

Cambridge Language Surveys

The Germanic Languages

Wayne Harbert

CAMBRIDGE

CAMBRIDGE

www.cambridge.org/9780521808255

This page intentionally left blank

THE GERMANIC LANGUAGES

Germanic – one of the largest subgroups of the Indo-European language family – comprises 37 languages with an estimated 470 million speakers worldwide. This book presents a comparative linguistic survey of the full range of Germanic languages, both ancient and modern, including major world languages such as English and German (West Germanic), the Scandinavian (North Germanic) languages, and the extinct East Germanic languages. Unlike previous studies, it does not take a chronological or a language-by-language approach, organized instead around linguistic constructions and subsystems. Considering dialects alongside standard varieties, it provides a detailed account of topics such as case, word formation, sound systems, vowel length, syllable structure, the noun phrase, the verb phrase, the expression of tense and mood, and the syntax of the clause. Authoritative and comprehensive, this much-needed survey will be welcomed by scholars and students of the Germanic languages, as well as linguists across the many branches of the field.

WAYNE HARBERT is Professor and Director of Undergraduate Studies in the Department of Linguistics, Cornell University. He has published extensively on syntactic topics, with a particular emphasis on historical syntax. His work on Germanic languages covers a wide range of problems in historical Germanic syntax and phonology, drawing on data from Gothic, Old English, Old High German, Old Saxon, Old Norse, and Modern German.

CAMBRIDGE LANGUAGE SURVEYS

General editors

P. Austin (*SOAS, London*)

J. Bresnan (*Stanford University*)

B. Comrie (*Max Planck Institute for Evolutionary Anthropology, Leipzig*)

S. Crain (*University of Maryland*)

W. Dressler (*University of Vienna*)

C. Ewen (*University of Leiden*)

R. Lass (*University of Cape Town*)

D. Lightfoot (*University of Maryland*)

K. Rice (*University of Toronto*)

I. Roberts (*University of Cambridge*)

S. Romaine (*University of Oxford*)

N. V. Smith (*University College, London*)

This series offers general accounts of the major language families of the world, with volumes organized either on a purely genetic basis or on a geographical basis, whichever yields the most convenient and intelligible grouping in each case. Each volume compares and contrasts the typological features of the languages it deals with. It also treats the relevant genetic relationships, historical development, and sociolinguistic issues arising from their role and use in the world today. The books are intended for linguists from undergraduate level upwards, but no special knowledge of the languages under consideration is assumed. Volumes such as those on Australia and the Amazon Basin are also of wider relevance, as the future of the languages and their speakers raises important social and political issues.

Volumes already published include

Chinese Jerry Norman

The Languages of Japan Masayoshi Shibatani

Pidgins and Creoles (Volume I: Theory and Structure¹; Volume II: Reference Survey) John A. Holm

The Indo-Aryan Languages Colin Masica

The Celtic Languages edited by Donald MacAulay

The Romance Languages Rebecca Posner

The Amazonian Languages edited by R. M. W. Dixon and Alexandra Y. Aikhenvald

The Languages of Native North America Marianne Mithun

The Korean Language Ho-Him Sohn

Australian Languages R. M. W. Dixon

The Dravidian Languages Bhadriraju Krishnamurti

The Languages of the Andes Willem Adelaar with Pieter Muysken

The Slavic Languages Roland Sussex and Paul Cubberley

The Germanic Languages Wayne Harbert

THE GERMANIC LANGUAGES

WAYNE HARBERT



CAMBRIDGE UNIVERSITY PRESS

Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo

Cambridge University Press

The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org

Information on this title: www.cambridge.org/9780521808255

© Wayne Harbert 2007

This publication is in copyright. Subject to statutory exception and to the provision of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published in print format 2006

ISBN-13 978-0-511-26690-4 eBook (Adobe Reader)

ISBN-10 0-511-26690-1 eBook (Adobe Reader)

ISBN-13 978-0-521-80825-5 hardback

ISBN-10 0-521-80825-1 hardback

ISBN-13 978-0-521-01511-0 paperback

ISBN-10 0-521-01511-1 paperback

Cambridge University Press has no responsibility for the persistence or accuracy of urls for external or third-party internet websites referred to in this publication, and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.

to Myrlyn, Diane, and Jenny

CONTENTS

List of tables x

List of language abbreviations xi

Acknowledgments xii

1 Introduction 1

- 1.1 Some remarks on the organization of this volume 1
- 1.2 Divergence and convergence in the Germanic languages 6
 - 1.2.1 Germanic languages and Standard Average European 9
 - 1.2.2 Typological classification 12
- 1.3 A survey of the Germanic languages 13
 - 1.3.1 East Germanic 14
 - 1.3.2 West Germanic 15
 - 1.3.3 The North Sea Coast languages 17
 - 1.3.4 North Germanic 19

2 The Germanic lexicon 21

- 2.1 Loanwords 22
- 2.2 Derivation 26
 - 2.2.1 Compounding 29
- 2.3 Discourse particles 32
- 2.4 Phrasal verbs 36

3 The sound systems of Germanic: inventories, alternations and structures 41

- 3.1 Segmental inventories and alternations 41
 - 3.1.1 The obstruents: place and manner of articulation 41
 - 3.1.2 The sonorants 53
 - 3.1.3 The vowels 56
- 3.2 The suprasegmental phonology of the Germanic languages 65
 - 3.2.1 Syllable structure and sonority 65
 - 3.2.2 Length 74
 - 3.2.3 Lexical stress and the foot in Germanic phonology 79
 - 3.2.4 Word tone in Germanic 84

4	The Germanic nominal system: paradigmatic and syntagmatic variation	89
4.1	Nominal inflection	89
4.1.1	Historical prelude: Indo-European heritage and Germanic innovation	89
4.1.2	Inflectional categories of the noun	91
4.2	The internal structure of the nominal phrase	122
4.2.1	Nominal phrases without nominal heads	122
4.2.2	Nominal phrases with pronoun heads	124
4.2.3	Nominal phrases headed by lexical nouns	126
4.3	Determiners	137
4.3.1	Weak quantifiers	137
4.3.2	Definite articles	141
4.4	Genitive phrases	148
4.4.1	Genitive phrases and determiners	150
4.4.2	Proper name possessors and pronominal possessors	155
4.4.3	The English 'double genitive' construction	158
4.4.4	The prenominal periphrastic possessive (<i>Jan se boek</i>) construction	158
4.4.5	Special developments of the genitive in possessive constructions	161
4.4.6	Meronymic constructions	164
4.5	Predeterminers	167
4.6	Discontinuous nominal phrases	168
4.7	Adjective phrases	171
4.7.1	Comparison	174
4.8	Pronouns in the Germanic languages	175
4.8.1	Personal pronouns	175
4.8.2	The referential properties of pronouns	196
4.9	The external syntax of noun phrases: subjects	214
4.9.1	Subject agreement	214
4.9.2	The typology and distribution of "expletive" arguments	224
4.9.3	Subjects in imperative clauses	236
4.9.4	Derived subjects and the syntax of voice	237
4.9.5	Raising constructions	256
5	The verbal systems of Germanic: paradigmatic and syntagmatic comparison	270
5.1	Historical prelude: the Indo-European heritage and Germanic innovations	270
5.1.1	Types of verbal inflection	270
5.1.2	Categories of verbal inflection: the Germanic tense/mood system and its Indo-European antecedents	272
5.2	Modal auxiliaries	285
5.3	Developments in the expression of tense in Germanic	292
5.3.1	Identifying periphrastic tense/aspect constructions	293
5.3.2	Future	297

5.3.3	Perfect tenses and past tenses	301
5.3.4	The meaning of the (present) perfect and the past	307
5.3.5	Progressive	315
5.4	Voice inflections and voice auxiliaries	316
5.4.1	Passive	317
5.4.2	Middle voice	322
5.4.3	The Scandinavian <i>s</i> -passive	327
5.5	Nonfinite verbal forms	329
5.5.1	The infinitive	329
5.5.2	The present participle and the English gerund	341
5.5.3	The past participle and the supine	345
5.6	Verbal valency	347
5.7	Head, complement and adjunct placement in the verb phrase	350
5.7.1	The relative order of the verb and its complements within the VP	353
5.7.2	The order of objects	362
5.8	Phrasal verbs	366
6	The syntax of the clause	369
6.1	Sentence adverbs	370
6.2	The syntax of negation	376
6.2.1	Some definitions: negation, scope and polarity	376
6.2.2	Scope, polarity and syntactic position	377
6.2.3	Polarity items, negative concord and multiple negation	379
6.2.4	Negative complementizers and pleonastic negation	382
6.2.5	Constituent negation with sentential scope	383
6.2.6	On formal differences between sentential negators and constituent negators	392
6.2.7	On the typology of sentential negation in Germanic	394
6.3	The syntax of the left periphery: topics, verb-second and subject/verb inversion	398
6.3.1	Main-clause/subordinate clause asymmetries in verb position	400
6.3.2	The triggers of verb-second order and the typology of V-2	404
6.4	Complementizers	415
6.4.1	That-clauses	415
6.4.2	Infinitive complements	417
6.5	Relative clauses, questions and other fronting constructions	420
6.5.1	A typology of Germanic relative constructions	420
6.5.2	Questions	473
6.5.3	Topic constructions and left dislocation constructions	478

References 482

Index 505

TABLES

- 4.1 Adjective endings in selected Germanic languages 99
- 4.2 Definite articles/demonstrative adjectives in selected Germanic languages 106
- 4.3 Second person pronouns in selected Germanic languages 193
- 5.1 Inflectional endings of active strong verbs in selected Germanic languages 285

LANGUAGE ABBREVIATIONS

AF	Afrikaans	ME	Middle English
BN	Bokmål Norwegian	MHG	Middle High German
DA	Danish	NF	North Frisian
DU	Dutch	NN	Nynorsk Norwegian
EF	East Frisian	NO	Norwegian
EME	Early Middle English	OE	Old English
EN	English	OHG	Old High German
FA	Faroese	ON	Old Norse
FR	Frisian	OS	Old Saxon
GE	German	PG	Pennsylvania German
GMC	Germanic	SAE	Standard Average European
GO	Gothic	SW	Swedish
IC	Icelandic	WF	West Frisian
IE	Indo-European	YI	Yiddish

ACKNOWLEDGMENTS

I would like to thank Abby Cohn, Sarah Fagan, Roger Lass and Paul Roberge for their willingness to read and comment on portions of this work. It has benefitted substantially from their attention and advice. Thanks are also due to Jessica Bauman for her valuable assistance in the early stages, and to Steve Barganski for his thorough and expert copy-editing. I owe a special debt of gratitude to Diane Harbert, who, in addition to her constant support throughout this long project, devoted many hours to proofreading the manuscript, catching a multitude of errors in the process. I gladly accept responsibility for the probably numerous errors which remain.

Introduction

1.1 Some remarks on the organization of this volume

No single volume can adequately address a topic area as broad as “The Germanic Languages” in all of its aspects. It is necessary to single out a particular dimension on which to focus. Languages can be looked at in their societal context, for example, with attention to such questions as their use and significance in the communities of speakers who employ them, their relationship with the associated cultures (including, for example, literary uses), their demographics and their variation along geographical and demographical dimensions. One can alternatively regard language from a historical perspective, as chronological sequences of divergences and convergences, states and transitions. Each of these points of view has provided the organizational framework for successful volumes on the subject. It is also possible, abstracting away from their social, geographical, cultural and temporal contexts, to examine the languages of the family as assemblages of grammatical units, rule systems and constructions. This is the perspective which I will adopt here. The present volume is aimed primarily at those who are interested in how the Germanic languages are put together – what they have in common in terms of their linguistic organization and how they differ from each other structurally. That choice in turn determines several other features of the organization of the volume. In particular, I will not adopt the standard and often successful approach of covering the territory by means of a series of self-contained descriptions of individual languages. That encyclopedic approach is an ideal format for describing languages in their socio-cultural setting, since the demographic, historical, cultural and geopolitical situation of every language is unique. When the focus is on the grammatical structures, patterns and inventories of the languages, though, such an organizational model becomes less ideal. For one thing, it necessarily leads to a large amount of repetition. The Germanic languages are, after all, more alike than they are different, and this becomes increasingly true the farther one descends the genetic tree. Once one has read about the structure of the noun phrase in Swedish, for example, a description of the noun phrase in Danish will present few surprises. Such a format is also not

conducive to side-by-side comparison of the ways in which the languages accomplish particular tasks, and so does not present a ready picture of structural commonalities and differences across the family. I have therefore decided to organize the discussion according to linguistic constructions and subsystems, rather than by languages. For example, there is a section on vowel systems, a section on the expression of future tense, and a section on relative clause formation, in each of which an individual Germanic language may be mentioned or not, depending on whether it offers something of particular interest in connection with the grammatical phenomenon under investigation. These decisions will no doubt make the volume less useful for readers with certain purposes. In particular, since it does not include chapters on individual Germanic languages, it does not provide a sense of how the grammars of individual languages work as integrated systems. Fortunately, there are other volumes suited to the interests of readers who want to inform themselves about the shape of individual Germanic languages. König and van der Auwera 1994 is particularly to be recommended. There are also volumes which approach these languages from a historical perspective – most notably, the recent volume by Howell, Roberge and Salmons (forthcoming). It is hoped that what is lost in the present treatment in terms of coherent pictures of individual languages is compensated for by a clearer family portrait.

A further practical consideration in favor of the present format is that it allows us to sidestep the thorny question of how many Germanic languages there are, and which varieties to devote chapters to. In volumes on language families in which the main aim is the exhaustive description of particular languages it is usual to single out a particular variety of each language as the object of that description. Most often, the written standard variety is chosen (even though linguists recognize the privileged position of standard languages to be largely a matter of historical accident), and nonstandard dialects are given relatively short shrift. Such an *a priori* limitation would simply not work in a study in which the main focus is the range of grammatical phenomena found in the Germanic languages, since, as we will be seeing, the family abounds in highly interesting and sometimes widespread linguistic developments which happen only to be found in nonstandard varieties. The standard languages show a relatively high degree of homogeneity, in part the result of their centuries of contact with each other and other Western European standard languages as languages of high culture and literature. The range of structural variation among these varieties is thus relatively small in comparison with that found when nonstandard varieties are taken into account. The division of the territory into, for example, a chapter on Dutch (represented by standard Dutch) and a chapter on German (represented by standard German) is arbitrary not only because of the substantial variation that exists within the individual languages, but also because of the famous fuzziness of the boundaries between languages in some cases. The Germanic languages include two notable dialect

continua – the West Germanic dialect continuum, encompassing Belgium, the Netherlands, Germany, Switzerland and Austria, and the Scandinavian dialect continuum, encompassing Denmark, Sweden, Norway and parts of Finland – in which it is impossible to draw non-arbitrary language boundaries between neighboring varieties at any point (see Crystal 1987: 25).

There are also some more theoretical reasons for adopting the construction-by-construction approach followed here. While I have attempted as much as possible to keep linguistic theory in the background, this book is very much informed by the spirit of recent “principles and parameters” approaches to linguistic variation – the idea that languages are not free to differ from each other arbitrarily and without limit, but rather that linguistic variation is constrained by general parameters of variation, and that therefore structural differences across the languages of the family may be expected to be patterned, rather than random. The construction-by-construction, side-by-side format of the volume serves to highlight such patterns of variation as are found.

Once the decision was made to organize the presentation around patterns, paradigms and constructions, rather than around languages, no principled reason remained for including only the modern members of the family. If the book were an examination of languages in context, partitioning the Germanic languages according to the salient demographic property of having or lacking native speakers might make sense. Once the focus is on structure, though, separating them on the basis of this criterion seems plainly more arbitrary, since few structural differences correlate with this distinction. Whether a particular variety belongs to early Germanic or modern Germanic is not entirely unrelated to its structural characteristics; there are several features which unite the postmedieval members of the family, and distinguish them from the early medieval varieties. Some such differences arise by virtue of the fact that the later languages, but not the earlier ones, were around to participate in pan-European diffusions of such features as the distinction between formal and familiar forms of second person pronouns and indefinite articles. Others arise because the later languages, but not the earlier ones, participated in late parallel developments such as open syllable lengthening or the rise of medial negators. More often, though, the linguistic features which turn out to have general predictive value – whether a language is O(bject) V(erb) or VO, for example, or whether it has lexically case-marked arguments or not – are ones which crosscut the early Germanic/modern Germanic distinction. Therefore, it was decided for present purposes to treat both the pre-modern members of the family and the living members of the family on a par, to the extent possible – as different variants on a common theme. In treating Gothic and Old High German, for example, side-by-side with Afrikaans and Faroese, this volume differs from most other treatments of the Germanic languages.

The descriptions offered here are not theory-neutral; I doubt that it is possible to do linguistic description in a truly theory-neutral way. My particular training and

inclination as a syntactician (working within the Government-Binding/Principles and Parameters tradition) necessarily informs my descriptions of particular phenomena, the kinds of explanations I offer for instances of variation (in which the assumption of parameters will play a prominent role, for example), and even, to an extent, the kinds of structures and phenomena singled out as worthy of description and explanation, as well as those left out of consideration. A scholar with a different theoretical orientation would imaginably have produced a somewhat different work. Nonetheless, an effort has been made to keep theoretical assumptions in the background in order to make the descriptions accessible to all readers with a background in linguistics, and to deploy theory-specific terminology only when it substantially contributes to the efficiency of the description.

The goals of the work are fundamentally synchronic: to identify and describe structural similarities and differences across the Germanic family. Nonetheless, it will be seen that discussion of linguistic history intrudes with some frequency. There are various reasons for this. For one thing, many of the accounts offered for the distribution of features across these languages are typological in nature. Many claims are made of the following sort: languages in a subgroup of the Germanic languages share a feature Y because they share a linked property Z from which the presence of Y follows. The validity of such typological linkages is supported by showing that they vary together over time – that when Y arises by historical change, Z appears too. Second, there are some shared features of the family or subgroups within it whose appearance and distribution can only be explained in historical terms: features which exist only because of historical facts of inheritance or borrowing. Some of the more interesting cases involve differences in the uses to which inherited “junk” (Lass 1988) are put. See, for example, the discussion of the weak/strong adjective contrast in Section 4.2.3.3 and the discussion of the development of the reflexive/nonreflexive possessive distinction in German (Section 4.8.2.1.1).

Discussions of phonology and the lexicon are accorded less space than the discussion of morphology and syntax. The particular choices made with respect to how much attention to give to each of these topics reflect, besides space limitations and the particular interests of the author, the fact that morphosyntactic aspects of the grammar are more amenable to the systematic contrastive treatment adopted here; their side-by-side investigation holds out the most promise of helping us to answer one of the central questions of the volume: In what systematic ways are the Germanic languages alike and in what ways different? Aside from prosodic phonology, it is difficult to make typological statements about the sounds or the vocabulary of the Germanic languages of comparable generality to those possible for the morphosyntax of these languages.

This book does not separate the treatment of morphology (or “accidence”) from the treatment of syntax in the way that is familiar from most handbooks of Germanic

languages. This fact, too, is related to its central goal of systematic, side-by-side comparison of the Germanic languages; there are numerous cases in which some languages in the family use inflectional morphology to encode particular structural relations among elements, while others avail themselves of syntactic means for this purpose. So, for example, Gothic and some of the Scandinavian languages have passive affixes – a matter of morphology – but these are functionally equivalent to the periphrastic passive markers of other languages in the family, which are syntactic in nature since they are free morphemes. Similarly, some of the languages of the family exhibit case inflections – a matter of morphology – but the same grammatical relationships which are encoded by means of these are encoded by means of free morphemes (in the form of prepositions) in other languages. Treating the two separately because one is a morphological phenomenon and the other a syntactic phenomenon would, of course, obscure the fundamental point of their functional equivalence. Instead, I have chosen a different organizational scheme, based on lexical classes. Chapter 4, for example, treats the morphosyntax of nouns and the other lexical categories (adjectives, determiners and pronouns) with which they are associated, and the syntax of the phrases in which these categories participate. Chapter 5 is devoted to verbs and their phrases. Within each of these discussions, there is a secondary division into a discussion of the paradigmatic properties of these lexical items followed by a discussion of their syntagmatic properties. Paradigmatic relationships are the relationships obtaining between an expression and other expressions which are substituted for it in different contexts. Case paradigms, and their prepositional phrase equivalents in languages without case (e.g., *the man*, *to the man*, *of the man*. . .) are instances of paradigmatic relationships. The syntagmatic relationships of a linguistic expression are the relationships which hold between it and non-equivalent expressions with which it is concatenated in forming larger linguistic expressions. The relationship between a subject and a verb, for example, is of this type. To a certain extent, this distinction overlaps with that between morphology and syntax, since, for example, case paradigms are a matter of morphology, and putting together a noun phrase and a verb phrase to form a sentence is a matter of syntax, but the two are not entirely isomorphic.

By its nature, a survey of this sort consists largely of reports of previous scholarship. This work owes a great debt to the centuries-long tradition of description of Germanic languages, and, in particular, to a recent spate of reference grammars and grammatical sketches of high quality for individual languages. The reader may find the following to be of particular interest: Allan *et al.* 1995; Bandle 2002, 2005; Booij 1995, 2002b; Braunmüller 1991; Collins and Mees 1996; Donaldson 1981, 1993; Engel 1988; Haugen 1982; Holmes and Hinchcliffe 1994; Jacobs 2005; Katz 1987; König and van der Auwera 1994; Kristoffersen 2000; Lass 1994; Lindow *et al.* 1998;

Lockwood 1995; Mitchell 1985; Tiersma 1999; Zifonun *et al.* 1997. However, the exercise of creating a construction-by-construction comparison of all of the Germanic languages (and attempting to fill in the considerable gaps in the available descriptions of the older languages in particular, as required by that exercise) has turned up occasional patterns and generalizations which had not been observed before.

1.2 Divergence and convergence in the Germanic languages

Germanic (hereafter, GMC) is, in the first order, a genetic concept. The GMC languages share many properties and constructions by virtue of common ancestry. Common inheritance is the reason, for example, that they signal inflectional contrasts by a mixture of suffixation and alternations in root vowel. It is also the reason that they have only a single inflectional past tense, and do not distinguish between preterite and imperfect, for example. The GMC languages share the first of these characteristics with other languages with which they are more remotely related, including the neighboring Celtic and Romance languages. The second, however, is a GMC innovation, which sets GMC apart from these other branches of Indo-European (IE), the larger family to which it belongs, including Celtic and Romance. Among the other distinguishing characteristics of the GMC languages which set them apart from their IE ancestor are:

- the fixing of the accent on the root or first syllable of the word (Section 3.2.3) and the possibly related tendency to reduce final syllables
- the incorporation of verbal nouns and verbal adjectives into the verbal paradigms as infinitives and participles
- the reduction of the system of inflectional tenses to a simple contrast between non-past and past and the possibly related tendency to introduce new periphrastic constructions for the expression of tense and aspect
- the introduction of a class of weak verbs, with “dental preterites”
- the systematization and restructuring of vowel alternations (ablaut) in the signaling of tense contrasts in the inherited strong verbs (Section 5.1)
- the reduction of the IE inventory of moods by conflation of the subjunctive and optative (5.1.2.2.1)
- the reduction of the inherited system of inflections for verbal voice and the consequent introduction of periphrastic passives (Section 5.4)
- the introduction of a strong/weak inflectional distinction in adjectives (4.2.3.3)
- the introduction of definite articles

- the reduction of the IE case system to four core cases (nominative, accusative, dative, genitive, with occasional survivals of other cases) (4.1.2.3.1)
- the development of a productive class of so-called weak nouns, based on IE *n*-stems
- the introduction of relative pronouns (6.5.1.1) based on demonstrative and (secondarily) interrogative pronoun paradigms
- the introduction of verb-second word order (Section 6.3).

In some cases, the common inheritance of the GMC languages has taken the form of an inherited dilemma, to which the individual languages have responded with individual and original solutions. This is illustrated in the interesting example of the varied treatment of weak and strong adjective endings, for example, as discussed in 4.2.3.3.

Some of the GMC languages are more closely related than others. The precise nature of these genetic relationships has historically been a matter of dispute (see Nielsen 1989 for an overview of early GMC), but a very widely accepted hypothesis is that GMC first split into a Northwest GMC branch and an East GMC branch (represented almost solely by Gothic). The differences between the East GMC group and the Northwest GMC group are partly matters of regional variation. So, for example, IE final **-ō* became *-a* in Gothic (*giba* ‘I give’), but *-u* in Northwest GMC (Old High German *gibu* ‘I give’). To some extent, though, they are matters of chronology. Many differences between East GMC and Northwest GMC reflect the fact that East GMC separated from the rest of GMC early and was recorded early, and so retains archaic features lost in the remaining languages (such as passive inflections and reduplicated verbs), and fails to participate in the later innovations in which those other languages took part.

Northwest GMC in turn is hypothesized to have split into a North GMC and a West GMC branch. The existence of a North GMC subgroup is beyond dispute, given the strong familial resemblance of its member languages to each other; these languages are the products of a very robust heritage of common innovation in all areas of grammar, which sets them apart from the rest of GMC, and the resemblances have been further reinforced by subsequent sustained contacts, with the result that there is still today a high degree of mutual intelligibility among them. The evidence for a West GMC genetic subgroup is more problematic, and has been called into question (though see Voyles 1971). On surer ground is the existence of a strongly innovative subgroup of West GMC languages, the North Sea Coast, or Ingvaenic group, consisting of Anglo-Saxon, Frisian and Old Saxon, which share a number of features to the exclusion of German, their sister West GMC language. Among these features, perhaps the ones with the greatest systematic significance are the loss of person distinctions in the

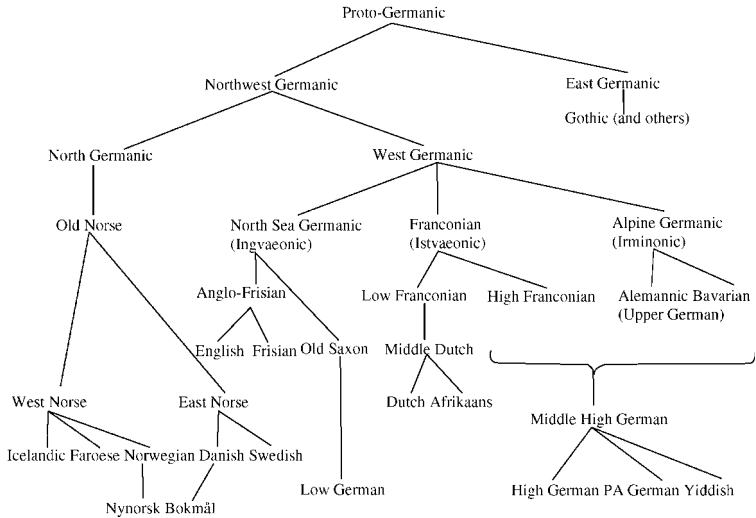


Figure 1.1 *The Germanic Family Tree*

plural verb, loss of case contrasts in part of the pronominal paradigms, and loss of GMC reflexive pronouns. Each of these will be discussed in Chapter 4.

The tree shown in Figure 1.1 gives a widely accepted, though not uncontroversial, picture of the genetic relationships among the GMC languages.

This tree sets forth a hypothesis about genetic relatedness, its branches graphically representing the order of divergence from a common ancestor (“the tree model”). Such tree diagrams do not give a complete picture of the interrelationships among them, though, and must be supplemented by another graphic device, such as the curly brackets used here. Similarities between languages are not always the result of common ancestry. Rather, originally separate varieties can converge over time through borrowing/areal spread of linguistic features across geographical space and linguistic boundaries (the “wave model”). In addition to shared ancestry, the GMC languages have remained geographically contiguous, creating the constant possibility of linguistic borrowing, mutual influence, and consequent convergence. For example, note that there is no single branch of the tree which dominates “German”; the German language (to the extent that it is a unitary language at all) is the product of centuries of mutual influence between originally separate West GMC linguistic groups. In a similar way, the varieties that we label “Low German,” regarded now (in part for political reasons) as a variety of German, originated as a variant of Ingvaenic West GMC – Old Saxon – which originally had more in common with Old English, but which has been

“Germanized” by successive waves of linguistic influence from the south. For early GMC, Rosel (1962) and Nielsen (1989) have reconstructed a complex history of periods of waxing and waning linguistic and cultural affinity between GMC subgroups, in order to account for the pattern of shared features. In later GMC, besides the interactions which gave rise to modern German, we can mention the mutual influence among the Scandinavian languages (particularly during the period of Danish hegemony beginning in the fourteenth century and lasting, in the case of Norwegian and Faroese, into the twentieth century), which resulted in a high degree of homogeneity at all levels, the possibly profound Norse influence on English beginning in the Old English period, which has been implicated in many of the features of Modern English but whose effect on the grammar of English is still awaiting a full evaluation, and the strong influence of Low German in late medieval times on the Scandinavian languages during the period of the Hanseatic league. The effect of the latter appears to be particularly strong in Danish, which in some respects (including phonology (Section 3.2.2) and syntax (Sections 4.9.4.1.2.2 and 4.9.5.2), for example) resembles German more than the other Scandinavian languages. In some cases, the effect of contact has been claimed not to be limited to direct borrowing, but to appear in grammar simplification/constructional loss, as a result of disrupted transmission of the language between generations (e.g., Norde 2001: 243; McWhorter 2002).

1.2.1 Germanic languages and Standard Average European

Such convergence by diffusion of linguistic features across boundaries is possible even when the languages in question are not related, or only remotely related. Vennemann, in a series of papers (Vennemann 2003a,b,c), has hypothesized such external influences from the very earliest period of GMC prehistory (see also Schrijver 2003). As a result of such contacts with neighboring languages, the GMC languages in modern times have been claimed to have become, in greater or lesser degrees, part of a group of “Standard Average European (SAE) languages,” which share with other languages of north central Europe (notably Romance languages) a cluster of linguistic constructions to the exclusion of geographically more distant languages on the European periphery. Haspelmath (1998) discusses the eleven most compelling features of SAE, though suggesting that there are other, weaker ones. These are:

- a. Definite and indefinite articles. This is, in fact, problematic as an SAE feature. While the GMC languages all exhibit definite articles, at least in an embryonic form (see Section 4.3), they share these not only with Romance, but with the Celtic languages, which are not part of the SAE cluster. On the other hand, while most of the modern GMC languages

have indefinite articles in common with the Romance languages, these are demonstrably of late origin. They are not found in the earliest attested versions of these languages, nor yet in Modern Icelandic.

- b. *Have*-perfects. All of the extensively attested GMC languages except Gothic have a periphrastic perfect formed with *have* plus a past participle – which they share with all of the Romance languages, as well as Czech and some Balkan languages.
- c. All GMC languages, including Gothic, have a periphrastic passive formed with the past participle plus a verb of being or becoming (Section 5.4.1). They share this feature exclusively with the Romance and Slavic languages, according to Haspelmath.
- d. Anticausative prominence. Languages make use of various means for deriving verbs from other verbs while changing their valency. In Gothic, for example, transitive/causative verbs sometimes involve additional morphology, relative to their intransitive counterparts (*wakan* ‘to be awake’ / *wak-j-an* ‘to waken someone’), but sometimes intransitives are morphologically more complex than their transitive/causative counterparts. Included here are inchoatives with the inherited *-nan* suffix (*gaskaidnan* ‘to divorce (intrans)’ ~ *skaidan* ‘to separate (trans)’)) and the apparently innovated reflexive middle verbs (*sik laisjan* ‘to learn’ – literally, ‘to teach oneself’). Haspelmath claims that the derivation of intransitives from transitives (through the addition of “anticausative” morphology) is most frequent in German, French, Romanian, Russian, Greek and Lithuanian, while “causativization” – the derivation of transitives from intransitives – is more usual in neighboring non-SAE languages.
- e. Nominative experiencers. In some languages the semantic argument roles of agent or actor are assigned to nominative subjects, and for semantic roles other than agent/actor (including roles such as experiencer and possessor) are represented by non-nominative nominal phrases (see Section 4.2.1.4.2). Thus, for example, in Scottish Gaelic, *I have a book* is expressed as *Tha leabhar agam* ‘Is a book with-me’, and *I like the book* is expressed as *Is toil leam an leabhar* ‘Is pleasing with-me the book’. According to Haspelmath, the SAE languages, to a greater extent than neighboring languages, tend to realize these experiencer and possessor arguments, too, as nominative subjects, as English does. The fit of this feature with other hypothesized SAE features is quite loose, however. On the one hand, Basque and Turkish – not SAE languages by other standards – have a high proportion of nominative experiencers. On the other hand, Icelandic and Faroese have low ratios, and are thus excluded from the SAE fold.

- f. Dative external possessors. In Romance, continental West GMC, Gothic, Balto-Slavic, Hungarian, Greek and Armenian, the possessor is expressed as a dative in certain kinds of possessive constructions (Gothic *afmaimait imma auso* ‘he-cut-off him-Dat ear’). This construction is missing in non-SAE languages, but also in Modern English (see Section 4.4.6 and McWhorter 2002). Among the IE languages, it is lacking only in English and Insular Celtic, according to Vennemann (2003b: 356).
- g. Many of the SAE languages, including Romance languages, GMC and Albanian, allow negated nominal phrases to carry the force of sentential negation (Gothic *ni waihts im* ‘I am nothing’). The distribution of this possibility in GMC is explored in Section 6.2.5. It is found in all GMC languages except Old Saxon. Non-SAE languages tend to require a sentential negator, typically attached to the verb.
- h. Particle comparatives. The SAE languages – including GMC (Section 4.7.1), Romance, Balto-Slavic, Balkan languages, Hungarian, Finnish and Basque, according to Haspelmath – characteristically make use of a “comparative particle” which is not a preposition, and thus does not affect the case of the following nominal (Gothic *frijondans wilja seinana mais þau gub* ‘loving their own will more than God-Acc’). Non-SAE languages use a variety of different devices.
- i. A-and-B conjunction. Most European languages, according to Haspelmath, including the SAE languages, make use of an A-and-B construction for conjunction, as opposed to a variety of other devices (e.g., A-and B-and, A B-and) found elsewhere.
- j. Relative clauses. Among the many strategies available for forming relative clauses (Section 6.5.1.5.1), the use of a strategy involving relative pronouns which occur at the front of the relative clause and encode the case number and person features of the relativized argument (Gothic *sunus meins, in þuzei waila galeikada* ‘my son, in whom I am well pleased’) is claimed by Haspelmath to be unique to SAE languages.
- k. Verb fronting in polar questions. In GMC, Romance and Slavic languages and modern Greek, yes/no questions are formed by fronting the verb (Section 6.3.2), rather than by intonation or the use of a question particle.

This SAE hypothesis must be regarded with caution for a number of reasons. First, it is clear that the different hypothesized SAE features do not pick out the same subsets of languages. Second, they are attributable to widely different time periods. One of them – the external dative construction – is probably of IE date, as Haspelmath notes.

Others date from different prehistoric eras (Gothic, for example, has a periphrastic passive but no periphrastic perfect), while at least some others date from postmedieval times. This is true, for example, of indefinite articles, and the events which led to Icelandic and Swedish being on opposite sides of the fence with respect to nominative experiencers. The SAE constructions also sometimes differ in detail in the languages that share them, in ways which undermine the likelihood of common origin. GMC, for example, uses ‘become’ as the cardinal passive auxiliary, while Romance uses ‘be’, and it is not unimaginable that the two could have originated independently, starting out from copular sentences in which the predicate was a participle functioning as a stative adjective. Similarly, the core GMC relative pronouns are based on demonstrative pronouns while the relative pronouns of Romance are based on interrogative pronouns, and the syntactically comparable relative pronouns of ancient Greek appear to have been relative pronouns from the beginning (Fortson 2004: 130). Thus, we might be dealing here with three independent responses to the breakdown of the inherited correlative syntax of IE (see Section 6.5.1), especially in view of the difficulty of defending on other linguistic or historical evidence the existence of one or more eras of contact sufficient for diffusion of the construction to have taken place.

Other attempts have been made as well to account for features of subgroups of GMC, by diffusion of constructions across major language boundaries. Lindstedt (2000: 371) claims that the loss of simple past tense in some varieties of spoken German and the extension of present perfect to the expression of (imperfective) past tense is an areal development, encompassing southern German as well as northern Italian and spoken French.

1.2.2 Typological classification

In addition to genetic groupings, it is sometimes profitable to group the GMC languages typologically, according to structural properties – groupings which often crosscut genetic groupings, as well as each other. For example, it is proposed in Section 4.9.5 that those GMC languages which have developed V(erb)–O(bject) syntax have a wider range of “noun-phrase raising” constructions than OV languages, and in Section 5.7.1.1 it is observed that “verb-raising” is characteristic of the OV languages of GMC. It also appears that those languages which have lost morphological case contrasts in the noun phrase, whether VO (English, Swedish, Danish, Norwegian) or OV (Dutch, West Flemish, Frisian, Afrikaans), also have some properties in common to the exclusion of GMC languages with richer morphological case (see Section 5.7.2.1). Such instances are often viewed as the results of implicational linkages between grammatical properties, of the sort pursued in “principles-and-parameters” models of linguistic variation. Thus, in addition to sharing features by

virtue of common inheritance, and by virtue of language contact/areal spread, some features shared by subsets of the GMC languages arguably arise as by-products of other typological commonalities.

English is in many respects a typological outlier in the GMC family. McWhorter (2002: 265) claims that “where a subset of GMC languages have departed sharply from the original GMC ‘typology,’ English never fails to be a member,” and enumerates ten features with respect to which English deviates from all or virtually all the rest of GMC. These include the loss of inherent reflexives (in middle voice constructions, for example), the absence of external possessors (Section 4.4.6), the absence of gender (Section 4.1.2.2), the loss of GMC prefixes, the absence of a perfect construction with ‘be’, the absence of passives with ‘become’, differences in verb-second syntax (Section 6.3), the absence of a distinctively singular form of the second person pronoun (Section 4.8.1.5), and the loss of the indefinite pronoun *man*. McWhorter attributes these changes to disruption of intergenerational transmission during the Danish conquest.

1.3 A survey of the Germanic languages

Another dimension along which the GMC languages can be divided is the distinction between state and non-state languages. The GMC languages, relative to other language families of the world, include a high proportion of national languages. These include some varieties of English (hereafter EN), German (GE), Dutch (DU), Swedish (SW), Danish (DA), Norwegian (NO), Faroese (FA), Icelandic (IC), Afrikaans (AF) and Luxembourgish, as well as some English-based creoles (Tok Pisin in Papua New Guinea and Bislama in Vanuatu). Other GMC languages are non-state languages. These include Yiddish (YI), Pennsylvania German (PG), Schwyzertütsch, East Frisian (EF) and all of the varieties regarded as nonstandard “dialects” of the state languages (including some sufficiently remote from the standard variety that they would count as separate languages under different political circumstances). Intermediate in status is West Frisian (WF), which is officially the second language of the Netherlands, though according to Gorter *et al.* (2001: 111), this status has “only entailed moderate promotion by the state.” Some varieties, such as Limburgisch in the Netherlands and North Frisian (NF) in Schleswig-Holstein, have official standing as “regional languages.”

This accidental distinction between state and non-state languages is not unconnected with the internal shape of the language, since official recognition as a standard for public purposes usually brings with it greater normalization and regulation, the development of varying degrees of stylistic divergence between written and spoken varieties, and the articulation of the vocabulary in certain domains (e.g., technology and bureaucracy). It is also, of course, of very profound importance for their external

histories and their viability over the long term; all of the non-state languages mentioned above are at considerable risk of extinction. With few exceptions (YI and PG), the situation of non-state GMC languages is made yet more problematic by the fact that they are sufficiently similar to the surrounding dominant languages that they tend to be dismissed as “mere dialects” of those languages, and denied recognition as independent languages, along with such prestige and consideration as comes with that status.

1.3.1 East Germanic

East GMC has no surviving members. It is represented mainly by Gothic (GO), preserved in sixth-century manuscripts of portions of a fourth-century Bible translation, eight leaves of a piece of biblical exegesis (named the *Skeireins* by its first editor) and scattered minor documents.¹ GO apparently survived as a liturgical language along the Danube into the ninth century, and a quite remarkable survival of GO into modern times was documented in the 1560s by Ferdinand Ogier de Busbecq, a Flemish diplomat, who recorded a word list of about a hundred items from informants from the Crimea (see Stearns 1978). GO is written in its own alphabet, based largely on that of Greek with possible influence from Latin and GMC runes, apparently devised by the bishop Wulfila, to whom the Bible translation is also attributed. Other East GMC languages are known only through meager onomastic evidence. The received view is that since the GO corpus consists mainly of highly faithful translations from Greek originals, it is of no value for the recovery of GMC syntax – a view which the present discussion will call into question.

Antonsen (1975: 24ff.) takes the position that Northwest GMC is recorded at roughly the same time (starting in the third century) in a broad area of northern continental Europe and southern Scandinavia, in the earliest GMC runic writing, though the identification of the language of these inscriptions as Northwest GMC has been disputed. Runic writing slightly antedated the first GO writing, and it is possible that Wulfila was familiar with it and made marginal use of it in devising his alphabet. The runic writing system, known as the *futhark* because of the canonical order of its first six characters, was based on western Mediterranean models. The corpus of inscriptions in the earliest runic writing system, the older *futhark*, though scant, is of great importance for our understanding of GMC. Although GO is frequently cited for its archaic nature, it has in fact deviated from common GMC relatively far in comparison to the roughly contemporary language recorded in the

¹ In the present work, all citations from the Gothic Bible, as well as their Greek equivalents, are taken from the edition of Streitberg 1960.

brief epigraphic documents in the older *futhark* – particularly with respect to the preservation of vowels in unstressed syllables. I will cite here where possible only examples in the earliest period of the older *futhark*, which Antonsen labels Northwest GMC, though I will refer to them more neutrally as Runic GMC. The runic *futhark* writing tradition persisted long after the separation of the Northwest GMC languages, surviving in modified forms in the Anglo-Saxon *futhorc* in which the earliest EN literature is recorded, and the younger *futhark*, which remained in use in Scandinavian countries until modern times.

1.3.2 West Germanic

The individual languages of the West GMC group are the next attested, Old English (OE) and Old High German (OHG) being documented from roughly the eighth century, and Old Saxon (OS) in the ninth century, Frisian (FR) and DU from still later periods.

1.3.2.1 *German and its siblings*

Standard GE is conventionally referred to as (New) High GE. The qualifier ‘High’ is a topographical term, referring to those varieties spoken in the relatively higher elevations of the south, as opposed to the low-lying plains of northern Germany. Linguistically, the two varieties are most saliently distinguished by the fact that High GE varieties participated in all or part of the consonant changes known as the High German consonant shift (Section 3.1.1.3). In fact, the various dialects of GE can be arrayed on a cline according to the degree to which they participated in these changes. They are most fully implemented in the Alpine dialects in Swabia, Bavaria, Austria and Switzerland, but carried through less consequently in the Franconian dialects, and not at all in the Low GE dialects. Thus, dialectologists speak broadly of Upper GE, Middle GE, and Low GE dialects. Standard GE is based largely on Middle GE (Franconian) dialects, whose phonology it reflects. The Upper GE dialects and the Franconian dialects were originally separate branches of West GMC, which have coalesced over the centuries into a (partial) linguistic unity through mutual linguistic influence. Low GE arose from OS, a North Sea Coast language, and still reflects this origin in some ways – for example, in the single form of verb endings for all persons in the plural. Partly in consequence of its heterogeneous origins, GE is much more dialectally diverse than any of the other modern GMC languages. Gordon (2005) reports that the Mainfrankish dialect and the Swabian dialect, for example, are only about “40% inherently intelligible with Standard GE.”

GE is conventionally divided into three chronological periods: Old High German (OHG), from the earliest attestation (eighth century) to the end of the eleventh century,

Middle High German (MHG), from the eleventh century to the end of the sixteenth century, and Modern GE (NHG), from the end of the sixteenth century. The original literature for the first of these is relatively small, becoming quite voluminous in later periods. The various varieties of GE have about 100,000,000 speakers, constituting the largest linguistic community in Western and Central Europe. Standard GE is the official language in Austria (7.5 million), Germany (78 million), Liechtenstein, one of four official languages in Switzerland (4 million), where a majority speaks Swiss GE (Schwyzertütsch), and one of four official languages in Luxembourg. It is also spoken among minority populations in central and eastern Europe and northern Italy. The largest number of GE speakers outside Europe is in the USA, where there are as many as 1.5 million speakers.

Two other West GMC languages, YI and PG are basically of Middle GE origin, as demonstrated by the degree of their participation in the High German Consonant shift, though the facts of YI are complex. The precise date and dialect of the GMC source of YI, and the question of whether it is most appropriately characterized as a West GMC language or as a “fusion language” are addressed in Jacobs 2005: 9ff. A main dialect division arose between those varieties spoken in GMC-speaking countries (Western YI) and those spoken in the Slavic countries to the east (Eastern YI). Western YI declined due to linguistic assimilation to GE, and the vast majority of YI speakers speak Eastern varieties. There are four major modern dialects of YI: Western, Central, Northeastern and Southeastern, Standard YI reflecting features of the latter two (Jacobs *et al.* 1994: 390). YI is written in the Hebrew alphabet, and is thus the only modern GMC language not written in Roman letters. (Examples in the present work are given in transliteration, following the conventions of the particular source from which the examples are cited.) YI is spoken by about 5 million people, in New York, Israel, Melbourne, Montreal, Mexico City and Buenos Aires (Jacobs 2005: 3). It is not an official language in any of the countries in which it is spoken, and not a majority language in any geographical area larger than individual neighborhoods, though it has achieved stable bilingualism in many of these areas, as the language of ultraorthodox Jewish communities (Fishman 2001).

PG (Pennsylvanisch, sometimes referred to in EN colloquially and inaccurately as Pennsylvania Dutch) was introduced into North America beginning in the early eighteenth century. It has maintained itself most robustly as the community language of Old Order Mennonite and Old Order Amish communities, in which spheres it has also achieved a relatively stable bilingualism. It resembles most the Frankish dialects of GE. There are 2 to 3 million speakers of PG in Canada, the United States, Central and South America.

Luxembourgish (Luxembourgeois, Letzeburgisch), also a Moselle Franconian dialect in origin, is the national language in Luxembourg, and spoken natively by a

large majority there (300,000), with small populations in Belgium and France. It is not much used for written communication, in which domain French and GE are used instead.

Limburger, or Limburgs Plat, a Middle Franconian language, more or less mutually intelligible with Riparian dialects of GE, according to Gordon (2005), is spoken by around 1.5 million people in the provinces of Limburg in the Netherlands and Belgium, having official status as a regional language in the former.

1.3.2.2 *Dutch and its siblings*

DU (Nederlands) is the descendant of Old Low Franconian – that is, that portion of Franconian West GMC which did not undergo the High German consonant shift – though it also shows the effects of contact with Ingvaemonic languages. Reflecting the linguistic complexity of northwestern coastal Europe, DU has a wide range of dialects, with sometimes rather low mutual intelligibility. Chronologically, DU is divided into Old DU (to 1150), of which only scant records exist, Middle DU, from 1150 to 1500, which is represented by a very substantial original literature, and Modern DU, since 1500. DU is currently spoken by about 21 million people in Europe – 15 million in the Netherlands and 6 million in Belgium. The variety spoken in Belgium is often called Vlaams (Flemish), and listed by Gordon (2005) as a separate language, though it is not generally so treated. DU is also an official language in former DU colonies in Surinam and the DU Antilles, and is spoken in Indonesia as well.

AF originated from DU dialects, transplanted to southern Africa in the seventeenth century, though it has been transformed by massive effects of language contact, and is, besides EN, the GMC language which deviates grammatically the farthest from the others. It is spoken by about 6 million people, and is one of the official languages of the Republic of South Africa (co-official with EN, Ndebele, Southern and Northern Sotho, Swazi, Tsonga, Tswana, Venda, Xhosa and Zulu), and Namibia.

1.3.3 The North Sea Coast languages

FR (Frysk), an Ingvaemonic or North Sea Coast language closely akin to EN in its origins, is spoken by about 750,000 people in the Netherlands and Germany, constituting the language of about 4% of the inhabitants of the former (Gorter *et al.* 2001: 103). FR has four dialects with low mutual intelligibility. WF, in northern coastal Holland, is spoken by about 400,000 people, and has a literary standard (Hoekstra and Tiersma 1994: 506). Northern FR is spoken in Schleswig-Holstein and adjacent islands in northern Germany by about 10,000 people. Eastern FR is spoken in a small region in northwestern Germany near the Dutch border by a very few speakers (1000, according to Hoekstra and Tiersma 1994: 505, though Gordon (2005)

gives a much higher figure). The three varieties are not mutually intelligible, and, for that matter, there is not full mutual intelligibility among the various dialects of NF. The earliest FR texts are from the second half of the thirteenth century. FR is officially recognized as the second language of the Netherlands, where it has status in law and education. Northern FR has some official status as a regional language in Schleswig-Holstein.

OS is represented by the important ninth-century epic poem *Heliand*, based on the New Testament, and a probably related fragmentary poetic retelling of Genesis, as well as minor documents. In the Middle Ages, one of the successors of OS, Middle Low GE, developed considerable prestige as the language of the Hanseatic League, and had significant linguistic influence on the continental Scandinavian languages. It has, however, declined quite markedly in status since then; none of the modern offshoots of OS is a state language, and they are, in fact, now widely regarded as dialects of the surrounding dominant languages, in spite of their separate origin and the fact that they are often not intelligible to speakers of the standard varieties of those dominant languages. Because of the lack of official status and perception as autonomous languages, they are highly subject to attrition in favor of the surrounding language. According to Lindow *et al.* (1998: 20), the ability to speak Low GE (Niederdeutsch, Plattdüütsch) ranges from 46% to 64% in different areas of Low GE territory. The “Low Saxonian” varieties of northeastern Holland, in which Ingvaemonic was historically a major constituent, are also regarded as DU dialects (de Schutter 1994: 441), though there has been an effort to achieve official recognition for them like that accorded to Limburger (see the listings for Stellingwerfs, Twents and Veluws in Gordon 2005).

EN, also a West GMC language in origin, stands out from the rest of the family in a number of ways, lexically, grammatically and demographically. In the lexicon, EN has absorbed massive amounts of non-native vocabulary, particularly from Romance languages (French and Latin), and it has developed a habit of forming neologisms from non-native elements to a much greater extent than other languages of the family (see Chapter 2). In terms of grammar, as noted in Section 1.2.2, it has also innovated to a much greater extent than other GMC languages. EN also holds a unique demographic position within the family and, indeed, among the languages of the world, in that it is more widely spoken than any other language (though exceeded by Mandarin Chinese in the number of native speakers). Exact numbers are hard to come by. According to one estimate, 350 million people speak EN natively, and another 350 million make significant use of it as a second language. Other estimates are provided in McArthur 2002: 3. It is also a global lingua franca in a number of domains. For example, air traffic controllers use EN as their common language of communication worldwide. Crystal (2003: 109) lists seventy-five countries in which EN plays an important role,

as the majority language, an official language or as a major language in education, commerce or some other domain.

1.3.4 North Germanic

The North GMC languages have remained relatively homogeneous for a longer time than West GMC languages, and differences tend to be gradual, rather than abrupt. Haugen (1982: 3) expresses the view that “Scandinavia may be looked on as a single speech continuum.” Characteristics of Common Scandinavian, distinct from those of West GMC, begin to emerge around 550 AD, according to Haugen (1982: 5), and the linguistic differentiation of the individual dialects begins to be evident from around 1150 AD, the earliest division being between a West Norse group, represented by IC, FA and NO, and an East Norse group, represented by SW and DA. The linguistic separation of SW and DA began as recently as 1300 AD. Though Scandinavian manuscripts are relatively late compared to those of West GMC, the runic inscriptions in the older *futhark* after c. 550 AD and the younger *futhark* after c. 800 AD provide a continuous though scant record of linguistic development. In later medieval times, the North GMC branch is represented by the very rich original literature of Old Norse (ON) (a normalized version of Old IC and Old NO), which, in spite of its name, is contemporary with the middle period of such West GMC languages as EN and GE. On typological grounds, the modern North GMC languages can be divided into a conservative Insular North GMC group, represented by IC and FA, and an innovative group, represented by the continental Scandinavian languages – SW, NO and DA. This classification crosscuts the earlier division into Western and Eastern North GMC; NO, which, like IC, is a Western North GMC language in origin, has come to show a marked typological resemblance to SW. The continental group in general has converged to a remarkable degree, and investigators remark on the high degree of mutual intelligibility among them (Haugen 1982: 18), though Haberland (1994: 316) notes that the sound system of DA presents something of an obstacle to this possibility of “inter-Scandinavian semi-communication.”

IC (Íslenska), the official language of Iceland, with 250,000 speakers, is the most grammatically conservative of the modern North GMC languages – indeed, of the modern GMC languages in general. Such dialect differences as exist are very slight.

FA (Føroysk) is spoken by 50,000 people in the Faroe Islands, where it is an official language, joining DA in that status in 1948. It resembles IC more than any other language, though it has changed phonologically and inflectionally to a greater extent than the former. It is also much more varied in its spoken forms.

SW (Svensk) is the native language of over 8 million people in Sweden, and 300,000 in southern and western parts of Finland. It is an official language in both countries, sharing that status with Finnish in Finland.

NO (Norsk) has two different, co-official written standard varieties – Bokmål ‘book language’ (hereafter, BN) or Riksmål (earlier Dano-Norwegian (Haugen 1982: 16)) and Nynorsk ‘New Norwegian’, or Landsmål (hereafter, NN) – the former based on written DA and the DA-influenced speech of NO upper classes and the latter on rural dialects of NO little influenced by DA (Haugen 1982: 16). 83% of Norwegians receive primary education in BN, and 17% in NN (Askedal 1994: 221). The total population of NO speakers is about 4 million. All NO dialects are mutually intelligible (Askedal 1994: 221).

DA (Dansk) is an official language in Denmark, the Faroe Islands (together with FA) and Greenland (together with Kalaallisooq), with about 5 million speakers. Gordon (2005) attributes the view to Norbert Strade that the Jutish dialect, spoken in Jutland near the GE border, is sufficiently different from Standard DA that “from the viewpoint of intelligibility, it could be considered a separate language.”

The Germanic lexicon

The focus of this chapter is in fact considerably narrower than the title might suggest. The study of the lexicons of languages is a multifaceted pursuit. The words of a language are the interface between its internal aspect, as a cluster of linguistic systems, and its external aspect, as a way of encoding and cataloguing the experiences of its speakers (Lass 1994: 178). Study of the lexicon thus straddles the study of purely linguistic aspects of language and the more general study of culture, since the vocabularies of languages are shaped by and reflect the intellectual and material culture in which their speakers function. Both of these are worthwhile undertakings. Much can be learned about the developmental histories of societies by studying the ways in which their vocabularies change over time. Indeed, in the case of cultures no longer extant, language often provides prospects for a reconstruction of the life of the mind which is finer-grained and more nuanced than what is possible through the more ambiguous and indirect evidence of physical remains and artifacts. The GMC languages in particular possess rich and multilayered lexicons which, when carefully examined, can generate an intricate picture of contacts, events, influences and cultural trends spanning millennia, even in the absence of a direct historical record. However, the focus of the present volume is on the GMC languages as systems and structures, and in keeping with that focus I will forego discussion of the possibilities for linguistic archaeology which the GMC lexicon affords, and concentrate exclusively on its more purely linguistic aspects – on the processes of word formation and the ways in which the shape of the lexicon and changes in the lexicon affect the other linguistic subsystems of the language. Loanwords, for example, will be of interest here not for their cultural and historical import, but for the ways in which the linguistic systems of individual languages have responded to them – the degree to which they are integrated into the phonological and morphological patterns of the language, the degree to which they carry over characteristics of their origin, and, in the latter case, the effects they have had on the shape of the language (for example, in adding new segments to the sound system, in altering patterns of stress assignment, or in disrupting the transparency, productivity or generality of word-formation processes).

Even within this narrowly defined domain, only certain word classes and word-formation processes have been singled out for description, selected on the basis that they are in some way typical of the family. A thorough examination of the structure of the lexicon in GMC languages would take far more space than is available here. Instances of one particular kind of lexical development, that of “grammaticization” (the rise of grammatical function words), are treated severally under the appropriate chapters on morphosyntax. So, for example, case-marking prepositions, definite and indefinite articles are treated in the chapter on nouns and noun phrases, while modal verbs and infinitive markers are treated in the chapter on verbs and verb phrases.

The chapter will conclude with a discussion of two categories of lexical items particularly characteristic of the GMC languages: discourse particles and phrasal verbs.

2.1 **Loanwords**

Much of the core vocabulary of the GMC languages is of IE origin, as are their basic modes of word formation. However, from the beginning, borrowing from other languages has been among the ways in which the GMC languages have acquired new words. A substantial portion of the common GMC vocabulary is not transparently of IE origin. A partial list of such words is given in Lass 1994: 181–182. It is possible that some of these words simply have non-obvious IE etymologies (see, for example, Orel 2003), but several of them belong to specific areas of endeavor (*boat, sea, sail, ship, sheep, lamb, calf, plow*) and may mark the borrowing of new words along with technologies. These words, if in fact they are borrowings, were fully integrated into inherited inflectional paradigms. The verb *drink*, for example, listed by Lass as non-IE, nonetheless assumed the inherited mode of tense formation through alternation in internal vowels (ablaut). See Bammesberger 1986, 1990, Heidermanns 1993 and Casareto 2004 for a discussion and exhaustive listing of the membership of major word classes in Common GMC.

From the time of these early loans, all of the GMC languages have continued to take new vocabulary from neighboring languages, at different rates, in different forms and from different sources. Main episodes of borrowing have included the following:

- a. Common GMC adopted a number of words from its Italic and Celtic neighbors in prehistoric times, and the individual languages adopted relatively large amounts of Latin vocabulary in early medieval times with the rise of monastic culture. The GO Bible contains numerous words taken from Greek, though it is unclear to what extent these in fact achieved the status of loanwords.

- b. OE adopted large numbers of Scandinavian vocabulary items during the period of Danish invasions beginning in the ninth century, though these do not usually appear in writing until the Middle English (ME) period because of the way they were distributed geographically. These Scandinavian loans are unusual in that they involve items which are not typically borrowed – everyday, low-prestige terms, and even some function words and grammatical formatives.
- c. Subsequent to the Norman Conquest, vast numbers of Norman French words entered the EN lexicon. Other GMC languages, too, including GE, have drawn heavily on French, beginning in medieval times. The items in question are typically high-prestige vocabulary. Belonging to a different register are the numerous French borrowings for everyday words appearing in West Flemish, as discussed in Haegeman 1992 and the references discussed there, which, like the Scandinavian loans in EN, evidence a different sort of language-contact situation.
- d. The Scandinavian languages drew heavily on Middle Low GE vocabulary beginning in the late twelfth century and reaching a peak during the period of the Hanseatic League, from approximately the thirteenth to the sixteenth centuries (including a number of Latinate words which came into these languages through Middle Low GE) – see especially Bandle *et al.* 2002: 955. This was followed during the Renaissance by a period of borrowing from High GE, under the influence of the Reformation.
- e. FA has adopted large numbers of DA words. This tendency is resisted officially; the DA borrowings are discouraged in writing and not recognized by dictionaries, for example, giving rise to a wide discrepancy in vocabulary between spoken and written registers (Barnes 1994: 216; Bandle *et al.* 2005: 1583).
- f. YI has drawn extensively on the vocabularies of Hebrew and Slavic languages.
- g. All of the GMC languages have taken part in the diffusion of the increasingly international vocabulary of such fields as technology, commerce and politics.
- h. In the twentieth and twenty-first centuries, all of the other GMC languages have been under the expanding influence of the EN lexicon.

Particularly in modern times, the borrowing of foreign vocabulary has become ideologically loaded, leading to alternating periods of receptiveness to direct borrowing and resistance to it. Representing the poles of the spectrum among the modern languages are EN, on the one hand, which has developed a vocabulary considerably more extensive than that of the other GMC languages by massive overt

borrowing, and which seems inclined to build its neologisms by using formatives from the classical languages (Greek and Latin), and, on the other, IC, in which there is strong resistance to the introduction of foreign elements into the lexicon. Thráinsson (1994: 188) characterizes the IC situation as follows:

The Modern Icelandic lexicon is relatively free of unassimilated loanwords . . . there is a strong and conscious effort to create new words from Icelandic for new concepts, for example in science and technology. Many professional societies have their own language committees that meet regularly to discuss proposed neologisms or to try to come up with new ones.

The methods used in coining new words include translation of the foreign word bit by bit, compounding of existing nouns, derivation by productive derivational affixes, creation of new roots and assimilation of foreign words to the Icelandic sound and inflectional systems. Sometimes old words are also given new meanings.

In the latter category, Thráinsson cites *sími* 'telephone' ← *síma* 'thread', *tölva* 'computer' ← *tala* 'number', *útvarp* 'radio' ← *út* 'out' + *varpa* 'throw'.

It should be pointed out that the cases just cited from IC do in fact involve a type of borrowing; languages need to add new vocabulary as new concepts come to require linguistic expression, and among the strategies deployed at various times in the GMC languages by way of balancing this need for new words with the resistance to foreign elements has been such covert borrowing – borrowing the lexical concept, as IC, has done, but not the phonological shape of the word which would identify it as foreign. Such covert borrowing takes the form of calques, or loan translations, in which the template for formation of the word is adopted, but native morphemes substituted for foreign ones. This phenomenon has existed (side-by-side with straightforward borrowing) from the earliest times. So, for example, GO borrows *aíwaggeli* 'gospel' and *aíwlaugia* 'benefaction' directly from Greek *eu-angélion* and *eu-logía*, but translates *eu-genés* 'well born' as *goda-kunds*. OE (and other early GMC languages) had a decided preference for loan translations. This can be seen by comparing the OE and the Modern EN counterparts to the following Latin words (examples in part from Lass 1994).

(2.1)	<i>Latin</i>	<i>OE</i>	<i>Mod EN</i>
	<i>miserecordia</i>	mildheortness 'mildheartness'	mercy
	<i>circumcidere</i>	ymbsnīpan 'around-cut'	circumcise
	<i>illuminatiō</i>	inlīhtnis 'inlight-ness'	illumination/enlightenment
	<i>trinitas</i>	þrīnis 'three-ness'	trinity
	<i>longanimus</i>	long-mōd 'long-mood'	patience
	<i>unanimus</i>	añ-mōd 'one mood'	unanimous

Other languages, too, exhibit historical fluctuation between direct borrowing and covert borrowing. GE, for example, which borrowed extensively from French in previous eras, experienced a period of relative linguistic purism in the seventeenth and eighteenth centuries, borrowing the names of new chemical elements during that era, for example, as loan-translations (*Wasserstoff* ‘water-stuff’, from French/neo-Greek *hydrogène*, *Sauerstoff* ‘sour-stuff’, from French *oxygène*). In Modern DA the EN terms *landslide victory*, *nuclear family* and *think tank* are calqued as *jordskredsejr*, *kernefamilie* and *tænketank* respectively (Allan *et al.* 1995: 556).

Calquing is one of the means languages make use of for nativizing borrowings, by integrating them into native lexical, phonological and inflectional patterns. In addition, direct borrowings containing non-native sounds are typically remade by way of substituting the closest native sound (“phonological nativization,” Hock 1986: 390). Hock, for example, cites the substitution of /k/ for /x/ in EN borrowings from GE. The /u:/ of GO *Rumoneis*, from Latin *Rōmāni* is claimed to result from the fact that, at the time of the borrowing, GO had no closed /o:/ sound like that of Latin, and the GO /u:/ was perceived as the closest substitute. Phonological nativization also manifests itself in the adaptation of borrowings to the prosodic patterns of the language, as reflected, for example, in the accent shift which has taken place in IC *stúdent*, *prófessor*. Borrowed forms are also typically nativized inflectionally, by combining them with the (least marked) inflectional affixes of the language. Aside from possible early instances like *drink*, for example, which form their tenses by ablaut, verbs are almost invariably borrowed into the GMC languages as weak verbs, which form their past tenses by the more usual means of adding a dental suffix. Inflectional nativization is normally quite systematic; among inflectional affixes, only plural markers are sometimes carried over from the source language, as in EN *stratum*~*strata*, GE *Thema*~*Themata*, SW *gangster*~*gangsters*, DU *museum*~*musea*. Even in these cases, the borrowed plural formation tends to be replaced by an unmarked native one in time. Thus, GE also has *Themen*, SW also has *gangstrar-na* ‘the gangsters’ (Holmes and Hinchcliffe 1994: 27), and DU also has *museums* (Booij 2002b: 29).

One particular problem faced in the nativization of nouns borrowed into some GMC languages is what gender to assign to them. This problem has been particularly thoroughly studied with respect to loanwords in GE (Reed 1942; Arndt 1970; Carstensen 1980; Gregor 1983; Schlick 1984; Köpke and Zubin 1995; Hickey 2000). Some general strategies emerge. Interestingly, even when the source language has a system of grammatical gender, the gender of the noun in the source language does not seem to play a role in determining its gender in the borrowing language. Hock (1986: 402), for example, notes that GE *Garage* is feminine (as are all nouns ending in *-age*), even though its French source is masculine. Typically, borrowed nouns are assigned their gender on the basis of the gender of semantically similar native nouns. Berger *et al.* (1972: 206) cite GE

Chanson, feminine in French but neuter in GE because of the synonymous *das Lied* ‘the song’, *High Society*, feminine in GE because of *die Gesellschaft*, and *der Star*, masculine because of *der Stern*. Holmes and Hinchcliffe (1995: 12) cite SW *ett team, ett jobb, ett game*, all neuter because of the neuter gender of the synonymous native nouns *ett lag, ett arbete, ett spel*. In other cases, the assignment is based on the shape of the noun, not always in a way that has a basis in native principles of gender assignment. For example, in SW, foreign loans ending in *-ek, -em* and *-iv* are neuter (*ett bibliotek* ‘the library’, *ett system, ett adjektiv*) (Holmes and Hinchcliffe 1995: 6). In GE, nouns in *-age* or *-tion* are feminine (*die Frustration*), while nouns in *-ing* are neuter (*das Happening*). Where none of these strategies leads to a single result, there can be uncertainty and fluctuation in gender assignment (GE *der* or *das Joghurt*, *der, die* or *das Dschungel*), though there seems to be some evidence for default gender preferences, which may vary across the GMC languages (see Section 4.1.2.2).

On the other hand, borrowings are sometimes not fully nativized, particularly in instances in which they are very numerous, or where the retention of foreign features is favored by prestige. Such cases can lead to importation of non-native linguistic features, and consequent modification of the linguistic system. Most of these effects have involved marginal, local introductions of new sounds. GO, for example, exhibits a high front rounded vowel, represented as <w>, but only in Greek words (*swnagoge* ‘synagogue’, *hwssopon* ‘hyssop-Dat’). ME adds /ʒ/ from French, as does MHG. GE /dʒ/ also occurs only in foreign loans (*Dschungel*), and some varieties of YI have adopted palatalized consonants from Slavic (Jacobs *et al.* 1994: 394). Most of the time, these sorts of developments remain restricted to the forms with which they are borrowed.

Borrowing of foreign words has also introduced new derivational affixes into the GMC languages. While many of these too remain confined to the non-native vocabulary, they can on occasion become more productive, combining with native roots. This is the case with the GE verbal suffix *-ieren*, as in *buchstabieren* ‘to spell’ and the GO agent noun suffix *-areis*, from Latin *-arius*, as in *bokareis* ‘scribe’.

Some of the effects of wide-scale borrowing are less specific. In EN, for example, Romance influence (on literary tradition, as well as word formation) is linked to the loss in ME of the metaphorical compounds (*kennings*) which so characterized OE literary language. To the massive borrowing of Romance forms may also be traced the current preference for borrowing or building new words from foreign elements, rather than native elements, with the consequent loss of the transparency in the lexicon to which Lass has alluded (see the next Section).

2.2 Derivation

The inherited devices for word derivation in GMC consisted of prefixes and suffixes, sometimes working in conjunction with internal modifications of the root (ablaut – see

Section 5.1.1). Since words in GMC (and IE) are right-headed, suffixes determine the category of the word. Thus we use suffixes to effect a change of lexical category – for example, from a verb to a noun. Prefixes can be grammatically active too, though, in the sense that they can change the valency and case-government of the forms to which they are attached. So, for example, GO *gaggan* ‘to go’ is an intransitive verb, while *mibgaggan* ‘to accompany’ takes a dative object. In GE, the prefix *be-* is found frequently as a transitivity marker (*leben* ‘to live’, *beleben* ‘to animate’).

In general, the inherited derivational prefixes of GMC have survived better than the suffixes, in large part for phonological reasons connected with the fixing of the stress accent of GMC on syllables at the beginning of the word (see Section 3.2.3). Thus, for example, most of the prefixes found in GO (*ga-*, *fra-*, *un-*, as well as a large inventory of prepositions which function as prefixes) are cognate to prefixes still found in modern GE and DU. The North GMC languages and EN, however, for reasons that probably do not have to do with phonological change, have lost many of these; as Heusler (1967: 40) points out, true (unstressed) prefixes were generally lost in ON. McWhorter (2002: 232) observes that in EN the former prefixes *be-*, *ge-*, and *for-* (as in OE *sēon* ‘to see’, *besēon* ‘to look at’, *winnan* ‘to toil’, *gewinnan* ‘to win’) whose counterparts in other West GMC languages are frequent and partly productive (AF *verafrikaans* ‘to Afrikaansify’), have been given up as well.

IE had a very extensive inventory of derivational suffixes. These were added to roots, rather than to words. That is, the base-forms to which they were attached did not appear in an unaffixed form as independent words. The subsequent history of this inventory in the GMC languages, as Lass (1994: 199) notes, has been one of fairly steady decline. Suffixes have undergone phonological reduction and modification, which, even when not resulting in total loss, have often rendered them unrecognizable as derivational morphemes. This effect has often been reinforced by the loss of contrasting, derivationally related forms whose existence would have served to make it easier to identify the affixes. The fact that *mind*, *birth*, *seed*, *need*, *deed*, *gift*, *might* all share a common (pre-)GMC suffix, for example, and *death*, *flood*, *ground*, *greed*, *lust*, *thought*, *thirst* another one (Casaretto 2004: 496ff., 522ff.) is no longer part of our sense of relatedness among lexical items.

Among verbs, at least, the system of inherited derivational affixes, if no longer productive, was at least partially transparent in the earliest attested GMC.¹ In GO, for example, there were such clear alternations as that among a strong durative verb *wakan* ‘to be awake’, the intransitive inchoative weak verb (*ga-*)*wak-nan* ‘to become awake’ and the transitive causative weak verb (*us-*)*wak-jan* ‘to wake someone’. A similar

¹ Clear evidence for productivity, such as might be provided by the extension of these suffixes to recent borrowings, is slight; Gothic does borrow Latin *accumbere* ‘to recline’ as *anakumb-jan*, with a native affix.

triplet is found in the paired verbs *ga-qiu-nan* ‘to come to life’, *ga-qiu-jan* ‘to vivify’ and their root adjective *qiu-s* ‘alive’. Such transparent correspondences are much less in evidence in later medieval GMC languages, however; the inchoative *-nan* verb class is entirely lost in West GMC, though still in evidence in Scandinavian (SW *mörkna* ‘become dark’ from *mörk* ‘dark’), and the causative *-jan* verbs, while still a distinct class, are no longer identifiable by virtue of containing a single suffix with a particular shape. We recognize the OHG counterpart to GO *wakjan*, for example, as a member of the class of ‘*-jan*-verbs’ because it exhibits a geminate consonant in the present (*weckit* ‘he wakes s.o.’) and a preterite (*wahta*) in which the past tense ending is added directly to the root, without a linking vowel. In still other verbs of this class, the causative affix shows up as a linking element *-i-* between the root and the past tense marker: *dhecch-i-dôn* ‘they covered’. But after the late medieval simplification of geminates and the elimination of contrasts in unstressed vowels, even these characteristics can no longer be relied upon. Resultantly, the class largely loses its coherence, and verbs of this former class are identifiable only in cases such as GE *röten* ‘reden’ from *rot* ‘red’, *stärken* ‘strengthen’ from *stark* ‘strong’ by virtue of the umlauting effect of the former *-jan* suffix on the root vowel. Here, we are no longer dealing so much with a verb class as with a word-formation rule.

Lass (1994: 191ff.) notes that one of the consequences of this long-term trend has been a progressive erosion in the speaker’s sense of relatedness among forms in the lexicon. He cites the fact (p. 199) that there are 139 noun- and adjective-forming suffixes of IE or GMC date which are identifiable in Old GMC words, of which only 42 remain reasonably productive in OE. That is, about two-thirds of them had been lost or had become opaque by the time of late medieval GMC. For all of that, the lexicon of OE is still characterized by relatively highly productive use of affixes in the derivation of new vocabulary, in comparison with some modern languages. Lass states that

One of the most striking features of the Old English lexicon is the extensive involvement in [word formation], not only of transparent affixation, compounding, and conversion, but of other devices of varying ages: ancient ones like ablaut, and newer ones like i-umlaut. This results in what Kastovsky (1992: 294) calls “large morphologically related word-families”; considerable portions of the lexicon “cohere” in a rather special way, characteristic of older IE and to some extent more archaic modern languages like German, but quite alien to Modern English.

(Lass 1994: 191)

The history of derivational morphology in GMC, however, has not been only one of loss. Also in evidence is a countervailing tendency to introduce new derivational affixes. In place of the obsolete and opaque inherited suffixes for nominal derivation

has arisen a new set of more salient suffixes which derive from originally independent lexical items. Kluge (1913: 227), who refers to these as “Kompositionssuffixe” (roughly, “compound-suffixes”), notes that they are more numerous and more usual in West GMC than in GO, which lacks *-haid* ‘-hood’, *-dōm* ‘-dom’ and *-skapi* ‘-ship’, though it has *-leiks* ‘-like’ and *-sams* ‘-some’. He concludes that this sort of formation was only in its beginning stages in common GMC. Many of these new suffixes belong to the class of “non-cohering” suffixes, which form their own prosodic domains, separate from that of the root (Kristoffersen 2000: 43).

This development seems to have been limited to nouns and adjectives, no similar developments being found in verbs – no doubt connected with the fact that compounds headed by verbs are uncommon in GMC. Of the new verbal derivational affixes arising to replace the inherited ones which had been lost or become opaque, only one, the (Middle) EN *-en* in *redden*, *blacken*, is in part a native development. Even this one is apparently also partly borrowed, reflecting the influence of Scandinavian *-nan* verbs, according to the *OED*. The majority of new verb-deriving affixes in the GMC languages are borrowings from Greek or the Romance languages, as in GE *-ieren*, DA *-ere*, EN *-ate*, *-ize* and *-ify*. Some of these have come to be sufficiently native to attain productivity and be usable with native as well as foreign roots (e.g., GE *stolzieren* ‘to swagger’, based on *stolz* ‘proud’, and EN *bowdlerize*).

Forms like *freedom* and *friendship* also reflect a typological departure from the mode of word formation in early GMC. These new suffixes are of a different sort from the earlier ones, in that they do not attach to stems but to words – that is, to forms which are able to appear independently. Lass (1994: 200) sees in this typological shift from a stem-based to word-based lexicon “a major evolutionary direction” in GMC. A late phenomenon which can perhaps be regarded as the ultimate outcome of this trend is the employment of category shift or zero-derivation in modern EN – the frequent and prescriptively decried tendency to create new verbs, for example, by using nouns as verbs with no derivational affixes (*to impact*, *to interface*).

2.2.1 Compounding

Compounding is the process of forming new words by conjoining two or more root morphemes. Compounds in GMC, like other words, are right-headed, in the sense that it is the rightmost element in the compound which determines the category of the whole, as well as such grammatical features as gender and the choice of plural affixes.²

² English *passers-by* represents a possible counterexample. If it is a compound, it is clearly the leftmost element which is the head, since the whole is understood to be a (plural) noun. Jacobs (2005: 146) observes that YI has, in addition to compounds formed on the GMC model, with

In principle, members of almost any lexical category can serve as the first element of the compound, and any of the major categories, noun, adjective or verb, can appear as the second element. In fact, compounds headed by verbs are not common in GMC, and not equally possible in all of the modern languages. Booij (2002b: 141) reports them to be unproductive in DU, while Josefsson (1998: 68) reports that they are reasonably productive in SW. There is some direct evidence of their non-productivity in early GMC: as Eythórsson (1995: 24) points out, GO ‘resolves’ such Greek compounds as *hydrotéō* ‘water-drink’ into a separate verb plus object (*driggkan wato*). Noun- and adjective-headed compounds, on the other hand, have from the beginning been markedly productive, and figure prominently as a poetic device in older GMC poetry, in the form of the frequent metaphorical compounds known as *kennings* (Later Runic *walha-kurna* ‘foreign corn’ = ‘gold’, Tjurkø bracteate, Antonsen 1975: 79).

The development from a stem-based lexicon to a word-based lexicon discussed in the preceding section is also visible in the history of compound formation in GMC. Lass (1994: 194) notes that a distinction is traditionally made between “genuine compounds,” in which the first element is a stem (a root plus the associated noun-class marker) and “non-genuine” compounds, in which it is not. The former is the older of the two, and the one inherited from IE. Lass cites GO *fot-u-baurd* ‘foot-board (stool)’ as an example, and formations of this sort are abundant in the compound names of the older runic inscriptions (*Hle-wa-gastiz* ‘Fame-Guest’). In each case, the first element contains a noun root and a noun class marker (-u, -wa) but it is not a free-standing form, lacking a case affix. There can be other differences as well between the first element of a compound and the same lexical item in its free-standing form. Kluge (1913: 229) cites such examples as GO *mana-seps* ‘man-seed’, but *mann* ‘man’.

These stem compounds later became increasingly less frequent, according to Lass, in consequence of the breakdown of the noun-class system, and were replaced increasingly by newer “non-genuine” compounds. These new types of compounds come in three varieties. The earliest of them, apparently, are what we may characterize as “root” compounds, differing from stem compounds in that the first element consists solely of a lexical root without an accompanying class marker. These are reflected in a

the modifying element preceding the head (and bearing stress), a second model, patterned after Hebrew, in which the head is first and the modifier second, the main stress falling on the latter. Both of these patterns are highly productive, and either may combine elements of Hebrew and GMC origin (*sōf zác* ‘final sentence’, *sōf zác* ‘end of sentence’). Booij (2002b: 143) points out that compounds like *redskin*, sometimes characterized as “exocentric,” or headless, are in fact right-headed compounds. It is the right-hand member which determines the category of the whole. What sets them apart is the fact that they are used as metonyms, referring in this case, for example, to ‘someone who has red skin’. On the other hand, Booij points out some more complex cases of metonymic compounds in DU in which the rightmost element does not fully determine the grammatical properties of the whole. *Het oor* ‘the ear’ is neuter, but *de domoor* ‘the dumb-ear’ = ‘idiot’ is common gender.

small number of cases even in GO, for example, *gud-hus* ‘temple’ (‘God-house’), *piu-magus* ‘servant-boy’, *hauh-puhts* ‘high-seeming’ (Braune 1973: 69), and become increasingly frequent in later GMC. The second type of innovative compound, the genitive type, in which the first element is a noun with a genitive case suffix, is apparently of later occurrence. According to Kluge (1913: 229) and Lass (1994: 195), it is uniquely represented in GO by *baurg-s-waddjus* ‘city’s wall’. It is also rare in other early GMC languages, according to Kluge, appearing first in names for the days of the week (*Wodnes-dæg*) as a calque on the Latin names, but it becomes more frequent later.

It is not clear that such genitive compounds are still in use in the modern GMC languages. There are indeed some in which the two nouns are separated by a morpheme which stems from, and looks like, a genitive suffix. Such an element, for example, is found in GE forms like *Liebe-s-gedicht* ‘love poem’ and *liebe-s-krank* ‘lovesick’. Lass (1994: 195) notes that the *-s-* cannot be analyzed synchronically as a genitive marker even though it is one by origin since, among other things, *-(e)s* as a genitive case marker appears only on masculine and neuter nouns in GE, and *Liebe* ‘love’ is a feminine noun. These compounds thus represent yet a third type of innovative compound, in which the two components are joined by a fixed and meaningless “linking” morpheme. Lass cites as an early instance, the OE *stān-e-gella* ‘stone-yeller (kestrel)’, in which the *-e-* appears as a semantically empty linking element. Each of the modern languages, in addition to “root” compounds, has a small inventory of such linking elements. DU, for example, has three such elements (Booij 2002b: 178): *-s-*, of genitive origin (*schaap-s-kop* ‘sheep’s head’), *-er-*, of plural origin, and *[-ə]* (sometimes written <en> in official orthography) (*schaap-e(n)-vlees* ‘sheep-flesh’ = ‘mutton’), which in some instances is a case marker in origin and in others a word-final *-e-* which was apocopated in non-compound forms. The inventory in DA is similar (*dreng-e-skole* ‘boys’ school’, *dag-s-rejse* ‘daytrip’).

Booij (2002b: 179f.) presents a number of arguments for viewing such elements as “stem-extensions” of the first member of the compound, rather than as linking elements which are just inserted between two members of the compound. First, he notes that they belong to the first constituent prosodically, as evidenced by syllabification. Second, in constructions involving “gapping,” such as *schap-en- of vark-en-vlees* ‘mutton or pork’ (lit., sheep-EN or pig-EN meat), in which two conjoined first constituents of the compound share a final constituent which occurs only after the second of them, each of them is nonetheless equipped with its own “linking element.” Further, it is always the first element of the compound which determines whether a particular linking element can occur in the compound. Finally, some of these linking elements can occur outside of compounds, in instances of derivation involving certain non-cohering suffixes, as in *arbeid-s-loos* ‘work-less (unemployed)’. Booij concludes

that they are therefore suffixes on the first member of the compound (see also Josefsson 1998: 61). Viewed in this way, these compounds bear a certain resemblance to the original stem compounds of GMC. In languages offering a choice of such stem extensions, there are typically some partial generalizations about which ones can be used in a given instance – all of them making reference, as noted, to properties of the first noun. These are described for DU by Booij (2002b: 180) and for DA by Allan *et al.* (1995: 544), for example. Josefsson (1998: 60) reports a partial generalization in SW about the distribution of compounds with the linking element *-s* and those without: the linking element must be used if the lefthand member of the compound is itself a compound: *barn+bok+klubb* ‘book club for children,’ but *barn+bok+s+klubb* ‘club for children’s books’.

The linking element, or stem extension, can sometimes have the same shape as plural morphemes.³ In most such cases, however, the morpheme has no plural force. There are, though, occasional compounds in the GMC languages in which the first element of the compound is semantically as well as formally plural. Allan *et al.* (1995: 33) cite in DA, for example, such cases as *børnebørn* ‘grandchildren’, plural of *et barnebarn* ‘the grandchild’, either *barnefødsler* or *børnefødsler* ‘childbirths’ as the plural of *en barnefødsel*, and *mødrehjem* ‘home for mothers’. For DU, Booij (2002b: 147) cites *huiz-en-rij* ‘row of houses’ and notes such contrasts as that between *bedrijf-s-terrein* ‘company area’ and *bedrijv-en-terrein* ‘companies’ area’ (2002b: 179).

2.3 Discourse particles

One highly distinctive word category shared by (perhaps) all of the GMC languages and apparently rare outside that family, is a class of short, unstressable, typically monosyllabic words with at most vague propositional meaning which signal something about the speaker’s relation to the utterance or to the hearer. These can, for example, indicate the degree of the speaker’s commitment to the truth of the proposition, the source for the speaker’s knowledge, the assumptions the speaker makes about the degree to which his/her knowledge or belief is shared or not shared by the hearer, and (particularly in the case of commands) the degree of imperative force the hearer is invited to see in the utterance. Particles fitting this description have been reported for all of the modern GMC languages. For example, they are described for GE by Abraham (1991, 2000, 2001b) and Zifonun *et al.* (1997: 903ff., 1206ff.), for DU by Foolen (1995), for DA by Davidsen-Nielsen (1996), for SW by (Aijmer 1996),

³ Booij (2002b: 179) notes that the officially promoted practice of spelling the stem extension [-ə] as <en> in DU just in cases where the plural suffix of the first element would also be *-en* reinforces such an identification.

for IC by Thráinsson (1994: 188), for NO by Andvik (1992), for FR by Hoekstra and Tiersma (1994: 528), for YI by Jacobs (2005: 207) and for AF by Donaldson (1993: 213). The particles in question are sometimes referred to generally as “modal particles” (e.g., Arndt 1960; Abraham 2000; Foolen 1995), while others (Aijmer 1996) seem to restrict this label to the subset of them which convey epistemic modality (that is, information about the speaker’s beliefs concerning the proposition). Other labels attached to them are “discourse particles” or “discourse markers,” “metacommunicative particles,” “filler words” and “shading particles” (*Abtönungspartikeln*: Zifonun *et al.* 1997). In the present treatment, I will refer to them as “discourse particles” (glossed as PART in the examples).

Discourse particles can be categorized in a variety of different ways, and the reader is referred to the references cited for various categorization schemes. Davidsen-Nielsen (1996: 285) distinguishes between a group of particles in DA which are oriented toward the hearer (2.2a) and a second set, illustrated in (2.2b), which are oriented toward the speaker, “in the sense that they reflect the speaker’s conception of, or attitude to, his own knowledge of the state of affairs referred to” (Davidsen-Nielsen 1996: 285).

- (2.2) a. John er da/ jo/ nu/ skam en flink fyr (DA)
 John is PART a nice guy
 (DA, Davidsen-Nielsen 1996: 285)
- b. John er nok/ vel/ vist en flink fyr
 John is PART a nice guy

In the first case (2.2a), that of hearer-oriented particles, *jo* signals the speaker’s expectation that the knowledge expressed in the proposition is shared and accepted by the hearer, while *da* and *nu* signal, to varying degrees, the speaker’s expectation that the hearer does not accept or agree with the truth of the proposition. *Skam* is used to reassure the hearer of the truth of the proposition. In the second group (2.2b), the speaker-oriented particles *nok*, *vel* and *vist* serve to distance the speaker from a commitment to the truth of the proposition, and to convey information about the source of the knowledge. *Vel*, for example, expresses uncertainty and invites confirmation from the hearer, functioning more or less equivalently to tag questions. *Vist* suggests that the knowledge is not direct, but based on report. *Nok* suggests that the probability of the truth of the proposition is a conclusion of the speaker himself. This group is the one that Aijmer (1996), writing on SW, labels “modal particles,” and she compares them with the “evidential” morphemes used in other languages of the world to convey information about the source and reliability of the knowledge embodied in the proposition.

The discourse particles of the other GMC languages overlap in functions with these examples from DA. Hearer-oriented particles are, for example, reflected in GE (2.3a),

where the *ja* signals the assumption that the knowledge is shared and agreed to by the hearer, and (2.3b), where the *doch* signals that the hearer is being reminded of something known but forgotten. To this class may perhaps be added particles which serve as instructions to the hearer about how to interpret requests or commands, rendering them softer and more polite in some instances, as in SW (2.3c), GE (2.3d.), DU (2.3e.) (“down-graders” or “down-toners”), and more emphatic in others, as in GE (2.3f) (Abraham’s “up-graders”). In commands, as Zifonun *et al.* (1997: 1219) note, such particles anticipate the degree of receptiveness on the part of the hearer to the command, reflecting the speaker’s assumption about the extent to which the hearer is prepared to carry out the requested action. A further instance of a discourse particle affecting illocutionary force is the YI example in (2.3g), in which the presence of the *den* signals that the question is offered rhetorically, and a negative answer is presupposed.

- (2.3)
- a. Wir sind ja alle Freunde
we are PART all friends [as you know]
(GE, Wauchope 1991: 1)
 - b. Er war doch lange krank gewesen
he was PART long sick
‘He was sick for a long time [don’t you remember?]
(GE, Karagjosova 2003: 339)
 - c. Men du kan väl köpa litte köttfärs hos Märta
but you can PART buy some mince at market
(SW, Aijmer 1996: 402)
 - d. Könntest du mal eben’n bißchen rücken?
could you PART PART a bit move
‘Couldn’t you move over just a bit?’
(GE, Wauchope 1991: 14)
 - e. Komm maar binnen!
come PART in
‘Please come in’
(DU, Foolen 1995: 59)
 - f. Hör’ schon auf!
PART
‘Stop that [for God’s sake]’
(GE, Abraham 2000: 335)
 - g. Ikh hob den gelt?
I have PART money?
(YI, Lockwood 1995: 61)

Further subject-oriented cases are given in (2.4). AF *glo* in (2.4a) indicates that the proposition is merely opinion. SW *visst* in (2.4b) signals an inference.

- (2.4) a. Hy het glo nie 'n sent van die erfenis gekry nie
 he has PART not a cent from the inheritance gotten not
 'Apparently, he didn't get a cent of the inheritance'
 (AF, Donaldson 1993: 214)
- b. Han förstår visst allt man säger till honom
 he understands PART all one says to him
 'He seems to understand everything one says to him'
 (SW, Aijmer 1996: 394)

Traditionally, these particles are treated as adverbs and it is often the case that they are translational equivalents of full lexical adverbs with clear propositional meaning, as in (2.4a). However, they differ from such adverbs in a number of respects, enumerated in Wauchope 1991: 48, for example. They are typically monosyllabic, weakly stressed, indeclinable (that is, lacking comparative forms), unable to stand alone as an utterance (as adverbs like *apparently* can), and have a fixed position of occurrence, after the finite verb (never initial) but before the rheme of the sentence. Etymologically, they often differ from lexical adverbs in being related to/homophonous with function words – such as conjunctions, answering particles and interjections – rather than adjectives. Thus, for example, GE *denn*, *aber* function as both conjunctions and discourse particles, and *ja* is at once a discourse particle and a question-answering particle ('yes'). They also differ semantically from more canonical adverbs, providing information that is sometimes quite hard to state propositionally about how the sentence is situated in context, including the speaker's presuppositions, expectations of the hearer, attitudes toward the truth of the utterance or its illocutionary intent.

Among the modern languages, EN makes far less extensive use of such discourse particles than the others. It has a few: *just*, for example, functions as one such particle with a variety of contextually determined meanings, ranging from softening of commands to expression of the fact that the state of affairs described in the utterance is contrary to the speaker's expectations. EN does not have the extensive system of discourse particles found in the others, though, and typically employs other means, such as lexical adverbs, tags and parentheticals instead. Davidsen-Nielsen (1996: 287) notes in particular that DA hearer-oriented discourse particles are often rendered in EN with tags like 'You know', 'You see', while speaker-oriented ones are often rendered with parentheticals such as 'I guess', 'I suppose', 'It seems'. Andvik gives the following EN translational equivalents for NO modal particles.

- (2.5) nok ‘all right, probably, to be sure’
 vel ‘I suppose, no doubt, of course’
 da ‘certainly’
 jo ‘you know’
 nå ‘after all, really’
 visst ‘surely, I believe’
 (NO, Andvik 1992: 3)

All of the other modern languages have elaborated systems of discourse particles. The status of such particles in early GMC awaits fuller exploration. In GO, only two possible cases emerge. These are the forms *ibai*, which in its occurrence in questions presupposes a negative answer (Streitberg 1920: 219), and *an*, which appears to invite an affirmative response, both of them having the effect of making the questions rhetorical:

- (2.6) a. *ibai* mag blinds blindana tiuhan
 PART can the-blind the-blind lead?
 (GO, Luke 6:39)
 b. *an* nuh þiudans is þu?
 PART now-and king are you?
 ‘So you are a king, then?’
 (GO, John 18:37)

At least some of the modal particles of modern GMC languages are of demonstrably late origin (Wauchope 1991: 1). On the other hand, Wauchope argues that at least some of them, including *thoh*, *ia* and *thanne*, had such functions already in OHG times, and since the cognates of these forms also function as discourse particles in modern continental Scandinavian languages, it is possible that the discourse particle system reaches in its roots at least back to common Northwest GMC.

- (2.6) c. *Thes nu gidua thu nu unsih wis, wer thoh manno thu sis*
 Of this make you now us wise, who PART of-men you are
 ‘Make this known to us, just who you are among men’
 (OHG, Otfrid I:27-35-37, cited in Wauchope 1991: 82)

2.4 Phrasal verbs

Another highly (and increasingly) characteristic lexical type in the GMC languages is the class of so-called phrasal verbs.⁴ These verbs, represented by GE *nachschlagen*

⁴ These are also variously known as particle verbs, verb-particle combinations, separable-prefix verbs and stressed-prefix verbs (Dehé *et al.* 2002: 1). The label separable-prefix verb is used in describing them in the verb-final languages, such as GE and DU, where, as a side-effect of the

'look up (a word/fact)', EN *look up*, for example, consist of a lexical verb plus a "particle," the two of them forming a semantic unit, conveying a single verbal meaning. The present section will focus on their lexical properties. Their syntactic properties, also a matter of considerable interest, will be discussed in Section 5.8.

Such two-part verbs are often used where other languages would use a single, morphologically complex verb. So, for example, in GO, phrasal verbs with the directional particles *inn* and *ut* are often used in combination with a verb of motion or transfer to convey the meaning of a Greek verb with a directional prefix, such as *eis-* 'in' or *ex-* 'out', as in the following examples:

- (2.7)
- a. atgaggands inn
going in = Greek *eis-elthōn*
(GO, Matt. 9:25)
 - b. saei inn ni atgaggaip þairh daur
who in not goes through door
= Greek *ho mē eis-erchómenos* 'the one not entering'
(GO, John 10:1)
 - c. atiddja ut Peilatus
went out Pilate = Greek *ex-êlthen*
(GO, John 18:29)
 - d. biþe ut usiddjedun eis
as out departed they = Greek *ex-erchoménōn*
(GO, Matt. 9:32)

Both the GO particle and the Greek prefix convey the same meaning, but they have different morphological status. The prefix is a bound morpheme, always occurring in a fixed position in the word, whereas the particle in the GO examples is an autonomous morpheme, variously preceding and following the verb, and even occurring separate from it, as in (2.7b). The inclination to prefer such phrasal verbs over verbs with directional prefixes is no doubt connected with the general preference in GMC for packaging motion and manner together in lexical verbs, to the exclusion of direction (French *entrer dans la maison en courant* vs. EN *run into the house*), as discussed in Section 5.6.

fact that nonfinite verbs, including infinitives, occur in final position, the verbal particle precedes the infinitive. The infinitive is used as a citation form in dictionaries (GE *anrufen* 'call up'), and therefore in the citation form the particle looks like a prefix. The view that they are a sort of prefix is also conveyed in the orthography. It is nonetheless clear that they are not prefixes in the usual sense, on phonological or morphological grounds. They have their own stress, and they are separable from the associated verb in a way prohibited to genuine affixes. They in fact exhibit effectively the same degree of syntactic autonomy that particles do in such VO languages as SW and EN.

In GO and other early GMC languages, phrasal verbs seem to have been limited to this directional type, and modern GMC still reflects this history, the most common sort of verbal particle typically being homophonous with directional prepositions. This homophony makes the identification of phrasal verbs in languages like GO somewhat uncertain. Since *inn* and *ut* have fixed directional meanings, and since they are clearly not bound morphemes, one might alternatively suppose that these constructions involve two separate lexical items, and that *inn* and *ut* are not instances of a special class of verbal particles but are simply directional adverbs or intransitive prepositions with spatial meanings, in free association with the verb, rather than forming a single lexical item with it. Compelling evidence distinguishing these two views is hard to come by for the historical languages. However, the fact that *inn* and *ut* occur in GO exclusively in association with such verbs is consistent with an identification as verbal particles. Clear syntactic evidence that directional particles are not just adverbs/intransitive prepositions is found in the modern languages, where they are distributionally distinct from these other types of elements. For example, Dehé *et al.* (2002: 2) note that, while adverbials in SW and EN (including some which seem to be functionally identical to particles) are prohibited from intervening between a verb and its direct object, particles belonging to phrasal verbs may do so freely:

- (2.8) a. I carried in/ off/ *inside/ *into the house the boxes
 b. Johan skrev upp numret
 Johan wrote down number-the
 (SW, Dehé *et al.* 2002: 3)
 c. *Johan ställer dar glaset
 Johan puts there glass-the
 (SW, Dehé *et al.* 2002: 3)

Thus, they have privileges of distribution not shared by adverbs. Conversely, Dehé *et al.* note that in verb-final languages, such as GE, adverbs can appear in positions which are prohibited to particles.

- (2.9) a. wenn ein Licht aufleuchtet / *wenn auf ein Licht leuchtet
 if a light up-lights *if up a light lights
 b. wenn ein Licht plötzlich leuchtet/ wenn plötzlich ein Licht....
 if a light suddenly shines if suddenly a light
 (GE, Dehé *et al.* 2002: 4)

The identification of phrasal verbs has also become easier in some instances on semantic grounds beginning in late medieval times, with the development of some new types of particle verbs which appear not to have existed earlier. While the core cases of verbal particles are directional prepositions in origin, and had transparent directional

interpretations in the early languages (Elenbaas 2003), they have in many instances lost their original lexical semantic content entirely, and are only interpretable in combination with the verb. We may distinguish two subtypes. In one of them, idiomatic verb-particle combinations, the combination has a noncompositional meaning. That is, the meaning of the whole cannot be computed from the meanings of its component parts. In cases like GE *nach=schlagen* (literally, ‘after-strike’) or EN *look up*, the meaning ‘consult a reference source for’ is simply assigned to the whole phrasal verb, and no component of it can be specifically attributed to *up*, for example. In this feature, idiomatic phrasal verbs are like compounds, whose meanings also tend toward noncompositionality (*egghead*, *flatfoot*), and it is clear at this point that we are dealing with items which constitute a single lexical item in some sense, in spite of their apparent syntactic complexity. Since it is not computable from its parts, as Jackendoff (2002: 73) notes, the meaning of *look up* must be listed as an entry in our mental lexicons. The second innovative type of phrasal verb is of a more regular sort; the verbal particle has again lost its original directional meaning, but has developed a new meaning relating to the shape of the action described by the verb. Denison (1981) identifies as a cardinal instance of this type the “completive *up*,” represented in *eat up*, *break up*, *fill up*, for example. Denison notes that this use belongs to the semantic domain of *Aktionsart*, or verbal aspect; the particle alters the contour of the action denoted by the simplex verb, indicating in this case that that action took place up to the point of logical, irreversible conclusion. Wedel (1997) refers to this as “complexive aspect.” Denison notes that the aspectual use of verbal particles is also exemplified, though less productively, by the particles in *kill off*, *chatter away*. Jackendoff (2002: 77) observes that in some instances the addition of an aspectual particle also has the effect of changing the valency of the verb; durative phrasal verbs with the particles *on* or *away* cannot take direct object arguments, even though formed from transitive verbs: *fight* (**the army*) *on*, for example.

Both the idiomatic type of phrasal verb construction and the aspectual type seem to have their roots in the later medieval languages. The development of completive *up*, for example, is discussed in detail in Denison 1981, who assigns it to the twelfth century in EN and notes that it appears to have arisen earlier yet in the Scandinavian languages (ON *brjóta upp* ‘break up’) from which in fact the EN construction may derive (cf. also Kastovsky 1992: 320). It finds only isolated representation in the continental West GMC languages (GE *aufessen* ‘eat up’, *aufgeben* ‘give up’), where Denison assigns a later date. Some observers have noted that this development reflects a basic change in the way the categories of *Aktionsart* found expression in the GMC lexicon. In earlier GMC, *Aktionsart* had been encoded either by means of derivational suffixes or by means of prefixes. As an example of the role of suffixes in signaling aspectual contrasts, we can cite GO *dreiban* ‘drive’, a primary strong verb, and the corresponding derived,