1

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

1

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 some-
thing 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 recom-
mended. 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 (repre-
sented 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 individ-
ual 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 Ingvaeonic 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
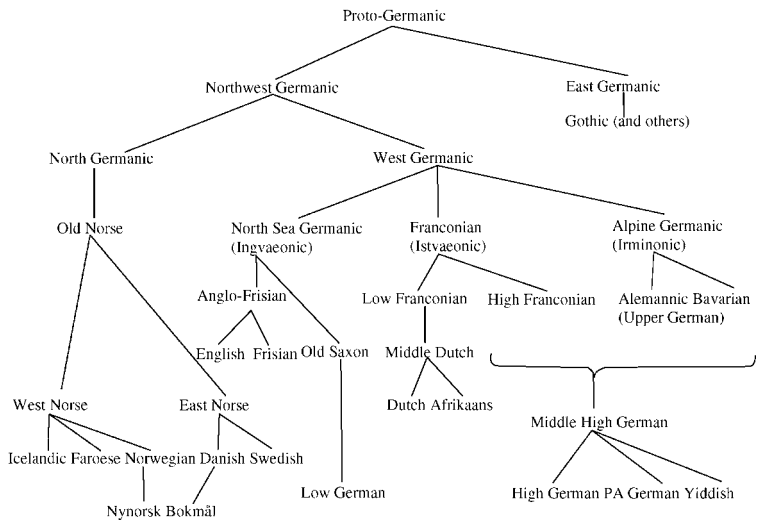
Figure 1.1 *The Germanic Family Tree*

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 Ingvaeonic 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 sub-
groups, 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 been 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 influ-
ences 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 construc-
tions 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 coun-
terparts. Included here are inchoatives with the inherited *-nan* suffix
(*gaskaidnan* 'to divorce (intrans)' $\sim$ *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 transi-
tives 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 argu-
ments, 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.