For my parents

The Oxford Guide to Etymology

Philip Durkin

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS
Great Clarendon Street, Oxford OX2 6DP
Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide in
Oxford New York
Auckland Cape Town Dar es Salaam Hong Kong Karachi
Kuala Lumpur Madrid Melbourne Mexico City Nairobi
New Delhi Shanghai Taipei Toronto
With offices in
Argentina Austria Brazil Chile Czech Republic France Greece
Guatemala Hungary Italy Japan Poland Portugal Singapore
South Korea Switzerland Thailand Turkey Ukraine Vietnam
Oxford is a registered trademark of Oxford University Press
in the UK and in certain other countries
Published in the United States
by Oxford University Press Inc., New York
© Philip Durkin 2009
The moral rights of the author have been asserted
Database right Oxford University Press (maker)
First published 2009
All rights reserved. No part of this publication may be reproduced,
stored in a retrieval system, or transmitted, in any form or by any means,
without the prior permission in writing of Oxford University Press,
or as expressly permitted by law, or under terms agreed with the appropriate
reprographics rights organization. Enquiries concerning reproduction
outside the scope of the above should be sent to the Rights Department,
Oxford University Press, at the address above
You must not circulate this book in any other binding or cover
and you must impose the same condition on any acquirer
British Library Cataloguing in Publication Data
Data available
Library of Congress Cataloging in Publication Data
Data available
Typeset by SPI Publisher Services, Pondicherry, India
Printed in Great Britain
on acid-free paper by
CPI Antony Rowe, Chippenham, Wiltshire
ISBN 978–0–19–923651–0
1 3 5 7 9 10 8 6 4 2

Contents

Acknowledgements viii
About this book ix

1 Introduction 1
1.1 What is etymology? 1
1.2 Some basic concepts: two example etymologies 3
1.3 Why study etymology? 22
1.4 What an etymologist does 31

2 What is a word? Which words need etymologies? 34
2.1 What are words? 34
2.2 How new words arise 43
2.3 Lexicalization 49
2.4 Examples of lexicalization 51
2.5 Apparent reversals of the process 56
2.6 Cranberry morphs 56
2.7 Which words need etymologies? 58

3 Are words coherent entities? 61
3.1 Variety in form and meaning: *poke* 'bag, sack' 62
3.2 Do we know precisely when a word's history begins? Can we assume continuity of use? 68
3.3 Homonymy and polysemy 74
3.4 How polysemy–homonymy relations can change 76
3.5 Merger (or near-merger) in form and meaning 79
3.6 Splits in word form 83
3.7 A case of merger followed by a split 86
3.8 Homonymic clash 88

4 Word formation 94
4.1 Issues concerning affixation 95
4.2 Synonyms, nonce formations, and blocking 103

with synonyms of different origin: the difference in the receptivity of various languages to loans of various origins may often be a question of avoidance of the institutionalization of loans rather than avoidance of initial, nonce borrowing.[18]

[18] A very interesting comparison could be made with the complex history of the reception of English loanwords in Japanese, particularly the massive borrowing of English words since the end of the Second World War. For a very useful recent account of this topic (albeit largely from the standpoint of second language learning) see Daulton (2008), and compare also Shibatani (1990).

6

The mechanisms of borrowing

6.1 Perspectives from contact linguistics 155
6.2 What is basic vocabulary? 157
6.3 Language shift 161
6.4 Borrowing within and between languages 164
6.5 Borrowings from more than one language 165
6.6 Continuing semantic influence and interference 167
6.7 Multiple borrowings from the same source 169
6.8 How can we tell that borrowing has occurred? 169
6.9 Lexical borrowing and code-switching 173
6.10 Some conclusions from chapters 5 and 6 177

6.1 Perspectives from contact linguistics

In the preceding chapter we looked at some of the circumstances and causes of lexical borrowing. We saw that a satisfactory account of a borrowing will not simply assert that a borrowing has occurred, but will also provide some plausible context for it to have occurred in. We also saw that close investigation of such etymologies can reveal a great deal about linguistic and cultural history. Our focus has been largely on how individual speakers of languages adopt new lexical items. This reflects a major focus in the field of contact linguistics, well characterized by the Middle English dialectologist Angus McIntosh (1994: 137):

Fundamentally, what we mean by 'languages in contact' is 'users of language in contact' and to insist upon this is much more than a terminological quibble and has far from trivial consequences.

In the past several decades, work on contact linguistics has brought a certain amount of attention to bear on lexical borrowing. Sometimes, admittedly, the focus has been on borrowing of grammatical or phonological features, and lexical borrowing has been investigated more for the light it can throw on such phenomena than for its own sake. A good deal of attention has rightly been given to bilingualism. Only extremely limited borrowing is possible in a contact situation if neither the speaker of the donor language nor the speaker of the borrowing language knows anything of the other's language: someone pointing at an object and speaking a word is possibly giving its name, but the potential for confusion is enormous. Only a little more borrowing is possible if the speaker of the donor language knows something of the borrowing language, but the speaker of the borrowing language knows nothing of the donor language: the speaker of the borrowing language may ask 'what do you call this?' and receive in reply a word from the donor language, but again confusion may very easily result.[1] For any more extensive borrowing to occur we must have either a situation in which two dialects or languages are at least in part mutually intelligible, or one in which at least one speaker of one language has at least enough knowledge of another language to apprehend a word in that language and adopt it in her/his own language. To this limited extent, most borrowing will involve some degree of either mutual intelligibility or bilingualism. (In linguistic use the term 'bilingual' is often used in a very much broader sense than its everyday meaning 'having fluency approaching that of a native speaker in more than one language'.) A distinction is often made between basic and non-basic vocabulary, basic vocabulary being taken to be much

[1] For an example of confusion which probably arose in such a situation see the etymology in *OED* or in Corominas and Pascual (1981) 690–1 of the Spanish word *pulque*, the name of a kind of drink made from the fermented sap of the agave or maguey. This is probably borrowed from a Nahuatl word *puliuhki* which in fact means 'decomposed, spoiled'. The drink is called *octli* in Nahuatl, and the Spanish name probably results from misapprehension of the phrase *octli puliuhki* 'spoiled pulque', which would have been heard frequently since pulque spoils easily if not drunk within twenty-four to thirty-six hours.

more resistant to borrowing in normal borrowing situations than non-basic vocabulary.[2]

6.2 What is basic vocabulary?

Any assessment of borrowing on the basis of a distinction between basic and non-basic vocabulary begs the question of how these terms are defined. Here there is a good deal of opacity in much of the scholarly literature. A useful sketch of what is usually meant by basic vocabulary is provided by Trask:

> There is clear evidence that certain semantic classes of words are much less likely to be borrowed than other words. These are chiefly the items of very high frequency which we would expect to find in every language: pronouns, lower numerals, kinship terms, names of body parts, simple verbs like *go*, *be*, *have*, *want*, *see*, *eat*, and *die*, widespread colour terms like *black*, *white*, and *red*, simple adjectives like *big*, *small*, *good*, *bad*, and *old*, names of natural phenomena like *sun*, *moon*, *star*, *fire*, *rain*, *river*, *snow*, *day*, and *night*, grammatical words like *when*, *here*, *and*, *if*, and *this*, and a few others.
>
> (Trask (1996) 23; reprinted Millar (2007) 27)

The classic codification of this approach is in the lists of basic items devised by the linguist Morris Swadesh in the 1950s, especially a short list consisting of 100 items and a longer one with 200 items. A full listing with supporting discussion is given in McMahon and McMahon (2005) 33–9 (which provides an excellent introduction to this field), and also in Millar (2007) 483–4. To give an indication of the sorts of items that are included, the first ten items in alphabetical order in the 100-meaning list are 'all', 'ashes', 'bark', 'belly', 'big', 'bird', 'bite', 'black', 'blood', 'bone'; in the 200-meaning list 'and', 'animal', 'at', 'back', 'bad', 'because', 'blow' are added in the same section of the alphabet. I have given these items in quotation marks, because as items on the list they represent meanings, not words. However, if the language that we are considering is modern (standard) English, then the words

[2] A very influential set of generalizations about what happens when one language is in contact with another in a maintenance situation (i.e. where language A shows borrowing from language B, but where neither language is being abandoned by its speakers) is found in Thomason and Kaufman (1988: 74–6) in the form of a 'borrowing scale', which is revised considerably in Thomason (2001). On language maintenance and language shift see further section 6.3.

which denote these meanings are the same as those that occur in the list.[3] We can therefore use the Swadesh lists to tell us at least a little about the numbers of words ultimately borrowed from various sources that can be found in the most basic levels of the vocabulary of modern English. If we consider the origins of the words in just the sample that I have already given, then *animal* stands out as the only word which is a borrowing from French and/or Latin, and in the rest of the 200-meaning list we find additionally only *count*, *flower*, *fruit*, *mountain*, *person*, *push*, *river*, *round*, *turn* (originally an Old English borrowing from Latin), *vomit*, and (debatably) *lake*, plus *because* as a probable loan translation.[4] Out of these only *mountain*, *person*, and *round* occur in the shorter 100-meaning list. Interestingly, the first dates recorded for these words in English (excluding *turn*) range between the early thirteenth and late fourteenth centuries, with the largest concentration being in the fourteenth century.[5] The vast majority of the words in both lists are part of the lexical inheritance of Old English, although not all have secure Germanic etymologies, an extreme instance being *bird*, which has no known cognates outside English. One item in the 100-meaning list, *dog*, stands out as an English word of quite unknown and much disputed origin which is first recorded in the eleventh century; the usual word for a dog in Old English is *hund*, modern English *hound*. There are a number of words of Norse origin in the 100-meaning list, *bark* (of a tree), *egg*, *root*, *skin*, (partly) *give*, (perhaps) *big*, *die*, and additionally in the 200-meaning list *dirty*, *hit*, *husband*, *leg*, *near*, *rotten*, *sky*, *they*, *wing*, (perhaps) *fog*.[6] There are also a

[3] Although see McMahon and McMahon (2005) 41 on the problems that can arise where more than one lexical item could fill the same slot in the list: such as *little* or *small* in English, for example.

[4] For convenience I use here simply the etymologies offered in the *Oxford Dictionary of English Etymology*. Among these basic vocabulary items there are many words of uncertain etymology, and some of very hotly disputed etymology, and any more detailed analysis would need to take account of some of the major areas of uncertainty.

[5] The first edition of the *OED* lists unambiguous evidence for *animal* only from the sixteenth century onwards, but the *Middle English Dictionary* offers convincing evidence from the end of the fourteenth century.

[6] Thomason and Kaufman (1988: 365 note 22) also use the Swadesh 200-item list to assess borrowing of basic vocabulary from Norse and French, but with slightly different resulting totals from mine. As they do not list the items taken to be of Norse or French origin, it is impossible to see whether this results from using slightly different modern English words in the semantic slots of the Swadesh list, or from assuming different etymologies, or from some other cause.

few words probably borrowed from Middle Dutch or Middle Low German: *dull*, *split*, and (perhaps) *rub*. (For a slightly different approach based not on Swadesh lists but on lists of the highest-frequency items in corpora of contemporary English see Minkova and Stockwell (2006).)

What this very sketchy survey does not tell us is some very crucial information about each of these borrowings:

- When did the initial borrowing occur?
- How long did each item compete with an earlier synonym, and what estimates can we arrive at for the frequency levels of each?
- When did each item become the usual term for this meaning in everyday use?
- What factors, if any, can be identified which favoured its adoption?
- Is it the usual term in all stylistic levels, registers, and regional varieties even today?

Such questions are not all readily answered even for such common words as these in such a comparatively well-documented and thoroughly studied language as English.[7]

In the list I gave of English borrowings from Old Norse, the third person plural personal pronoun *they* is of particular interest. Even among the Swadesh list items, some items are more susceptible to borrowing than others, and it is generally held that personal pronouns are among those grammatical closed-class items which are least likely to be borrowed. Thomason (2001: 83–4, quoting work by Christopher Court posted on an electronic list) draws an interesting comparison with some languages of Southeast Asia, where borrowing of pronouns is fairly common, but where pronouns constitute less of a closed-class group, with numerous alternative forms occurring which can be exploited in marking different social relationships; compare the use of distinct second person pronoun forms for intimacy/informality and distance/respect/formality in many European languages from the Middle Ages onwards, as English *thou* and *you*, French *tu* and *vous*, etc. The English adoption of *they* from Norse is interesting in that the native form is simply replaced by a borrowed form. (Although it has been suggested that the native forms of the demonstrative pronoun may also have had some input.) Additionally, the borrowed form spread (gradually)

[7] For an interesting recent discussion of frequency of occurrence as a factor influencing rates of lexical replacement see Pagel, Atkinson, and Meade (2007).

from areas in which there was direct contact between speakers of Norse and English to areas in which there would have been little or no direct contact with Norse speakers; indeed, the spread to many areas occurred after the end of the period in which Norse was likely to have been in use in any part of mainland Britain. It has long since been shown that in fact language-internal factors played a crucial role in this process: borrowed *they* provided a much clearer contrast with the singular form *he* than did the inherited forms *hi*, *heo*, *he*, etc. This is sometimes referred to as a 'therapeutic' process, restoring important contrasts in the grammatical system which had become obscured as a result of phonetic change.[8]

Prestige is often offered as the explanation for the borrowing of basic vocabulary, but in some cases one may suspect that this is simply because it is clear that need will not work as an explanation, and prestige is the most readily available alternative. Lass (1997: 186–8) offers some very interesting examples of borrowing of fairly basic vocabulary items, such as names of parts of the body and of common foods, into Yiddish both from Hebrew and from Slavonic languages, in a sociolinguistic situation where Yiddish speakers in Eastern Europe would have been very unlikely to have regarded the language of neighbouring Slavonic speakers as particularly prestigious. He also provides a useful selection from among the many early borrowings into Finnish (a non-Indo-European, Uralic language) from Germanic languages and from other branches of Indo-European, including some numerals, names of body parts, and days of the week. These certainly show us that contact occurred between speakers of Finnish and speakers of various Indo-European languages (and incidentally, they often provide very useful evidence for earlier stages in the histories of these languages), but it is probably unwise to try to reconstruct any scenario for the type of contact which may have taken place.

Additionally, it must be recognized that the Swadesh lists cover only a tiny slice of the very most basic vocabulary, compiled explicitly in order to focus on those words thought least likely to be replaced as a result of borrowing over time. Even if we extend the list to include for instance larger numbers of names of body parts or of foodstuffs which are basic in a particular area, we will still only be looking at a fragment of the lexis

[8] See Samuels (1972) 71–2 for a daring but controversial account of the subsequent competition between the borrowed pronoun *thei* and *thei*, the southern reflex of Old English *þēah* 'although'.

of any language. If we are interested in the impact of borrowing on the whole of the lexis of a language, we may feel that there are distinctions just as significant to be drawn between much larger slices of the vocabulary. For instance, McMahon and McMahon (2005: 7) speak about borrowing from 'prestigious neighbours' being likely to include 'religious, cultural, or technological vocabulary'. In such contexts, we might suspect that there is an important distinction between the vocabulary which is at least in the passive competence of almost all adult speakers, and vocabulary which is largely confined to the competence of certain groups within society. We may also suspect that very many words which now form part of the passive competence of most speakers appeared first in one of the specialist vocabularies within a speech community, and only gradually spread more widely, although there has been surprisingly little research on this topic.

In the following section we will look at a rather different process which may also have a major impact on basic vocabulary, and which often cannot easily be distinguished from cases of borrowing.

6.3 Language shift

The situations that we have looked at so far all involve language maintenance, where the speakers of the borrowing language continue to speak the same language. However, lexis (as well as other features) can also enter a language through the distinct process of language shift. Here, the speakers of one language (let us call it language A) abandon that language in favour of another (B), but in the course of this process some features from language A are transferred into or imposed on language B. The process is typically inter-generational, when a bilingual generation is succeeded by one which retains only one of the languages, but with some transfer of vocabulary from the other. Townend, in the course of a wide-ranging investigation of the degree of mutual intelligibility likely to have existed between English and Old Norse in England, examines the possibility that at least some of the words of Norse origin which we encountered in the previous section could in fact show the result of shift-induced imposition (or lexical transfer) rather than borrowing in a situation of language maintenance:

> Traditionally, it is a linguistic commonplace that, generally speaking, words are borrowed from one language to another on account of either need or prestige... The

Norse loans in English have posed problems for linguists, as it is not clear which of these two causations is operative, or whether there are diachronic and diatopic variations in causation... The Norse loans in Old English tend to be need-based borrowings, denoting new objects (particularly nautical and legal terminology), whereas many of the Norse loans in Middle English can in no way be regarded as need-based borrowings as they constitute so-called core vocabulary. Normally, this would imply that Norse enjoyed greater prestige in the Middle English period than it did in the Old English, but this seems impossible, since it was in the Viking Age (if ever, and only in certain areas) that the Norse-speaking population was in authority over the English-speaking. To regard the Norse core vocabulary items appearing in Middle English as the result of imposition through shift rather than of borrowing would appear to remove this problem.

(Townend (2002) 203–4)

Such ideas present a powerful challenge to traditional notions of the causation of borrowing in terms of need and prestige. However, it can be difficult to identify transfer of lexis resulting from language shift with any certainty. As already noted in section 5.6, it is sometimes assumed that language shift from Anglo-French to Middle English was a major cause in the borrowing (or importation) of French lexis into English, and further that Anglo-French constituted a superstratum (rather than an adstratum or substratum) on the basis of the large amount of vocabulary which entered English, particularly in specialist fields such as legal language.[9] (A superstratum is a language of a dominant group which influences that of a subordinate group. An adstratum is a language which influences that of a neighbouring group without any such relationship of social dominance. A substratum is either a language of a subordinate group which influences that of a dominant group, or a language formerly spoken by a group which influences their subsequent acquisition of another language through grammatical, lexical, or phonological features.) Such a claim would be greatly reinforced if one could first pinpoint the likely period in which French speakers shifted to English, then identify words which entered English from French in this same period, and finally demonstrate why they are less likely to show the results of general processes of borrowing. At the very least, one would want to demonstrate some difference between this period and those preceding and following it, e.g. in total numbers of words entering English from French, or in the registers to which they belong. It is not clear how well this has been demonstrated in the case of English and

[9] See especially Thomason and Kaufman (1988) 116.

French, although there is rather better evidence in the case of English and Norse, as set out by Townend. As already noted in section 6.2, even the basic vocabulary items which entered English from French (or Latin) span a considerable historical period, which in itself rather argues against the hypothesis that transfer rather than borrowing played a significant part in this process.

Additionally, while language shift may account for words entering a language, it does not explain their subsequent spread into more general usage, particularly when, as in the case of Anglo-French speakers in medieval England, the language-shifting population made up only a small proportion of the total population, or when, as in the case of Norse speakers, they were largely confined to particular geographical areas of settlement. Whether lexis initially entered a language through contact-induced borrowing or shift-induced transfer, we must still account for its subsequent spread and competition with any pre-existing native synonyms. In the case of the Norse-derived element in Middle English, this issue has been addressed very carefully by Dance (2003) 311, 313:[10]

Full acceptance of a lexical item as the dominant expression within its field can take centuries to occur, if it ever does. Some lexical redundancy, or variant ways of expressing the same concept within the same lexical field, is natural within a system... just as is variation in terms of pronunciation and morphology. And, while it is proper in historical terms to describe such variation as contributing towards the process of change, nevertheless from a synchronic perspective it need not be seen as having such an 'effect' at all: the variation merely exists, and is available to be conditioned by factors such as social/stylistic level, perceived dialectal flavour, or, put more generally, simply according to the contexts and uses with which the different forms in question happen to have become associated by a particular speaker...

Lexical borrowing can be seen simply as adding to variation in the first instance, a predictable consequence of the increase in weak social ties that results from a contact situation, and not as a drastic imposition on the core of a language's vocabulary that needs to be accounted for by tremendous pressures of 'prestige' attaching to the source tongue.

Ideally, etymologies of borrowed items will account for such factors, explaining not only the initial adoption of a word, but its subsequent spread within the lexical system. In many cases we may lack sufficient evidence to trace this process, and even where the evidence is available such an investigation will be outside the scope of all but the most adventurous

[10] See also in this connection the important discussion in McIntosh (1978) and the essays on Middle English word geography in Laing and Williamson (1994).

etymological projects. However, the example of research such as Dance's should at least caution us against drawing over-hasty conclusions about the causation of borrowing in any particular instance, and whether this applies to initial adoption or subsequent spread, or to both.

6.4 Borrowing within and between languages

There is an intimate connection between borrowing of vocabulary from one language to another and the spread of words from one person's vocabulary (or idiolect or personal linguistic system) to another's. The latter process is sometimes called internal borrowing. There are certainly similarities between the two processes, but there are also important distinctions to be made. Differences between the grammatical systems of any two languages may well have a significant impact on borrowing, or even prevent it from happening at all. If we consider also borrowing between different dialects of a language, as well as borrowing between closely related languages, especially those with at least some degree of mutual intelligibility, we can place different types of borrowing on a rough cline, indicating difficulty of borrowing:

Less difficulty

- Between individuals who speak the same dialect of the same language, and have similar social status, profession, interests, etc.
- Between different specialist registers
- Between dialects, especially where there are significant differences in the phonology and/or grammar of the two dialects
- Between languages with some degree of mutual intelligibility
- Between languages with a long history of contact
- Between closely related languages
- Between unrelated languages

More difficulty

As Samuels points out, in both interlinguistic and intralinguistic borrowing the communicative needs are the same (1972: 97):

> In theory, the processes of spread could be regarded as the same, irrespective of whether the contact is between dialects of the same language or between different languages. This is because the dispositions and attitudes of those who have something to communicate are parallel; in both, there is a common tendency for speakers to adjust their speech to bring it nearer to that of their interlocutors.

Additionally, as we have seen, the two processes of interlinguistic and intralinguistic borrowing very often work in succession to one another, with the initial adoption of a word from another language being followed by intralinguistic spread. Furthermore, what have traditionally been offered as causal explanations for interlinguistic borrowing may in a good many cases apply more properly to subsequent intralinguistic spread. There are also cases where a word history shows successive waves of interlinguistic borrowing, either through subsequent direct contact with the original donor (or a related word), or through intralinguistic merging of words of distinct origins, as we will investigate further in the next section. In doing so we will turn again to English borrowing from French and Latin, since the relatively rich documentation available for all three languages enables us to build up an especially detailed picture of various borrowing phenomena.

6.5 Borrowings from more than one language

Borrowing may play a part in the type of composite word origin that we encountered in chapter 3. For instance, English *mien* 'the look, bearing, or manner of a person' (first recorded in the sixteenth century) has normally been explained as showing a clipped form of the word *demean* (ultimately of French origin, but long established in English by this date) merged with a loan from the totally unrelated French word *mine* 'appearance'. The second word was borrowed during the course of the Great Vowel Shift (see section 7.2.3), and evidently its vowel was identified with a variety of different English sounds, as is reflected by sixteenth- and seventeenth-century spellings such as *mine*, *meine*, *miene*; it would appear that identification with *meane* (from *demean*) was crucial in blocking development to /maɪn/, although the modern spelling *mien* reflects consciousness of the (partial) origin from French *mine*.[11]

A distinct phenomenon is shown by words which appear to have been borrowed partly from a word in one language and partly from a cognate word (which is usually either identical or very similar in form) in another language. A good test case is provided by words borrowed into English originally from Latin before the Norman Conquest, and subsequently either

[11] For some very interesting similar examples from Jamaican Creole see Cassidy (1966).

reborrowed from or reinforced by the corresponding (Anglo-)French words in the centuries after the Norman Conquest:

- *hellebore* was borrowed in the Old English period from Latin *elleborus*, but Middle English forms of the type *ellebre* point very strongly to secondary borrowing from (Anglo-)French forms of the type *ellebre*.
- *purpure* was borrowed in the Old English period from Latin *purpura*; disyllabic forms partly result from Old English inflected forms with syncope of the medial vowel, but it is likely that their later frequency owes a good deal to the variation in (Anglo-)French between forms of the type *purpure* and forms of the type *porpre*, *pourpre*.
- *pease* (of which modern *pea* is an inferred singular form) shows Old English borrowing from post-classical Latin *pisa* (variant of *pisum*), but (Anglo-)French influence in the Middle English period is demonstrated by forms with a diphthong (e.g. *peise*).

In other cases such as *passion* there is no very conclusive formal or semantic evidence pointing to reborrowing, although the frequency of Middle English forms such as *passioun* or *passiun* would be most easily explained as resulting from direct Anglo-French influence. *pelican* shows no distinctive formal or semantic influence from French in the Middle English period, but the survival of what was originally an Old English borrowing from Latin was probably reinforced by the formally and semantically corresponding French word, and in the early modern period we find the Middle French word used in metaphorical senses denoting an alembic and a type of device for extracting teeth earlier than we find the corresponding senses in English, suggesting secondary semantic borrowing.[12] *turn* is another similar example which we encountered among the Latin and/or French items in the Swadesh lists in section 6.2.

In the period immediately after the Norman Conquest, and beyond into the early modern period, we can find examples where clear evidence of word form, or meaning, or date of first attestation points to borrowing from French and others where it points to borrowing from Latin. But we also find many other examples where we cannot be certain: *manifest* (first attested in English in the fourteenth century) could equally well be from either French *manifeste* or Latin *manifestus*; similar cases are presented by for instance *negotiation*, *opposition*, *opulent*, or *pedagogy*.[13] These are mostly words which appear in French either as learned or semi-learned borrowings from Latin, or as remodellings of words which had earlier undergone phonological or morphological change during the transition from Latin to French (see examples in section 6.7). In either case the result is a French word which shows an exact formal correspondence with its Latin etymon, making it impossible to tell which is the etymon of the corresponding English word. Should such words be regarded as borrowings from French, or from Latin, or from both? Close attention to the particular circumstances of many such examples, and to parallel cases where a word shows formal or semantic influence from each language, suggests that the best course is generally to assume borrowing partly from the one language and partly from the other. This accords well with a linguistic environment where many speakers of English had a good knowledge of both French and Latin and made use of both languages (either actively or passively) at one time or another for various different technical functions. It is hardly surprising that words which were identical or near-identical in form and meaning in both Latin and French should have affected English partly through the one route of transmission and partly through the other. It is likely that we are seeing the results of multiple acts of interlinguistic borrowing, some from French, some from Latin. In some cases even the initial adopter may have had little notion which of the two languages she or he was accessing. What subsequently becomes generalized in the lexis of English is a composite of these various acts of borrowing, open to further ongoing influence (in form or meaning) from either or both of the donors.

[12] In fact Old English *pellican* only occurs in glosses to Latin *pellicanus* in Psalm 101 in the Vulgate, where (as also in many later examples in English) it refers to some sort of bird of uncertain identity found in the wilderness. For a very instructive account of the difficulties encountered by Old English glossators in glossing this line, and of the difficulties posed for modern philologists attempting to interpret this material, see Lass (1997: 83–8).

6.6 Continuing semantic influence and semantic interference

A borrowed word may continue to show semantic influence from its donor for centuries after the date of the original borrowing. For instance:

- English *presence* is a Middle English borrowing from French and Latin, but its use with reference to the external appearance or the impressive or handsome bearing of a person (from 1570) appears to follow slightly earlier use in this sense in French in the sixteenth century.

[13] See Durkin (2002a, 2002b, 2006a, 2008) for more detail on these and similar words.

- *present* in the sense 'to stage or put on (a play)' appears to follow French use in the sixteenth century, although the word was borrowed in a number of other senses over the preceding three centuries.

We may also find semantic influence from a related word in another language. English *popular* is on morphological grounds clearly to be identified as a borrowing of Latin *populāris* 'of or belonging to the people as a whole, belonging to or used by ordinary people, available to the whole community, of the common people, supporting or professing to support the interests of the common people, liked or admired by many people', and it largely corresponds to the Latin word in its earliest meanings in English. However, it also seems to show semantic influence from the morphologically distinct but transparently related French word *populaire* 'of, relating to, or consisting of ordinary people, current among the general public, seeking the favour of the populace, known and liked among the people, vulgar, coarse, democratic, (of a disease) epidemic'.[14]

If we were to classify these cases in terms of our typology in section 5.1, we might postulate a particular kind of semantic loan, where the foreign-language model happens to be either identical to the original donor form or related to it. We cannot assume that all of the senses shown by a borrowed word and shared with its donor were borrowed at the time of the original borrowing.

It is also interesting to consider such cases in the light of the phenomenon of code-alternation identified in research on contact linguistics, where interference from a source language on a target language has been identified as a result of speakers alternating between the use of one language and another. Thomason (2001: 138–9) summarizes several such instances, including that of a native speaker of Italian who spent the majority of her education in the United States and found subsequently that there were interference features from English in her use of Italian, such as using Italian *libreria* 'bookshop' in the sense 'library' (Italian *biblioteca*) as a result of association with English *library*. In the case of *library* and *libreria* there are no senses actually in common, just equivalent word forms with meanings which both have to do with books. We can see how much more readily interference phenomena may come into play with a word like *popular*, where English already shared many senses with the French word.

[14] For fuller discussion of all of these examples see Durkin (2008).

6.7 Multiple borrowings from the same source

Sometimes we find what are commonly referred to as etymological doublets, where two borrowings occur which are ultimately from the same source. For instance, the English verb *affect* shows a late Middle English borrowing from French *affecter*, but earlier in Middle English we find (now obsolete) *afaite* < French *afaiter*. Both French verbs, *affecter* and *afaiter*, have the same etymon, Latin *affectāre*: *afaiter* shows regular development of the Latin word, while *affecter* shows a later learned borrowing from Latin. In English we see a similar phenomenon with the verbs *provide* and *purvey*: the first of these is borrowed directly from Latin *prōvidēre*, while *purvey* is from French *purveier*, which shows the French development of the Latin word. In cases like these the distinct forms of the donor words lead in the borrowing language to formally distinct words which show partial synonymy. We also find cases where the identical donor form is borrowed in two different periods, giving rise to two distinct words in the borrowing language. For instance, French *artiste* is borrowed into English in the sixteenth century, giving rise to modern English *artist*. However, in the late eighteenth century the same French word was borrowed again, this time giving English *artiste*, a word form distinct from *artist* in both spelling and pronunciation, and at least sometimes employed with a semantic distinction: an *artist* typically being someone practising the fine arts, especially painting, whereas an *artiste* is typically a performing artist.

The occurrence of such repeated borrowings from the same ultimate source should alert us to the likelihood of multiple inputs also having existed in cases like those discussed in sections 6.5 and 6.6, where there is no significant change in word form in the donor language(s).[15]

6.8 How can we tell that borrowing has occurred?

If we are to be sure that a borrowing has occurred, ideally we will find an exact correspondence in word form, meaning, and date. Our supposed donor form will precisely explain the word form of the supposed borrowing, it will be recorded in a meaning or meanings which give

[15] For further examples compare: *attack* and *attach*, *cadence* and *chance*, *marchpane* and *marzipan*, *master* and *magister*, *maugre* and *malgré*, *minion* and *mignon*, *peasant* and *paysan*, *ransom* and *redemption*.

an adequate explanation of the meaning or meanings of the supposed borrowing, and it will be attested at an earlier date. Obviously, if our supposed donor also has a well-established further etymology, that will help eliminate the possibility that the relationships between the donor and borrower might in fact have been the other way around. Similarly, if one language shows only a figurative or narrowed sense development from what is demonstrably a more basic sense in another language, as in the case of English borrowing of *friar* in chapter 1, we may feel confident about the direction of borrowing, although many cases are much less clear-cut than this one, and it is often advisable to take a cautious approach to assumptions about the direction of semantic change.

However, very often we will be working in situations where we have less data, or where there is more uncertainty of other kinds. For instance, when we are looking at two closely related languages, it can often be hard to tell which language a word may have originated in. Norse borrowings into English present such a problem, and here the method usually applied is to look for either formal or semantic innovations in either language: if a word shows a sound change found in Norse but not in English, it is a near certainty that it shows Norse influence of some sort; if it shows a semantic innovation known to occur in Norse but not known to occur in Old English, it is also probable (although much less certain) that we have a borrowing.[16]

Sometimes the intralinguistic and extralinguistic data simply are not conclusive: *pack* 'bundle, package' is first recorded in English in an occupational surname *Pakbyndere* 'packbinder' at the end of the twelfth century. It first occurs only eight years later in Middle Dutch, a difference so slight as to be entirely trivial in this period. It subsequently occurs in Middle Low German, and thence in a number of other Germanic languages and in several of the Romance languages. The word has no further etymology, and its origin is a mystery. Similar mystery surrounds a number of words in Germanic languages with initial /p/, since it is normally thought that initial /b/, which would give /p/ by Grimm's Law, was either very rare or perhaps did not occur at all in proto-Indo-European.[17] The connection of *pack* with trade doubtless explains its almost simultaneous appearance on both sides of the English Channel. It has normally been assumed that the word was

[16] See Björkman (1900) and Dance (2003) for discussion of the methodology; compare also Lass (1997) 203–5.

[17] Compare also discussion of *plough* in section 8.10. Other notoriously difficult etymologies involving words with initial *p-* in Germanic languages include *park*, *pot*, and *path*.

borrowed into English from Dutch, which is not unreasonable given that there are very few Dutch documentary records earlier than this date, but it is also perhaps possible that Dutch borrowed the word from English, or both languages from a third source.

In some cases it can be very hard to tell whether a borrowing has occurred at all, particularly with words which are not the base word in a morphologically related group. English *ravine* is recorded from the seventeenth century onwards, and is a borrowing from French *ravine*. From the mid nineteenth century we also find in English a verb *ravine* 'to score (earth etc.) with ravines', earliest in 1858. This could easily be a conversion from the English noun *ravine*, and the existence of a verb *raviner* in the same meaning in French could be purely coincidental. Alternatively, particularly since the earliest example of the English verb is in a book about the geology of central France, we might think that a borrowing from French is likely: either the English word is a loanword from French *raviner*, or it is a conversion of the English noun but on the model of the French verb. However, the earlier existence of *ravined* and *ravinement* in English, which could both be explained as formations from the noun *ravine* but which could alternatively be analysed as formations from the verb, might help shift the balance back towards a derivation within English. Certainty is likely to continue to elude us. Such doubtful cases are very common, and can probably be found in any family of words in which the base word has originally been borrowed from a foreign language and there has been a subsequent history of contact with that language.

We may see just how finely balanced decisions can be in this area from an example where new information has led to a change in assumptions about whether a loan has occurred. I will take an example from the new edition of the *OED*. Modern English *ravenous* 'very hungry' shows a broadened use of the (still current) sense '(of an animal) given to seizing other animals as prey'. In early modern English it also shows a sense 'given to plundering', and is part of a small family of words together with the verb *raven* (also in the forms *ravin*, *ravine*, etc.) and the noun *ravin* (also in the forms *raven*, *ravine*, etc.) in similar senses. In the first edition of the *OED* (in a fascicle first issued in 1903) all three were given as borrowings from (Old) French, thus:

ravin, noun < French *ravine* (ultimately < Latin *rapīna* 'rapine')
raven, verb < French *raviner*
ravenous, adjective < French *ravineux*

These etymologies were based largely on information from the main source of information on Old and Middle French then available, Godefroy's *Dictionnaire de l'ancienne langue française* (1880–1902). However, of these three etymologies, only that given for the noun *ravin* seems entirely satisfactory today. So far as the verb *raven* is concerned, an Old French verb *raviner* is indeed recorded in the sense 'to take off by force', but only in the twelfth century, several hundred years before the first appearance of the English word. In later use the French verb has only the meanings 'to stream, rush' and 'to furrow (the earth etc.) with gullies or ravines', ultimately giving rise to the English noun *ravine* 'deep narrow gorge or cleft' encountered in the previous paragraph. The adjective *ravenous* is first recorded in English in the late fourteenth century in the sense '(of an animal) given to seizing other animals as prey'. There is a corresponding Old French adjective *ravineux*, *ravinos*, *rabinos*, recorded from the twelfth to the fourteenth centuries, but with the meanings 'rapid, impetuous' (although the related word *rapineux* is recorded with a meaning much closer to that of the English word). Thus in the case of the verb and the adjective the supposed French etymons do not provide a very good fit with regard to date or meaning, and it is advisable to look for another etymology if possible. One is ready to hand, since both words can be explained as formations within English from the noun *ravin*. Thus instead of the picture given above we now have:

ravin, noun < French *ravine* (ultimately < Latin *rapīna* 'rapine')
raven (also *ravin*, *ravine*, etc.), verb < *ravin* (also *raven*, *ravine*, etc.), noun
ravenous, adjective < *ravin* (also *raven*, *ravine*, etc.), noun + *-ous*

It is important to note that this picture could change again if new information on usage in French comes to light: it is very possible that this might emerge from the new edition of the *Anglo-Norman Dictionary* currently in preparation. Such etymologies are highly dependent on the nature and quality of the available data. The assumptions made in the first edition of the *OED* seemed sensible on the basis of the much sketchier information on Old French and Middle French lexis then available, and it is possible that some new lexicographical finds might come to light in the future that challenge our assumptions once again – although obviously the better and fuller our lexicographical resources become, the less likely it is that we will have quite so many such surprises.

The importance of cultural and historical background was shown earlier by the example of *tomato* in section 5.5. We know that the Nahuatl word was probably borrowed first into Spanish, because Spanish speakers were the first speakers of a European language to be in contact with speakers of Nahuatl. Hence it is very satisfactory to find Spanish is the first European language in which the word is recorded, and that the forms in other European languages also support the hypothesis of transmission originally via Spanish.

6.9 Lexical borrowing and code-switching

In section 2.2.2 we looked at nonce borrowing of foreign-language words, and at how the early history of a borrowing may be like that of *focaccia* in English, showing a number of separate introductions of the word, followed eventually by more general adoption. We have also seen that this sort of more general adoption will not happen for all words, and that most will not progress beyond the stage of nonce borrowing. In section 5.1.5 we looked at the difficulty of trying to distinguish between *Fremdwörter* and *Lehnwörter*. Morphological and phonological adaptation are only tendencies. Phonological adaptation in particular operates on a cline which makes it very difficult to say that adaptation has or has not occurred in a particular instance. If phonological adaptation consists of no more than pronouncing a foreign word with an accent, then this sort of adaptation will surely be shown also by many instances of the nonce use of a foreign word. If an early user of the word *focaccia* in an English sentence pronounces it with final /ə/ rather than /a/, this certainly reflects adaptation to the usual phonological structure of English words, but it does not necessarily show a naturalized borrowing: the speaker may simply be accommodating to the speech of interlocutors (perhaps in order to appear less pretentious), or may have a poor knowledge of Italian phonology.[18] It is perfectly possible for such a pronunciation to occur in a sentence of the type 'In Italian they call this type of bread *focaccia*', where the word is explicitly identified as belonging to Italian rather than English. We may find a similar situation with morphology, e.g. 'In Poland they eat stuffed dumplings resembling ravioli, which they call *pierogis*', where the Polish plural form *pierogi* is suffixed with the English plural ending *-s*, but where the established Italian borrowing *ravioli* retains the Italian plural form. It is likely that there is a

[18] See Haugen (1950) 215–17 for some further discussion and examples.

cline here, from explicit identification of a word as belonging to a foreign language, e.g.:

(i) 'I had some of a type of bread which they call in Italian *focaccia*' (even though the form in such a sentence may show intentional or unintentional phonological or morphological adaptation)

to use of a word with reference to a foreign culture but without explicit identification of its foreign-language identity:

(ii) 'On our holidays we had sandwiches made with focaccia every day' (where italics may well be used in print)

to uses where there is no flagging of foreign status:

(iii) 'I always think that focaccia is the best type of bread to have with salad'

We may decide on a pragmatic basis that the first type of sentence does not show borrowing, even if the quoted foreign-language word may show some phonological or morphological adaptation, whereas the other two types of sentence do show borrowing. Many historical dictionaries adopt this approach. However, any systematic distinction between more and less naturalized borrowings is very difficult to apply with any consistency. Also, if we are working solely from the evidence of historical spelling forms, many of the finer details are likely to be irrecoverable.

A rather difficult complication is found in determining the relationship between lexical borrowing and the phenomenon known as code-switching, where bilingual speakers switch between use of one language and use of another, in the knowledge that they are addressing others who also have some knowledge of each language, and who are hence to at least a very limited extent bilingual. Code-switching may occur at sentence boundaries, i.e. intersententially, or at the level of the word, phrase, or clause, i.e. intrasententially. (In fact, some linguists reserve the term code-switching for switches which occur intersententially, and refer to those which occur intrasententially as code-mixing, but I will not adopt this distinction here.) The relationship between code-switching and lexical borrowing is much disputed, but it seems a reasonable assumption that code-switching within a bilingual community at least sometimes results in lexical borrowing.[19]

For instance, a word from language A may be used within sentences of language B by bilingual speakers as a code-switch. This may happen for any of various reasons, such as cultural associations, group identity, or because it expresses a concept not so easily expressed within the existing resources of language B. Subsequently this word may become adopted into the wider speech community of monolingual speakers of language B.

Such issues can often lead to practical difficulties in assessing written evidence from the past. As we noted in section 5.6, in the multilingual culture of later medieval England, English was the language of everyday life for all or nearly all speakers, but Latin and Anglo-French had important roles in many areas of professional and cultural life. In this context we often encounter situations where it is very difficult to tell whether a word ultimately of foreign origin occurring in a given document shows a borrowing or a code-switch. Many documents have a basic grammatical framework which is Latin, or at least they show Latin grammatical endings and concord, but they also contain many words of vernacular (English or Anglo-French) origin. Some of these vernacular words show Latin grammatical endings, and hence could be regarded as loanwords into Latin. Others do not, and hence we could argue that they show code-switches, from Latin to one of the vernacular languages. In other documents the basic framing language may be English or Anglo-French, but we find many words which belonged originally to the other vernacular language. In this situation the general similarity of inflectional endings, especially in nouns, often makes it impossible to identify words as belonging on morphological grounds to either Middle English or Anglo-French. In such a case, do we have a text which shows an extremely high incidence of lexical borrowing, or do we have a text which shows a great deal of code-switching? Some of the implications of this sort of text (of which my sketch represents only the crudest summary) are picked out by Hunt, a scholar who has done a great deal to draw attention to material of this kind:[20]

> If language acquisition ... takes place in a context of competing codes in a multilingual situation in which individuals accommodate their linguistic behaviour to that of groups with which they wish to be identified, or contrariwise, then the situation

[19] For a useful overview and references to the relevant literature see Thomason (2001) 131–6, and compare also Mahootian (2006). For accounts broadly in favour of the proposition that single-word code-switching often leads to lexical borrowing see Myers-Scotton (2002) or Thomason (2003). For the alternative view that most singly occurring foreign-language words should be regarded not as code-switches but as borrowings, see Poplack, Sankoff, and Miller (1988), Poplack and Meechan (1998), Poplack (2004).

[20] For a summary of other recent work in this area see Pahta and Nurmi (2006).

is inevitably one of great complexity, which will not surprise anyone familiar with contact linguistics and the world of bi-dialectalism, bilingualism, diglossia, borrowing, transfers, interference, shift, relexicalisation, pidginisation, and creolisation. Linguists have frequently sought to identify borrowings in the languages of medieval Britain, but in the context of multilingual societies it can be unrealistic to attempt to distinguish code-switching from borrowing.

(Hunt (2000) 131)

Such issues as these pose problems for anyone who wishes to establish what the earliest example is of a particular borrowing. Historical dictionaries tend generally to take a pragmatic approach: if a word subsequently becomes frequent enough to merit inclusion, early examples found within the context of an utterance in the borrowing language will generally be taken as showing the borrowing, rather than an independent code-switch. More problematic can be cases where the word is found in a foreign-language or multilingual context. For instance, *plane* meaning either a bricklayer's or a carpenter's tool, is a Middle English borrowing from (Anglo-)French. In the evidence presented for this word in the new edition of the *OED*, a first example in an English context is found from a little before 1425 in a Latin–English glossary:

*a*1425 *Medulla Gram.* (Stonyhurst) f. 37v, Leuiga: a leuor or a plane.

An example from 1404 in a Latin context with the English plural morpheme *-ys* probably shows an earlier example of the English borrowing (on the grounds that *-ys* is not normally found as a plural morpheme in Anglo-French):

1404 in J. T. Fowler *Extracts Acct. Rolls Abbey of Durham* (1899) II. 397 In custodia Plumbarii, 2 planys.

Earlier still a vernacular word is found in other Latin documents, but in these cases there is nothing to tell us whether it is the English word or its French etymon:

1350 in J. T. Fowler *Extracts Acct. Rolls Abbey of Durham* (1899) II. 550 Uno Plane et aliis instrumentis pro officio plumbar', emptis, 2 s. 5 d. 1399 in J. Raine *Fabric Rolls York Minster* (1859) 18 Instrumenta carpentariorum... Item, j plane de ere.

Here the approach of different dictionaries in fact differs: the *OED* places these ambiguous examples in its etymology section, but the *Middle English Dictionary* places them in its main documentation for the Middle English word, without comment, in line with its general policy in dealing with

vernacular words occurring in Latin documents. On a purely practical level, examples such as this remind us of the importance of looking closely at what the earliest examples presented in a dictionary are, and they should be a salutary reminder to any etymologist that dates alone, unsupported by a reference, are often not very helpful.[21]

Foreign-language phrases can also be ambiguous as to whether they are lexical borrowings or code-switches. They may show grammatical characteristics of the donor or source language which are alien to the borrowing or receiving language, such as agreement features or word order, although such features are not always reproduced faithfully, in which case we can probably safely assume that borrowing has occurred. This is typical of borrowing of French food terms in English. In French *moules marinières* the adjective *marinier* 'marine' is postposed after *moule* 'mussel', and shows feminine plural concord. The *OED*'s evidence suggests that the written form *moules marinière* is more common in modern use in English than *moules marinières*; this probably results simply from the final *-s* being silent, but it could also reflect influence from either French *moules à la marinière* 'mussels in the marine style' or the rare French blended form *moules marinière*. Substitution of one or more native elements is often found, as in *beef bourguignon* beside earlier *boeuf bourguignon*: we could say that the borrowed phrase has been remodelled with substitution of *beef* for French *boeuf*, or we could say that beside the original borrowing we have a subsequent loan blend. Extensive naturalization in form is often found in borrowed phrases, such as English *billy doo* 'love letter' < French *billet doux*. Such naturalization shows that we have a borrowing rather than a code-switch, but it does not tell us whether this was via an initial code-switch. Interestingly, *OED*'s first quotation for *billet-doux* suggests at the very least an imitation of code-switching: 'He sings and dances en François, and writes the billets doux to a miracle' (1673: John Dryden *Marriage à la Mode* II. i. 261).

6.10 Some conclusions from chapters 5 and 6

Some generalizations about etymologies which involve borrowing can be drawn from the topics we have examined in the last two chapters:

- Borrowed words are often subject to processes of accommodation to the phonology or morphology of the borrowing language, either at the

[21] On the source of the dates used in this book see chapter 1, footnote 6.

time of borrowing or subsequently. They may also form compounds or derivatives in the borrowing language.
- Not all components of the meaning of a word need be borrowed.
- Borrowed words are subject to change (semantically, phonologically, or morphologically), just like any other words.
- Borrowing between languages (interlinguistic borrowing) is not necessarily a once-and-for-all process, just as borrowing within languages (intralinguistic borrowing) is not either.
- Initial interlinguistic borrowing is typically followed by intralinguistic borrowing, as a word spreads to different registers or varieties of a language and to the usage of different speech communities.
- After the date of initial borrowing, borrowed items frequently show further influence from the donor language, through the borrowing of additional senses, or through formal remodelling after the donor form. Loanwords may in time become either less like the corresponding form in the donor language (through internal processes of change in either the borrowing language or the donor language), or more like it (through remodelling of a previously naturalized form after the form in the donor language).
- A good etymology which involves borrowing will have a working hypothesis as to how and why (as well as when and where) borrowing occurred, and also as to how and why the borrowed word, sense, etc. has subsequently spread within the borrowing language.

Of course, we will not always have sufficient data to address all of these issues, but we should not assume, simply because we have only limited data available to us, that the reality is likely to have been any less complex than in instances where we do have abundant data.

7

Change in word form

7.1 Two Germanic sound changes 180
7.2 Examples of English sound changes 182
7.3 Sporadic sound changes 195
7.4 Associative change in word form 197
7.5 Metanalysis 207
7.6 How regular are regular sound changes? 208
7.7 Examples of arguments based on word form 211

Nearly all etymological research involves analysis of change in word form. This analysis also raises some of the most theoretically complex areas which we will encounter in this book. In the first part of this chapter we will examine examples of:

- regular sound changes, both isolative ones, occurring irrespective of any particular phonetic environment, and conditioned ones, triggered by a particular conditioning environment
- sporadic sound changes
- analogy, folk etymology, and other types of associative change in word form

After this, we will turn to one of the key issues in historical linguistics, namely just how much regularity is in fact shown by what are called regular sound changes, and what the implications are for various different types of etymological research.