ROUTLEDGE STUDIES IN LINGUISTICS

Language Contact and the Origins of the Germanic Languages

Peter Schrijver

ROUTLEDGE

Language Contact and the Origins of the Germanic Languages

History, archaeology, and human evolutionary genetics provide us with an increasingly detailed view of the origins and development of the peoples that live in northwestern Europe. This book aims to restore the key position of historical linguistics in this debate by treating the history of the Germanic languages as a history of its speakers. It focuses on the role that language contact has played in creating the Germanic languages, between the first millennium BC and the crucially important early medieval period. Chapters on the origins of English, German, Dutch, and the Germanic language family as a whole illustrate how the history of the sounds of these languages provide a key that unlocks the secret of their genesis: speakers of Latin, Celtic, and Balto-Finnic switched to speaking Germanic and in the process introduced a 'foreign accent' that caught on and spread at the expense of types of Germanic that were not affected by foreign influence. The book is aimed at linguists, historians, archaeologists, and anyone who is interested in what languages can tell us about the origins of their speakers.

Peter Schrijver is professor of Celtic languages and culture at the University of Utrecht. He is a historical linguist working on ancient and medieval languages in Europe. His publications include *Studies in British Celtic Historical Phonology* and *Celtic Influence in Old English*.

Routledge Studies in Linguistics

1 **Polari—The Lost Language of Gay Men**
Paul Baker

2 **The Linguistic Analysis of Jokes**
Graeme Ritchie

3 **The Irish Language in Ireland**
From Goídel to Globalisation
Diarmait Mac Giolla Chríost

4 **Conceptualizing Metaphors**
On Charles Peirce's Marginalia
Ivan Mladenov

5 **The Linguistics of Laughter**
A Corpus-assisted Study of Laughter-talk
Alan Partington

6 **The Communication of Leadership**
The Design of Leadership Style
Jonathan Charteris-Black

7 **Semantics and Pragmatics of False Friends**
Pedro J. Chamizo-Domínguez

8 **Style and Ideology in Translation**
Latin American Writing in English
Jeremy Munday

9 **Lesbian Discourses**
Images of a Community
Veronika Koller

10 **Structure in Language**
A Dynamic Perspective
Thomas Berg

11 **Metaphor and Reconciliation**
The Discourse Dynamics of Empathy in Post-Conflict Conversations
Lynne J. Cameron

12 **Relational Semantics and the Anatomy of Abstraction**
Tamar Sovran

13 **Language Contact and the Origins of the Germanic Languages**
Peter Schrijver

Language Contact and the Origins of the Germanic Languages

Peter Schrijver

Routledge
Taylor & Francis Group
NEW YORK AND LONDON

First published 2014
by Routledge
711 Third Avenue, New York, NY 10017

Simultaneously published in the UK
by Routledge
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

Routledge is an imprint of the Taylor & Francis Group,
an informa business

Library of Congress Cataloging-in-Publication Data

Schrijver, Peter.
Language contact and the origins of the Germanic languages / Peter Schrijver.
pages cm. — (Routledge Studies in Linguistics ; 13)
Includes bibliographical references and index.
1. Languages in contact. 2. Germanic languages—Etymology. 3. Germanic languages—Grammar, Comparative. 4. Germanic languages—Grammar, Historical. I. Title.
PD582.S37 2013
430—dc23
2013017509

ISBN: 978-0-415-35548-3 (hbk)
ISBN: 978-0-203-00191-2 (ebk)

Typeset in Sabon
by Apex CoVantage, LLC

Contents

Preface

This book is written for anyone who wants to know more about the earliest history of one of the most successful language families in the world, both in terms of numbers of speakers and in terms of the ideas expressed by those speakers during the last 1300 years: Germanic. The idea behind this book is that the English, Dutch, and German languages, and indeed the Germanic family as a whole, are founded on the input of people who did not originally speak Germanic but switched to it in the course of time. Additionally, I hope to show how studying language can contribute to our knowledge of the history of its speakers, and in this sense the book is intended not only for an audience of linguists but also for historians and archaeologists. Readers are not required to have any previous knowledge of linguistics, for all important concepts and the methodology of language reconstruction will be explained to them. This does not mean, however, that this book is an easy read throughout. Although I have aimed at maximum clarity, complex matters – and there are some to be found here – cannot be made simpler than they are, although they can be presented more simply than they usually are. It is my hope that anyone with genuine interest in language history and a little bit of time on their hands can understand everything I have written.

I have not striven to present the current consensus on language contact and the rise of the Germanic languages, first of all because there is none and, secondly, because presenting consensus in historical linguistics is a dreary and sterile business. Instead, I have concentrated on full and coherent argumentation regarding the theme of the book, so that readers who agree or disagree with what I write will be able to understand and formulate why. This effort entails that the book is not at all comprehensive: many ideas that over the years have been expressed in print about the origins of the Germanic languages are left unmentioned, not necessarily because I find them incorrect, but because they are not germane to the issues raised in the book.

This book has been long in the making, and I have done my utmost to test the patience of some of my colleagues and of the publisher. There are various reasons for the delay, apart from my inveterate optimism in planning

ahead and the fact that academic life itself has a habit of interfering with work. In spite of the crushing weight of published scholarship on the history of the Germanic languages, where an article published in 1870 is usually as relevant as one published last year, there is actually very little accumulated knowledge on which to fall back if one wishes to find out about the role that language contact has played in the early history of the Germanic languages. Another reason for the delay is the vastness and complex nature of the linguistic material involved. Anyone who has tried to master the historical phonologies of Old English or Dutch will know what I am talking about, and the reader will get a bitter taste of it in the chapter about Dutch. By definition, language contact involves more than one language, and in the case of early Germanic, the contact languages lie outside Germanic. Hence, one may spend the best part of one's life studying Germanic philology and not be able to write one sensible word about the theme of this book. Latin, the earliest stages of the Romance languages, Irish, British Celtic, Finnish, and Saami are the contact languages that will make an appearance in this book, and more than a glancing acquaintance with all of them was required in order to assess their contribution to Germanic.

A major advantage of a long gestation period is that it has given me the opportunity to try out, in various talks and in specialist publications, some of the ideas that will be presented here (see in particular Schrijver 2002, 2009, 2011a). All reactions, which varied from matter-of-fact criticism to mild enthusiasm and roaring silence, have been taken on board to the best of my abilities.

My thanks are due to Lisette Gabriëls, who read the manuscript before publication and suggested many improvements, and to Willem Vermeer, who has been my mentor and subsequently my partner in crime in ancient language contact studies.

I. Introduction

1. WHAT THIS BOOK IS AND IS NOT ABOUT

In recent decades, a wealth of scholarly literature about language contact has seen the light of day. Ever since Uriel Weinreich (1953) not only made the study of language contact respectable again but also restored it to a position of central importance in the linguistic enterprise in general, it has become clear that language contact is one of the most important triggers for language change. From among the vast literature on that subject, I wish to mention in particular Thomason and Kaufman (1988) and Thomason (2001), which are influential and eminently accessible general surveys.

This book aims to find out the extent to which language contact has been involved in the emergence of the individual Germanic languages and of the Germanic language family as a whole. Pride of place is given to the origins of English (chapter II), German (chapter III), and Dutch (chapter IV), while the origin of the Germanic family itself as well as the earliest histories of its East, West, and North Germanic branches are the subject of chapter V. Since it is relevant to the origins of English, the origin of Irish is discussed at length in chapter II, section 8.

Since the main interest of the book lies with fundamental structural changes caused by contact, I shall have next to nothing to say about the borrowing of words. Loanwords usually form a thin varnish and do not affect grammar to any noticeable extent, and so it is with the Germanic languages, however packed they may be with loanwords from many sources. Language contact causes structural changes if it affects the sound system (phonology and phonetics) and syntax. For practical reasons, this book is almost completely about sound change rather than syntactic change. That has to do with the time frame it covers and with the nature of the available source material. Because of the emphasis on origins, the chosen time frame is between the first millennium BC (chapter V) and the period spanning late Antiquity and the earliest Middle Ages (approximately the first to eighth centuries AD; chapters II, III, and IV). In this period, contemporary written sources dealing with the languages involved are very scanty, if not altogether absent. Therefore, much of the argument will revolve around

reconstructed stages of Germanic languages. This is not necessarily a great disadvantage because the state of the art in the linguistic reconstruction of the Germanic languages is so evolved that there is a wealth of information available, and the methodology used for reconstructing language, the so-called comparative method, is one of the most robust methodologies available to science (see I.3). But this emphasis on reconstruction, and consequently our dependence on the comparative method, does allow a much firmer grasp on the sound systems than on syntax. While the comparative method enables us to reconstruct sound changes (I.3) and the chronological order in which those changes operated (II.4.2), there is no reliable method for reconstructing syntactic changes (II.4.1), let alone putting those changes in a chronological order. For that reason, this book will largely ignore syntactic change and concentrate on sound change through contact.

Although the theme of the book is evidently linguistic, dealing with language contact inevitably entails dealing with communities of speakers of different languages that interact with one another. Since I shall have a lot to say about communities shifting from one language to another, the question arises who those communities were and why they shifted to a new language. Language shift is always the result of strong political or socio-economic pressure, and this is where linguistics and history meet. It is the specific aim of this book to explain how linguistics can contribute to our knowledge of history. I contend that language is one of the most important sources for the history of the 'Dark Ages', the period spanning late Antiquity and the early medieval period.

Since using language to reconstruct history is by no means straightforward, I have been as explicit as possible in explaining the lines of linguistic thought that underpin the conclusions of the book. I have tried to avoid the use of linguistic jargon, and since I do not assume that the reader has any previous knowledge of linguistics, I have explained all concepts and terminology that are necessary in order to follow the line of thought.

2. LANGUAGE CONTACT AND LANGUAGE CHANGE

I follow Thomason and Kaufman (1988) in differentiating between two ways in which languages in contact may influence one another: change through borrowing and change through shift. The difference is best explained from the point of view of a speaker. Let us use the example of someone who speaks English as his mother tongue and is exposed to Welsh because he lives in a community in which Welsh is widely spoken. He might easily adopt Welsh language features in his English, such as words for notions that do not have an English equivalent, e.g. *Eisteddfod*, the name of the most famous Welsh cultural festival. Once he starts doing this, he is changing his

English through borrowing. He has set foot on a slippery slope, which theoretically may lead him to borrow ever more words from Welsh, including words for notions for which English does have an equivalent, and ultimately even for basic vocabulary items, such as words for body parts, the sun and moon, water and fire, and basic verbs of the type 'do, make, go'. In cases of intensive borrowing relations, our English speaker may also adopt structural features, such as pronunciation (think of the characteristically Welsh intonation patterns) and items of Welsh syntax. Essentially, our English speaker is a fully competent speaker of English, who borrows from Welsh irrespective of whether he is almost completely ignorant of Welsh or a fluent speaker of that language. What he does is change English (his own English, that is) by borrowing from Welsh.

Now suppose that our English speaker starts learning Welsh, in the hope of becoming part of the local Welsh community. He experiences what every learner of a second language is faced with: English will interfere with the Welsh he is trying to learn. Whereas Welsh words will be acquired relatively easily, it is much more difficult to adopt a perfect Welsh pronunciation and perfect Welsh syntax. As a result, even a relatively successful learner of Welsh will betray himself as a non-native, in particular because of his English accent and his imperfect ability to master the rules of Welsh syntax. The variety of Welsh that he speaks has arisen as a result of change through shift, that is, because the speaker has shifted from a language he speaks fluently (in this case English) to a language he knows imperfectly (Welsh). His way of not quite mastering Welsh is called *imperfect learning*.

So, by and large, borrowing betrays itself by the introduction of loanwords, to which borrowing of sound structure and syntax may or may not be added, while change through shift betrays itself through the introduction of sound structure and syntax, to which borrowing of words may or may not be added. Borrowing and change through shift are useful concepts that discipline our thinking about language contact, but the model to which those concepts belong is just that: a model. Reality is often more complex. What is actually borrowing may look like shift. Thomason (2001: 11, 80–81) points out that speakers of Montana Salish, in the northwestern United States, for cultural reasons refuse to adopt English loanwords; if English influence on Montana Salish were to betray itself (all speakers are bilingual), it would be through borrowing of sound structure or syntax, not through borrowing of words. So the contact situation would look like change through shift (i.e. speakers changing Montana Salish by shifting from English to Montana Salish) rather than borrowing (i.e. speakers changing Montana Salish by borrowing features from English). That would be an inaccurate account of the real situation, in which speakers of Montana Salish are in the process of shifting to English.

Another complication is that what looks like borrowing may actually be shift. In bilingual Welsh-English communities, the local variety of English

often betrays the influence of Welsh intonational patterns: an English stressed syllable adopts low tone, while the immediately following syllable adopts high tone. The phenomenon arose when speakers of Welsh learned English but stuck to Welsh intonation (change through shift). But a native monoglot speaker of English who lives in a bilingual Welsh-English community and is confronted with English spoken with a Welsh accent on a daily basis may well import a Welsh intonational pattern into his own English (change through borrowing). Hence a feature that arose through shift can spread through borrowing.

In view of such complications, my use of the terms *change though shift* and *borrowing* comes with a health warning. In this book, *change through shift* always means that the change started as a result of people who spoke language A becoming bilingual in A and B, and then shifting to language B, in the process introducing elements of A into B. Subsequently, the change that entered B from A may (and usually does) spread through B by borrowing. Nevertheless, I shall take pains to find out which of the two terms (*change through shift* or *borrowing*) is appropriate for the individual cases of change through contact that are discussed throughout the book.

3. LANGUAGE CONTACT IN DEEP TIME

While the literature on language change through contact is vast, it is also one-sided in that it focuses on the collection and analysis of examples for which there is little doubt that a contact situation is involved. Thus, it is possible to study speakers of Scots Gaelic shifting to English, German immigrants in the United States speaking English, or an ancient text in Greek written by an Egyptian in the first century AD. What has been studied much less is how, in the absence of extralinguistic evidence for contact, a contact situation can be reconstructed on the evidence of linguistic data alone. This will be one of the main preoccupations in the following chapters. When we are dealing with linguistic changes that happened long ago, precise information about language contact situations that may be implicated in the changes is usually absent. For instance, we might want to assume that during the early medieval period a Celtic-speaking population in England switched to speaking Anglo-Saxon. In the process, the Celtic from which people switched vanished, which would make it very difficult to determine how the original Celtic might have influenced Anglo-Saxon, for all that is left of the original Celtic is a so-called substratum in Anglo-Saxon: presumedly Celtic features that survived in the speech of Celts who switched to Anglo-Saxon and that then spread in the Anglo-Saxon community to a lesser or greater extent. What we are assuming is that a vanished substratum influenced Anglo-Saxon, and we assume that it did only because we think we can prise apart the substratum from the Anglo-Saxon in which it has nested itself. Before we know it, we are involved in a circular argument.

As an antidote against substratomania, Thomason and Kaufman (1988: 111–112) write:

> In order to be able to make educated guesses in this area [i.e. interference through language shift], we must be able to identify a substratum language or language group (some of) whose speakers shifted to the target language [i.e. the language to which people switched] at the relevant time period; we must have information about its structure; and we must have information about the structure of the target language before the shift. These methodological prerequisites have frequently been ignored by substratum enthusiasts. . . . It is possible, for instance, that Celtic languages of the British Isles owe their un-Indo-European-like system of initial-consonant lenition, and other features too, to a pre-Indo-European substratum; but since we have no information about what language(s) the pre-Indo-European inhabitants spoke, we cannot establish such a cause for these changes (even if we were to agree that an external explanation is needed).

These so-called methodological prerequisites are brought into position in order to chase serious linguists away from exploring language shift in deep time. That is because the impossible is demanded: Thomason and Kaufman convict a murderer only if they have seen him commit the murder. That stance is perhaps understandable in the case of linguists who are not used to the subtlety of the detective work that goes into language reconstruction, but I was surprised to find that an eminent Indo-Europeanist and Celticist recently embraced this point of view, too (McCone 2005: 406). It is precisely the job of historical linguists to unearth the subtle and indirect clues that point to contact situations in deep time, as I hope to show.

4. THE COMPARATIVE METHOD

All of historical linguistics, including this book, would be reduced to a very learned form of informed speculation if it were not for the comparative method, which has been with us since 1878, when Hermann Osthoff and Karl Brugmann published their famous Neogrammarian manifesto in the preface to volume 1 of their *Morphologische Untersuchungen auf dem Gebiete der indogermanischen Sprachen*. This was the most important of a number of breakthroughs in the nineteenth century which secured historical linguistics a place at the forefront of science. Typical of the status of historical linguistics was that it provided inspiration for Darwin's theory of evolution.[1] The backbone of the comparative method is how it deals with sound change, as any good textbook on historical linguistics explains in detail.[2] The present book prides itself on not requiring any previous knowledge about linguistics

on the part of the reader, so an explanation is in order. In fact, since every educated human being should be aware of the method and hardly anyone actually is (and this includes not a few professional linguists), it would be irresponsible of me not to explain it, however briefly. Let me do so by giving a practical example.

Welsh and Breton are closely related languages that belong to the Celtic branch of the Indo-European language family. Both languages enter the written record in patchy sources from the end of the first millennium AD, which multiply considerably during the later medieval period. Being related means that Welsh and Breton stem from a common ancestor, Proto-British, which was spoken approximately during the sixth century AD in western Britain. Apart from a few names in Latin sources, no texts in Proto-British survive, so we have no direct access to that long-lost language. Yet historical linguists feel confident that they would be able to keep up a simple conversation with a sixth-century Proto-Brit, not by bombarding him or her with thirteenth-century Welsh or fifteenth-century Breton in the idle hope that a basic rapport might be struck up in this way, but by actually speaking a reconstructed variant of sixth-century Proto-British. How is this done?

Consider the following list of words in Medieval Welsh and Medieval Breton. Since we shall concentrate on the vowel sounds, note that Welsh *aw* is pronounced as [au] (as in English *house*), and Breton *eu* is pronounced as [ø] (as in French *feu* 'fire'). In all words, stress is on the first syllable.

	Medieval Welsh	**Medieval Breton**	**meaning**
1.	*brawd*	*breuzr*	'brother'
2.	*mawr*	*meur*	'big'
3.	*diawg* (2 syllables)	*dieg* (2 syllables)	'lazy'
4.	*marchawg*	*marcheg*	'horseman'
5.	*ofer*	*euver*	'vain'
6.	*bore* (2 syllables)	*beure*	'morning'
7.	*brodyr*	*breuder*	'brothers'
8.	*trindawd*	*trinded*	'trinity'
9.	*llawn*	*leun*	'full'

All Welsh words look similar to their Breton counterparts, and their meanings are more or less the same. Hence it is not too hard to believe that the Welsh and Breton words are cognates; that is, they derive from the same ancestral, Proto-British forms. What we shall do now is to figure out how the Proto-British forms were pronounced.

The **first step** is to align each individual phoneme (i.e. each sound, keeping things simple) of a Welsh cognate word with its counterpart in the Breton cognate word:

	Welsh		Breton
1	*b*	~	*b*
	r	~	*r*
	aw	~	*eu*
	d	~	*z*
	∅	~	*r*
2	*m*	~	*m*
	aw	~	*eu*
	r	~	*r*
	etc.		

Wherever a Welsh sound corresponds to the same sound in Breton, as in the case of *b* and *r* in the word for 'brother' and *m* and *r* in the word for 'big', things are simple: the words for 'brother' and 'big' in ancestral Proto-British will have had the same sounds in the same positions in the word. But wherever Welsh and Breton have different sounds, it is unclear what to reconstruct:

	Welsh		Breton	Proto-British reconstruction
1	*b*	~	*b*	**b*
	r	~	*r*	**r*
	aw	~	*eu*	**?*
	d	~	*z*	**?*
	∅	~	*r*	**?*
2	*m*	~	*m*	**m*
	aw	~	*eu*	**?*
	r	~	*r*	**r*

Reconstructed forms, whether they be individual phonemes or whole words, are conventionally written with a star in front of them, in order to indicate their hypothetical nature. What we shall now do is to follow a procedure that will enable us to fill in the question marks involving the Welsh vowels *aw* and *o*. The same procedure would allow us to fill in the other question marks, but following that through would lead us too far astray.

So we wish to find out which Proto-British sounds are hiding underneath Welsh *aw* and *o*. The **second step** in the procedure is to establish so-called regular correspondences. In the cognate pair 1, *brawd* ~ *breuzr*, we observe that where Welsh has *aw*, Breton has *eu*. This correspondence recurs in 2, *mawr* ~ *meur*, and in 9, *llawn* ~ *leun*. We have found that the correspondence Welsh *aw* ~ Breton *eu* is not limited to a single word but is recurrent: there is a regular correspondence between Welsh *aw* and Breton *eu*. It is typical of related languages that cognate words are made up of phonemes that correspond regularly to one another.

So we have found a regular correspondence Welsh *aw* ~ Breton *eu*, which occurs in words 1, 2, and 9. But Breton *eu* is involved in another regular correspondence as well, that is, with Welsh *o*: see words 5, 6, and 7. A third regular correspondence, between Welsh *aw* and Breton *e*, can be found in words 3, 4, and 8:

regular correspondence a: Welsh *aw* ~ Breton *eu*
regular correspondence b: Welsh *o* ~ Breton *eu*
regular correspondence c: Welsh *aw* ~ Breton *e*

If the word list were extended to include more words containing Welsh *aw* and Breton *eu*, more examples for each regular correspondence would be found, confirming the recurrence of the patterns involved.

We are now ready to take the **third step**, which concerns establishing the number of phonemes in Proto-British that gave rise to the Welsh and Breton vowels. Although only two different Welsh (*aw*, *o*) and two different Breton (*eu*, *e*) phonemes (sounds) are involved in the three regular correspondences, the number of phonemes that may be reconstructed for Proto-British is maximally three, not two. That is because the number of regular correspondences rather than the number of different sounds involved in them equals the maximum number of Proto-British phonemes:

regular correspondence a: Welsh *aw* ~ Breton *eu* < Proto-British *X
regular correspondence b: Welsh *o* ~ Breton *eu* < Proto-British *Y
regular correspondence c: Welsh *aw* ~ Breton *e* < Proto-British *Z

X, Y, and Z are cover symbols for whatever sounds are hiding behind them. The symbol < is a shorthand for 'develops from' (its counterpart > means 'develops into').

A crucial aspect of the comparative method is that the maximum number of three phonemes in Proto-British can be reduced to two or one if they show a so-called complementary distribution. If, say, correspondence a, *aw* ~ *eu*, always occurs in a sound context in which correspondence b, *o* ~ *eu*, never occurs, a and b are said to be in complementary distribution. The consequence is that a and b go back to just one rather than two sounds in

Proto-British. It so happens that all three regular correspondences are in complementary distribution:

a. Welsh *aw* ~ Breton *eu*:	only in words consisting of one stressed syllable (*brawd, mawr, llawn*)
b. Welsh *o* ~ Breton *eu*:	only in the first (stressed) syllable of words that have more than one syllable (*ofer, bore, brodyr*)
c. Welsh *aw* ~ Breton *e*:	only in the second (unstressed) syllable of words that have more than one syllable (*diawg, marchawg, trindawd*)

Consequently, the three regular correspondences can be reconstructed as one Proto-British phoneme (rather than three phonemes), which we shall call *X.

The **fourth step** is to turn our findings so far into so-called sound laws: rules that govern the development of a phoneme on its way from the protolanguage to the daughter languages. Sound laws are formulated in such a way that they do not allow exceptions. Formulating the sound laws that govern the development of *X in Welsh and Breton is easy now that the complementary distributions have been found:

Proto-British *X	> Medieval Welsh *aw* in final syllables (e.g. *brawd, marchawg*)
	> Medieval Welsh *o* in non-final syllables (e.g. *ofer, brodyr*)
Proto-British *X	> Medieval Breton *eu* in initial syllables (e.g. *meur, beure*)
	> Medieval Breton *e* in non-initial syllables (e.g. *dieg, trinded*)

It can easily be seen that this formulation of the sound laws involving *X completely accounts for the developments of *X in Welsh and Breton (at least as far as can be judged on the basis of the word list). If Welsh words were to turn up that have *aw* in a non-final syllable, or Breton words with *eu* in a non-initial syllable, they violate the sound laws. In that case the sound laws need to be refined until they account for the exceptions as well, or the exceptions have to be explained in a different way (as loanwords or new formations, for instance).[3]

The **fifth and final step** is filling in *X with phonetic content. This is the least exact step in the procedure, for it is based on educated guesses. Deciding what *X may have sounded like is possible by studying the phonetics of its children, i.e. Welsh *aw* and *o* and Breton *eu* and *e*. What all four have

in common is that they are vowels (to be more exact, *aw* is a combination of two vowels, together forming one syllable, a so-called diphthong). So *X probably was a vowel. Three out of four of its offspring are simple vowels (*o*, *eu* = [ø], *e*) rather than diphthongs (*aw*), so *X probably was not a diphthong. Three out of four of its offspring are produced with lip rounding (the *w* in Welsh *aw*, the *o*, and Breton *eu* [ø], so it is reasonable to assume that *X was a rounded vowel. Vowels can be close (produced with the mouth almost closed, as in the case of [u]), open (produced with the mouth relatively wide open, as in the case of [a]), or mid (in-between, as in [o]). Three out of four children of *X are mid, so *X is bound to have been a mid vowel. So *X probably was something like a rounded mid vowel, i.e. a sound like [o] or [ø]. This best guess can be further refined on the basis of other considerations. The oldest sources of British Celtic are names in Latin texts, which cover the period between the first century AD and the end of the first millennium. Around 600 *X was spelled as <o> (Welsh *mawr*, Breton *meur* 'big' appears in names as <mor>, for instance); earlier it was spelled <a>. The latter agrees with the way in which *X was pronounced in the closest cognate languages of British Celtic, viz. Irish *á* [ɑː] and Gaulish *a* (exact pronunciation unclear). Currently, the best guess as to the nature of *X is [ɔ], that is, an open-to-mid rounded back vowel (back means the vowel is pronounced with the tongue slightly retracted), approximately as in British English *dog* [dɔg].

This five-step procedure can be performed for all individual sounds in all of the Welsh and Breton words cited earlier, and indeed for all sounds that make up the entire lexicon and grammar of Welsh and Breton. In this way, the lexicon and most of the grammar of Proto-British can be reconstructed. Since British Celtic is closely related to Irish, the common ancestor of those two can be reconstructed as well, and so on, until we reach Proto-Celtic, the common ancestor of all Celtic languages; Proto-Italo-Celtic, which is the common ancestor of Celtic and the Italic language family (whose best-known member is Latin, itself the ancestor of the Romance language family); and finally Proto-Indo-European, the common ancestor of all Indo-European languages.

The power of the comparative method is based on the fact that sounds in languages change in such a way that their behaviour can be captured by rules which ideally allow no exceptions (sound laws). Its most powerful effect is that it enables us to reconstruct protolanguages, even though those protolanguages may look nothing like their descendants. We have seen one simple example of this effect: based on the Welsh sounds *aw* and *o* and the Breton sounds *eu* and *e* we were able to reconstruct one single phoneme in Proto-British (not two, as are attested in Welsh or Breton), and we were able to ascribe a probable sound value [ɔ] to that sound, which differs phonetically from all of its offspring in Breton and Welsh (*eu*, *e*, *aw*, *o*). This effect becomes more noticeable when deeper reconstructions are made. For

instance, we can be certain that the Proto-Celtic form from which Welsh *pump* 'five' and Irish *cúig* 'five' are descended was **k^{w}enk^{w}e*, even though not one single sound of **k^{w}enk^{w}e* survives unchanged in its descendants. Similarly, and at a deeper level, English *I* [aɪ] and Russian *ja* 'I' can be reconstructed with equal confidence as Proto-Indo-European **eǵHom* (where **ǵ* sounds like *eggyolk* and the **H* is a consonant whose exact phonetic value has not been determined apart from the fact that it was a fricative produced at the back of the mouth). The comparative method is the backbone of all linguistic reconstructions (starred forms), which appear abundantly in this book.

II. The Rise of English

1. LANGUAGES COMPETING FOR SPEAKERS: ENGLISH AS A KILLER LANGUAGE

It is common for languages with expanding populations of speakers to grow at the expense of other languages. The better a language manages to increase the number of its speakers, the more aggressively it behaves towards competing languages. If we were to trace back the pedigrees of all monolingual speakers of English in the modern world, we would find that the majority stem from ancestors that one or two centuries ago spoke a different language. That is especially true for areas in which English is a newcomer, such as North America, Australia, and New Zealand, where it ousted the languages of immigrants from other parts of the world than England. Those languages, such as Norwegian, Dutch, German, Russian, and Italian, survive in good health today in their old homelands. The hundreds of native languages of North America and Australia have fared much worse, however, and many of them are on the brink of extinction, if not already beyond. English is one of a handful of particularly successful killer languages in the modern world, together with Spanish and Portuguese in Latin America, and Russian and Chinese within their respective state boundaries. The scale on which those killers operate nowadays may be unprecedented, but there is nothing new about languages expanding at the cost of others. It is simply a condition of survival that is as old as linguistic diversity itself.

Although in this book I shall continue to speak about language competition and killer languages, there are three ways in which the imagery is potentially misleading. Firstly, a language, defined as a chain of mutually comprehensible neighbouring dialects, is not the basic unit at which the competition takes place, although it is the unit on which this book focuses. Within a language, dialects (language varieties spoken in a particular area) and sociolects (varieties spoken by a particular layer of society) compete in the same way: think of the loss of dialectal variety over the last century in most European countries as a result of the spread of one or a few varieties. On the most elementary level, the tiniest difference between the speech

of two speakers engaging with one another is subject to the same type of pressure, as one of the speakers may choose to conceal or highlight that difference in order to avoid or create social distance.

That brings us to the second possible confusion to which the imagery may give rise: it is not so much languages that compete with one another as people, who, consciously or unconsciously, instrumentalize language as one of the means to express social relations. In no way is the English language intrinsically fitter than the languages it replaces. It is the behaviour of its speakers that is responsible for the rate at which a language spreads or contracts. English spread across Ireland from the sixteenth century onwards not because English is a better language than the native Celtic language Irish that it replaced but because it was the language of the political, economic, and social elite as well as of colonist farmers who immigrated from Britain. Since a good command of English became a condition for upward social mobility, the Irish-speaking population felt pressure to become bilingual Irish-English speakers. When in the nineteenth century the Great Famine hit the rural poor in particular, many of whom happened to be Irish speakers, and Irish became stigmatized as the language of the destitute as well as the backward, the stage was set for a large-scale switch from Irish-English bilingualism to English monolingualism, and Irish political independence and the active promotion of Irish fluency through the school system and civil service have managed to achieve very little in their efforts to turn the tide.

Language hitches a ride on the back of human history. The mechanisms responsible for a language's spread or contraction are as complex as the factors determining the course of history itself. Any ploy between the extremes of genocide and the successful flogging of undubbed television programmes is capable of promoting one language at the expense of another. A language may spread because its speakers decide to slaughter all male speakers of another language, take their land, and enslave female speakers, for reasons that have nothing to do with language. Or a language may spread because people wish to master that language as one way of becoming part of a society that offers its speakers the opportunity to climb the social ladder more effectively. Perhaps surprisingly, the historical linguist has a hard time telling such extremes apart on the basis of linguistic evidence alone. The expansion of a language is usually just a concomitant effect of particular socio-political or economic changes. Consequently, saying that languages rather than speakers compete is a metaphor. But since this is a book about linguistics and early European history, it is also a useful abstraction, because when we say about Ireland, for instance, that English is outcompeting Irish, that statement subsumes all possible historical scenarios responsible for the fact and allows me to postpone, sometimes indefinitely, answering the question of what exactly went on in Ireland between 1600 and the present day. The deeper we delve down into history and prehistory, the murkier will be the historical record and the greater the need for the abstraction.

The third possible misconception to which talk about competing languages may give rise is that whenever languages meet, there will inevitably be competition between them. Such is not the case. Bilingualism and even multilingualism are the norm in many parts of the world up to the present day. For instance, many people living in the northeastern Caucasus are at least bilingual and more often tri- or quadrilingual. They would often speak Russian and the East Caucasian language Avar, which is the lingua franca of Dagestan, as well as one or more local languages of East Caucasian, Turkish, or Iranian extraction. This situation has been relatively stable for centuries, apart from Russian, which is a nineteenth-century newcomer in the area. Each language has its own particular niche in which it is used, and people switch from one language to the other accordingly. High-ranking officers and civil servants in the Austro-Hungarian Empire before 1918 would be fluent in German, Hungarian, and one or more Slavic languages, such as Croat, Slovenian, or Czech. In those settings, multilingualism is a matter of survival. For many people living in Wales, western Scotland, the Finnmark, Frisia, Brittany, the Lausitz, Graubünden, South Tyrol, or the Basque Country, to mention just a few European examples, bilingualism has been a fact of everyday life for generations. So languages can coexist relatively peacefully within a single community and even within the confines of a single skull. Why, then, highlight languages struggling with one another? Firstly, because this happens to be a book about people shifting from one language to another. Such shifts inevitably go through stages at which people are bilingual, but those stages may not last more than one or two generations. Therefore, the focus here is on unstable bilingualism that results in shift. And, secondly, the nature of the material discussed in this book involves looking at language within a span of a couple of centuries. That wide a time frame increases the chances that what at one point was stable bilingualism will become destabilized and give way to language shift.

English's history as a killer language has a respectable pedigree in Britain, too. It has made heavy inroads into Scots Gaelic and Welsh. It finished off Manx on the Isle of Man at the beginning of the twentieth century and Cornish about a century before that. Scots Gaelic, Welsh, Manx, and Cornish are all Celtic languages. Before the modern period, the rise of English likewise matched the decline of Celtic languages. English also managed to swallow up languages that were introduced into Britain by erstwhile conquerors. One of them is the French dialect of Anglo-Norman that was introduced as an upper-class language from the eleventh century onwards. Danish in the east and Norse in the northwest of England and in Scotland were imported in the period of Scandinavian expansion between the late eighth and eleventh centuries, and they, too, ultimately fell victim to English, sometimes within a generation or two, sometimes after many centuries.

2. THE ANGLO-SAXON SETTLEMENTS

The origin of English is famously tied up with an extinction event, as Anglo-Saxon settlers moving from their homelands in present-day northern Germany and Denmark brought along the dialects ancestral to English and gradually destroyed the fabric of Roman British society in a colonization movement that started after 400.

> In the archaeological view . . . the sequence of the transition from Roman Britain to Anglo-Saxon England is remarkably clear and comprehensible. Within the latest levels of Roman-period sites of virtually all types, including villas, villages, towns, forts, cemeteries and temples, a new, intrusive Germanic material element is often found. . . . This phase is the beginning of a subsequently unbroken sequence of Germanic cultural presence, soon a dominance, in Britain. It is a phase of radically different character, sequentially unconnected, to earlier isolated finds of Germanic character in Britain. . . . Thus this phase is rationally to be identified as the inception of the Anglo-Saxon Period. From this point onwards the new Germanic sites regularly outlive the earlier Roman sites, on which any late intrusive Germanic element is always ephemeral. There is no known case of any continuing, hybrid Romano-Germanic site emerging from this meeting of cultures.[1]

When by the seventh century the dust begins to settle, Anglo-Saxon kingdoms cover most of England and southern Scotland. The written record is dominated by Latin—the language of the Church and of learning rather than everyday speech—and by Old English. The latter is an array of dialects that presumably directly continue the dialects imported by the Anglo-Saxon settlers in the preceding centuries. They replace the British Celtic language, which was widespread before 400 but in the course of the medieval period managed to survive only at the western and northern fringes of Britain, in Cornwall, Wales, and Cumbria, although initially some pockets probably still remained elsewhere, e.g., in Devon, Dorset, and the Fens and in the former northern British kingdoms between Strathclyde and Edinburgh.

Latin must have been spoken widely in late Roman Britain as well, probably not so much the Classical, literary variety based on the works of Caesar and Cicero, which early post-Roman British authors like Gildas and St Patrick strove to write, after it had become the language of the Church, but rather the grammatically much simplified Late Spoken Latin that flourished throughout the Roman Empire from Libya to Hadrian's Wall and that is ancestral to all modern Romance languages. Its fate in Britain between 500 and 700 is not clear, but it is reasonable to think that a form of Spoken Latin survived well into this period (see section II.5.2).

At a conservative estimate, the population of Britain on the eve of the Anglo-Saxon migrations amounted to two to four million[2]—that is, two to four million speakers of Celtic and/or Late Spoken Latin. We know that the number of speakers of those languages shrank dramatically in the course of the medieval period, withdrawing as they did to Wales, Cornwall, and, across the Channel, Brittany, where they survived as the medieval written languages Welsh, Cornish, and Breton, respectively. Spoken Latin became extinct in Britain. But where did the people who spoke those languages go? Many possible scenarios have been entertained: they fell prey to genocide, inflicted either by the sword or by gradual starvation; they mass-migrated to areas beyond the immediate grasp of the invaders; or they were enslaved to become a vast underclass of mainly agricultural labourers. They may even have thrown in their lot with the new powers so successfully that they became as Anglo-Saxon as the Anglo-Saxons themselves, both culturally and linguistically. Circumstances may have varied from one period to the next and from one place to the next, so that multiple scenarios may have come about. What all these scenarios have in common is that they are quite drastic: they are geared to explain the almost unimaginable: how, in the course of just a few centuries, what began with a few boatloads of Anglo-Saxon mercenaries managed to transform a land populated by millions into a linguistically and culturally Anglo-Saxon society.

3. THE VANISHING OF THE CELTS AS SEEN BY LINGUISTS

To a large extent, it is linguistics that is responsible for thinking in terms of drastic scenarios. If a large Celtic-speaking indigenous population shifted to speaking the language of the Anglo-Saxon conquerors, linguists would expect to find certain traces of that shift, which in this particular case they do not. In order to understand the role of linguistics in the debate about Anglo-Saxon settlement, we need to introduce a few general concepts.

3.1. Absent Traces of Language Shift: Sound System and Syntax

When people adopt a second language, they find it difficult to acquire it so perfectly as to be indistinguishable from native speakers of that second language. That is because their first language provides a matrix into which the second language tends to be squeezed. The phenomenon is well known to anyone who has tried to learn a foreign language. Native speakers of English have difficulty pronouncing, say, German and French *r*-sounds because they tend to substitute the standard German and standard French uvular trill [ʀ] by their native standard English alveolar approximant [ɹ]. The converse holds for native speakers of French and German. Such interference by one's first language is especially prominent in the sound system, as this simple example illustrates, and also in syntax. English has a fairly rigid Subject-Verb-Object