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

First published 2009

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

5

Lexical borrowing

5.1 Basic concepts and terminology 132

5.2 What constitutes a *borrowing*? 140

5.3 Motivation for borrowing: traditional explanations 142

5.4 Examples of borrowing because of 'need' 143

5.5 Borrowing of a new word for a new product 145

5.6 Patterns of borrowing in the history of a language 149

5.1 Basic concepts and terminology

Borrowing is the usual term for the process by which a language (or variety) takes new linguistic material from another language (or variety), usually called the donor. In keeping with the focus of this book, I will look mostly at borrowed words, but it is important to note that other units such as morphemes or phonemes, or even syntactic features, may also be borrowed. Borrowing occurs in situations of language contact, and is indeed an almost inevitable consequence of it, although the levels and the types of borrowing which are found differ greatly in different types of contact situation.

The term *borrowing* is conventional and is in almost universal use, but it is no new observation that the metaphor of 'borrowing' is not entirely apposite. The relevant item is not taken away from the 'donor' language as a result of the 'borrowing'; rather, it spreads from one language to another, with the result that it is subsequently found in both. Furthermore, there is no assumption that anything will be 'given back' to the 'donor', precisely because nothing has been given away in the first place. Crucially, a word which has been borrowed will very likely change and develop in different

ways in the donor language and in the borrowing language, or it may very possibly die out in either or both.

Perhaps the easiest way to illustrate this is by looking at some of the relatively rare instances when a word or phrase which has been borrowed from one language into another is either borrowed back subsequently into the original donor language, or affects the meaning of the donor:

- French *prêt-à-porter* 'designer clothing sold ready to wear' (1951) is formed < *prêt* 'ready' + *à* 'for, to' + *porter* 'to carry, to wear' on the model of (or as a loan translation of) English *ready-to-wear*. The French expression is then borrowed into English (in unnaturalized or semi-naturalized form) as *prêt-à-porter* (in 1957), hence as a synonym or near-synonym of the existing expression *ready-to-wear*, but with the advantage of the perceived prestige of French terminology in the world of fashion.
- English *milord* 'an English nobleman in Europe, an Englishman travelling in Europe in aristocratic style' (1607) is a borrowing < French *milord* (earlier *milourt*), but this is itself a borrowing (with conversion from form of address to noun, and narrowing of meaning) < English *my lord*. In this instance, English speakers adopting French *milord* in its restricted sense were probably aware of its origin, and were making something of an ironic joke.
- English *panchway* 'light rowing boat used on rivers in Bengal' (1737) is < Bengali *pānsui*, variant of *pānsi* 'pinnace' < English *pinnace* + Bengali *-i*, suffix forming adjectives. (The form of the word in Bengali is perhaps influenced by various words in local vernaculars with meanings connected with water and sailing which have initial *pān-*.)
- English *mama-san* (1904) is < Japanese *mama-san* 'honoured mother, madam, proprietress, manageress of a bar, etc.', which is itself < *mama* 'mother' (an early-twentieth-century borrowing < English *mama*) + *-san*, an honorific title. (We might perhaps wonder how far such a word, which is used only when referring to Japanese cultural contexts, can be said to have been borrowed into English at all: this is a topic which we will turn to in detail in chapter 6.)
- *phase* was borrowed into English from French in the seventeenth century. In English, it developed specific senses in the fields of physics and chemistry, which were then borrowed back into French as semantic loans. In other words, the French word retained the same word

form, including its distinct pronunciation, but adopted new senses from English.

- English *pioneer* is < Middle French *pionnier* (French *pionnier*) 'labourer employed in digging' (*a*1230; earlier in senses 'foot soldier', 'pedestrian'), 'soldier employed to dig trenches and mines' (*c*1380); but subsequently in the nineteenth century the French word shows the senses 'an early colonist', 'an innovator' as semantic loans from English.
- English *plumber* is < Anglo-French *plummer*, *plomner* and Middle French *plommier*, *plombier*. In the 1970s, English *plumber* comes to have a specific metaphorical meaning (originally in the context of the Watergate scandal) 'a person employed to investigate or prevent "leaks" of information from a government office, department, etc.', and (in spite of the divergence in word form) this specific sense is borrowed by French *plombier* as early as 1973.

In these examples we have already seen several different types of lexical borrowing, and it is clear that we need some sort of typology and terminology to distinguish between them. A typology which is often employed makes the following main divisions:[1]

- Loanwords
- Loan translations
- Semantic loans
- Loan blends

5.1.1 Loanwords

Loanwords show borrowing of a word form and its associated word meaning, or a component of its meaning. Usually there is some degree of accommodation to the sound system of the borrowing language, e.g. English *phase* /feɪz/ (or when borrowed in the late seventeenth century /fe:z/ or /fɛ:z/) < French *phase* /faz/. Loanwords may show adaptation to the inflectional morphology of the borrowing language; for instance, many nouns borrowed into English show a regular plural in *-s* or *-es* in place of whatever plural morpheme is found in the donor language. However, many scholars draw a distinction between loanwords and words which show complete

[1] For an important, and more detailed, analysis see Haugen (1950). See also Fischer (2003) for an overview of different approaches.

replacement of a morph in the stem of the borrowed word with a morph from the borrowing language: for such cases see below on loan blends.

5.1.2 Loan translations

Loan translations (or calques) show replication of the structure of a foreign-language word or expression by use of synonymous word forms in the borrowing language, e.g. French *prêt-à-porter* is a calque on English *ready-to-wear*. We might be tempted to define loan translation as the use of 'the corresponding word forms' in the borrowing language, but this begs many questions, as there is seldom a precise one-to-one correspondence between any part of the lexicon of two languages. Even in the example of *prêt-à-porter* it is not completely certain that the French expression is modelled on English *ready-to-wear* rather than being a less exact loan translation of the synonymous *ready-for-wear*.

We cannot always be sure whether a particular formation is a loan translation, or simply a coincidental parallel in another language. For instance, English *Middle Europe* 'a loosely defined region of central northern Europe, extending roughly from Germany in the west to Poland and Hungary in the east' is probably formed on the model of German *Mitteleuropa*. The German term is recorded earlier with the same meaning, and in the culturally dominant language in the relevant geographical area, but we lack any evidence to prove that a loan translation has occurred. Clearer cases occur when we encounter a highly lexicalized (possibly encyclopedic) meaning which is very unlikely to be coincidental, such as English *New Christian* '(in medieval and early modern Spain) a Christianized Jew or (less frequently) Moor, especially one who converted only nominally in order to escape persecution or expulsion'. This is clearly modelled on Spanish *cristiano nuevo* in the same meaning and attested considerably earlier. Sometimes the historical record indicates the existence of a parallel in another language which is unlikely, on grounds of semantic probability, to be the result of coincidence, but one may have no clear way of telling which direction the influence has taken. However, if both languages have a well-documented historical record for the period in question, then dates of first attestation alone may sometimes be sufficient to create reasonable certainty about the direction of borrowing. For instance, *Nile green* 'a pale bluish green colour supposedly resembling that of the Nile' probably shows a calque on French

vert du Nile. There is nothing in the contexts of the earliest examples in English or French which would preclude the reverse being the case, but the relative dates of first attestation, 1871 in English, 1830 in French, are probably enough for reasonable certainty that the French usage came first. In such cases, exhaustive searching in documentary sources might provide convincing support for a French origin of the term (or the reverse), but any definite proof on purely linguistic grounds is impossible.

5.1.3 Semantic loans

These show extension of the meaning of a word as a result of association with the meaning of a partly synonymous word in another language. The two words may be ultimately related, as in the cases above of French *phase* and English *phase* or French *plombier* and English *plumber*. They may have a formal resemblance to one another, but in fact not be related at all historically: for instance English *manage* and *management* were influenced semantically by French (unrelated but similar-sounding) *ménager* and *ménagement*. In other cases, the words involved may be unrelated and also bear no significant formal resemblance to one another: for instance, English *manner* shows considerable semantic influence from both Latin *modus* and Latin *mōs*; it occurs as a conventional translation equivalent of both of these from an early date. Similarly, classical Latin *ratiō* probably meant originally 'count, account', but acquired numerous other senses (such as 'reason') by association with ancient Greek *lógos*, which also had the meanings 'count, account'.

As with loan translations, it can often be difficult to differentiate cases of semantic borrowing from coincidental semantic development in two languages.[2] An additional concern is that it may sometimes be hard to tell apart (a) cases where the meaning of a word has been influenced directly by association with the range of meanings of a foreign-language word with which it shows some semantic overlap, and (b) cases of (not specifically linguistic) cultural influence in the development of concepts. For instance, the meanings of words denoting such concepts as 'god', 'heaven', 'hell' in English and other Germanic languages are profoundly influenced by contact both with Christian culture and with the paganism of Roman and

[2] For discussion and exemplification of this issue see Hoad (1993).

Greek antiquity, but it is often difficult to tell whether this shows a linguistic process of influence of Latin and Greek words on the meaning development of partial synonyms in English, or whether the influence is an extralinguistic one on the development of the concepts which these words denote. The development of the meaning of the word *hell* in Old English and Middle English was greatly influenced by both Christian and pagan Roman and Greek conceptions of the afterlife, but in lexical terms English *hell* corresponds to at least two different, semantically non-overlapping, groups of words in Latin, Greek, and Hebrew: in Latin, *Orcus*, *īnferī*, and *īnferna* all denote the abode of the dead (or *Hades* in the Greek tradition, in biblical use corresponding to Hebrew *šĕ'ōl*, literally 'grave'), while hell as a place or condition of punishment for sin is denoted by *gehenna* (a word ultimately of Hebrew origin via Greek). Is the influence that we have in this instance primarily lexical, or cultural? And how viable is it to make a distinction between the two?

In some cases we may wonder whether there is any continuity at all with the existing word, or whether we do not in fact have an independent borrowing, hence a new loanword, which happens to be homonymous with an existing word. English *milord* presents just such a difficult case: is it a borrowing of French *milord* (with a naturalized pronunciation, with final /d/ based either on the spelling or on association with *lord*), or does it show the existing reduced form of the form of address *my lord* (as in *You rang, milord?*), in (semantically narrowed) use as a noun on the model of the French word?

In cases of semantic loan, and perhaps also in cases of loan translation, we may prefer to say that we do not have borrowings at all but (in the case of semantic loan) semantic change or (in the case of loan translation) new words or phrases occurring as a result of influence from another language. We might indeed choose to explain the process in terms of analogy (see section 7.4), and say that what all three categories have in common is that they show the influence of one language on the lexis of another.

5.1.4 Loan blends

The three categories already described provide a useful framework for considering different types of lexical borrowing, but, as already noted, the dividing line between them is often unclear. We may have difficulty in assigning a particular example to a particular category: some examples

may seem to sit rather awkwardly between categories. Many scholars in fact identify an intermediate category between loanwords and loan translations: loan blends. These show borrowing of a complex word with substitution of one or more native morphs for morphs in the borrowed word. English *neurotize* 'to provide with new nerve fibres or nerves' shows a borrowing of French *neurotiser* with substitution of English *-ize* or *-ise* for (ultimately related) French *-iser*. I have chosen a very rare word, because it has the advantage of being a very clear example of this phenomenon: *neurotiser* is a coinage by the French scientist Vanlair from 1882; its *-t-* is unexplained (it is perhaps after French *névrotique* 'neurotic') and is carried over faithfully into the English word. Another fairly clear case is shown by the example of *pioneer* which we encountered in section 5.1. The recorded French forms all show the ending *-ier*, and the English forms all show morphological substitution of either *-eer* or (in early forms such as *pioner*) *-er*. The French word is a derivative of *pion* 'foot soldier', which did not exist in English in this form at the date when *pioneer* was borrowed. (*pawn* does show borrowing of a variant of the same French word, and *pion* was itself later borrowed into English in some specialist uses.)

In very many other possible instances of loan blends (in English, certainly the majority) there are other available analyses. For instance, *martyrize*, *moralize*, *naturalize*, *neutralize*, *organize* all certainly show at least some degree of French influence. However, we cannot confidently eliminate the possibility that they may not be loan blends but loan translations, from a previously borrowed or otherwise related root word with the English suffix *-ize* added, on the model of the French word. Thus we could analyse *martyrize* as showing borrowing of French *martyriser* with remodelling of the ending after *-ize*, or we could analyse it as a formation < *martyr* + *-ize* on the model of French *martyriser*. In particular instances etymological dictionaries may make decisions in favour of, or have a policy of opting for, one possibility or the other, but this is a different matter from demonstrating without doubt that a particular word shows either a loan blend or a loan translation.

To take a couple of further examples, English *nosology* 'treatise dealing with diseases, classification of diseases' (1721) could readily be interpreted either as a formation from the neo-classical combining forms *noso-* and *-logy* (both of which are productive in English at this date) on the model of Latin *nosologia* (i.e. as a loan translation), or as a borrowing of the Latin word with substitution of *-y* for final *-ia* (i.e. as a loan blend). Similarly,

South African English *moderature* 'the executive council of a synod of the Dutch Reformed Church in South Africa' clearly shows some sort of influence from Afrikaans *moderatuur* in the same sense, but it is less clear whether the English word shows a loan blend with substitution of *-ure* for Afrikaans *-uur*, or a formation < *moderate* (verb) + *-ure*, on the model of the Afrikaans word (i.e. a loan translation).

5.1.5 Lehnwörter and Fremdwörter

An important and influential tradition in linguistics in the German-speaking world makes a further distinction within the category of loanword, distinguishing *Lehnwörter* 'loanwords' from *Fremdwörter* 'foreign words'. In this tradition, a *Lehnwort* shows accommodation (where appropriate) to native phonology and morphology and may give rise to new derivatives within the borrowing language, while a *Fremdwort* retains (broadly) its foreign-language pronunciation and may show non-native morphology (especially plural inflections which are not found in native words), and does not give rise to new derivatives within the borrowing language. This distinction has been very influential in many aspects of linguistic work in the German-speaking world, including lexicography: all but the most clearly assimilated and frequently used loanwords are often excluded from historical or etymological dictionaries of German, and find their place instead in separate dictionaries of *Fremdwörter*. However, in practice the distinction is hard to maintain consistently. Where different variants of a particular borrowed word show differing degrees of naturalization in pronunciation, or where the plural morphology shown by a word differs between naturalized and non-naturalized patterns, the distinction between *Lehnwörter* and *Fremdwörter* cannot easily be used as a criterion for determining how words will be treated lexicographically. For instance, in English the plural of *appendix* is sometimes *appendixes*, following the usual pattern of English plurals, and sometimes *appendices*, as in Latin (although with different pronunciation from in Latin). It would be very difficult to distinguish on that basis between a Lehnwort *appendix* with a plural *appendixes* and a Fremdwort *appendix* with a plural *appendices*, and if *appendix* was encountered in the singular, how could one tell which of the two it was? The distinction between *Lehnwörter* and *Fremdwörter* will not be used in this book, although it is interesting to note that it has some points of connection

with current debates about code-switching which we will consider at the end of chapter 6.[3]

5.2 What constitutes a borrowing from language X into language Y?

In this book whenever I say that a word (or phrase etc.) was borrowed from one language to another, this means that, so far as we are able to ascertain, the borrowing was direct, unless it is specified that it was via the intermediary of another language. In the latter case, strictly speaking we have two separate acts of borrowing, from the first language into the intermediary, and thence into the destination language. However, if there has been no change in word form or meaning in the intermediary language, it may be difficult to demonstrate that this intermediate stage has actually occurred. Furthermore, we may suspect that perhaps the borrowing has been partly via an intermediary and partly direct from the original language (see section 6.5).

Some studies attempt to identify that component in the lexis of a language which shows a distinctive trace of origin in a certain other language. Terms ending in *-ism* such as *Anglicism* and *Gallicism* are frequently used to denote such lexis. The large collaborative project headed by Manfred Görlach which gave rise to the *Dictionary of European Anglicisms* and the accompanying set of studies *English in Europe* (Görlach 2001, 2002a) is a good illustration of this approach. Görlach and his collaborators looked at words ultimately of English origin in sixteen different contemporary European languages. They did not pay particular attention to the immediate mode of transmission. In many cases suitable information would anyway not have been available for them to identify this, at least not for a large and consistent wordlist across a wide range of languages. All words which were 'Latinate or neo-Greek' in composition were omitted from the study, unless 'an English pronunciation was attested in at least one language, making the word an Anglicism and forcing its inclusion' (Görlach (2001) xix). Additionally, 'words not known to the general educated reader' were omitted. (We will return to this problematic area at several points in this

[3] Chambers and Wilkie (1970: 70–1) make the point that these terms *Lehnwort* and *Fremdwort* are to a large extent artificial constructs of the debate about borrowing in German cultural history, which we will touch on again in sections 5.4 and 5.6.

and the following chapter.) Words formed from names were also omitted on the grounds that 'the process by which names become words is very different in individual languages and it was impossible to make clear-cut decisions'. The resulting study offers the reader an interesting perspective on the relative spread across a range of European languages of words found in general use which an Anglophone might spot as being ultimately of English origin. What it does not do, and would not purport to do, is to give an accurate impression of the extent of borrowing directly from English into each of these sixteen languages.[4] The two research questions concerned are different, and each demands a different approach.

To take another perspective, Dance (2003) makes a detailed study of lexis of Norse origin occurring in early Middle English texts from the south-west Midlands. This is an area which saw little or no Scandinavian settlement, so most lexis ultimately of Norse origin in texts from this region is likely to be the result of at least two stages of borrowing: initial borrowing from Norse into English in areas where speakers of the two languages were in direct contact, and subsequent internal borrowing into the dialect of the south-west Midlands. This is a point that Dance is careful to observe, describing such lexis as 'Norse-derived'.[5] (There is a further terminological problem whenever we speak about borrowing into English from 'Norse', since the forms commonly cited as 'Old Norse' are in fact predominantly Old Icelandic, because that is the earliest Scandinavian variety to have extensive written records, but these records are later than the period of greatest influence of Scandinavian languages on English, and also show significant dialectal differences from the varieties which were in contact with English. For a useful recent discussion of some of the main issues see Coates (2006a).)

Thus it is crucially important in etymologies involving borrowing to be clear what sort of event we think we are describing, or more often, what range of possible events we think our etymology might describe. The fact that a word ultimately appears to originate in a particular language need not mean that it was borrowed immediately from that language. Similarly, borrowing is not a simple, once-and-for-all process. We will look in detail at some more complex cases later on, but it is as well to be aware that

[4] An interesting methodological comparison is provided by Brown (1999), an investigation of the names found for European cultural importations in native American languages, in a context where there are very few early linguistic records on which to base an analysis.

[5] We will return to this study in section 6.3.

any model is naive which assumes that we can pick out a point at which borrowing into 'a language' occurs: the process of adoption and spread shown by borrowings tends to be just as gradual and incremental as that shown by any other new lexis.

5.3 Motivation for borrowing: traditional explanations

The commonest motivations for lexical borrowing have traditionally been identified as need and prestige. Typically, borrowing because of need is said to occur when a new thing or concept is encountered which already has a name in the donor language but not in the borrowing language, or at least not one known to the borrower. Borrowing because of prestige is sometimes said to occur when a speaker perceives that there is greater social cachet attached to a word from another language. (In the previous sentence *cachet* is a good example of a prestige borrowing from French; it is a near-synonym of the earlier French loan *prestige*.) Another way of putting the same distinction would be to say that borrowing for need is necessary borrowing, because there is a lexical gap, and borrowing for prestige is unnecessary borrowing, because an adequate means of expressing the same concept already exists.[6] Unnecessary borrowing is often an important source of stylistic variants in a language.

There are some difficulties with both concepts. Need is probably the less problematic of the two. As we will see, newly imported traded items, newly encountered products or features of the natural world, new scientific discoveries or intellectual concepts, will all have an effect on the lexis of a language: put simply, they all require names. When a new thing is first encountered through the agency of speakers of another language, or in or near an area in which they live, they will very likely already have a name for it, and this name is likely to have an influence on the name adopted in the language of the people encountering this thing for the first time. The foreign-language name is likely either to be borrowed as a loanword, or to form the basis for a loan translation or a semantic loan. (However, as we will see in the next section, this is far from inevitable.)

Borrowing for prestige is a more difficult concept, and can sometimes lead to oversimplification of complex sociolinguistic situations. Typically

[6] See for example Mahootian (2006) 513. For some further perspectives on this issue see also Ross and Durie (1996) 21.

the term is used to describe borrowings which occur in a context where the donor language has a particular status in any of various social or cultural situations: for instance, as a language of learning or science, as the language of a politically or socially dominant class, or as the language associated with a particular social activity. In some cases, dominance of one language in a particular function, field of discourse, etc. may seem a more apposite conception than prestige. Additionally, it is often necessary to distinguish between the processes responsible for the initial occurrence of a word in utterances in another language, and its subsequent adoption by increasing numbers of speakers and in an increasing range of contexts. (Compare sections 6.3, 6.4.)

5.4 Examples of borrowing because of 'need'

A frequent type of borrowing for 'need' occurs in the language of science when a new entity, process, concept, etc. is named in one language and that name is transferred to other languages. In section 5.1.4 we encountered *neurotize* and *nosology*, which either entered English from or were modelled on words in French and Latin, although in both cases the elements from which the words are formed are ultimately of Greek origin. Scientific naming of new entities and concepts normally remains restricted to technical registers, and it is common to speak of such vocabulary as belonging to an international 'language' of science. Within this scientific register, the boundaries between individual languages as regards lexis may be particularly fluid, and the composition of new words is often transparent as the result of the use of a shared set of word-forming elements which are for the most part ultimately of Latin and Greek origin (see section 4.3.1). Some languages may show slightly more resistance to the adoption of such vocabulary, or may have done so in earlier historical periods: compare for instance *oxygen*, French *oxygène*, etc. (from elements ultimately of Greek origin) with the loan translations German *Sauerstoff* and Dutch *zuurstof* (the names all ultimately reflect Lavoisier's conception of the nature of the substance). However, in a relatively small alphabetical sample of English words we can find *ommatin*, *ommatophore*, *ommin*, *ommochrome*, *omphacite*, *oncosine*, *onofrite*, *onomasiology*, *onomatopoesis*, *ontogenesis*, *ooblast*, *ooid*, *ombrophilous*, *ombrophily*, all of which are either borrowed from or modelled on German words (*Ommatin*, *Ommatophor*, *Ommin*,

Ommochrom, etc.), which in turn were formed from elements ultimately of Latin or Greek origin which are common to the technical vocabularies of many modern languages.[7] Such items would almost certainly be omitted from a study of 'Germanisms' in English of the sort we looked at in section 5.2, but they all first appear as German words in German sentences, and have entered English from German.

Occasionally a newly named scientific category comes to be part of a fairly basic level of vocabulary. The word *petal* is first recorded in English in 1712, denoting what the *OED* defines as 'each of the modified leaves, typically distinctly coloured, which form the segments of the corolla of a flower'. Its prior history can be summarized as follows:

Greek *pétalon* 'leaf'
> post-classical Latin *petalum* (mid 17th cent.), in various technical senses (alongside the usual Latin word for a leaf, *folium*)
> English *petalum* (1687), *petal* (1712)

Before the word *petal* was borrowed into English, petals were not distinguished from other kinds of leaf by any special name. Even the specifying compound *flower-leaf* is only recorded from the early eighteenth century, although it is evident that some particular collocations with *leaf* referred conventionally to the petals of particular plants: e.g. *rose leaf* (first recorded in the Middle English period) refers most frequently to the brightly-coloured leaves of the rose's flowers and not to the green waxy leaves of its stems. Today *petal* seems to be a name for an obvious category in the natural world, and few children will have difficulty in identifying the petals of at least those plants, such as a rose or a daisy, which have brightly-coloured flowers with well-defined individual petals, even though in some cases what is identified by the child or layman as a *petal* will be differently classified by a botanist. (In section 8.2 we will look at prototype semantics, a framework which explains this sort of situation very well.) Yet both the word and (it seems) the concept were borrowed into English from Latin as used by early scientists, and then within English we might say that a further borrowing occurred, from the language of science to the more general language. From the perspective of the meaning relations found in modern English, we might say that this borrowing of the word *petal* helped

[7] See Durkin (2006a) for full details of this sample, and for discussion of why an English historical dictionary can sometimes be the easiest place to find information about such words even when they do not originate in English.

fill a lexical gap, in providing a word for this distinctive part of a flower, but we might also note that speakers up to this point seem not to have perceived this as an important distinctive category. Hence it might be better to say that the borrowing led to the creation of a new semantic category rather than filling a gap in the lexicon.

5.5 Borrowing of a new word when a new product of the natural world is encountered

The word *tomato* shows a relatively simple case of borrowing a new word to denote a newly encountered thing. The Spanish word *tomate* is first recorded in 1532, soon after the Spanish conquest of the Aztec empire. The word is a borrowing of Nahuatl *tomatl* denoting the same plant. Nahuatl was the language of the Aztecs, as well as other peoples of the region, and is still spoken in parts of Mexico today; the word *tomatl* may ultimately be a derivative from *tomau* 'to grow'. The English word *tomato* first appears in the form *tomate* in 1604, and is a borrowing from Spanish *tomate*. Similarly, French *tomate* occurs in the late sixteenth century in an isolated early example in a translation from Spanish, although it does not become frequent until the eighteenth century; German *Tomate* is first recorded in the seventeenth century, and Portuguese *tomate* in the early eighteenth century. We so far have a very simple picture: Spanish has borrowed the word from Nahuatl, almost certainly close to the time when Spanish speakers first encountered the plant. The Spanish word shows minimal formal adaptation of the Nahuatl word in order to replace the final consonant cluster /tl/, not found in native Spanish words, with the much more familiar combination of consonant plus vowel /te/. English, French, German, and Portuguese all borrow the Spanish word, either directly or via one another, although in some instances with loss of the final vowel in the spoken form. The modern English form *tomato*, first recorded in the middle of the eighteenth century, poses the only slight difficulty in the story presented so far: it probably arose as an alteration of earlier *tomate* by association with the name of a different plant which also happened to have originated in the Americas, *potato*.

However, this pattern of borrowing was far from inevitable, as becomes clear when we consider for a moment the modern Italian word for the tomato, *pomodoro*, a (rather fanciful) descriptive name compounded from Italian elements, and meaning literally 'apple of gold' (1544). Likewise

French earlier had *pomme dorée* and *pomme d'or*, and also *pomme d'amour*, literally 'apple of love' (1549; still found in the south of France), which gave rise to English *love-apple* as a loan translation of the French word. English *love-apple* is recorded as a name of the tomato in 1578, a quarter of a century before the occurrence of any form of the word *tomato* in English. Similarly *apple of love* is found from 1597. However, the evidence of corpora of historical texts suggests that neither term was ever very common in English. In modern-day Austrian German the usual name is *Paradeiser*, reflecting earlier *Paradiesapfel*, literally 'apple of paradise', a word found in the fourteenth century denoting the pomegranate and alluding to the fruit in the biblical story of the Garden of Eden, subsequently transferred in meaning after the arrival of tomatoes from the Americas. Thus even in a small selection of the major languages of western Europe we have several different strategies for naming the tomato, and a variety of different outputs.

A. Nahuatl *tomatl* > Spanish *tomate* (1532) > French *tomate* (late 16th cent.), German *Tomate* (17th cent.), Portuguese *tomate* (18th cent.), English *tomate* (1604), later (with remodelling after *potato*) *tomato* (18th cent.)
B. Italian *pomodoro* (1544), French *pomme dorée* (16th cent.), *pomme d'or* (17th cent.)
C. French *pomme d'amour* (16th cent.), model for English *love apple* (1578), *apple of love* (1597)
D. Austrian German *Paradeiser*, earlier †*Paradiesapfel*

We have already touched on the word *potato*. This shows some similarities to the history of *tomato*, but also some further complexities, involving the word's meaning as well as its form. The word is first recorded in English in 1565, denoting the edible root of the plant *Ipomoea batatas*, a plant of tropical American origin (in fact the product of cultivation by the peoples of the Americas) which is now usually referred to as the *sweet potato*. This plant was brought back to Spain from the Americas by Columbus after his voyage of 1492, and became widely cultivated in Europe, especially southern Europe. In Spanish this plant is called *batata*. This is a borrowing from an American Indian language, probably Taino. Borrowing of *batata* from Spanish is shown by Dutch *bataat* and German *Batate*, and also by obsolete English †*batata*. In the early sixteenth century a variant *patata* occurs in Spanish, probably arising from association with the name of a

quite different plant *papa*, which we will discuss further in a moment. The Spanish word was borrowed into French as *patate* and Italian as *patata*, in both cases denoting the sweet potato. The word was also borrowed into English, appearing in the sixteenth century as *patata* but also in a variety of other forms, including *potato*, which gradually became the form in general use. The *o* in the first syllable of this form probably arose from confusion over the value of an unstressed vowel in an unfamiliar borrowed word; the final *-o* in English *potato* lacks any obvious explanation.[8] (For simplicity of presentation, I have not mentioned early variation in word form in Dutch, German, French, Italian, or related forms in other languages.) Thus:

'sweet potato' (*Ipomoea batatas*)

Spanish *batata* (< Taino?), later (probably after *papa* 'potato') *patata* > Dutch *bataat*, German *Batate*, French *patate*, Italian *patata*, English †*batata*, †*patata*, *potato*

In English the word *potato* was also used to denote many other edible tubers, especially those originally imported from the Americas. In particular, it was used from the end of the sixteenth century as the name of the plant *Solanum tuberosum* and its edible tubers. This was another cultivated plant species, this time of South American origin, which was first encountered by Europeans during Spanish exploration of the Andes in the 1530s. In Spanish this plant was called *papa*, a borrowing from Quechua, and that remains its usual name in the Spanish-speaking Americas. In Britain, as elsewhere in temperate parts of Europe, it became a major food source. As such it came to seem the obvious referent of the word *potato* in its broadened meaning 'plant (from the Americas) with edible tubers', with the result that the 'original' *potato*, *Ipomoea batatas*, came to be distinguished as the *sweet potato* or sometimes as *yam* (more usually the name of yet another plant from the Americas with edible tubers, *Dioscorea*). However, there was nothing inevitable about the transfer of the name from the one

[8] One possibility is perhaps that the word was identified with words of the type *meadow*, *pillow*, which had variants with both reduced and unreduced final syllable, and hence the form *potato* arose by analogy with these, although if so it is surprising that spellings with *-ow* or *-ou* are not more common. On the small number of (mostly learned) words in English in this period with final /ə/ see Britton (2007) 527. Another possibility is that final *-o* was perceived as typically Spanish, which would be supported by the frequent alteration of the ending of words in *-ade* or *-ada* as *-ado* in this period (see *OED* at *-ado* suffix).

plant to the other. In other parts of Europe where potatoes were much grown, such as the Netherlands and the German-speaking countries, different names were adopted. The usual word in modern Dutch is *aardappel*, and similarly in Austria and parts of Germany it is *Erdapfel*, in both cases a compound from the words for 'earth' and 'apple', found much earlier denoting various other edible products of plants which are found either on or in the earth, and transferred in meaning to denote the newly encountered South American plant. French *pomme de terre*, literally 'apple of the earth', is again found denoting other plants from an early date, long before the period of European contact with the Americas, although its use denoting *Solanum tuberosum* is probably modelled on either Dutch *aardappel* or German *Erdapfel*, since the cultivation of this plant in France probably spread from Holland or Germany. The standard German term is *Kartoffel*, a word of complex history borrowed originally from Italian *tartufolo*, itself from an unattested Latin **territūberum*, literally 'earth tuber', originally denoting a truffle. Meanwhile, in Spain, perhaps under the influence of English, *patata* has in fact been found in the sense *Solanum tuberosum* from the beginning of the nineteenth century or earlier. Thus:

'potato' (*Solanum tuberosum*)

A. Quechua *papa* > Spanish *papa*
B. Dutch *aardappel*, Austrian German *Erdapfel*, models for French *pomme de terre*
C. German *Kartoffel*, originally < Italian *tartufolo* (< an unattested Latin **territūberum*, literally 'earth tuber', originally denoting a truffle)
D. English *potato*, a transferred use of the name of the sweet potato; hence also Spanish *patata*

In several other major European languages the plant has names which were current as the names of different plants before the advent of the potato (ultimately) from the Americas. If we were to extend our survey to include names for this plant in non-standard and regional varieties of these languages, we would find a yet more varied and complex picture. An even more complex set of ultimately related vegetable names can be traced by pursuing *aubergine* and *brinjal* in a good etymological dictionary. For a very complex example, and a classic etymological tour de force, see Ross (1952) and (1958) 146–8 on *ginger*.

5.6 Patterns of borrowing in the history of a language

We have touched on cultural considerations already, and have seen that the study of lexical borrowing is often closely interconnected with cultural history and external, extralinguistic factors. A good example is provided by the history of borrowing from French into English.

Although communication is known to have occurred across the English Channel between the Anglo-Saxons and the French, there are barely any borrowings from French into English which can be dated reliably to before the Norman Conquest in 1066.[9] *proud* is one of the very few secure examples: phonology and semantics both point clearly to borrowing from Old French rather than Latin, and the phonology points more precisely to borrowing from a western variety of Old French, and to a date of borrowing probably not earlier than the ninth century. Its recorded meanings in English show pejoration (see section 8.6.3) of the Old French meanings 'courageous, valiant, good, noble, just, prudent, wise, profitable, advantageous'. We must be cautious here, since the Old French word is in fact not recorded until considerably later than the earliest records of the word in Old English, but the French word's Latin etymon *prode* 'profitable, advantageous, useful' supports the originally positive meaning. (We will return to *proud* and its derivative *pride* in section 7.2.4.)

In the post-Conquest period, large numbers of borrowings are found, including some items of basic vocabulary (on this difficult concept see further section 6.2). It has been argued that some of these words entered English as a result of members of the Norman governing class switching from French to English as their language of everyday use, but this is very uncertain.[10] What is more certain is that in the later medieval period French was the first language of very few people in England, but it remained in daily use in many branches of professional and intellectual life, including the law and parliamentary business, alongside Latin, which had enormous importance, especially as the language of the church and much secular

[9] On early borrowing from French compare Kastovsky (1992) 337–8, Burnley (1992) 429, von Mengden (1999), Dietz (2003); for a view of *proud* slightly different from the one from *OED3* presented here, see von Mengden (2001).

[10] Compare for example Thomason and Kaufman (1988) 68, but see further section 6.2 below.

administration, from the Old English period onwards. Crucially, Latin and French were both much more developed as instruments of literate activity than was English, which only begins to develop any sort of (post-Conquest) supra-regional literary status in the second half of the Middle English period. English was clearly the dominant vernacular in everyday use, at least if we ignore for the moment areas where there may have been competition from Celtic languages or in the early Middle English period from Scandinavian languages, but both French and Latin had well-established roles in the life of society, particularly in written use and in the performance of various official, technical, and economic functions. The variety of French in question was Anglo-French, the lexis of which showed numerous formal and semantic differences from the French of the continent. The situation is well summarized by William Rothwell, editor of the *Anglo-Norman Dictionary*:

> Anglo-French ... was for centuries one of the two languages of record as used in government, the law, commerce and education in medieval England as well as of a wide-ranging literature. Insular French evolved in parallel and in constant contact with Middle English on the soil of England; it was not some sort of foreign decoration lightly superimposed on the native idiom. The officials of all ranks and their clerks who drafted and copied records all day in Latin and French were in large measure English and moved freely from one language to another according to the nature of their work and the company in which they found themselves.
>
> (Rothwell (1998) 159–60)

In the late Middle English period, and especially in the early fifteenth century, the use of (Anglo-)French in these technical and economic functions within England showed considerable decline,[11] but in this same period French culture was gaining in importance and dominance throughout the rest of Europe. French first became the principal language through which the Renaissance, and hence that part of the inherited classical learning which was in Latin and so most accessible to Western scholars, was conveyed to northern parts of Europe. The considerable borrowing of French lexis in Older Scots also reflects this. Then in the early modern period French began to outshine Italian as the leading vernacular language of culture and learning even in more southerly parts of Europe as well. Consequently, the level of borrowing from French remains high throughout the

[11] See the detailed sketch in Rothwell (2005). Compare also Machan (2003).

late Middle English and early modern English periods, although its relative importance declines in comparison with increasing numbers of new words of all origins, and particularly borrowings directly from Latin.[12] The actual numbers of borrowings from French do not show a significant decline until the eighteenth century, with a further steep drop in the twentieth century.[13] Much of the vocabulary borrowed from French in this period belongs to learned or literary discourse or to other specialist registers, or shows a notably high stylistic level.[14]

The influences of Latin and French on the lexis of English work largely in tandem: a large proportion of the borrowed French words are not only ultimately of Latin origin, but show a transparent correspondence in word form with their Latin etymons. Indeed, very many of these French words are not the regular reflexes of Latin words via proto-Romance, but are instead learned borrowings from Latin into French from the Old French period onwards. The impact on the lexis of English is enormous, as witnessed by the fact that ultimately many originally French or Latin affixes became productive in English (compare section 4.1.1), although there is scant evidence that this had happened before the end of the Middle English period.[15]

This process of borrowing of affixes ultimately facilitated further borrowing of more French and Latin lexis, since in many cases the composition of a newly encountered French or Latin word would be transparent to an

[12] For numerical analysis see Durkin (2008), and references to further literature given there; see also Dekeyser (1986), based on data from the *Middle English Dictionary*.

[13] For a preliminary discussion see Durkin (2006b); see also Mair (2006) 54, and further references given there.

[14] For a classic account of some of the results of this process in the lexis of modern English, where (near) synonyms of native and French/Latin origin often coexist (e.g. *brotherly* and *fraternal*, *heavenly* and *celestial*), and where a noun of native origin often has a corresponding adjective of Latin/Romance origin (e.g. *oral* beside *mouth*, *urban* beside *town*) see Ullmann (1962) 106–10, 145–51, who also offers a stimulating comparison with German and French. For an examination of the use of vocabulary of different origins in different literary styles in the early modern period see Adamson (1999).

[15] Dalton-Puffer (1996) finds very little evidence within the Middle English period for hybrid formations with native bases and Romance suffixes; there is also very little evidence for English formations from Romance elements which are not paralleled in French. Prefixes, which are excluded from Dalton-Puffer's study, might add a few further examples of hybrids, such as *renew*.

English speaker. In the subsequent centuries we find countless instances where it is almost impossible to judge whether an English word shows the result of word formation within English and just happens to have parallels in French and Latin, or whether it is modelled on French and/or Latin words, or whether it is in fact a borrowing from French and/or Latin. (See examples in section 5.1.4, and see further section 6.5 on words borrowed partly from French and partly from Latin.)

Many questions remain unanswered about even such a relatively well-investigated area as French borrowing in Middle English. Rothwell has done an enormous amount to demonstrate the continuing uses of Anglo-French in later medieval England, and to illustrate Middle English borrowing of distinctively Anglo-French lexis (compare *poke* in section 3.1 < Anglo-French *poke* as opposed to continental French *poche*). What we lack is any detailed study of just how frequently Middle English borrowing from French shows forms or meanings which are unique to either Anglo-French or continental French. Such a study would need to be backed up by an analysis of such factors as date of first occurrence, linguistic register, subject field, etc., in order to determine whether we can identify trends in borrowing from either Anglo-French or continental French in particular sub-periods or areas of social or intellectual activity. The new edition of the *Anglo-Norman Dictionary* now in preparation will make this much easier, although it may be that ultimately our surviving records of Anglo-French are not sufficient to allow us to gain a reliable picture of which words and senses may never have belonged to Anglo-French, and may instead have entered Middle English through direct contact with continental French, e.g. through literary contact.

Leaving aside the specific issue of borrowing from either Anglo-French or continental French, we can begin to sketch out some of the main factors which would need to be taken into account in any ideal, detailed analysis of borrowing from French into English:

- frequency (rather than just absolute dates of first attestation)
- later borrowing of specific senses (compare section 6.6)
- geographical variation and spread within English (compare section 6.4)
- linguistic register of the items borrowed
- how far we can estimate whether borrowings belonged to the 'general' vocabulary or only to more specialist vocabularies

Questions of the register and degree of currency of borrowings are crucial. As noted in Smithers's short essay 'Early Middle English' in Bennett and Smithers (1968) lii:

> It was probably not only an author's audience, but also his own background, endowments, and tastes that determined the number of adoptions from O[ld] F[rench] that he used. This is one of the reasons why the first record of a French word in M[iddle] E[nglish] should not necessarily be assumed (as is commonly done) to imply that it was, or even soon became, generally current in the 'language'. In fact, so long as we are dealing with any one M[iddle] E[nglish] work, the influence of French vocabulary on the 'language' is an abstraction: such a notion applies only to words which are found, on analysis of many works, to recur in several of them.

The resulting receptivity of English to French (and Latin) borrowings is also a subtle matter. The extent of borrowing provoked some negative comment in the early modern period, although the overall picture is somewhat mixed.[16] There were also calls for linguistic purism in early modern Germany, but there they much more frequently had an actual impact on the shape and composition of the lexis of German. We might speculate that this was because there had been comparatively little integration of Latinate and Romance lexis into German up until this date, and also because language had a crucial role in defining identity before the unification of Germany in the late nineteenth century. To take two simple examples from the world of languages and linguistics:

- *Wörterbuch* (1631; in early use also *Wortbuch*) 'dictionary' (literally 'book of words') was adopted by linguistic purists as an alternative to borrowed *Lexicon* or *Nomenclator*[17]
- *Mundart* (1641) 'dialect' (< *Mund* 'mouth' and *Art* 'manner, type') was adopted as an alternative to borrowed *Dialect* (1634; now *Dialekt*)

Both *Wörterbuch* and *Mundart* have become the usual words in modern German, largely replacing the earlier borrowed terms. When assessing borrowing into a language it can be very important to examine the subsequent frequency of use of borrowed terms, and in particular how they compete

[16] Compare Nevalainen (1999) 358–60, Görlach (1999) 479–80.

[17] German *Wörterbuch* is in fact itself a calque of a word in another language, albeit in this instance a closely related language in which both parts of the compound are cognate with (and easily recognizable as being equivalent to) those in the German word: Dutch *woordboek* (1599; now *woordenboek* (1648)).

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):