries than there is for the later contact situations, but if such a shift took place, it might be the accelerant needed to explain why the loss of categories proceeded so much more quickly in English than other Germanic languages. MLG speakers used continental Scandinavian varieties as an L2 starting around 1250

and may have represented a significant portion of the population in Scandinavian cities connected with the Hanseatic League. This was around the time when divergences among the North Germanic languages began to occur, with more innovations occurring in the varieties that had the closest contact with MLG. Later, there was contact between the more innovative Scandinavian varieties, such as Danish, and more conservative varieties, such as Norwegian. Finally, asymmetrical bilingualism of Faroese speakers with the prestige language Danish over the last two centuries has had lexical and possibly structural effects on Faroese.

The Balkan Sprachbund languages are defined by their structural convergences as a result of sustained contact. This began as early as the 7th century and no later than the 11th or 12th century and continued under Ottoman rule until the mid-19th century. The nature of the various contact situations is less clear, but Bulgarian and Macedonian were in the middle of the prestige scale during the Ottoman period, meaning they likely had high numbers of both L2 speakers and bilingual L1 speakers. Modern Slavic languages in established contact situations, such as Molisean Croatian with Italian, have undergone similar changes to those in Bulgarian and Macedonian, providing further evidence for the role of contact. Finally, speakers of languages in the (former) Roman Empire, such as Frankish, shifted to Latin and later Romance varieties, but it is unclear if this had any significant structural effects.

It is very unlikely that any of these contact situations had effects as drastic as creolization. Therefore, contact situations interacted with the internal factors described above, contributing to simplification and other developments in nominal morphosyntax. In a general sense, contact may best be seen as accelerating tendencies already present in a language or resulting in a preference for particular strategies for resolving morphosyntactic inconsistencies. Specific interactions among internal and external factors are discussed in the following subsection.

In Faroese, one potential result of the asymmetrical bilingualism with Danish is the increased acceptability of prepositional constructions to mark recipients instead of the Dat suffix. This would constitute structural convergence with Danish. In fact, the Dat may be gradually disappearing in spoken Faroese; there is no sound change that can account for this, so if the Dat is eventually lost, it will confirm the role of analytic tendencies due to contact. I propose that a similar process was at least partially responsible for the loss of a productive Gen in spoken Faroese, as well as for some of the case loss in the Continental Scandinavian languages that

cannot be explained by sound change and number profiling alone, such as the loss of the Dat in the standard varieties, in contrast to its survival in more isolated inland varieties of Norwegian and Swedish.

A high number of L2 speakers may contribute to functional narrowing. Insights can be gained from the study conducted by Vakareliyska (1991) on Wernicke's aphasics, if L2 speakers are assumed to resemble Wernicke's aphasics in their difficulty accessing peripheral semantic features, which would differ from those in their L1 even if the two languages had similar case systems, then they would sometimes be expected to use a different morphological case. This is just one mechanism by which L2 speakers could motivate case loss. It involves more subtle changes than simply failing to acquire case forms, which might explain why L1 speakers are willing to adopt these changes so readily, a question often left unanswered in proposals of case loss due to L2 speakers.

The role of contact in gender development is even less clear than in case loss. However, it is possible that mutual reinforcement among the Balkan Sprachbund languages has preserved certain conservative features, including a productive ambigeneric class derived from the neuter in Romanian. All Balkan Sprachbund languages have retained the neuter in some form, in contrast to the Western Romance languages. In addition, a similar ambigeneric class has developed in Albanian, potentially due to influence from Romanian; however, it has not yet fully displaced the inherited neuter, providing insight into how the earlier development in Romanian occurred. Other shared conservative features include a reduced case system in all of the Balkan Sprachbund languages except CSB and CSM, which have lost the case system entirely in the nominal paradigms. Thus, despite the sharing between or among these languages of functional mergers and analytic constructions, contact among them may have inhibited total case loss, while a higher number of L2 speakers led to almost complete case loss on nouns in CSB and CSM.

Koineization may account for some early simplification in Greek and, more tentatively, Germanic languages such as English, but it cannot account for case loss on its own. As mentioned above, this could explain why CSM experienced the same case loss as in CSB, despite developing a weaker dynamic stress and not undergoing the same vowel reduction as a result.

6.4. Concluding Thoughts

It is unlikely that any individual factor can account for the degree of case and gender loss in the Germanic, Romance, and Balkan Sprachbund languages. A combination of factors is necessary to account for all of the different outcomes regarding case and gender. Authors who argue that any one factor, whether sound change or external influence, can fully explain these developments in morphosyntax gloss over the nuances involved. Ultimately, speakers must be motivated to simplify or stop using an inflectional paradigm since sound changes with similar effects on distinctions in nominal inflection can lead to very different outcomes. For example, Russian and the other East Slavic languages have retained robust case systems despite undergoing many sound changes similar to those that contributed to the almost complete loss of case distinctions on nouns in Bulgarian and Macedonian. The difference must lie in how speakers react to these neutralizations in response to the accidental change in the paradigmatic system. Their motivations for reacting in different ways are best sought in the morphological interactions among the categories expressed together in portmanteau inflections. My study expanded this approach. I have also tested this approach using original measures of number syncretism on nouns and gender syncretism on agreement targets, in addition to Kürschner and Nübling's gender-declension Types. The results indicate that phonologically and morphologically driven factors can explain many of the divergences among languages, but there is still a role for syntactically driven and language-external factors. Even if these were not the primary cause of case and gender loss, they likely impacted the path these developments took and the rate at which they occurred. Nevertheless, measuring the effects of phonological and morphological changes on syncretism is an integral approach to understanding the patterns of case and gender loss in IE languages and other languages with IE-type portmanteau paradigms, as long as they have the long-written tradition necessary for tracing these developments.

APPENDIX A
GRAMMATICAL DESCRIPTIONS OF EME TEXTS EXAMINED IN CHAPTER V: BRUT AND THE OWL AND THE NIGHTINGALE

Grammatical information about each text except Beowulf, whose grammar is consistent with standard descriptions of OE, is drawn from editions and other relevant publications. Forms obtained from these sources were corroborated and supplemented through searches of an online transcription version of each text. To determine the morphological case marking of forms found in the text, occurrences in unambiguous syntactic contexts were considered first: the subject for the Nom, the DO of normal transitive verbs for the Acc, the object of certain prepositions for the Dat, and adnominal possession for the Gen. When multiple alternate forms were found for a given case/number/gender combination, the most common is listed first, unless indicated otherwise, followed by the next most common, and so on. Letters listed in parentheses are optional: forms occur with them sometimes, but without them other times. The symbol "*" before a form indicates that it is not attested for that case but is expected, based on the forms of other words that are present in the text and/or established syncretism. The symbol "—" indicates that a form is not attested, and there is not enough information to determine what it would be. Alternate forms that only occur once in a text are generally not listed because the numerous slight variations from the much more common forms would clutter the paradigms. An exception is made when none of the alternate forms occur with much frequency.

A1 Layamon's Brut

Grammatical information about the Cotton Caligula manuscript of Brut is drawn from Bøgholm (1944) and the inflections chapter of Burrow and Turville-Petre (2007:19-37). Forms obtained from these sources were corroborated and supplemented through searches of an online transcription version of the manuscript, using the same procedure as the other grammatical descriptions.

Nouns, adjectives, pronouns, and demonstratives could generally distinguish two numbers, singular and plural, and four morphological case markings: Nom, Acc, Gen, and Dat. The status of grammatical gender in Brut is uncertain. Nouns sometimes behave as if they were a

different gender than in OE. For example, some endings, such as Gen singular *-(e)s*, originally confined to the masculine and neuter, are sometimes found on feminine nouns.

A1.1 Nouns

Nouns in the Cotton Caligula manuscript of Layamon's Brut are subject to phonological reduction in unstressed syllables, meaning that *e* is the only vowel letter that appears in endings. However, there are still several case and number endings. Overall, this text is more likely than the other EME texts investigated to use endings with the letter *n* derived from the OE weak declension. Dat singular *-e* and Dat plural *-en* are common on nouns of all genders, although sometimes *-e* is found in the Dat plural as well and *-en* in the Dat singular. The endings *-es* and *-en* are both common for the Nom/Acc plural of all genders and are sometimes used to mark other cases in the plural as well. Neuters also use zero endings in the Nom/Acc plural sometimes, although these are less common than *-es* and *-en*. The distinct Gen plural ending *-ene*, derived from the weak ending *-ena*, and *-e*, from the strong ending *-a*. At least two paradigms for each gender are provided in the following tables: one ending in a consonant in the uninflected form, another ending in *-e*. The paradigm of the common masculine noun *mon* 'person, man', which still shows the effects of *i*-mutation, is also given.

Table 109. Masculine Noun Declension in Brut

Case	'warrior'		'man'		'honor'
	Sg.	Pl.	Sg.	Pl.	Sg.
Nom	cniht	cnihtes, cnihten, cnihte	mon, man	men	wurðscipe
Acc	cniht	cnihtes, cnihten, cnihte	mon, man	men	wurðscipe
Gen	cnihtes	cnihtene, cnihten, cnihte	monnes, mannes	monnen, monne, monnene	*wurðscipes
Dat	cnihte	cnihten, cnihtes	manne, monne	monnen, mannen, monne, men(nen)	wurðscipe, wurðscipen

Only singular forms are attested for *wurðscipe.*

Table 110. Neuter Noun Declension in Brut

Case	'land'		'kingdom'	
	Sg.	Pl.	Sg.	Pl.
Nom	lond	londes, lond	riche	riche, richen
Acc	lond	londes, lond	riche	riche, richen
Gen	londes	londe	*riches	*richen
Dat	londe, lond, londen	londe, lond, londen	riche, richen	*richen

Table 111. Feminine Noun Declension in Brut

Case	'city'		'sorrow'	
	Sg.	Pl.	Sg.	Pl.
Nom	burh	bur(h)ȝen, bur(h)ȝes	sor(h)ȝe, seorwe	sorȝen
Acc	burh, bur(h)ȝe	bur(h)ȝen, bur(h)ȝes	sorȝe, sor(h)ȝen, seorwe	sorȝen
Gen	bur(h)ȝe	bur(h)ȝe	*sorȝe(n)	*sorȝe(n)
Dat	bur(h)ȝe, burh, bur(h)ȝen	bur(h)ȝen	sor(h)ȝe, sorȝen, seorwen	sor(h)ȝen

A1.2 Adjectives

Adjectives still maintain a distinction between strong (indefinite) and weak (definite) declensions, but the endings for both declensions are significantly reduced from OE: vowels are generally reduced to the letter *e* and final consonants are often omitted. Weak adjectives end in *-e*, *-en* or have a zero ending, regardless of gender, number, or morphological case.

Table 112. Strong Adjectives in Brut

Case	Singular			Plural
	Masc.	Neut.	Fem.	All genders
Nom	-Ø	-Ø	-e	-e
Acc	-ne	-Ø	-e	-e
Gen	-e(s)	-e(s)	-(e)re	-(e)re, -ene
Dat	-e(n)	-e(n)	-(e)re	-e(n)

The letter *e* was rarely replaced by *a* or, very rarely, by *æ* in many of the forms. Even strong adjectives sometimes drop the *-e* ending in the feminine and plural, Nom and Acc.

A1.3 Personal Pronouns

Besides having different forms for the four morphological case markings, the personal pronouns also vary by person, as in the other texts investigated. Of the EME texts investigated, Brut is the only one in which pronouns also retain distinct dual forms in addition to singular and plural. The

third-person singular also has distinct forms for masculine, feminine, and neuter. A number of alternate forms are employed for certain case/number/person combinations.

Table 113. Personal Pronouns in Brut

Case	1st Person			2nd Person			3rd Person			
							Sg.			Pl.
	Sg.	Dual	Pl.	Sg.	Dual	Pl.	Masc.	Neut.	Fem.	All genders
Nom	ich, i	wit, wet	we	þu, tu, þeo	ȝit, ȝet	ȝe	(h)e	hit	heo	heo, (h)a
Acc	me	unc(k)	ous, us	þe	inc	eow	(h)ine, hene	hit	heo	heo, (h)a
Gen	*min	unker(e)	ure	*þin	incker(e), unker	*eowre	his	his	(h)ire, heore	he(or)e, hire
Dat	me	unc(k)	ous, us	þe	inc	eow	him, heom	him	he(o)re, hire	h(e)(o)m, ham

This table also highlights the syncretism found in the personal pronoun paradigms. The first- and second-persons use the same form for the Acc and Dat. In the third-person, only the masculine singular distinguishes all four cases. As in OE, the third-person plural and neuter singular each use a single form for the Nom and Acc. In a change from OE, the feminine singular also uses the same form for the Nom and Acc. As in OE, the feminine singular has the same basic forms for the Dat and Gen, although different variations were found in each. The Gen forms in parentheses are not found in non-possessive uses in this text, but the base possessive form is included for reference. Possessives are treated as a separate category in the next section. There is some evidence for the beginnings of a merger between Dat and Acc forms in the third-person, in the dialect of the scribes but not the original author.

A1.4 Possessives

Possessives are similar and often the same as the corresponding Gen pronoun. They vary by person and number.

Table 114. Possessives in Brut

1st Person			2nd Person			3rd Person		
						Sg.		Pl.
Sg.	Dual	Pl.	Sg.	Dual	Pl.	Masc.	Fem.	All genders
mi(n), mine, mines, mire	unker(e)	ure, oure	þi(n), þine, þines, þire	incker(e), unker	eo(w)re, eower	his, hes	(h)ire	he(o)re

The first- and second-person singular possessives have a few alternate forms. Some of these are reflexes of the OE possessive pronouns, which declined like strong adjectives, but others are phonologically conditioned. Forms spelled without a final *n* are used optionally before words that begin with a consonant. Before vowels, only the forms with *n* are used. The forms ending in *-ne* are also used as uninflected forms, whether followed by a consonant or a vowel, and regardless of gender and morphological case. There are two specific inflected forms: one ending in *-re*, that can be used when modifying nouns in the Gen plural or feminine Dat/Gen singular; and another ending in *-nes,* that can be used when modifying masculine/neuter Gen singular nouns. None of the dual and plural possessives, nor the third-person singular possessives, change to agree with the morphological case of the noun they modify. Variations in spelling, including the presence or absence of a final *-e*, do not have grammatical significance.

A1.5 Relative Pronouns

Three indeclinable relative pronouns are used: *þe*, *þa*, and *þat*. The form *þe,* the most common in this manuscript, reflects the indeclinable relative found in OE, while *þa* and *þat* reflect the Nom/Acc plural and neuter singular Nom/Acc, respectively, of the OE demonstrative/relative *se* 'that'. All three forms can be used with antecedents of any gender and number. Another relative derives from the OE interrogative/indefinite pronoun *hwa* 'who' and corresponds to ModE 'who', but this can also have neuter antecedents in this manuscript. The relative *whilc* has an inflected Acc form and is often used as a correlative or indefinite relative.

Table 115. Relative Pronouns in Brut

Case	Forms	
Nom.	wha, w(h)æ	whilc, wulc
Acc	whan, wham	w(h)ulcne
Dat	wham, whan	whulchen

As seen in the table above, there is some confusion between what were originally distinct Acc and Dat forms in OE. The indeclinable form *w(h)at* is also used, including with animate antecedents. For these relatives and the interrogatives below, forms with *h* are generally more common and sometimes the only ones that occur, but there are cases where the situation is reversed and forms without *h* are more common.

A1.6 Interrogatives

Three interrogatives are used in this manuscript: *wha* 'who/what' is an interrogative pronoun; *whulc* 'which' is an interrogative adjective; *whaðer* 'who/what(ever)' is rare and is also found as an indefinite. All three have some declined forms**.**

Table 116. Interrogatives in Brut

Case	Forms		
Nom.	wha, w(h)at	whulc	whaðer, whæðer
Acc.	whan	w(h)ulcne	—
Gen.	whes	w(h)ulche, w(h)ulchere	whaþere
Dat.	whan	w(h)ulcne	—

The form *what* can be used for *wha*, even when referring to people. The Instr forms are only used for the neuter. The gender of the modified noun determines which Dat form of *whulc* is used: *whulce* is used with masculine and neuter nouns while *whulchere* is used with feminine nouns.

A1.7 Demonstratives

The reflexes of the OE demonstrative paradigm *se* 'that' have been adopted as the definite article and is described in section A1.8 below. However, the neuter form *þat* is sometimes used with a demonstrative sense. See section A1.2.6 above for more information about the grammaticalization of demonstratives in EME. It is generally indeclinable but is found with inflected forms such as *þan/þon* for the Dat. The masculine form *þe* can be used as a resumptive pronoun:

(55) *þe ærche-biscop of Lundene eo-d-e an his riht hond-en and bi his luft side*
the Archbishop of London walk-PST-3SG on his right hand-DAT and by his left side

þe of Eouerwik-e
the of York-DAT
'the Archbishop of London walked on his right hand, and by his left side, the same of York' (Brut, ll. 12207-8)

The *þe* before *ærche-biscop* 'archbishop' is a definite article, but the second *þe* is not: it takes the place of *þe ærche-biscop* 'the archbishop' in the second clause.

The reflexes of the OE demonstrative paradigm *þes* 'this' continue to be used with this same meaning, but with fewer gender distinctions. Distinctions in case marking are maintained to a greater extent than the other EME texts investigated.

Table 117. Proximal Demonstrative in Brut

Case	Sg.	Pl.
Nom/Acc	þe(o)s, þis, þus, þas	þas, þeos, þæs
Acc	þesne, þisne	þisse(n)
Gen	þisse(s), þis, þeos, þissere	þissere
Dat	þisse(n), þissere	þisse(n)

Many of the alternate forms are used interchangeably. However, some are distinguished on the basis of the historical gender of the noun they modify. For example, the Acc forms *þesne/þisne* are only used with masculine singular nouns, otherwise one of the Nom forms is used. In the Dat singular, *þisse(n)* is used primarily with masculine and neuter nouns, while *þissere* is used with feminine nouns. Likewise, *þissere* is used with feminines in the Gen singular, while the other forms are used for masculines and neuters.

A1.8 Articles

As mentioned above, the reflexes of the OE demonstrative paradigm *se* 'that' have taken on the role of definite article by this stage of the language.

Table 118. Definite Articles in Brut

Case	Singular			Plural
	Masc.	Neut.	Fem.	All genders
Nom.	þe	þat, þet, þæt	þa, þæ, þeo	þa, þæ, þeo, þaie
Acc.	þane, þene, þ(e)one	þat, þet, þæt	þa, þæ, þeo	þa, þæ, þeo, þaie
Gen.	þas, þes, þæs	þas, þes, þæs	þare, þere, þære	þare, þere
Dat.	þan, þ(e)on, þen	þan, þ(e)on, þen	þare, þere, þære	þan, þ(e)on, þen

For the most part, the same distinctions in case marking are maintained as in OE, but the forms vary in spelling, mainly in terms of vowels, but also including Dat forms with *-n* instead of *-m*. One distinction is lost: feminine singular Nom and Acc are now syncretic, as in the third-person pronoun paradigm. Forms are sometimes used in the incorrect morphological case or as uninflected forms, e.g., *þet* and *þone*. This manuscript uses the reflexes of OE *an* 'one' as an indefinite article:

Table 119. Indefinite Articles in Brut

Case	Singular		
	Masc.	Neut.	Fem.
Nom.	a(n)	a(n)	a(n)
Acc.	an(n)e, æn(n)e, enne, one	a(n)	ane, æne, enne
Gen.	anes	anes	are, ære
Dat.	an(n)e, æn(n)e, enne, one	an(n)e, æn(n)e, one	are, ære

In the uninflected form, the letter *n* can be dropped optionally before words that begin with a consonant. Before vowels (or the letter *h*), only the form with *n* is used.

A2 The Owl and the Nightingale (O&N)

Grammatical information about O&N is drawn from the Inflexions chapter of Burrow and Turville-Petre (2007:19-37). Forms obtained there were corroborated and supplemented through searches of an online transcription version of the manuscript, using the same procedure as the other grammatical descriptions.

A2.1 Nouns

Nouns in this text are subject to phonological reduction in unstressed syllables, meaning that *e* is the only vowel letter that appears in endings. However, there are still several case and number

endings. The ending *-e* is used for the Dat singular of all genders and is sometimes found in the Dat plural as well. Of the paradigms listed, Dat plural *-en* only appears on *sunne* 'sun'. The ending *-es* is the most common for the Nom/Acc plural of all genders and is often used to mark other cases in the plural as well. The historically neuter *word* 'word' is attested in an Acc plural context with a zero ending, but *-es* is more common. A paradigm representative of each gender is given. The paradigm of the common masculine noun *mon* 'person, man', which still shows the effects of *i*-mutation, is also given.

Table 120. Masculine Noun Declension in O&N

Case	'song'		'man'	
	Sg.	Pl.	Sg.	Pl.
Nom	song	*songes	mon, man	men
Acc	song	songes	mon, man	men
Gen	songes	*songes	monnes, mannes	monne, manne
Dat	songe	*songes	men, manne	monne, manne, men

Table 121. Neuter Noun Declension in O&N

Case	'word'	
	Sg.	Pl.
Nom	word, worde	words
Acc	word	wordes, word
Gen	*wordes	*words
Dat	worde, word	worde

Table 122. Feminine Noun Declension in O&N

Case	'sun'	
	Sg.	Pl.
Nom	*sunne	sunne
Acc	*sunne	sunne
Gen	*sunne	*sunne
Dat	sunne	sunnen

The form *tale* is better attested in the singular, *sunne* in the plural.

A2.2 Adjectives

Adjectives in this text also maintain a distinction between strong (indefinite) and weak (definite) declensions. However, the strong declension has far fewer distinctions in form than Brut. In fact,

it only has two forms: *-e* for all forms except the Nom singular of all genders and the Acc singular of masculine and neuter, all of which have a zero ending. These are also the only forms where the weak declension is distinct from the strong, since the former has *-e* for all forms.

A2.3 Personal Pronouns

The personal pronouns found in this text are very similar to those in Brut, with the exception of a few spelling differences and alternate forms. However, unlike Brut, most of the dual forms are not present. Only *unker*, the Gen/possessive of the first-person dual, is found.

Table 123. Personal Pronouns in O&N

Case	1st Person		2nd Person		3rd Person			
	Sg.	Pl.	Sg.	Pl.	Sg.			Pl.
					Masc.	Neut.	Fem.	All genders
Nom	ich, i	we	þu	ȝe	he	hit	ho, he(o), hi	hi, h(e)o
Acc	me	us	þe	ow	hine	hit	hi, heo	hi
Gen	mi(n)	ure	þin	ower	his	his	hire, hore	h(e)ore
Dat	me	us	þe	ow	him	him	hire	h(e)om

A2.4 Possessives

Possessives are similar and often the same as the corresponding Gen pronoun. They vary by person and number.

Table 124. Possessives in O&N

1st Person			2nd Person		3rd Person		
Sg.	Dual	Pl.	Sg.	Pl.	Sg.		Pl.
					Masc.	Fem.	All genders
mi(n), mine, mire	unker	ure	þi(n), þine, þire	ower	his	hire, hore	h(e)ore

As in Brut, the first- and second-person singular possessives have a few alternate forms, mostly phonologically conditioned. Forms spelled without a final *n* are used optionally before words that begin with a consonant. Before vowels, only the forms with *n* are used. The forms ending in *-ne* are also used as uninflected forms, whether followed by a consonant or a vowel. There is one specific inflected form ending in *-re*, which is attested with feminine Dat singular nouns. None of the dual and plural possessives, nor the third-person singular possessives, change to agree with the morphological case of the noun they modify.

A2.5 Relative Pronouns

Two indeclinable relative pronouns are used: *þat* and *þe*. The form *þat*, the more frequent variant in this text, reflects the neuter singular Nom/Acc of the OE demonstrative/relative *se* 'that' but can be used with antecedents of any gender and number. The form *þe*, which reflects the indeclinable relative found in OE, is only attested once, used with a plural antecedent.

A2.6 Demonstratives

The demonstrative derived from the paradigm of OE *þes* 'this' has fewer distinct forms in O&N than in Brut. The demonstrative *þis* is used as an uninflected form. In the first example below, it is used with the historically neuter Dat singular noun *spelle* 'story', while in the second example, it is used with the historically feminine Gen singular noun *worldes* 'world':

(56) *her n-is na-more of þis spell-e*
here NEG-is NEG-more of this story-DAT.SG
'Here there is no more of this story' (O&N, l. 1794)

(57) *al þis world-es blis*
all this world-GEN.SG joy
'all (of) this world's joy' (O&N, l. 1280)

The demonstrative *þes* is used as a masculine form (Acc in the one attestation in this text). The demonstrative *þos/þeos* is used as a plural and feminine singular form. Finally, *þisse* is used as a Dat form and is only attested with neuter nouns in this manuscript.

A2.7 Articles

The forms of the definite article that are found in this text are similar to those in Brut, with spelling differences analogous to the personal pronouns and demonstratives. However, certain forms are not attested but are assumed based on the patterns of syncretism found in Brut.

Table 125. Definite Article in O&N

Case	Singular			Plural
	Masc.	Neut.	Fem.	All genders
Nom	þe	þat	þo	þeo
Acc	þane, þene	þat	*þo	þeo
Gen	þas	*þas	þare	*þare
Dat	þan, þon, þen	þan, þon	þare	þan, þare

By far the most common form of the definite article is *þe*, which can be considered uninflected as it is used without reference to gender or morphological case. This text also has an indefinite article, but with far fewer forms than the one in Brut. The articles *a*, *an* and *one* are used as uninflected forms. The form *a* can be used before words that begin with a consonant. Before vowels (or *h*), only *an* and *one* are used. There are two specific inflected forms that are used on occasion: *an(n)e* for the masculine Acc and *ore* for the feminine Dat.

APPENDIX B
LIST OF ABBREVIATIONS

1	1st person
2	2st person
3	3st person
Abl	ablative
Acc	accusative
AG	Ancient Greek
Amb	Ambiguous
BChS	Bulgarian Church Slavonic
BCMS	Bosnian-Croatian-Montenegrin-Serbian
CL	Classical Latin
COMP	Complementizer
CSA	Contemporary Standard Albanian
CSB	Contemporary Standard Bulgarian
CSDan	Contemporary Standard Danish
CSDu	Contemporary Standard Dutch
CSF	Contemporary Standard French
CSHG	Contemporary Standard High German
CSI	Contemporary Standard Italian
CSM	Contemporary Standard Macedonian
CSR	Contemporary Standard Russian
CSRm	Contemporary Standard Romanian
CSS	Contemporary Standard Swedish
Dat	dative
DO	direct object
EME	Early Middle English
EModSw	Early Modern Swedish
ES	Early Spanish
EVL	Eastern Vulgar Latin
MLG	Middle Low German
MNw	Middle Norwegian
ModE	Modern English
ModGr	Modern Greek
ModLG	Modern Low German
ModSp	Modern Spanish
MSrb	Middle Serbian
MSw	Middle Swedish
MyG	Mycenaean Greek
NEG	negative
NEUT	neuter
NNw	Neo-Norwegian
Nom	nominative
NP	noun phrase
OCS	Old Church Slavonic
ODan	Old Danish
ODu	Old Dutch
OE	Old English
OF	Old French
OHG	Old High German
OIc	Old Icelandic
ON	Old Norse
ONw	Old Norwegian
OS	Old Saxon
OSc	Old Scandinavian
OSw	Old Swedish
PDE	Present Day English
PFV	Perfective Aspect of Verbs

FEM	feminine
Gen	genitive
IE	Indo-European
INF	infinitive
Instr	instrumental
IO	indirect object
KG	Koine Greek
LME	Late Middle English
Loc	locative
LOE	Late Old English
LPS	Late Proto-Slavic
MASC	masculine
MB	Middle Bulgarian
MC	Molisean Croatian
MDan	Middle Danish
MDu	Middle Dutch
ME	Middle English
MF	Middle French
MHG	Middle High German
PGmc	Proto-Germanic
PIE	Proto-Indo-European
PL	plural
PMB	Pre-Middle Bulgarian
PP	past participle
Prep	prepositional
PRES	present
PRESP	present participle
PS	Proto-Slavic
PST	past
SG	singular
SUBJ	subjunctive
VL	Vulgar Latin
Voc	vocative
WVL	Western Vulgar Latin
?	a form of uncertain identity
??	a form of more uncertainty
???	unreconstructable form
—	a form is not attested

APPENDIX C

TRANSLITERATION OF OLD CYRILLIC

Old Cyrillic	Roman transliteration
А а	a
Г г	g
Д д	d
О о	o
Р р	r
С с	s
Т т	t
Х х	x
Ѡ ѡ	o
ъ	ŭ (*back jer*)
ь	ĭ (*front jer*)
Ꙗ ꙗ	ja
Ѧ ѧ	ę, ję (*jus minor*)
Ѫ ѫ	ǫ, jǫ (*jus major*)

APPENDIX D
LIST OF OE AND EME TEXTS

1. Beowulf

Beowulf was written between the first half of the 8th century and 11th century. The only extant copy is found on folios 129-198 of the British Library MS Cotton Vitellius A.xv, (Klaeber 2008., xxv). Nowell Codex is a second manuscript of the Cotton Vitellius A.xv that contains Beowulf. This manuscript was part of the Cotton collection before being donated to Great Britain in 1702 by Sir John Cotton, the grandson of the collections creator, Sir Robert Cotton, as the beginning of the British Library. It consists of 3182 lines written in the accent- and alliteration-based meter used for all OE poetry.

Edition:

Klaeber, Friedrich, John D. Niles, Robert E. Bjork, and R. D. Fulk. (2008). *Klaeber's Beowulf and the Fight at Finnsburg*. (4th ed). Toronto: University of Toronto Press.

2. Layamon's Brut

Layamon's Brut was likely written in the second half of the 13th century, after the death of King Henry II in 1189 (Burrow and Turville-Petre 2007:96). It is part of the British Library MS Cotton Caligula A.ix, which was part of the Cotton collection, as was the surviving copy of Beowulf. The manuscript is a small vellum quarto that contains 259 folios, 192 of which form the Brut (Le Saux 1994:1).

Layamon's Brut is an English adaptation and expansion of the French poem Roman de Brut, which was written in octosyllabic couplets and completed by the Norman poet Wace in 1155. Roman de Brut itself was based on Geoffrey of Monmouth's Latin History of the Kings of Britain, which was first issued around 1136. It describes the kings of Britain beginning with their legendary founder Brutus, including the hero Arthur, and ending with Cadwallader, who lost Britain to the Saxons (Burrow and Turville-Petre 2007:96). Brut is a poem of 16,096 lines, written in a style combining native alliterative poetry with French/Latin rhymed poetry: lines are divided into half-lines that are usually coupled by alliteration and sometimes also rhyme or half rhyme (ibid., 97).

Editions:

Brook, G. L. and R. F. Leslie. (1963). *Laȝamon: Brut*, vol. 1 (lines 1-8020). London/New York/Toronto: Oxford University Press.

Brook, G. L. and R. F. Leslie. (1978). *Laȝamon: Brut*, vol. 2 (lines 8021-end). London/New York/Toronto: Oxford University Press.

3. The Owl and the Nightingale

O&N was written sometime after the death of Henry II in 1189 and probably before Henry III acceded to the throne in 1216. The version of the text investigated is based on the manuscript that is part of the British Library MS Cotton Caligula A.ix, which was copied in the second half of the 13th century and also contains Layamon's Brut (Burrow and Turville-Petre 2007:81). O&N consists of 14 folios near the end of the manuscript (Atkins 2014:xxii). It contains 1794 lines that form rhyming octosyllabic couplets, the primary metrical form used by French poets of the time and is a debate poem between an owl and a nightingale, a genre that Latin poets favored at that time (Burrow and Turville-Petre 2007:81).

Edition:

Cartlidge, Neil. (2001). *The Owl and the Nightingale: Text and Translation*. Exeter: University of Exeter Press.

REFERENCES CITED

Adrados, Francisco Rodríguez. (2005). *A History of the Greek Language: from its Origins to the Present*. Leiden: Brill.

Albright, Adam. (2008). "Explaining Universal Tendencies and Language Particulars in Analogical Change." In: Jeff Good (ed.), *Linguistic Universals and Language Change*. Oxford: Oxford University Press, 144-184.

Alkire, Ti and Carol Rosen. (2010). *Romance languages: A Historical Introduction*. Cambridge University Press.

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

Allen, Cynthia L. (2002). "Case and Middle English Genitive Noun Phrases." In David Lightfoot (ed.), *Syntactic Effects of Morphological Change*. Oxford: Oxford University Press, 57-80.

Allen, Cynthia L. (2008). *Genitives in Early English: Typology and Evidence*. Oxford/New York: Oxford University Press.

Allen, William Sidney. (1973). *Accent and Rhythm*. Cambridge: Cambridge University Press.

Anderson, George K. (1958). "*The Fifth Case in Old English." The Journal of English and Germanic Philology* 57(1): 21-26.

Andersson, Erik. (1994). "Swedish." In: Ekkehard König and Johan van der Auwera (eds.), *The Germanic Languages*. London*:* Routledge, 271-312.

Andreose, alvise and lorenzo Renzi. (2013). "Geography and Distribution of the Romance Languages in Europe." In: Martin Maiden, John Charles Smith and Adam Ledgeway. *The Romance Languages*, vol. 2. Cambridge: Cambridge University Press, 283-334.

Aristar, Anthony Rodrigues. (1997). "Marking and Hierarchy Types and the Grammaticalization of Case-Markers." *Studies in Language* 21(2): 313-368.

Arnhold, Anja. (2014). "Finnish Prosody: Studies in Intonation and Phrasing." PhD Dissertation, Univ.-Bibliothek Frankfurt am Main.

Asenova, Petja. (2002). Balkansko ezikoznanie: Osnovni problemi na balkanskija ezikov săjuz. (2nd edn.). Veliko Tărnovo: Faber.

Askedal, John Ole. (1994). "Norwegian." In: Ekkehard König and Johan van der Auwera (eds.), *The Germanic Languages*. London*:* Routledge, 219-270.

Atkins, John W. H. (2014). *The Owl and the Nightingale*. Cambridge: Cambridge University Press.

Audring, Jenny. (2006). "Pronominal Gender in Spoken Dutch." *Journal of Germanic Linguistics* 18(2): 85-116.

Aylett, Matthew and Alice Turk. (2006). "Language Redundancy Predicts Syllabic Duration and the Spectral Characteristics of Vocalic Syllable Nuclei." *The Journal of the Acoustical Society of America* 119(5): 3048-3058.

Baerman, Matthew, Dunstan Brown and Greville G. Corbett. (2005). *The Syntax-Morphology Interface: A Study of Syncretism*, vol. 109. New York: Cambridge University Press.

Balode, Laimute, and Holvoet, Axel. (2001). "The Lithuanian Language and its Dialects." *Circum-Baltic Languages: Typology and Contact*: 41-80.

Banniard, Michel. (2013). "The Transition from Latin to the Romance Languages." In: Martin Maiden, John Charles Smith and Adam Ledgeway (eds.), *The Romance Languages*, vol. 2. Cambridge: Cambridge University Press, 57-106.

Barðdal, Jóhanna and Leonid Kulikov. (2009) "Case in Decline." In: Malchukov and Spencer (eds.), *The Oxford Handbook of Case*. New York: Oxford University Press, 470-478.

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

Barðdal, Jóhanna. (2009). "The Development of Case in Germanic." In: Jóhanna Barðdal and Shobhana L. Chelliah (eds.), *The Role of Semantics and Pragmatics in the Development of Case*. (Studies in Language Companion Series 108). Amsterdam: John Benjamins, 123-159.

Barnes, Johnathan. (2006). *Strength and Weakness at the Interface: Positional Neutralization in Phonetics and Phonology*. Berlin/New York: Mouton de Gruyter.

Barnes, Michael P., and Eivind Weyhe. (1994). "Faroese." In: Ekkehard König and Johan van der Auwera (eds.), *The Germanic Languages*. London: Routledge, 190-218.

Baron, Naomi S. (1971). "A Reanalysis of English Grammatical Gender." *Lingua* 27: 113-140.

Behaghel, Otto. (1928). *Deutsche Syntax: Eine geschichtliche Darstellung*, vol. 3. Heidelberg: C. Winter.

Beito, Olav. (1957). "Har alle endringar i eit morfologisk system opphav i innhaldsplanet?." *Historisk språkvitenskap/Historical Linguistics*, 207-214.

Belić, Aleksandar. (1960). *Osnovi istorije srpskohrvatskog jezika. I: Fonetika*. Beograd.

Belić, Aleksandar. (1962). *Istorija srpskohrvatskog jezika*. Knj. II, Sv. 1: Reči sa deklinacijom. Beograd.

Bennett, Charles Edwin. (1913). *New Latin Grammar*. Boston and Chicago: Allyn and Bacon.

Berg, Ivar. (2015). "Stages in Deflexion and the Norwegian Dative." In: Dag T.T. Haug (ed.), *Historical Linguistics 2013: Selected papers from the 21st International Conference on Historical Linguistics, Oslo, 5-9 August 2013.* (Current Issues of Linguistic Theory 334). Amsterdam/Philadelphia John Benjamins, 179-194.

Berman, Ruth. A. (1985). "The Acquisition of Hebrew." In: Dan I. Slobin (ed.), *The Crosslinguistic Study of Language Acquisition*, vol. 1. Hillsdale, NJ: Lawrence Erlbaum Associates, 255-371.

Besters-Dilger, Juliane and Kurt Braunmüller. (2014). "Sociolinguistic and Areal Factors Promoting or Inhibiting Convergence within Language Families." In: Juliane Besters-Dilger, Cynthia Dermarkar, Stefan Pfänder and Achim Rabus (eds.), *Congruence in Contact-Induced Language Change: Language Families, Typological Resemblance, and Perceived Similarity* (Linguae & Litterae 27). Berlin/Boston: Walter de Gruyter, 390-410.

Bickel, Balthasar, Alena Witzlack-Makarevich and Taras Zakharko. (2014). "Typological Evidence Against Universal Effects of Referential Scales on Case Alignment." In: Ina Bornkessel-Schlesewsky, Andrej L. Malchukov and Marc D. Richards (eds.), *Scales and Hierarchies: A Cross-Disciplinary Perspective* (Trends in Linguistics: Studies and Monographs). Berlin/Munich/Boston: Walter de Gruyter, 7-44.

Blake, Barry J. (2004). *Case*. Cambridge: Cambridge University Press.

Bøgholm, Niels. (1944). *The Layamon Texts: A Linguistical Investigation*. Copenhagen: J. Jorgensen & Co.

Börjars, Kersti and Nigel Vincent. (2011). "Grammaticalization and Directionality." In: Bernd Heine and Heiko Narrog (eds.) *Oxford Handbooks Online*. Oxford/New York: Oxford University Press, 1-10.

Braunmüller, Kurt. (1989). "Voraussetzungen für die Übernahme mittelniederdeutscher Sprachstrukturen in die skandinavischen Sprachen." In: Hyldgaard-Jensen, Karl, Winge, Vibeke och Christensen, Birgit (eds.), *Niederdeutsch in Skandinavien II* (= Beihefte zur Zeitschrift für Deutsche Philologie 5), Berlin: Erich Schmidt Verlag: 9-29.

Braunmüller, Kurt. (1995). "Semikommunikation und semiotische Strategien. Bausteine zu einem Modell für die Verständigung im Norden zur Zeit der Hanse." In: Braunmüller, Kurt (ed.), Niederdeutsch und die skandinavischen Sprachen II. Heidelberg: C. Winter, 35-70.

Breu, Walter. (2008). “Der slavische Lokativ im Sprachkontakt. Ein Beitrag zur Binnendifferenzierung des Moliseslavischen.” In: Kosta, Peter and Daniel Weise (eds.), *Slavistische Linguistik* 2006/2007. München: Sagner, 59-101.

Brown, Dunstan. (1998). From the General to the Exceptional: A Network Morphology Account of Russian Nominal Inflection. PhD thesis, University of Surrey.

Brown, Wayles, and Theresa Alt. (2004). *A Handbook of Bosnian, Serbian, and Croatian*. Durham: Duke University.

Browne, Wayles. (1993). “Serbo-Croat.” In: Bernard Comrie and Greville Corbett (eds), *The Slavonic Languages.* London/New York: Routledge, 306-387.

Brozović, Dalibor and Pavle Ivić. (1988). *Jezik srpskohrvatski/hrvatskosrpski, hrvatski ili srpski*. Zagreb: Jugoslavenski leksikografski zavod.

Bubenik, Vit. (2006). “From Old to Modern Slavic.” *From Case to Adposition: The Development of Configurational Syntax in Indo-European Languages.* Amsterdam/Philadelphia: John Benjamins, 178-203.

Buk, Solomija, Ján Mačutek and Andrij Rovenchak. (2008). “Some Properties of the Ukrainian Writing System.” *Glottometrics* 16: 63-79.

Bungenstab, Edith. (1933). *Der Genitiv beim Verbum und sein Ersatz im Laufe der englischen Sprachgeschichte*. Friedrich-Wilhelms-Universität zu Breslau.

Burrow, John A. and Thorlac Turville-Petre. (2007). *A Book of Middle English*. Oxford: Joh Wiley & Sons.

Bybee, Joan L. (1985). *Morphology: A Study of the Relation between Meaning and Form*, vol. 9. Amsterdam/Philadelphia: John Benjamins.

Bybee, Joan L. (1995). “Regular Morphology and the Lexicon.” *Language and Cognitive Processes* 10(5): 425–55.

Bybee, Joan L. and Mary Alexandra Brewer. (1980). “Explanation in Morphophonemics: Changes in Provençal and Spanish Preterite Forms.” *Lingua*, 52(3-4), 201-242.

Bybee, Joan L., Paromita Chakraborti, Dagmar Jung and Joanne Scheibman. (1998). “Prosody and Segmental Effect: Some Paths of Evolution for Word Stress.” *Studies in Language. International Journal sponsored by the Foundation “Foundations of Language”* 22(2): 267-314.

Campbell, Alistair. (1959). *Old English Grammar*. Oxford: Clarendon Press.

Campbell, Lyle. (2013). *Historical linguistics*. Edinburgh: Edinburgh University Press.

Çanta, Agnesa. (2017). "The Category of Case in English and Albanian Nominal System: A Contrastive Analysis." *International Journal of English Linguistics* 7(1): 226-239.

Carstairs, Andrew. (1987). "Diachronic Evidence and the Affix–Clitic Distinction." In: Anna Giacalone Ramat, Onofrio Carruba, and Giuliano Bernini (eds.), *Papers from the 7th International Conference on Historical Linguistics*. Amsterdam and Philadelphia: Benjamins, 151–162.

Cassano, Paul. (1976). "Theories of Language Borrowing Tested by American Spanish Phonology." *Romance Philology* 30(2):331–42.

Catasso, Nicholas. (2011). "Genitive-Dative Syncretism in the Balkan Sprachbund: An Invitation to Discussion." *SKASE Journal of Theoretical Linguistics* 8(2): 70-93.

Civ'jan, Tat'jana. (1965). *Imja suščestvitel'noe v balkanskix jazykax: k strukturnotipologičeskoj xarakteristike balkanskogo jazykovogo sojuza*. Moscow: Nauka.

Clackson, James. (2007). *Indo-European Linguistics: An Introduction*. New York: Cambridge University Press.

Clark, Eve V. (1985). "The Acquisition of Romance, with Special Reference to French." In Dan I. Slobin (ed.), *The Crosslinguistic Study of Language Acquisition*, vol. 1. Hillsdale, NJ: Lawrence Erlbaum Associates, 687-782.

Corbett, Greville G. (1991). *Gender*. Cambridge/New York/Melbourne: Cambridge University Press.

Corbett, Greville G. (2013). "Number of Genders." In: Dryer, Matthew S. and Haspelmath, Martin (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at http://wals.info/chapter/30, Accessed on 2022-07-05).

Curzan, Anne. (2003). *Gender Shifts in the History of English*. New York: Cambridge University Press.

Cyxun, Genadij A. (1968). *Sintaksis mestoimennyx klitik v juznoslavjanskix jazykax*. Moscow: Nauka i texnika.

Dahl, Östen. (2004). *The Growth and Maintenance of Linguistic Complexity* (Studies in Language Companion Series 71). Amsterdam: John Benjamins.

Dahl, Östen. (2009). "Two Pathways of Grammatical Evolution." In: T. Givón, Masayoshi Shibatani (eds.), *Syntactic Complexity Diachrony, Acquisition, Neuro-Cognition, Evolution*. Amsterdam/Philadelphia: John Benjamins, 239-248.

Danylenko, Andriy. (2015). “On the Mechanisms of the Formation of the Polysemy of Comitative and Instrumental Categories in Slavic.” *Journal of Historical Linguistics* 5(2): 267-296.

Dawson, Hope C. (2003). “Defining the Outcome of Language Contact: Old English and Old Norse.” *Working Papers in Linguistics*, vol. 57: 40-57.

De Boel, Gunnar. (2008). “The Genesis of Clitic Doubling from Ancient to Medieval Greek.” In: Dalina Kallulli and Liliane Tasmowski (eds.), *Clitic Doubling in the Balkan Languages* 130. Amsterdam/Philadelphia: John Benjamins. 89-104.

De Schutter, Georges. (1994). “Dutch.” In: Ekkehard König and Johan van der Auwera (eds.), *The Germanic Languages.* London: Routledge, 439-477.

De Vos, Lien and Gunther De Vogelaer. (2011). “Dutch Gender and the Locus of Morphological Regularization.” *Folia Linguistica* 45(2): 245-281.

Delsing, Lars-Olof. (2002). “The Morphology of Old Norse II: Old Swedish and Old Danish.” In: Herbert Ernst Wiegand (ed.), *The Nordic Languages. An International Handbook of the History of the North Germanic.* Berlin/New York: Walter de Gruyter.

Demina, Evgenija I. (1971). *Tixonravovskij damaskin. Bolgarskij pamjatnik XVII v. Issledovanie i tekst. Čast’ II. Paleografičeskoe opisanie i tekst*. Sofia: Izdatelstvo na BAN.

Demiraj, Shaban. (1993). *Historische Grammatik der albanischen Sprache*. (Schriften der Balkan Kommission, 34). Vienna: Verlag der Österreichischen Akademie der Wissenschaften.

Dimitrescu, Florica, Viorica Pamfil, Elena Barborică, Maria Cvasnîi, Mirela Theodorescu, Cristina Călăraşu, Mihai Marta, Elena Toma, and Liliana Ruxăndoiu (eds.). (1978). *Istoria limbii române: Fonetică, Morfosintaxă, Lexic*. Bucharest: Editura Didactică şi Pedagogică.

Donaldson, Bruce. (1994). “Afrikaans.” In: Ekkehard König and Johan van der Auwera (eds.), *The Germanic Languages*. London: Routledge, 478-504.

Dresher, B. Elan and Aditi Lahiri. (1991). “The Germanic Foot: Metrical Coherence in Old English.” *Linguistic Inquiry* 22(2): 251-286.

du Nay, André. (1996). *The Origins of the Rumanians: The Early History of the Rumanian Language*. Toronto/Buffalo: Matthias Corvinus Publishing.

Duridanov, Ivan. (1956). “Kăm problemata za razvoja na bălgarski ezik ot sintetizăm kăm analitizăm.” *Godišnik na sofijskija universitet, Filologičeski fakultet*, 51(3): 87-257.

Dye, Melody, Petar Milin, Richard Futrell, and Michael Ramscar. (2017). "A Functional Theory of Gender Paradigms." In: Ferenc Kiefer, James Blevins and Huba Bartos (eds.), *Perspectives on Morphological Organization.* (Empirical Approaches to Linguistic Theory 10). Leiden: Brill, 212-239.

Eisenberg, Peter. (1994). "German." In: Ekkehard König and Johan van der Auwera (eds.), *The Germanic Languages*. London: Routledge, 349-387.

Enger, Hans-Olav. (2004). "On the Relation Between Gender and Declension. A Diachronic Perspective from Norwegian." *Studies in Language* 28(1): 51–82.

Enger, Hans-Olav. (2013). "Inflectional Change, 'Sound Laws' and the Autonomy of Morphology: The Case of Scandinavian Case and Gender Reduction." *Diachronica* 30(1): 1-26.

Faarlund, Jan Terje. (1994). "Old and Middle Scandinavian." In: Ekkehard König and Johan van der Auwera (eds.), *The Germanic Languages*. London: Routledge, 142-189.

Faarlund, Jan Terje. (2004). *The Syntax of Old Norse: With a Survey of the Inflectional Morphology and a Complete Bibliography*. Oxford: Oxford University Press.

Falaleev, Pavel. (2016). "Turkish loanwords in Macedonian: Speakers' Awareness and Attitudes." In: Maxim Makartsev and Max Wahlström (eds.), *In Search of the Center and Periphery: Linguistic Attitudes, Minorities, and Landscapes in the Central Balkans, Slavica Helsingiensia* 49. Finland: University of Helsinki, 115-124.

Fortson, Benjamin W. IV. (2004). *Indo-European Language and Culture: An Introduction*. West Sussex: John Wiley & Sons.

Friedman, Victor A. (1993). "Macedonian." In: Bernard Comrie and Greville Corbett (eds), *The Slavonic Languages.* London/New York: Routledge, 249-305.

Friedman, Victor A. (1999). "Evidentiality in the Balkans." In: Uwe Hinrichs (ed.), *Handbuch der Südosteuropa-Linguistik*. (Slavistische Studienbücher, Neue Folge 10.) Wiesbaden: Harrassowitz, 519–544.

Friedman, Victor A. (2006). "West Rumelian Turkish in Macedonia and Adjacent Areas." In: Hendrik Boeschoten, Lars Johanson and Vildan Milani (eds), *Turkic Languages in Contact* (Turcologica 61), 27–45. Wiesbaden: Harassowitz.

Friedman, Victor A. (2008). "Balkan Object Reduplication in Areal and Dialectological Perspective." In: Dalina Kallulli and Liliane Tasmowski (eds.), *Clitic Doubling in the Balkan Languages* (Linguistik aktuell vol. 130). Amsterdam: Benjamins, 35–63.

Friman, Helena and Göran Söderström. (2008). *Stockholm: En historia i kartor och bilder. Monografier utgivna av Stockholms stad*. Bonnier Fakta i samarbete med Stockholms stadsmuseum och Stockholmia förlag.

Fuchsbauer, Jürgen. (2018). "Some Observations on the Usage of Adnominal Genitives and Datives in Middle Bulgarian Church Slavonic." In: Björn Hansen, Jasmina Grković-Major and Barbara Sonnenhauser (eds.), *Diachronic Slavonic Syntax: The Interplay between Internal Development, Language Contact and Metalinguistic Factors*. Berlin/Boston: Walter de Gruyter, 13-28.

Fykias, Ioanis and Christina Katsikadeli. (2012). "The Form and Distribution of Some Neglected Morphosyntactic Categories in the History of Greek." *Selected Papers of the 10th International Conference of Greek Linguistics*. Komotini: Demokritus University of Thrace, 255-264.

Gallée, Johan Hendrik. (1891). *Altsächsische Grammatik*. Vol. 1. Halle: Niemeyer.

Garbacz, Piotr. (2014). "Morphological Dative in Norwegian Dialects." *Nordic Atlas of Language Structures Journal*, 1: 72-86.

Gasiński, Thaddeus Z. (1986) "English Elements in the Speech of the Croatian Immigrant Community of Santa Clara Valley, California." *Zborik matice srpske za filologiju i lingvistiku* (Novi Sad) 29(2): 31-45.

Georgiev, Vladimir. (1952). "Opit za periodizacija na istorijata na bălgarskija ezik." *Izvestija na Instituta za bălgarski ezik* 2: 71-116.

Getz, Robert et al. (ed.). Dictionary of Old English. Toronto: University of Toronto. https://doe.artsci.utoronto.ca/. Accessed 21 May 2022.

Gianollo, Chiara. (2010). "External Possession in New Testament Greek." In: Gualtiero Calboli and Pierluigi Cuzzolin (eds.), *Papers on Grammar* XI. Roma: Herder, 101-129.

Grandgent, Charles Hall. (1907). *An Introduction to Vulgar Latin*. Boston: D.C. Heath & Company.

Green, John N. (1988). "Spanish." In: Martin Harris and Nigel Vincent (eds.), *The Romance Languages*. London/New York: Routledge, 79–130.

Greenberg, Joseph. (1966). *Language Universals*. The Hague: Mouton.

Grimm, Scott. (2011). "Semantics of Case." *Morphology* 21(3): 515-544.

Gruber-Miller, Ann M. (1990). "Loss of Nominal Case Endings in the Modern Arabic Sedentary Dialects: Evidence from Southern Palestinian Christian Middle Arabic Texts." In: Mushira Eid (ed.), *Perspectives on Arabic Linguistics I: Papers from the First Annual Symposium on Arabic Linguistics.* (Current Issues in Linguistic Theory 63). Amsterdam: John Benjamins, 235-253.

Guentchéva, Zlatka. (2004). "La notion de médiation dans la diversité des langues." In: Régine Delamotte-Legrand (ed.), *Les médiations langagières Volume I: Des faits de langue aux discours*. Rouen: Publications de l'Université de Rouen, 11-33.

Gumperz, John J. and Robert Wilson. (1971). "A Case from the Indo-Aryan/Dravidian Border in India." In: Dell Hymes (ed.), *Pidginization and Creolization of languages*. Cambridge: Cambridge University Press, 151-168.

Gvozdanovic, Jadranka. (I989). "Defining Markedness." In: Miseska Tomic (ed.), *Markedness in Synchrony and Diachrony* (Trends in Linguistics Studies and Monographs 39). Berlin/New York: Mouton de Gruyter, 47-66.

Gyllin, Roger. (1991). *The Genesis of the Modern Bulgarian Literary Language*. (Acta Universitatis Upsaliensis – Studia Slavica Upsaliensia 30). Uppsala: Uppsala University.

Haberland, Hartmut. (1994). "Danish." In: Ekkehard König and Johan van der Auwera (eds.), *The Germanic Languages*. London: Routledge, 313-348.

Hamp, Eric P. (1982). "The Oldest Albanian Syntagma." *Balkansko ezikoznanie / Linguistique balkanique* 25:77–79.

Hansen, Aage. (1927). *Bestemt og ubestemt substantiv: bidrag til dansk substantiv-syntaks*. Kjøbenhavn: A. Busck.

Hansen, Aage. (1956). "Kasusudviklingen i dansk. " In: Erik Hansen and Lars Heltoft (eds.), *Festskrift til Peter Skautrup.* Aarhus: Aarhus University Press, 183-193.

Hare, Mary, Jeffrey L. Elman and Kim G. Daugherty. (1995). "Default Generalisation in Connectionist Networks." *Language and Cognitive Processes* 10(6): 601-630.

Härmä, Juhani. (2000) "Gender in French: A Diachronic Perspective." In: Terttu Nevalainen and Mirja Saari (eds.), *Gender in Grammar and Cognition II: Manifestations of Gender.* (Trends in Linguistics Studies and Monographs 124). Berlin/New York: Walter de Gruyter, 609-620.

Harmon, Zara and Vsevolod Kapatsinski. (2017). "Putting Old Tools to Novel Uses: The Role of Form Accessibility in Semantic Extension." *Cognitive Psychology* 98: 22–44.

Harris, Martin. (1988). "French." In: Martin Harris and Nigel Vincent (eds.), *The Romance Languages*. London/New York: Routledge, 209-245.

Harris, Martin. (1988). “The Romance Languages.” In: Martin Harris and Nigel Vincent (eds.), *The Romance Languages*. London/New York: Routledge, 1-25.

Haspelmath, Martin, Matthew S. Dryer, David Gil, and Bernard Comrie (eds.). (2013). *The World Atlas of Language Structures Online*. Munich: Max Planck Digital Library. http://wals.info.

Haugen, Einar and Thomas L. Markey. (1972). *The Scandinavian Languages: Fifty Years of Linguistic Research (1918-1968)*. The Hague: Mouton.

Haugen, Einar. (1976). *The Scandinavian Languages*. Cambridge, MA: Harvard University Press.

Hayes, Bruce. (1995). *Metrical Stress Theory: Principles and Case Studies*. Chicago: University of Chicago Press.

Heine, Bernd and Mechthild Reh. (1984). *Grammaticalisation and Reanalysis in African Languages.* Hamburg: Buske.

Heine, Bernd and Tania Kuteva. (2004). *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.

Heltveit, Trygve. (1958). “Attribute and Anaphora in the Gender System of English.” *Norsk Tidskrift for Sprogvidenskap* 18: 357–69.

Henriksen, Carol and Johan van der Auwera. (1994). “The Germanic Languages.” In: Ekkehard König and Johan van der Auwera (eds.), *The Germanic Languages*. London: Routledge, 1-18.

Hewson, John. (2006). “From Ancient to Modern Germanic.” *From Case to Adposition: The Development of Configurational Syntax in Indo-European Languages.* Amsterdam/Philadelphia: John Benjamins, 274-303.

Hilty, Gerold. (1968). “Westfränkische Superstrateinflüsse auf die galloromanische Syntax.” In: Baldinger, Kurt (ed.), *Festschrift für W. von Wartburg 1*. Tübingen: Niemeyer, 493–517.

Hinrichs, Uwe. (2004). “Ist das Bulgarische kreolisiertes Altbulgarisch?” In: Uwe Hinrichs (ed.), *Die europäischen Sprachen auf dem Wege zum analytischen Sprachtyp*. Wiesbaden: Otto Harrassowitz Verlag, 231-242.

Hjelmslev, Louis. (1935). *La catégorie des cas: étude de grammaire générale* (Acta Jutlandica VII/1). Aarhus: Universitetsforlaget.

Hock, Hans Henrich. (2021). *Principles of Historical Linguistics* (3rd edn.). Berlin/Boston: Mouton de Gruyter.

Hockett, Charles F. (1958). “A Course in Modern Linguistics.” *International journal of American Linguistics* 32(1): 59-73.

Hoeksema, Jack and Donna J. Napoli. (2008). “Just for the Hell of it: A Comparison of Two Taboo-term Constructions.” *Journal of Linguistics* 44(2): 347–378.

Hoekstra, Jarich and Peter Meijes Tiersma. (1994). “Frisian.” In: Ekkehard König and Johan van der Auwera (eds.), *The Germanic Languages*. London: Routledge, 505-531.

Hooper, Joan Bybee. (1980). “Child Morphology and Morphophonemic Change.” In: Jacek Fisiak (ed.), *Historical Morphology* 17. Berlin/Boston: De Gruyter Mouton, 157-185.

Horrocks, Geoffrey C. (2010). *Greek: A History of the Language and its Speakers*. Oxford: Wiley-Blackwell.

Horrocks, Geoffrey. (2007). “Syntax: From Classical Greek to the Koine.” In: Anastasios Phoivios Christidis (ed.), *A History of Ancient Greek. From the Beginnings to Late Antiquity*. Cambridge: Cambridge University Press, 618–631.

Hristov, Petko. (2008). “Mobilités du travail (gurbet), stratégies sociales et familiale: Une étude de cas dans les Balkans centraux.” *Balkanologie* 11(1-2): 2–14.

Humbert, Jean. (1930). *Le disparition du datif en grec.* (Collection linguistique publiée par la Société de linguistique de Paris 33). Paris: Société de linguistique de Paris.

Huntley, David. (1993). “Old Church Slavonic.” In: Bernard Comrie and Greville Corbett (eds), *The Slavonic Languages.* London/New York: Routledge, 125-187.

Hupchick, Dennis P. and Harold E. Cox. (2001). *The Palgrave Concise Historical Atlas of the Balkans*. New York: Palgrave.

Irvine, Susan (ed.). (2004). *The Anglo-Saxon Chronicle: A Collaborative Edition*. New York: Boydell & Brewer.

Ivanova-Mirčeva, Dora and Ivan Xaralampiev. (1999). *Istorija na bălgarskija ezik*. Veliko Tărnovo: Faber.

Jacobs, Neil G, Ellen F. Prince and Johan van der Auwera. (1994). “Yiddish.” In: Ekkehard König and Johan van der Auwera (eds.), *The Germanic Languages*. London*:* Routledge, 388-419.

Jagić, Vatroslav. (1894). [Review of] Lavrov, P. (1893). *Obzor zvukovych i formal’nych osobennostej bolgarskogo jazyka*. Moskva: Universitetskaja tipografija.

Jahr, Ernst Håkon. (1994). "Språkkontakt og språkforandring i Norden i hansatida." In: Ulla-Britt Kotsinas and John Helgander (eds.)*: Språkkontakt, dialektkontakt och språkförändring i Norden*. Stockholm: Institutionen för Nordiska Språk, 23-37.

Jakobson, Roman. (1932). "Zur Struktur des russischen Verbums. Charisteria Gvilelmo Mathesio oblata." Prague: Prazsky Lingvisticky Kroulek, 74-84. [Reprinted in *Selected Writings* II. The Hague: Mouton, 3-15.]

Jakobson, Roman. (1936/1984). "Contribution to the General Theory of Case: General Meanings of the Russian Cases." English translation by Linda R. Waugh, in: Roman Jakobson, *Russian and Slavic Grammar: Studies 1931-1981*. In: Linda R. Waugh and Morris Halle (eds). Berlin: Mouton, 59-103. (Originally published in 1936 as "Beitrag zur allgemeinen Kasuslehre: Gesamtbedeutungen der russischen Kasus," Travaux du cercle linguistique de Prague, Bd. 6(2), 240-288).

Jakobson, Roman. (1938). *Sur la théorie des affinités phonologiques des langues*. Copenhagen: Einar Munksgaard.

Jakobson, Roman. (1939). "Signe zéro." In: Bally, Charles (ed.), *Melanges de linguistique offerts à Charles Bally*. Geneve: Georg, 143-52. [Reprinted in *Selected Writings* II. The Hague: Mouton, 211-19.]

Janda, Laura A. (1996). "Figure, Ground, and Animacy in Slavic Declension." *Slavic and East European Journal* 40(2): 325-355.

Jendraschek, Gerd. (2009). "The Eternal Cycle of Morphological Change Condensation and Harmonization in Inflectional Paradigms." *Seoul National University*, 1-18.

Jespersen, Otto. (1894). *Progress in Language.* London: Swan Sonnenschein & Co.

Jespersen, Otto. (1935). *Growth and Structure of the English Language*. Leipzig: BG Teubner.

Johnson, D. J. L. (1972). "The Genesis of the Serbo-Croatian Genitive Plural in ā." *The Slavonic and East European Review* 50(120): 333-358.

Jónsson, Jóhannes Gísli. (2009) "Covert Nominative and Dative Subjects in Faroese." *Nordlyd* 36(2): *NORMS Papers on Faroese*, 142-164.

Joseph, Brian D. (2008). *On the Need for History in Balkan Linguistics*. New York: Beech Stave.

Joseph, Brian D. (2009). "Greek Dialectal Evidence for the Role of the Paradigm in Inflectional Change." *Morphology* 19(1): 45-57.

Juge, Matthew L. (2009). "Why Do Languages Lose Grammatical Categories? Latin and Romance Evidence." *Romance Quarterly* 56(1): 68-80.

Kapatsinski, Vsevolod, Shelece Easterday and Joan Bybee. (2020). “Vowel Reduction: A Usage-Based Perspective.” *Rivista di Linguistica* 32(1): 19-42.

Kapatsinski, Vsevolod. (2009). “Adversative Conjunction Choice in Russian (*no, da, odnako*): Semantic and Syntactic Influences on Lexical Selection.” *Language Variation and Change* 21(2): 157–173.

Kapović, Mate. (2017). “Proto-Indo-European Morphology.” In: Mate Kapović (ed.), *The Indo-European Languages.* (2nd edn.). New York: Routledge, 61-110.

Kastovsky, Dieter. (2000) “Inflectional Classes, Morphological Restructuring, and the Dissolution of Old English Grammatical Gender.” In: Terttu Nevalainen and Mirja Saari (eds.), *Gender in Grammar and Cognition II: Manifestations of Gender.* (Trends in Linguistics Studies and Monographs 124). Berlin/New York: Walter de Gruyter, 709-728.

Kaufman, Terrence and Thomason, Sarah. (1988). *Language Contact, Creolization and Genetic Linguistics*. Berkeley: University of California.

Kiparsky, Paul. (2012). “Grammaticalization as Optimization.” In: Dianne Jonas, John Whitman and Andrew Garrett (eds.), *Grammatical Change: Origins, Nature, Outcomes*. New York: Oxford University Press, 15-51.

Klaeber, Friedrich, John D. Niles, Robert E. Bjork, and R. D. Fulk. (2008). *Klaeber’s Beowulf and the Fight at Finnsburg*. (4th ed). Toronto: University of Toronto Press.

Knudsen, Trygve. (1967). *Kasuslaere*, Bind I. Oslo: Universitetsforlaget.

Koemeda-Lutz, Margit, Rudolf Cohen and Erwin Meier. (1987). “Organization of and Access to Semantic Memory in Aphasia.” *Brain and Language* 30(2): 321-337.

Koneski, Blaže. (1983). *A Historical Phonology of the Macedonian Language*. Heidelberg: Winter.

Krispel, Hiromi Mochida. (2013). Neuštokavische Innovationen und deren Reflexionen in den zentralsüdslavischen Standardsprachen. Diplomarbeit. universität Wien.

Kristoffersen, Kristian Emil. (2002). “The Morphology of Old Nordic I: Old Icelandic and Old Norwegian.” In: Herbert Ernst Wiegand (ed.), *The Nordic Languages. An International Handbook of the History of the North Germanic*. Berlin/New York: Walter de Gruyter, 911-925.

Kulikov, Leonid. (2006). “Case Systems in a Diachronic Perspective.” In: Leonid Kulikov, Andrej Malchukov and Peter de Swart (eds.), *Case, Valency and Transitivity.* (Studies in Language Companion Series 77). Amsterdam/Philadelphia: John Benjamins, 23-47.

Kulikov, Leonid. (2013). “Case Variation and Case Alternation in Indo-European and Beyond: A Diachronic Typological Perspective.” In: Gelderen Elly, Cennamo Michela and Jóhanna Barðdal (eds.), *Argument Structure in Flux*: *The Naples-Capri Papers.* (Studies in Language Companion Series 131). Amsterdam/Philadelphia: John Benjamins, 53-85.

Kürschner, Sebastian and Damaris Nübling. (2011). “The Interaction of Gender and Declension in Germanic Languages.” *Folia Linguistica* 45(2): 355-388.

Kurumada, Chigusa and Florian Jaeger. (2015). “Communicative Efficiency in Language Production: Optional Case-marking in Japanese.” *Journal of Memory and Language* 83: 152-178.

Kuryłowicz, Jerzy. (1964). *The Inflectional Categories of Indo-European*. Heidelberg: Winter.

Kurzová, Helena. (1993). *From Indo-European to Latin*. Amsterdam/Philadelphia: John Benjamins.

Labov, William. (1987). “The Overestimation of Functionalism.” In: René Dirven and Vilém Fried (eds.), *Functionalism in Linguistics*. *Linguistic and Literary Studies in Eastern Europe* 20. Amsterdam/Philadelphia: John Benjamins, 311-332.

Lahiri, Aditi, Tomas Riad and Haike Jacobs. (1999). “Diachronic Prosody.” In Harry van der Hulst (ed.), *Word Prosodic Systems in the Languages of Europe*. Berlin/New York: Mouton de Gruyter, 335-422.

Lasch, Agathe. (1914). *Mittelniederdeutsche Grammatik*. Halle: max Niemeyer.

Lass, Roger. (1992). “Phonology and Morphology.” In N. Blake (ed.), *Cambridge History of the English Language*, vol 2. Cambridge: Cambridge University Press, 23–155.

Lass, Roger. (1994). *Old English: A Historical Linguistic Companion*. Cambridge: Cambridge University Press.

Lass, Roger. (1997). *Historical Linguistics and Language Change*. Cambridge University Press.

Le Saux, Françoise H. M. (ed.). (1994). *Brut*. Woodbridge/New York: Boydell & Brewer Ltd.

Lee, Hanjung. (2006). “Parallel Optimization in Case Systems: Evidence from Case Ellipsis in Korean.” *Journal of East Asian Linguistics* 15: 69–96.

Leskien, August. (1976). *Grammatik der serbo-kroatischen Sprache. Lautlehre, Stammbildung, Formenlehre*. Heidelberg: Winter.

Lewis, Charlton Thomas and Charles Short. (1879). *A Latin Dictionary*. Oxford: Clarendon Press.

Lewis, Geoffrey. (2000). *Turkish Grammar*. Oxford: Oxford University Press.

Lewis, Robert E., et al. Ann Arbor (eds.). Middle English Dictionary. University of Michigan Press, 1952-2001. Online edition in Middle English Compendium. Ed. Frances McSparran, et al. Ann Arbor: University of Michigan Library, 2000-2018. <http://quod.lib.umich.edu/m/middle-english-dictionary/>. Accessed 21 May 2022.

Lightfoot, David. (1999). *The Development of Language: Acquisition, Change and Evolution*. Oxford/Malden: Blackwell.

Lightfoot, David. (ed.) (2002). *Syntactic Effects of Morphological Change*. Oxford: Oxford University Press.

Lindstedt, Jouko. (1998). "On the Balkan Linguistic Type." *Studia Slavica Finlandensia* 15: 91-103.

Lindstedt, Jouko. (2000). "Linguistic Balkanization: Contact-Induced Change by Mutual Reinforcement." In Dicky Gilbers, John Nerbonne and Jos Schaeken (eds.), *Languages in Contact* (Studies in Slavic and General Linguistics 28). Amsterdam: Rodopi, 231-246.

Lindstedt, Jouko. (2014). "Balkan Slavic and Balkan Romance: From Congruence to Convergence." In: Juliane Besters-Dilger, Cynthia Dermarkar, Stefan Pfänder and Achim Rabus (eds.), *Congruence in Contact Induced Language Change: Language Families, Typological Resemblance and Perceived Similarity.* Berlin/Boston: Walter de Gruyter. 168-183.

Lindstedt, Jouko. (2019). "Diachronic Regularities Explaining the Tendency Towards Explicit Analytic Marking in Balkan Syntax." In: Iliyana Krapova and Brian Joseph (eds.), *Balkan Syntax and (Universal) Principles of Grammar* (Trends in Linguistics Studies and Monographs, vol. 285). Berlin/Boston: Walter de Gruyter, 70-84.

Lomtev, G. P. (1953). "O sootvetstvii grammaticheskix sredstv jazyka potrebnostjam vzaimoponimanija." *Voprosy jazykoznanija,* 76-88.

Loporcaro, Michele, Vincenzo Faraoni, Francesco Gardani. (2014). "The Third Gender of Old Italian." *Diachronica* 31(1):1-22.

Luján, Eugenio R. and Cesar Abad R. (2014). "Semantic Roles and Word Formation: Instrument and Location in Ancient Greek." In: Silvia Luraghi and Heiko Narrog (eds.), *Perspectives on Semantic Roles* 106. Amsterdam/Philadelphia: John Benjamins, 241-270.

Lunt, Horace Gray. (1965). "On the Loss of Declension in Bulgarian and Macedonian." *Die Welt der Slaven* 10: 305–312.

Lunt, Horace Gray. (2001). *Old Church Slavonic* (7th edn.). Berlin/New York: Walter de Gruyter.

Luraghi, Silvia and Milena Krstić. (2018). “The Decay of Cases in Molise Slavonic.” In: Björn Hansen, Jasmina Grković-Major and Barbara Sonnenhauser (eds.), *Diachronic Slavonic Syntax: The Interplay between Internal Development, Language Contact and Metalinguistic Factors*. Berlin/Boston: Walter de Gruyter, 63-92.

Luraghi, Silvia. (1987). “Patterns of Case Syncretism in Indo-European Languages.” In: Anna Giacalone Ramat, Onofrio Carruba and Giuliano Bernini (eds.), *Papers from the 7th International Conference on Historical Linguistics*. Amsterdam: John Benjamins, 355–371.

Luraghi, Silvia. (1989). “The Relation Between Prepositions and Cases within Latin Prepositional Phrases.” In: Gualtiero Calboli (ed.), *Subordination and Other Topics in Latin.* Amsterdam: John Benjamins, 253-271.

Luraghi, Silvia. (2003). *On the Meaning of Prepositions and Cases. A Study of the Expression of Semantic Roles in Ancient Greek*. Amsterdam/Philadelphia: John Benjamins.

Luraghi, Silvia. (2004). “The Evolution of the Greek Nominal Paradigms from Mycenean to Modern Greek.” In: Ole Thomsen (ed.), *Classica et Mediaevalia,* vol. 55. Copenhagen: Museum Tusculanum Press, 361-380.

MacWhinney, Brian, Csaba Pléh and Elizabeth Bates. (1985). “The Development of Sentence Interpretation in Hungarian.” *Cognitive Psychology* 17(2): 178-209.

MacWhinney, Brian. (1978). “The Acquisition of Morphophonology.” *Monographs of the Society for Research in Child Development* 43(1/2): 1-123.

MacWhinney, Brian. (1984). “Grammatical Devices for Sharing Points.” In: R. Schiefelbusch (ed.), *The Acquisition of Communicative Competence*. Baltimore: Univ. Park Press.

Malić, Dragica. (1988). *Povaljska listina kao jezični spomenik*. Zagreb: Znanstvena biblioteka Hrvatskog filološkog društva.

Maling, Joan. (2002). “Það rignir flágufalli á Íslandi.” *Íslenskt mál og almenn málfræði* 24: 31–105. [Also published in a slightly different English version in *Working Papers in Scandinavian Syntax* 70: 1–60.]

Mallinson, Graham. (1988). “Romanian.” In: Martin Harris and Nigel Vincent (eds.), *The Romance Languages*. London/New York: Routledge, 391-419.

Mallory, James P., and Douglas Q. Adams. (2006). *The Oxford Introduction to Proto-Indo-European and the Proto-Indo-European World*. Oxford New York: Oxford University Press.

Mańczak, Witold. (1980). “Laws of Analogy.” In Jacek Fisiak (ed.), *Historical Morphology*, vol. 17 (Trends in Linguistics. Studies and Monographs). Mouton: The Hague, 283–288.

Marra, Antonietta. (1999). Il sistema dei casi nel nanaš dei Croati molisani. Processi di mutamento, decadenza e morte di una lingua. Ph.D. thesis, University of Pavia.

Marra, Antonietta. (2009). "L'uso del caso accusativo nei parlanti slavo-molisani." In: P. Desideri and C. Consani (eds.), *Alloglossie e comunita alloglotte nell'Italia contemporanea. Teorie, applicazioni e descrizioni, prospettive. Atti del XLI Congresso Internazionale di Studi della Societa di Linguistica Italiana*. Roma: Bulzoni, 95-119.

Marra, Antonietta. (2012). "Contact phenomena in the Slavic of Molise: Some Remarks about Nouns and Prepositional Phrases." In: Martine Vanhove, Thomas Stolz, Aina Urdze and Hotomi Otsuka (eds.), *Morphologies in Contact*. Berlin: Akademie Verlag, 265-282.

Marsden, Richard. (2010). *The Cambridge Old English Reader*. Cambridge: Cambridge University Press.

Matasović, Ranko. (2008). *Poredbenopovijesna gramatika hrvatskoga jezika*. Zagreb: Matica hrvatska.

Meiser, Gerhard. (1992). "Syncretism in Indo-European Languages: Motives, Process and Results." *Transactions of the Philological Society* 90(2): 187-218.

Meyer, Karl Heinrich. (1920). *Der Untergang der Deklination im Bulgarischen*. (Slavica 3). Heidelberg: Carl Winter.

Miletič, Ljubomir. (1890). "Staroto sklonenie v dnešnitě bălgarski naręčija." *Zbornik za narodna umotvorenija, nauka i knižnina*, Kniga II, 226–268. Sofia: Ministerstvoto na narodnoto prosvěštenie.

Miletič, Ljubomir. (1923). *Svištovski damaskin. Novobălgarski pametnik ot XVIII vek.* (=*Bălgarski starini* 7.) Sofia: Pridvorna pečatnica.

Miletič, Ljubomir. (1987). [1896–1900]. *Izsledovanija za bălgarite v Sedmigradsko i Banat*. Sofia: Nauka i iskustvo.

Minkova, Donka. (1983) "Middle English Final–e from a Phonemic Point of View." In: Michael Davenport, Erik Hansen and Hans-Frede Nielsen (eds.), *Current Topics in English Historical Linguistics*. Odense: Odense University Press. 191-209.

Minkova, Donka. (1991). *The History of Final Vowels in English: The Sound of Muting*. Berlin: Mouton de Gruyter.

Mirčev, Kiril. (1958). "Po văprosa za naj-rannite primeri ot analitičen datelen padež v bălgarskite pametnici." In: [ed.?] *Ezikovedski izsledvanija v čest na akademik Stefan Mladenov*. Sofia: Bălgarskata akademija na naukite, 37-46.

Mirčev, Kiril. (1978). *Istoričeska gramatika na bălgarskija ezik*. (3rd edn.). Sofia: Nauka i izkustvo.

Mišeska Tomić, Olga. (2006). *Balkan Sprachbund Morpho-Syntactic Features*. (Studies in Natural Language and Linguistic Theory 67). Dordrecht: Springer.

Mitchell, Bruce. (1985). *Old English Syntax: Concord, the Parts of Speech, and the Sentence*. Oxford: Oxford University Press.

Mladenova, Olga M. (2007). *Definiteness in Bulgarian: Modelling the Processes of Language Change* (Trends in Linguistics: Studies and Monographs 182). Berlin/ New York: Mouton de Gruyter.

Mladenova, Olga M. (2009). "On Morphosyntactic Change in Bulgarian: Case and Definiteness." *Diachronica*, 26(3): 408-436.

Mladenović, Aleksandar. (2008). *Istorija srpskog jezika*. *Odabrani radovi.* Belgrade: Čigoja štampa.

Moore, Samuel. (1928). "Earliest Morphological Changes in Middle English." *Language* 4(4): 238-266.

Mørck, Endre. (2005). "Morphological Developments from Old Nordic to Early Modern Nordic: Inflexion and Word Formation." In: Oskar Bandle et al. (ed.). *The Nordic Languages. An International Handbook of the History of the North Germanic Languages*, Vol. II. Berlin/New York: Walter de Gruyter, 1128-1148.

Morwood, James. (2002). *The Oxford Grammar of Classical Greek*. Sourcebooks, Inc.

Moscoso del Prado Martin, Fermin. (2014). "Grammatical Change Begins Within the Word: Causal Modeling of the Co-evolution of Icelandic Morphology and Syntax." *Proceedings of the Annual Meeting of the Cognitive Science Society* 36: 2657-2662.

Mussies, Gerard. (1971). *The Morphology of Koine Greek as Used in the Apocalypse of St. John: A Study in Bilingualism*. Leiden: Brill Archive.

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

Myers-Scotton, Carol. (2006). *Multiple Voices: An Introduction to Bilingualism*. Malden/Oxford/Victoria: Blackwell.

Nandriş, Grigore. (1969). *Old Church Slavonic Grammar*. London: University of London.

Nevis, Joel A. (1985). Finnish Particle Clitics and General Clitic Theory. Ph.D Dissertation, Ohio State University.

Newman, John G. (1999). “The Spread of the s-Plural in Middle English (1150-1420): A Corpus Study.” *Studia Anglica Posnaniensia XXXIV*. 73-89.

Newmark, Leonard, Philip Hubbard and Peter Prifti. (1982). *Standard Albanian: A Reference Grammar for Students*. Palo Alto: Stanford University Press.

Nielsen, Hans Frede. (1998). *The Continental Backgrounds of English and its Insular Development until 1154*. Odense: Odense University Press.

Nomachi, Motoki and Bernd Heine. (2011). “On Predicting Contact-Induced Grammatical Change: Evidence from Slavic Languages.” *Journal of Historical Linguistics* 1(1): 48–76.

Norde, Muriel. (1997a). The History of the Genitive in Swedish. A Case Study in Degrammaticalization. Ph.D. Dissertation, Universiteit van Amsterdam.

Norde, Muriel. (1997b). “Middle Low German - Middle Scandinavian Language Contact and Morphological Simplification.” *Multilingua* 16(4): 389-409.

Norde, Muriel. (2001). “The Loss of Lexical Case in Swedish.” In: Jan Terje Faarlund (ed.), *Grammatical Relations in Change,* vol. 56. Amsterdam/Philadelphia: John Benjamins*,* 241-272.

Nygaard, Marius. (1906). *Norrøn Syntax*. Oslo. Aschehoug.

O’Neil, Wayne. (1978). “The Evolution of the Germanic Inflectional Systems: A Study in the Causes of Language Change.” *Orbis* 27(2): 248-285.

Orel, Vladimir Ė. (2000). *A Concise Historical Grammar of the Albanian Language: Reconstruction of Proto-Albanian*. Leiden/Boston/Köln: Brill.

Paradis, Carole and Jean-François Prunet. (1991). *The Special Status of Coronals: Internal and External Evidence*. San Diego/London: Academic Press.

Pato, Enrique. (2018). “Indefinite Article+ Possessive+ Noun in Spanish: A Case of Refunctionalization?.” *Languages*, 3(44): 23-30.

Patseva, Mirena. (2017). “Bulgarian Word Stress Analysis in the Frame of Prosody Morphology Interface.” *Technical Report*. New Brunswick: Rutgers University.

Penny, Ralph (2002). *A History of the Spanish Language*. (2nd edn.) Cambridge: Cambridge University Press.

Perek, Florent. (2016). “Using Distributional Semantics to Study Syntactic Productivity in Diachrony: A Case Study.” *Linguistics* 54(1): 149-188.

Petersen, Hjalmar P. (2011). “Related Languages, Convergence and Replication: Faroese-Danish.” *International Journal of Bilingualism* 15(1): 101-120.

Pettersson, Thore and Sidney Wood. (1987). “Vowel Reduction in Bulgarian and Its Implications for Theories of Vowel Production: A Review of the Problem.” *Societas Linguistica Europaea* 21(2): 261-279.

Pharies, David A. (2007). *A Brief History of the Spanish Language*. Chicago: University of Chicago Press.

Philippi, Julia. (1997). “The Rise of the Article in the Germanic Languages.” In: Ans van Kemenade and Nigel Vincent (eds.), *Parameters of Morphosyntactic Change*. Cambridge: Cambridge University Press, 62-93.

Pierrehumbert, Janet. (2001). “Exemplar Dynamics: Word Frequency, Lenition and Contrast.” In: Joan Bybee and Paul Hopper (eds.), *Frequency and the Emergence of Linguistic Structure*. (Typological Studies in Language 45). Amsterdam: John Benjamins, 137-158.

Plank, Frans. (1979). “The Functional Basis of Case Systems and Declension Classes: From Latin to Old French.” *Linguistics*. 17: 611-640.

Poplack, Shana. (1980). “The Notion of the Plural in Puerto Rican Spanish: Competing Constraints on (s) Deletion. In: William Labov (eds.), *Locating Language in Time and Space*. New York: Academic Press, 55-67.

Popova, Marija. (1973). “Grammatical Elements of Language in the Speech of Pre-School Children.” In: C. A. Ferguson, D. I. Slobin (eds.), *Studies of Child Language Development*. New York: Holt, Rinehart, & Winston, 269-280.

Popović, Ivan V. (1960). *Geschichte der serbokroatischen Sprache*. Wiesbaden: Otto Harrassowitz.

Posner, Rebecca. (1996). *The Romance Languages*. Cambridge: Cambridge University Press.

Pott, August. (1836). *Etymologische Forschungen auf dem Gebiete der Indo-Germanischen Sprachen*, vol. 2. Lemgo: Meyer’sche Hof-Buchhandlung.

Preston, Dennis R. and Michael Turner. (1984) “The Polish of Western New York: Case.” In: Roland Sussex (ed.), *The Maintenance of the Slavonic Languages Abroad.* (Melbourne Slavonic Studies 18). Melbourne: University of Melbourne, 135-54.

Priestly, T. M. S. (1993). “Slovene.” In: Bernard Comrie and Greville Corbett (eds), *The Slavonic Languages.* London/New York: Routledge, 388–451.

Ramat, Paoplo. (1992). “Thoughts on Degrammaticalization.” *Linguistics* 30(3): 549-560.

Rapaport, Steven J. (2016). Prosodic Shift and Loss of Cases in Germanic, Romance and Hellenic Languages. Dissertation, Edinburgh: University of Edinburgh.

Reershemius, Gertrud. (2004). *Niederdeutsch in Ostfriesland: Zwischen Sprachkontakt, Sprachveränderung und Sprachwechsel*, vol. 119. Franz Steiner Verlag.

Reichenkron, Günter. (1962). "Der Typus der Balkansprachen." *Zeitschriftfür Balkanologie* 1: 91-122.

Reinhammar, Maj. (1993). "Levande språkhistoria." In: Lars Wollin (ed.) *Studier i svensk språkhistoria* 3: 183-191. Uppsala: Institutionen för nordiska språk.

Reinhart, Johannes. (2014). "Croatian Church Slavonic." In: Karl Gutschmidt, Sebastian Kempgen, Tilman Berger and Peter Kosta (eds.). *The Slavic Languages*, vol. 2. Berlin/New York: Walter de Gruyter.

Rickard, Peter. (2003). *A History of the French Language*. London/New York: Routledge.

Ringe, Don. (2006). *From Proto-Indo-European to Proto-Germanic*: *A Linguistic History of English*, vol. 1. Oxford: Oxford University Press.

Rinnert, Carol, and Harry A. Whitaker. (1973). "Semantic Confusions by Aphasic Patients." *Cortex* 9(1): 56-81.

Rosetti, Alexandru. (1986). *Istoria limbii române*. București: Editura Ştiinţifică şi Enciclopedică.

Rothstein, Robert. (1993). "Polish." In: Bernard Comrie and Greville Corbett (eds.). *The Slavonic Languages.* London/New York: Routledge, 686-758.

Sala, Marius. (2013). "Contact and Borrowing." In: Martin Maiden, John Charles Smith, and Adam Ledgeway (eds.), *The Cambridge History of the Romance Languages*, vol. 2. Cambridge: Cambridge University Press, 187-236.

Salvi, Giampaolo. (2011). "Morphosyntactic Persistence." In: Martin Maiden, John Charles Smith, and Adam Ledgeway (eds.), *The Cambridge History of the Romance Languages*, vol. 1. Cambridge: Cambridge University Press, 318-381.

Scatton, Ernest A. (1984). *A Reference Grammar of Modern Bulgarian*. Columbus, OH: Slavica.

Scatton, Ernest A. (1993). "Bulgarian." In: Bernard Comrie and Greville Corbett (eds.), *The Slavonic Languages.* London/New York: Routledge, 188-248.

Schenker, Alexander M. (1993). "Proto-Slavonic." In: Bernard Comrie and Greville Corbett (eds.), *The Slavonic Languages.* London/New York: Routledge, 60-121.

Schrijver, Peter. (2014). *Language Contact and the Origins of the Germanic Languages*. New York: Routledge.

Schulte, Kim. (2009). "Loanwords in Romanian." In: Martin Haspelmath and Uri Tadmor (eds.), *Loanwords in the World's Languages: A Comparative Handbook*. Berlin: Walter de Gruyter, 230-259.

Scott, Alan. (2014). *The Genitive Case in Dutch and German: A Study of Morphosyntactic Change in Codified Languages*. Brill.

Seefranz-Montag, Ariane. (1983). *Syntaktische Funktionen und Wortstellungsveränderung.Die Entwicklung 'subjektloser' Konstruktionen in einigen Sprachen* (Studien zur Theoretischen Linguistik 3). München: Wilhelm Fink Verlag.

Shapiro, Michael. (1986). "Language as Semiosis: A Neo-Structuralist Perspective in the Light of Pragmaticism." *Chinese Semiotic Studies* 18(1): 131-146.

Short, David. (1993). "Slovak." In: Bernard Comrie and Greville Corbett (eds.), *The Slavonic Languages.* London/New York: Routledge, 533-592.

Shrier, Martha. (1965). "Case Systems in German Dialects." *Language* 41(3): 420–438.

Sick, Bastian. (2005). *Der Dativ ist dem Genitiv sein Tod. Neues aus dem Irrgarten der deutschen Sprache* (Folge 2. 1. Aufl.). Köln: Kiepenheuer & Witsch.

Siegel, Jeff. (1985). "Koines and Koineization." *Language in Society* 14(3): 357-378.

Sihler, Andriew. (1995). *New Comparative Grammar of Greek and Latin*. New York: Oxford University Press.

Skaffari, Janne. (2017). "Language Contact: French." In: Laurel Brinton and Alexander Bergs (eds.), *The History of English: Middle English*, vol 3, Berlin: De Gruyter, 184-204.

Slobin, Dan I. and Thomas G. Bever. (1982). "Children Use Canonical Sentence Schemas: A Crosslinguistic Study of Word Order and Inflections." *Cognition* 12(3): 229-265.

Smoczyńska, Magdalena. (1985). "The Acquisition of Polish." In: Dan Slobin (ed.), *The Crosslinguistic Study of Language Acquisition Volume 2: Theoretical Issues*. Hillsdale, NJ: Lawrence Erlbaum Associates, 595-686.

Sobolev, Andrej N. (1991). "K istorii utraty balkanoslavjanskogo sklonenija." *Zbornik Matice srpske za filologiju i lingvistiku* 34(2): 7–41.

Sornicola, Rosanna. (2011). “Romance Linguistics and Historical Linguistics: Reflections on Synchrony and Diachrony.” In: Martin Maiden, John Charles Smith and Adam Ledgeway (eds.), *The Cambridge History of the Romance Languages,* vol. 1. Cambridge: Cambridge University Press, 1-49.

Spencer, Andrew. (2007). “The Possessum-Agreement Construction or Does Albanian Have a Genitive Case?.” In: Andrew Radford (ed.), *Martin Atkinson – The Minimalist Muse.* (Essex Research Reports in Linguistics 53). Department of Language and Linguistics: University of Essex, 219-250.

Stankovski, Miodrag, L’ubomír Ďurovič and Mijo Tomaševic. (1983). “Development Structures in the Family Language of Yugoslav Immigrant Children in a Swedish Language Environment.” In: L’ubomír Ďurovič (ed.), *Lingua in Diaspora: Studies in the Language of the Second Generation of Yugoslav Immigrant Children in Sweden*. *Slavica Lundensia 9.* Lunds universitet: Slaviska institutionen, 11-20.

Steinke, Klaus. (1968). *Studien über den Verfall der bulgarischen Deklination*. Munich: Otto Sagner.

Stieber, Zdzisław. (1969). *Zarys gramatyki porównawczej języków słowiańskich: Fonologia.* Warszawa: Panstwowe Wydawnictwo Naukowe.

Stojkov, Stojko. (1970). “Formele cazuale nominale în limba bulgară.” *Romanoslavica,* 17: 129–153.

Stolk, Joanne Vera. (2015). “Dative by Genitive Replacement in the Greek Language of the Papyri: A Diachronic Account of Case Semantics.” *Journal of Greek Linguistics* 15(1): 91-121.

Sumner, Meghan, Chigusa Kurumada, Roey Gafter and Marisa Casillas. (2013). “Phonetic Variation and the Recognition of Words with Pronunciation Variants.” *Proceedings of the Annual Meeting of the Cognitive Science Society* 35: 3486- 3491.

Sussex, Roland, and Paul V. Cubberley. (2011). *The Slavic Languages*. New York: Cambridge University Press.

Sussex, Roland. (1993). “Slavonic Languages in Emigration.” In: Bernard Comrie and Greville Corbett (eds.), *The Slavonic Languages*. London/New York: Routledge, 999-1036.

Suttle, Laura and Adele E. Goldberg. (2011). “The Partial Productivity of Constructions as Induction.” *Linguistics* 49(6): 1237-1269.

Svane, Olaf. (1958). *Die Flexionen in štokavichen Texten aus dem Zeitraum 1350-1400*. *Sprachgeschichtliche Untersuchungen über älteres Serbokroatisch*. Aarhus: Universitetsforlaget i Aarhus.

Terrell, Tracy. (1975). “Functional Constraints on Deletion of Word Final /s/ in Cuban Spanish.” *Annual Meeting of the Berkeley Linguistics Society* 1: 431-437.

Thomason, Sarah G. (2001). *Language Contact: An Introduction*. Edinburgh: Edinburgh University Press.

Thomason, Sarah G. (2008a). “Why Universals Versus Contact-Induced Change?.” In: Markku Filppula, Juhani Klemola, Heli Paulasto (eds.), *Vernacular Universals and Language Contacts*. New York: Routledge, 355-370.

Thomason, Sarah G. (2008b). “Social and Linguistic Factors as Predictors of Contact-Induced Change.” *Journal of Language Contact* 2(1): 42–56.

Thomason, Sarah G. (2009). “Why Universals Versus Contact-Induced Change?.” *Vernacular Universals and Language Contacts*, 355-370.

Thomason, Sarah G. (2010). “Contact Explanations in Linguistics.” In Raymond Hickey (eds.), *The Handbook of Language Contact*. Malden/Oxford/Victoria: Blackwell, 31-47.

Thompson, Rupert. (2010). “Mycenaean Greek.” In: Egbert J. Bakker (ed.), *A Companion to the Ancient Greek Language*. Malden/Oxford/West Sussex: John Wiley & Sons, 189-99.

Thráinsson, Höskuldur. (1994). “Icelandic.” In: Ekkehard König and Johan van der Auwera (eds.), *The Germanic Languages*. London*:* Routledge, 142-189.

Tiersma, Peter Meijes. (1982). “Local and General Markedness.” *Language* 58(4): 832-849.

Tilkov, Dimităr and Todor Bojadžiev. (2013). *Bălgarska fonetika*, Sofia: Nauka i izkustvo.

Tiraspol’skij, G. I. (1980). K voprosu ob uslovijax i pričinax utraty bolgarskogo substantivnogo sklonenija. *Sovetskoe slavjanovedenie* 4.

Torp, Arne. (1982). *Norsk og nordisk før og nå*. Oslo: Universitetsforlaget.

Trąba, Kamil. (2018). “The Development of the Greek Case System: Morphological Studies.” In: Krystyna Tuszyńska (ed.), *Miscellanea Novograeca Posnaniensia.* Warszawa: Wydawnictwo Naukowe, 107-136.

Traugott, Elizabeth Closs. (2011). “Grammaticalization and Mechanisms of Change.” In: Bernd Heine and Heiko Narrog (eds.) *Oxford Handbooks Online*. Oxford/New York: Oxford University Press, 1-8.

Trifonov, Jurdan. (1931). “Kăm istorijata na členните formi v bălgarskija ezik.” *Makedonski pregled* 7(1): 7–26.

Tristram, Hildegard. (2004). “Diglossia in Anglo-Saxon England, or What was Spoken Old English Like?” *Studia Anglica Posnaniensia* 40: 87–110.

Trudgill, Peter. (1989). “Contact and isolation in linguistic change.” In: Werner Winter (ed.), *Trends in Linguistics* (Studies and Monographs 43). Berlin/New York: Mouton de Gruyter, 227-238.

Trudgill, Peter. (2002). “Linguistic and Social Typology.” In J. K. Chambers, Peter Trudgill and Natalie Schilling-Estes (eds.), *The Handbook of Language Variation and Change*, 707–728. Malden, MA: Blackwell.

Trudgill, Peter. (2010). *Investigations in Sociohistorical Linguistics: Stories of Colonisation and Contact*. Cambridge University Press.

Trudgill, Peter. (2011). *Sociolinguistic Typology: Social Determinants of Linguistic Complexity*. Oxford: Oxford University Press.

Turbić-Hadžagić, Amira and Marica Petrović. (2011). “Promjena imenica u hrvatskoj listini Dubrovačkoga arhiva od 11. 6. 1253.” *časopis za jezikoslovna, književna i kulturna pitanja* 1(7):71-87.

Vakareliyska, Cynthia and Roger Gyllin. (in press). “Proto-Slavic and Old Church Slavonic.” In: Victor A. Friedman and Lenore Grenoble (eds.), *The Slavic Languages* (2nd edn.). London: Routledge.

Vakareliyska, Cynthia. (1990). The Dative-Accusative Case Opposition in Slavic Languages: Evidence from Aphasia. Ph.D. Dissertation, The Department of Slavic Languages and Literatures, Harvard University.

Vakareliyska, Cynthia. (1994). “A Model of the Dative/Accusative Opposition for Slavic Languages Based on Data from Aphasia.” *Supostavitelno ezikoznanie* (Sofia) 19: 7-13.

Vakareliyska, Cynthia. (1994). “Na-drop in Bulgarian.” *Journal of Slavic Linguistics* 2(1): 121-150.

Vakareliyska, Cynthia. (2002). “Na-drop Revisited: Omission of the Dative Marker in Bulgarian Double Dative Object Constructions.” In: Laura Janda, Steven Franks and Ronald Feldstein (eds.), *Where One’s Tongue Rules Well: Festschrift for Charles E. Townsend.* Bloomington, IN: Slavica, 165-192.

Vakareliyska, Cynthia. (2008). *The Curzon Gospel*, vol. 2., *A Linguistic and Textual Introduction.* Oxford: Oxford University Press.

Vakareliyska, Cynthia. (2018). “English Loanblend [N[N]]s and Related Constructions in Serbian..” In: Stephen M. Diceky and Mark Richard Lauersdorf (eds.), *V zeleni drželi zeleni breg*. Bloomington, IN: Slavica Publishers.

van der Wal, J. Marijke and Aad Quak. (1994). "Old and Middle Continental West Germanic." *The Germanic Languages*. London: Routledge, 72-109.

van Gelderen, Elly. (2000). *A History of English Reflexive Pronouns: Person, Self, and Interpretability*. Amsterdam/Philadelphia: John Benjamins.

Van Haeringen, Bernardus. (1962). *Gramarie: keur uit het werk van zijn hoogleraarstijd*. Assen: van Gorcum.

Van Kemenade, Ans. (1987). *Syntactic Case and Morphological Case in the History of English*. Dordrecht: Foris.

Van Kemenade, Ans. (1994). "Old and Middle English." In: Ekkehard König and Johan van der Auwera (eds.), *The Germanic Languages*. London: Routledge, 110-141.

van Reenen, P. and Schøsler, L. (2000). "Declension in Old and Middle French: Two Opposing Tendencies." In: John Charles Smith and Delia Bentley (eds.), *Historical Linguistics 1995 Volume 1: General Issues and Non-Germanic Languages*. Amsterdam/Philadelphia: John Benjamins, 327–44.

van Trijp, Remi. (2013). "Linguistic Assessment Criteria for Explaining Language Change: A Case Study on Syncretism in German Definite Articles." *Language Dynamics and Change* 3(1): 105-132.

Varvaro, Alberto. (2011). "Latin and the Making of the Romance Languages." In: Maiden, John Charles Smith and Adam Ledgeway. *The Romance Languages*, vol. 2. Cambridge: Cambridge University Press, 6-56.

Vincent, Nigel. (1988). "Italian." In: Martin Harris and Nigel Vincent (eds.), *The Romance Languages*. London/New York: Routledge, 279-313.

Vondrák, Václav. (1904). *O původu Kijevskych listů a Prazskych zlomků*. V Praze: Nákladem jubilejního fondu Krá.

Vondrák, Václav. (1912). *Altkirchenslavische Grammatik*. Berlin: Weidmann.

Voronina, I. I. (1976). "K voprosu o sisteme artiklej v albanskom jazyke." In: A. V. Desnickaja (ed.), *Grammatičeskij stroj balkanskix jazykov. Issledovanija po semantike grammatičeskix form*. Leningrad: Nauka, 126–161.

Wackernagel, Jacob. (1920). *Vorlesungen über syntax: mit besonderer berücksichtigung von griechisch, lateinisch und deutsch*. Basel: Philologischen Seminar der Universität Basel.

Wahlström, Max. (2015). The Loss of Case Inflection in Bulgarian and Macedonian. (Slavica Helsingiensia 47). PhD Dissertation, Humanities Faculty, Department of Modern Languages, University of Helsinki.

Wartburg, Walther von. (1950). *Die Ausgliederung der romanischen Sprachräume*. Bern: Francke.

Weigand, Gustav. (1888). *Die Sprache der Olympo-Walachen nebst einer Einleitung über Land und Leute*. Leipzig: Johann Ambrosius Barth.

Weinreich, Uriel. (1953). *Languages in Contact*: *Findings and Problems*. New York: Linguistic Circle of New York.

Weiss, Michael. (2010). "Morphology and Word Formation." In: Egbert J. Bakker (ed.), *A Companion to the Ancient Greek Language*. Malden/Oxford/West Sussex: John Wiley & Sons, 104-119.

Werner, Otmar. (1984). "Morphologische Entwicklungen in den germanischen Sprachen." In: Jürgen Untermann and Bela Brogyanyi (eds.), *Das Germanische und die Rekonstruktion der indo-germanischen Grundsprache.* Amsterdam/Philadelphia: John Benjamins. 181-226.

Wessén, Elias. (1968). *Svensk språkhistoria I: Ljudlära och ordböjningslära*. (Reprinted. Edsbruk: Akademitryck 1992).

Wheeler, Marcus and Boris Unbegaun. (2000). *The Oxford Russian Dictionary*. Oxford: Oxford University Press.

Windisch, Rudolf. (1973). *Genusprobleme im Romanischen: das Neutrum im Rumänischen*. Tiibingen: Narr.

Woodrow, Herbert. (1909). *A Quantitative Study of Rhythm: The Effect of Variations in Intensity, Rate and Duration*. New York: Science Press.

Wright, Joseph. (1906). *An Old High German Primer*. (2nd edn.). Oxford: Clarendon Press.

Wright, Joseph. (1917). *A Middle High German Primer: With Grammar, Notes, and Glossary*. Oxford: Clarendon Press.

Wurzel, Wolfgang U. (1984). *Flexionsmorphologie und Natürlichkeit*. Berlin: Akademie Verlag.

Zehentner, Eva. (2017). "Ditransitives in Middle English: On Semantic Specialisation and the Rise of the Dative Alternation." *English Language & Linguistics* 22(1): 1-27.

Zeschel, Arne. (2012). *Incipient Productivity. A Construction-Based Approach to Linguistic Creativity*. Berlin/Boston: De Gruyter Mouton.

Zipf, George Kingsley. (1949). *Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology*. Cambridge, MA: Addison-Wesley.

Zwicky, Arnold M. and Geoffrey K. Pullum (1983). “Cliticization vs Inflection: English n‘t’.” *Language* 59(3): 502–513.