

1 The Historical Evolution of English Pronunciation

JEREMY SMITH

Introduction

Since at least the nineteenth century, the study of sound-change has been at the heart of English historical linguistics and our current state of knowledge depends on the insights of generations of scholars. This chapter aims simply to give a broad outline of the current “state of the art”, confronting basic questions of historical explanation. What does it mean to “account for” or “explain” a sound-change? How far can sound-changes be “explained”? How does one practise English historical phonology?

It is held here that historical phonology is as much history as phonology, and this insight means that evidential questions need to be addressed throughout. To that end, evidential questions are addressed from the outset. The chapter proceeds through the examination of a series of case studies from the history of English, ranging from the period when English emerged from the other Germanic dialects to become a distinct language to residualisms found in present-day varieties.

Overall, the chapter invites readers to reflect on their own practice as students of historical phonology; the explanations offered are, it is held here, plausible ones but by no means closed to argument. Good historiographical practice – for academic disciplines are of course collective endeavours – demands that such explanations should always be contested, and if readers can come up with better, more plausible explanations for the points made here, that is a wholly positive development, indicating new ways forward for the subject.

A question of evidence

Present-Day English is full of phonological variation; this variation, which is the outcome of complex and dynamic interactions across time and space, is valuable evidence for past states of English. To illustrate this point, we might take

the varying British English pronunciations of the words (a) *good*, (b) *food*, and (c) *flood*: a Scot will commonly rhyme (a) and (b); speakers from northern England typically rhyme (a) and (c); southern British English speakers rhyme none of them. Another example: southern British English speakers have a phonemic distinction between /ŋ/ and /n/ in, for example, *sing*, *sin*; northern English speakers do not, since they retain a final plosive in *sing* and for them [ŋ] is environmentally conditioned (and thus an allophone of, and not a distinct phoneme from, /n/). Many speakers of Scots, the traditional dialect and accent of Scotland, as well as speakers from north-east England, will pronounce the vowels in words such as *cow*, *now*, *house* with a close rounded back monophthong rather than (as southern speakers do) with a diphthong (see further Wells 1982).

Those learning to read, or non-native speakers, might reasonably expect, in a supposedly *phonographic* language such as English, that words ending in the same three letters, viz. *-ood*, in the written mode, should rhyme when read aloud, but, as we have just observed, in many accents of English they do not. The reason for the variation, and for the mismatch between spelling and sound, is that sound-changes have occurred since the spelling-system of English was established and standardized, and that these sound-changes have *diffused* differently through the lexicon in different parts of the English-speaking continuum. Some changes have only been adopted in some varieties.¹

The outcome of such patterns of divergence and diffusion is a body of *residualisms*, i.e., older forms of the language that remain in some accents but have ceased to be used in others (see Ogura 1987, 1990; Wang 1969; Wells 1982). The Scots/north-eastern English monophthongal pronunciations, for instance, of *cow*, *now*, *house* reflect the monophthongal pronunciation that seems to have existed in English a thousand years ago, cf. Old English *cū*, *nū*, *hūs* respectively. These pronunciations are therefore residualisms.

Residualisms are one of the major sources of evidence for the *reconstruction* of past states of pronunciation. We might illustrate the process of reconstruction using residualisms by comparing the British, Australian, and US pronunciations of the word *atom*; British and Australian speakers pronounce the medial consonant as /t/ whereas US speakers characteristically use a voiced alveolar tap, meaning that in US English the word *atom* is a homophone with *Adam*. It is usual to consider the US pronunciation to be an innovation, whereas the other usages are residualisms, the evidence for this interpretation being that US speakers characteristically voice intervocalic sounds in derived forms, cf. US English intervocalic /d/ (however precisely realized) in *hitter* beside final /t/ in *hit*, beside /t/ in both environments in British and Australian usage. Such reconstructive processes are, of course, the basis of comparative linguistics.

However, deciding what is a residualism and what is not can be a difficult matter without further information. To take a large-scale example: the phenomenon known as Grimm's law (the "First Consonant Shift"), whereby a series of consonants in the Germanic languages seem to have undergone a comprehensive redistribution within the lexicon, is traditionally described as a Germanic innovation. Illustrative examples are given in Table 1.1.

Table 1.1 Grimm's law cognates in Germanic and non-Germanic languages.

	<i>Germanic examples</i>	<i>Non-Germanic examples</i>
/f/ - /p/	English <i>fish</i> , Norwegian <i>fisk</i>	Latin <i>piscis</i> , French <i>poisson</i> , Welsh <i>pysg</i>
/θ/ - /t/	English <i>three</i> , Icelandic <i>þrír</i>	Latin <i>trēs</i> , French <i>trois</i>
/h/ - /k/	English <i>hound</i> , German <i>Hund</i>	Latin <i>canis</i> , Welsh <i>ci</i> , Tocharian <i>ku</i>

However, some scholars, arguing that a similar process is also found in Armenian, like Germanic a “peripheral” language within the Indo-European group but at the eastern as opposed to the western end of that language-family’s extent, have argued that Grimm’s law represents a residualism rather than an innovation. This so-called “*glottalic*” theory is highly controversial, but that it has found purchase with at least some scholars indicates the nature of the problem (see Smith 2007: ch. 4).

The study of residualisms as evidence for the history of pronunciation, therefore, is – where possible – combined by researchers with other sources of evidence: sound-recordings, available since the end of the nineteenth century; contemporary comments on past pronunciation; past spelling-practices, given the mapping between speech and writing found in phonographic languages; and the practices of poets, in terms of rhyme, alliteration, and metre. Taken together, these various pieces of evidence allow scholars to develop plausible – though never, of course, absolutely proven – accounts of past accents, and sometimes even to offer plausible explanations for how particular accentual features emerged. A series of case studies follows, with special reference to the history of English, to illustrate the process of developing such plausible accounts and explanations.

Case study 1

Voiced and voiceless fricatives: development of new phonemic categories

The first of these case studies deals with the Present-Day English phonemic distinction between voiced and voiceless fricatives, a distinction that has emerged during the history of English and is reflected – albeit sporadically and unevenly – in Present-Day English spelling. The example also allows us to ask a certain key, and surprisingly neglected, question: what is a sound-change?

One such distinction, which often puzzles present-day learners of English, is to do with the pronunciation of the word *house*; when used as a verb, the word ends with /z/ but, when used as a noun, it ends with /s/. The usual historical explanation is as follows: in Old English, voiceless [s] and voiced [z] were allophones of the same phoneme, conventionally represented by /s/, and therefore in *complementary distribution* within the sound-system. It seems that /s/ was pronounced

voiced intervocalically, but voiceless when a word-final. The Old English word for “house” (noun) was *hūs*, while the Old English word for “house” (verb) was *hūsian*; when, in the transition from Old to Early Modern English, inflectional endings such as *-ian* were reduced and ultimately lost, a voiced sound emerged in final position in words such as “house” (verb), leading to the current pattern for the sound’s deployment. Since “house” (noun) and “house” (verb) now have distinct meanings marked by replacement of single word-final segments, the two words have come to form a *minimal pair* for the purposes of phonological analysis, and the phonemes /s, z/, now in *contrastive distribution*, may thus be distinguished.

Of course, the evidence we have for the initial complementary distribution can only be deduced; direct evidence, in the form of contemporary commentary or distinctive spellings from Old English times, is almost entirely lacking and the distribution of forms means that poetic evidence is not to be had. The issue is one of plausibility, in that the process of *phonemicization* just described aligns with known developments elsewhere in the linguistic system, notably inflectional loss.

Spelling evidence for sound change is really only available on a large scale from the Middle English period. Middle English is notoriously the period in the history of English when there is a closer alignment between spelling and pronunciation than before or since. Written English had a parochial rather than national function, used for initial or otherwise restricted literacy, while – following Continental practice – unchanging, invariant Latin was deployed as the language of record across time and space. Thus it made some sense to reflect English phonological variation in the written mode, since that made teaching reading easier. Only when English, towards the end of the medieval period, took on the role of a language of record did variation become inconvenient. The *standardization* of written English was a formal response to a change in linguistic function. That English spelling could remain fixed while pronunciation changed was first discussed by Charles Butler in his *English Grammar* (1633), who saw the development as regrettable and thus needing reform (Dobson 1968: 165), but the socially useful functionality, for record-keeping purposes, of a fixed spelling-system, despite a phonographic mismatch between spelling and widely attested pronunciations, has meant that comprehensive spelling-reform in English has never succeeded.

It is therefore possible – at least sometimes – to see reflections of sound-change in changes in spelling. As with the [s]/[z] distinction, Old English made no phonological distinction, it seems, between voiced and voiceless labio-dental fricatives and as a result the spelling <f> was used to reflect both, e.g., *fela* “many”, *hlāf* “loaf” (both with [f]), but *yfel* “evil” (with medial [v]). A phonological distinction seems to have emerged in the Middle English period largely as a result of the adoption of loan-words from French, e.g., *fine*, *vine*, and this distinction became sufficiently salient for a spelling-distinction, between <f> and <v>, to be adopted and even extended to native words, such as *evil*. The <f>/<v> distinction first emerged in Middle English and has been sustained ever since.

However, it is noticeable that even in Middle English conditions such developments do not always follow. Distinctions between other voiced and voiceless

fricatives, i.e., the alveolars /s, z/ (as we have just seen) and the dentals /θ, ð/, also emerged, but the spelling-evidence for such developments is uncertain. The letter <z> remains marginal in Present-Day English spelling, used in the initial position only in exotic words such as *zoo*, *zebra* and even replaced by other letters altogether in *xylophone*, *xerox*; in medial and final positions it is also in some sense “optional”, cf. the variation between *criticise*, *criticize*, or the fact that the word *ooze* is a homophone with the river-name *Ouse*. For Shakespeare, <z> was an “unnecessary letter” (*King Lear* II.2) and in Middle English <z> is witnessed only sporadically. It is noticeable that the only texts to use <z> consistently in the initial position are Middle Kentish ones, such as the *Ayenbite of Inwynt*, surviving in a manuscript localized to Canterbury in 1340, where a consistent distinction is made between, for example, *zom* (from Old English *sum* “a certain”) and *som* (from Old French *sum* “a sum (of money, etc)”). Initial voicing of fricatives seems to have survived in Kentish until the end of the nineteenth century though is now recessive (see Smith 2000 and references there cited).

Similarly marginal is the distinction in voiced and voiceless dentals. Present-Day English deploys <th> for both /θ/ and /ð/, except in specialist vocabulary such as *sandhi* or in forms made up for literary effect by philologists, such as the name *Caradhras* in J.R.R. Tolkien’s *The Lord of the Rings*; in both cases <dh> represents the voiced fricative sound. The reason for this limited reflection of a phonological distinction seems to be that there is only a limited set of minimal pairs, e.g., *thy*, *thigh*, and *that*, and at least in the initial position, the voiced dental fricative is restricted to “grammar words” such as *the*, *that*, *this*, *those*, *these*, *there*, *though*, or in certain pronouns such as *they*, *them*, *their*. In Middle and Early Modern English texts, there is some evidence that some scribes deployed <þ> – sometimes written in a manner indistinguishable from <y> – only in such words (e.g., the common use of <ye> for “the”). Such practice may reflect a sound-distinction, but equally plausibly it could be argued that it is simply a space-saving device, whereby a form largely predictable from context could be represented in abbreviated fashion (the custom of abbreviating forms such as “the” or “that” as <ye> or <y>, with superscript second letters, would support the latter interpretation).

The key point, of course, is that there is no necessary connection between what a medieval or renaissance scholar would have called the *figura* (written manifestation of a *littera* “letter”) with a particular *potestas* (sound-equivalent) (see Abercrombie 1949). To demonstrate this point, we might take, for instance, spellings of the words “shall”, “should”, common in the Middle English of Norfolk, viz. *xal*, *xuld*. In such cases, it is notoriously hard to establish the *potestas* of <x>. Is <x> in such words simply a local spelling for [ʃ] or does it represent a distinct sound? Its restriction to the words “shall”, “should” (until the very end of the Middle English period, when it is sporadically transferred to words such as *xuldres* “shoulders”) would suggest the latter, but there is no certainty as to the precise *potestas* to be assigned to it.

Support for a voiced/voiceless distinction in the fricatives, at least for the alveolar and dental sets, is suggested rather than proven by the spelling-evidence, and

other information is needed if we wish to establish the phonemicization in the history of English pronunciation. Unfortunately, there is no meaningful discussion of English pronunciation until the sixteenth century, when English became a respectable subject for intellectual study rather than simply a “vulgar” tongue; however, the evidence from then on becomes full. John Wallis’s *Grammar of the English Language* (1653), for instance, noted the distinction between what he called “hard *s*” and “soft *s*”, in which the latter was pronounced “per *z*” in a house, to house respectively (Kemp 1972: 178–179), and Wallis regretted the failure in English spelling to distinguish voiced and voiceless dental fricatives, which he regarded as “an unfortunate practice” (Kemp 1972: 176–177). Wallis states that the Welsh use <dd> for the voiced sound “though some maintain that *dh* would be a better way of writing it than *dd*; however they have not succeeded in getting the old established custom altered” (Kemp 1972: 177).

Interestingly, the labio-dental voiced/voiceless distinctions are not discussed to the same extent, possibly because the spelling-distinction was already accepted by early modern times. The spelling *hl̥uade* for the third-person preterite singular of *hl̥ifian* “stand tall, tower” appears in the late tenth century *Beowulf Manuscript* (MS London, British Library, Cotton Vitellius A.xv, *Beowulf* line 1799), beside the more common *hl̥ifade*. The spelling with <u> is usually taken as the earliest instance of an attempt to reflect a voiced–voiceless distinction in English spelling.

A good working definition of sound-change might be as follows:

Sound-change is a phenomenon whereby speakers adjust their phonologies, or sound-systems. The raw material for sound-change always exists, in the continually created variation of natural speech, but sound-change only happens when a particular variable is selected in place of another as part of systemic regulation. Such processes of selection take place when distinct systems interact with each other through linguistic contact, typically through social upheavals such as invasion, urbanization, revolution, or immigration.

However, two issues become fairly clear from the discussion so far. Firstly, as the form *hl̥uade* and the current restricted distribution of the voiced and voiceless dental fricatives suggest, sound-change is what might be termed an *emergent* phenomenon. That is, sound-changes are not sudden affairs but typically diffuse through time and space in a “sigmoid-curve” pattern, working their way through the lexicon. Diachronic discussion is not a matter of aligning a series of synchronic descriptions of phonological inventories at given points in time, i.e., a series of “maps”. It is a different kind of discourse (for the notion and importance of emergence, see especially the essays in Bybee and Hopper 2001).

Secondly, it is clear that, although almost all scholars accept a general narrative about the history of voiced and voiceless fricatives in the history of English, the evidence is indicative rather than conclusive. Potestates map on to figurae, but in complex ways, and without access to recorded sound from any period before the end of the nineteenth century it is not possible to offer any final, demonstrable proof of the structure of past sound-systems. The argument, as so often in historical study, is based on the *plausible interpretation* of fragmentary indicators.

Digraphs and diphthongs

The previous section focused on what is arguably the major phonological development in the history of English sounds: the emergence of a whole distinct category of phonemes. Changes in English vowels are more widespread, but making evidence a starting-point can also be most illuminating.

As with consonantal change, that potestates map on to figurae in complex ways can be illustrated with reference to the history of English vowels, and a Present-Day English example makes the point. In most modern accents, words with <ee> and <ea> commonly rhyme, e.g., *meet*, *meat*, although there are of course numerous exceptions, e.g., *greet*, *great*, and some alternative rhyming patterns, commonly, where the vowel is followed by /r/, e.g., *pear*, *pair* rather than *pear*, *peer* (although cf. the non-rhyming *fear*, *fair*), or by a dental or alveolar consonant, e.g., *breath* (rhyming with the personal name *Seth*) and *dead* (rhyming with *bed*). In some varieties, particularly conservative ones, what are clearly older patterns survive residually, e.g., in some accents of Irish English *meat* rhymes with *mate* rather than *meet*. The current complex distribution of <ea> spellings in relation to sound-systems is the result, as we might expect from the discussion so far, of sound-changes diffusing incompletely and irregularly across the lexicon subsequent to the standardization of the writing system.

It might be expected, in periods before the writing system became standardized, that the relationship between figurae and potestates might be closer, i.e., the language-variety in question would be more completely *phonographic*. However, despite a tradition of research of more than a century, very basic problems in the interpretation of vowel-potestates remain contested by scholars.

Anglo-Saxonists, for instance, still debate the existence of basic phenomena such as the nature of the diphthongal system and the interpretation of the spellings <ea, eo, ie>. Questions asked, still not conclusively answered, include:

1. Do these spellings really represent diphthongs?
2. Are they to be seen as equivalent to long monophthongs, i.e., VV?
3. How far are (as conventional wisdom holds) the “short diphthongs” <ea, eo, ie> to be seen as metrically equivalent to short vowels, i.e., V (vowels with which, historically, they tend to merge)?
4. How are the individual elements within these diphthongs (if that is what they are) to be pronounced?

These questions form a major conundrum in the study of Old English phonology.

Almost all scholars accept the existence in the West Saxon dialect of Old English of the long diphthongs spelt <ea, eo>, which represent the reflexes of Germanic diphthongs as well as the products of certain sound-changes. These diphthongs were “bimoric”, i.e., VV in terms of *metrical weight*, and thus equivalent to long monophthongs, sounds with which historically they tended to merge. The problem arises with the so-called “short diphthongs”, which were not the reflexes of Germanic diphthongs but arose as the result of sound-changes such as breaking or

“palatal diphthongization”, and have been believed by many scholars to be monomoric, i.e., V, and thus equivalent in metrical weight to a short monophthong. Richard Hogg sums up this view as follows: “... the traditional position holds that <ea, eo, io> always represented diphthongs both long and short except where the orthographic evidence suggests otherwise or the linguistic development is implausible ...” (1992: 17). The key problem is, as David White has pointed out (2004: *passim*), that such short diphthongs are vanishingly rare in world languages, and indeed not found in living languages at all; their presence in standard descriptions is the outcome in all cases of scholarly reconstruction.²

One argument offered originally by Marjorie Daunt (1939, 1952) and reiterated by White (2004) is that spellings such as <ea, eo>, when representing the “short diphthongs”, include a diacritic element, flagging the quality of the following consonant. Certainly it is generally accepted that such diacritic usages occur in Old English, e.g., spellings such *sēcean* “seek” (beside more common *sēcan*), or *geong* “young” (which would have yielded Present-Day English **yeng* if <eo> in this word had represented one of the presumed “short diphthongs”). It could therefore be argued that <ea, eo> in words such as *eald* “old”, *earn* “eagle”, *weorpan* “throw”, *eolh* “elk” represent /æ/ or /e/ followed by a “back (i.e., velarized) consonant”; <eo> in *heofon* “heaven” would be an attempt to represent /e/ “colored” by the back vowel in the unstressed syllable. Daunt pointed out that digraphs of various kinds were deployed by Old Irish scribes to flag the quality of neighboring consonants, and Old Irish scribal practice strongly influenced Old English usage.

However, there are problems with this analysis. Minimal pairs arose in West Saxon, subsequent to the operation of the sound-change that produced <ea> in *eald*, *earn*, etc., which seem to indicate that <ea> was perceived in West Saxon as distinct in quality from <æ>, e.g., *ærn* “house” beside *earn* “eagle”; despite suggestions to the contrary (e.g., White 2004: 80), it seems likely that, in the conditions of vernacular literacy obtaining in West Saxon, this difference indicates a real distinction in pronunciation. If there were no difference in pronunciation we would expect variation in spelling between **æld* and *eald* in West Saxon, and such a variation does not occur.

Although some languages (e.g., Scottish Gaelic) have a three-way length distinction, viz. V, VV, VVV (see Laver 1994: 442), it seems unlikely that Old English had the same system, with the short diphthongs to be interpreted as bimoric (VV) and the long diphthongs as trimoric (VVV). The “long diphthongs” of OE derive in historical terms from bimoric (VV) Proto-West Germanic diphthongs, and there does not seem to be any good reason to posit a lengthening, especially as, in later stages of the language, they tend to merge with long monophthongs (VV).

Perhaps the most economical explanation would be to see the “short diphthongs” as consisting of a short vowel followed by a so-called glide vowel, i.e., Vv in the environment of a following back consonant. Daunt herself argued that “there was probably a glide between the front vowel and the following consonant” (Hogg 1992: 18–19, and see references there cited). The distinction between monophthongs plus glides and diphthongs is a tricky one, but recent

experimental work on Spanish suggests that a robust distinction is possible (see Hualde and Prieto 2002). The spelling <ie> is used in Early West Saxon to represent the outcome of further sound-changes that affected <ea, eo>, and it therefore seems logical – if the Daunt/White interpretation is accepted – to assume that it, too, represents a diphthong, probably of the same kind (i.e., full vowel plus hiatus vowel).

Establishing the sound-equivalent (potestas) of a particular spelling (figura) is one thing; proceeding to explain the conditions under which a particular potestas emerged is another, and here we are on even more tenuous ground at such an early date in the history of English. The Old English spelling <ea> in *eald*, *earn*, etc., is a product of the sound-change known as “*Breaking*”, usually defined as a diphthongization in the environment of a following “back” (i.e., velar) consonant. Whether <ea> is to be interpreted as a diphthong or not is, as we have just seen, a complex question, but all scholars agree that the consonants <l, r>, etc., are “back” in terms of the Old English system. The question is, though, when did they *become* back consonants to induce the change?

One plausible possibility is that the precise realization of <l> in the Old English dialects manifesting breaking had undergone a change as the result of contact with other varieties, a change in consonantal realization that had a knock-on effect on the pronunciation of the preceding vowel. It is thus relevant to refer back to consonantal change when accounting for the evolution of vowels, flagging the dynamic interconnectedness of sound-changes. Breaking is the first sound-change that can be clearly located in Anglo-Saxon England after the so-called *Adventus Saxonum* (“the coming of the Saxons”), the period of transition between Romano-Celtic Britain and Anglo-Saxon England; earlier sound-changes, e.g. “*First Fronting*” (sometimes known as “*Anglo-Frisian Brightening*”), date from the period when the Angles and Saxons were still on the Continent of Europe. It thus developed, in West Saxon, at a time when Saxons were coming into contact with Angles in a condition of confused and complex social ties.

There is some evidence that, in Old Anglian, /l/ and /r/ were back consonants. Old Anglian was in origin the variety furthest north within the West Germanic-speaking area, being spoken in the area immediately abutting the most southern varieties of North Germanic, and the continual interchange between North and West Germanic, often commented on by linguists (see for instance Haugen 1976: *passim*), would clearly have impacted most upon it. Many of these southern varieties even now have a “dark /l/”, often referred to as “thick” or “cacuminal” /l/. It could therefore be argued that, when Anglian and Saxon varieties came into contact with each other as a result of the *Adventus Saxonum*, Saxons attempted to reproduce Anglian usage in situations of language contact; a “dark” form of /l/ would result. That Saxons would have imitated Anglians rather than vice versa is suggested by the evidence – admittedly somewhat tenuous – that Anglians dominated the early Anglo-Saxon polity: after all, the name “England” derives from “Angle”, and the name “Saxony” is applied to an area of present-day Germany (see further Smith 2007: ch 4, and references there cited).

The Great Vowel Shift

In the previous section, the explanation offered for change was in some sense sociolinguistic, but there were limits to such an approach, derived, quite simply, from the comparative paucity of evidence. The best that can be hoped for from such explanations is plausibility linked to certain arguments to do with similarities between past and present. In this section, greater evidence allows us to make such arguments more convincingly.

Such explanations as that just offered for the origins of Breaking, as the result of language contact in situations where one group might be considered more prestigious than another, may be tenuous, but they gain traction from the observable fact that such situations are observable in present-day language. As William Labov famously argued in what may be considered a foundational statement of the subdiscipline of *historical sociolinguistics*, the present can be used to explain the past (Labov 1974). Since the so-called “*uniformitarian hypothesis*”, accepted by linguists, holds that speakers in the past – like us – reflected their social structure in language (see, for example, Romaine 1982 and Machan 2003), it seems unarguable that the social setting of language-use in early times had an effect on linguistic development, specifically sound-change. The tenuousness of the explanation relates to the difficulty not of the principle but of our limited understanding of the precise social circumstances that obtained at the time.

It is therefore arguable that the more information we have about social structure the higher degree of plausibility there is about explaining a given sound-change. Thus a later change, such as the *Great Vowel Shift* of the fifteenth and sixteenth centuries, a process of raisings and diphthongizations that distinguishes the phonologies of Late Middle English period from those of the Early Modern English period and that may be described as a redistribution of sounds within the lexicon, can be explained fairly convincingly as the outcome of interaction between social groups in conditions of increasing urbanization.³

The origins of the Great Vowel Shift have, notoriously, been regarded by many scholars as “mysterious” (Pinker 1994: 250), an adjective that would seem to close down discussion. However, an interest in the Shift’s origins has persisted, particularly amongst scholars whose work engages with sociolinguistic concerns.

It is noticeable that the Shift took place at a key moment of transition in the history of English, when English ceased to be a language of comparatively low status in comparison with Latin and French and began to take on national roles, i.e., it underwent a process that Einar Haugen has referred to as *elaboration* (Haugen 1966; cf. also Hudson 1980: 32–34, and references there cited). The elaboration of English meant that prestigious varieties of that language began to emerge. The story of the Southern Great Vowel Shift relates, I have argued, intimately to that emergence. It seems that the Southern Shift derives from sociolinguistically-driven interaction in late medieval/early Tudor London, whereby socially mobile immigrant groups *hyperadapted* their accents in the direction of usages that they perceived as more prestigious. Such a process can be paralleled

in modern situations, whereby linguistic innovation is located in the usage of those who are weakly tied to their social surroundings (see Milroy 1992).

The origins of the Southern Shift correspond in date to four major – and, I would argue, linked – developments in the external and internal history of the English language. These developments are as follows:

- a. *The rise of a standardized form of English.* At the end of the fourteenth and the beginning of the fifteenth centuries, it is possible to detect, in the written mode and to a lesser extent in speech, the emergence of focused forms of language that are the precursors of Present-Day “standard” varieties.
- b. *The growth of London.* The end of the Middle Ages and the beginnings of the Tudor period saw the increasing significance of London as England’s major administrative and trading centre. From the fourteenth century onwards there was a major influx of immigration into the capital from the countryside as folk sought to improve their condition in the city. This is the age of the quasi-mythical figure of Dick Whittington, who moved to London, where the streets were (it was said) paved with gold, to make his fortune. The result was that London became, according to contemporaries, the only English city comparable in size and importance to continental centers such as Paris, Venice, and Rome (see, for a convenient account, Ackroyd 2002, and references there cited). London society, which (as nowadays) attracted incomers from elsewhere eager to take advantage of the opportunities it had to offer, may be characterized as one with *weak social ties* in comparison with those which obtained in the much more stable, less dynamic village society that existed elsewhere in England.
- c. *The loss of final –e.* The Shift corresponds in date to a grammatical development of considerable prosodic significance: the development of what is essentially the Present-Day English grammatical system with *the loss of inflectional –e*. Final –e was still in use in adjectival inflections in Chaucer’s time, as established (*inter alia*) by the poet’s verse practices, but the generations that followed Chaucer, from the end of the fourteenth century onwards, no longer recognized the form. The loss of –e had major implications for the pronunciation of English, whose core vocabulary became, to a large extent, monosyllabic in comparison with other major European languages.
- d. *Phonemicization of vowels affected by Middle English Open Syllable Lengthening in those accents where these vowels did not undergo merger.* This development was a consequence of the loss of final –e. There is good evidence, from contemporary rhyming practice in verse, that the comparatively prestigious form of speech represented by that of Geoffrey Chaucer distinguished carefully between the reflex of Old English *e* and *o*, which had undergone a quantitative change known as *Middle English Open Syllable Lengthening* and the reflex of Old English *ēa*, *ǣ*; with the loss of final –e, this distinction became phonemicized in Chaucer’s (more properly, Chaucer’s descendants’) variety and thus perceptually salient. However, in other varieties outside London, Middle English Open Syllable Lengthening-affected *e*, *o* merged with the reflexes of Old English *ēa*, *ǣ*, and *ā* > \bar{Q} respectively. These two systems may be characterized as System I and System II respectively.

With the rise of London and the perception of there being a prestigious form of speech that coincided with it, users of System II, whose social situation may be characterized as weakly tied, came into contact with users of System I. System I speakers distinguished phonemically between Middle English Open Syllable Lengthening-affected *e* and *o* and the reflexes of Old English $\bar{e}a$, \bar{a} , and $\bar{a} > \bar{Q}$, whereas System II speakers did not. Moreover, it seems likely that System I speakers, with a habit of pronouncing much of their stylistically marked vocabulary in a “French” way – see (a) – would have distinct ways of pronouncing mid-close \bar{e} and \bar{o} ; there is some evidence that French \bar{e} and \bar{o} were realized as somewhat higher in phonological space than the reflexes of English \bar{e} and \bar{o} , and adoption of French-influenced usages would have been encouraged by the presence of the extra phoneme, derived from Middle English Open Syllable Lengthening, in both front and back series of long vowels. R.B. Le Page has suggested that the aristocracy of the late fourteenth and fifteenth centuries were likely “to adopt affected forms of speech as a means of ‘role-distancing’ from the lower classes, from whom they had hitherto been differentiated by speaking French” (cited in Samuels 1972: 145–146). Further, if the raised “French” style pronunciations of \bar{e} and \bar{o} were adopted by System I speakers, it seems likely that diphthongal pronunciations of the close vowels \bar{i} and \bar{u} , which are attested variants within the phonological space of close vowels in accents with phonemic length, would have been favored by them, viz. [ī, ū], in order to preserve distinctiveness. Such a development would mean that a four-height system of monophthongal long vowels would be sustained, with Middle English /i:/ being reflected as a diphthong, albeit one with a comparatively close first element.⁴

We would expect in such circumstances that hyperadaptations would follow, and this is the basis of the argument for the origins of the Shift offered here. System II speakers, who may be characterized as weakly tied, socially aspirant incomers, encountered System I speakers whose social situation they wished to emulate. The process, it might be plausibly argued, would have worked somewhat as follows. System II speakers would have heard System I speakers using what they would have perceived as a mid-close vowel in words where they would use a mid-open vowel. Since final *-e* had been lost there would not be a grammatical rule to identify when such vowels should be used, and System II speakers, who formed the rising class of late medieval and early Tudor London, would replace their mid-open vowels (whether derived from Middle English Open Syllable Lengthening-affected *e*, *o* or from Old English $\bar{e}a$, \bar{a} , and $\bar{a} > \bar{Q}$) with mid-close ones. There would be phonological space for them to do so since they were also attempting to imitate the socially salient raised allophones of System I speakers’ “French” style raised /e:, o:/. Since these latter pronunciations were themselves not in the inventory of System II speakers, it seems likely that such pronunciations were perceived as members of the phonemes /i:, u:/ and would be reproduced as such (on hyperadaptation, see Smith 2007, and references there cited, especially Ohala 1993).

Of the remaining developments in the Shift, diphthongization of front vowels would derive from attempts by System II speakers to imitate System I speakers’ [ī, ū] allophones of /i:, u:/. Such selections would be encouraged by the need to

retain perceptual distance from the “French” style raised /e:, o:/, hyperadapted by System II speakers as /i:, u:/. As I have suggested elsewhere, the later development whereby Middle English /a:/ > /ɛ:/ probably derives from a distinct, sociolinguistically-driven process. Middle English phonemic /a:/ was comparatively new in most Southern English accents, being derived largely from Middle English Open Syllable Lengthening-affected /a/. The main accent in the South-East where phonemic /a:/ had existed beforehand was the Essex dialect, which seems to have been the “old London” usage characteristic of low-prestige speakers in the area. A raised pronunciation of Middle English /a:/, probably as [æ:], would have been another way of marking social distinction, which System I speakers would have been keen to make. System II speakers, attempting to replace their own realizations of /a:/ with System I’s [æ:], would have tended again to overshoot, identifying the System I [æ:] pronunciation with the next phoneme in their own series, viz. /e:/.

The outcome of all the developments just described was the distribution of vowels attested by the best writers on pronunciation in the sixteenth century. The developments just argued for, incidentally, also illustrate how sound-change is a processual, emergent phenomenon, not something that suddenly appears in saltatory fashion, as might sometimes appear to be the case from handbook accounts.

Explaining sound-change

We might now move to central issues raised by the case studies discussed. Historical explanations, such as those just provided for Breaking and the Great Vowel Shift, are necessarily exercises in plausible argumentation, and a plausible argument is not absolutely proven. In historical subjects, absolute proof is not to be had. The question, therefore, is: how can we assess the success of an historical explanation?

As I have argued elsewhere (Smith 2007), certain historical approaches, e.g., postmodernism, have emphasized the “observer’s paradox”, the way in which the frame of reference of the investigator constrains the enquiry. However, as I have suggested, the observer’s paradox should not be seen as disabling, but rather it places certain ethical requirements on historians: to be self-critical, to be open to other interpretations of events, and (above all) to be humble. Historians are (or should be) aware that their work is in no sense a last word on a topic but simply part of a continuing discussion in which their views may eventually come to be displaced. Explanations of sound change, like all historical explanations, are successful if they meet certain criteria of plausibility. As April McMahon has put it, “we may have to accept a ... definition of explanation at a ... commonsense level: explanation might ... constitute ‘relief from puzzlement about some phenomenon’” (1994: 45, and references there cited).

In assessing the plausibility of the accounts of the Shift just offered, it is perhaps a good idea to return to the notion of the uniformitarian principle, a notion that underpins what is probably the most fruitful current development in the study of

the subject, viz. historical sociolinguistics (see further Millar 2012 and references there cited), and a renewed focus on what has been called the “linguistics of speech”. Such a *parole*- (as opposed to *langue*)-based approach to linguistic investigation is informed by the close analysis of large bodies of data, both from the present-day and from the past, harnessing insights about the “dynamic” nature language derived from complexity science (for which see most importantly Kretzschmar 2009). The linking of present-day and past circumstances – as flagged by Labov back in 1974 – is crucial; if sound-changes in present-day circumstances take place because of certain social conditions, and if the phonetic processes that obtain in those circumstances (i.e., hyperadaptation) may be observed, then it seems at least plausible that similar processes governed sound-changes in the past. The study of past sound-changes, therefore, is a project that must be linked closely to an understanding of the dynamic and complex processes of social history. In so doing, we may be “relieved from puzzlement” – which is, in English historical linguistics, probably as good as it gets.⁵

NOTES

- 1 In a phonographic language there is, broadly speaking, a mapping between grapheme and phoneme. A *logographic* language, by contrast, is one where the mapping is between grapheme and notion. Written versions of Western European languages are largely phonographic; written Chinese is logographic. The difference may be illustrated by the symbols used for numbers; “8” is a logograph, corresponding to the written/spoken usages *eight* (in English), *huit* (in French), *otto* (in Italian), *acht* (in German), or indeed the spoken usages *bā* (Mandarin Chinese), *takwas* (Hausa), *siddeed* (Somali), or *walu* (Fijian). There are advantages to logographic languages; German speakers may not be able to understand Fijian speakers when they write in their native languages, but both Germans and Fijians will be able to understand each other’s mathematical symbols. Famously, Cantonese and Mandarin are not mutually intelligible when spoken, but since the writing-system commonly deployed in varieties of Chinese is in principle logographic it is possible for users of these varieties to understand each others’ writings. Logographic systems are problematized by their use of a very large number of symbols, and they are thus a challenge to the memorizing powers of those learning to read and write, but it is undeniable that they are useful as a language of record and transaction – which is why they emerged in Imperial China.
- 2 Richard Hogg was of course aware of the difficulty, although – appropriately in a handbook – he tended to the conventional view, and his qualification is therefore carefully expressed. A fuller quotation reads: “... the traditional position holds that <ea, eo, io> always represented diphthongs both long and short *except where the orthographic evidence suggests otherwise or the linguistic development is implausible ...*” (1992: 17; my italics).
- 3 Five-height systems of monophthongal phonemes are attested in the world’s languages, but are rare; three- and four-height systems are much more common (see Maddieson 1984: *passim*).
- 4 As well as a “full” Shift affecting both the long front and long back vowels of Middle English, characteristic of southern varieties, there was also a distinct Shift, affecting

primarily long front vowels, which is found in Northern accents. The discussion in this chapter focuses on the “full” or Southern Shift; for a discussion of both in much more detail, see Smith 2007: ch. 6, and references there cited. It is argued that the triggering of the “Northern” Shift was the result, like the Southern Shift, of socially-driven linguistic choices (i.e., it was a sociolinguistic phenomenon), whose outset related to earlier shifts in the back series of long vowels consequent on interaction with Norse.

- 5 For a similar attempt to use the present to explain the past, but with reference to a much more archaic set of sound-changes, see Jane Stuart-Smith’s discussion of the processes involved in ancient Italic accents (2004).

REFERENCES

- Abercrombie, D. 1949. What is a “letter”? *Lingua* 2: 54–63.
- Ackroyd, P. 2002. *London: The Biography*, London: Faber.
- Bybee, J. and Hopper, P. (eds.) 2001. *Frequency and the Emergence of Linguistic Structure*, Amsterdam: Benjamins.
- Daunt, M. 1939. Old English sound changes reconsidered in relation to scribal tradition and practice. *Transactions of the Philological Society* 108–137.
- Daunt, M. 1952. Some notes on Old English Phonology. *Transactions of the Philological Society* 48–54.
- Dobson, E.J. 1968. *English Pronunciation 1500–1700*, Oxford: Clarendon Press.
- Haugen, E. 1976. *The Scandinavian Languages*, London: Faber and Faber.
- Haugen, E. 1966. Dialect, Language, Nation. *American Anthropologist* 68: 922–935.
- Hogg, R. 1992. *A Grammar of Old English I: Phonology*, Oxford: Blackwell.
- Hualde, J.I. and Prieto, M. 2002. On the diphthong/hiatus contrast in Spanish: some experimental results. *Linguistics* 40: 217–234.
- Hudson, R. 1980. *Sociolinguistics*, Cambridge: Cambridge University Press.
- Kemp, J.A. (ed.) 1972. *John Wallis’s Grammar of the English Language*, London: Longman.
- Kretzschmar, W. 2009. *The Linguistics of Speech*, Cambridge: Cambridge University Press.
- Labov, W. 1974. On the use of the present to explain the past. In: *Proceedings of the 11th International Congress of Linguists*, L. Heilmann (ed.), 825–851, Bologna: Il Mulino.
- Laver, J. 1994. *Principles of Phonetics*, Cambridge: Cambridge University Press.
- McMahon, A. 1994. *Understanding Language Change*, Cambridge: Cambridge University Press.
- Machan, T.W. 2003. *English in the Middle Ages*, Oxford: Oxford University Press.
- Maddieson, I. 1984. *Patterns of Sounds*, Cambridge: Cambridge University Press.
- Millar, R.M. 2012. *English Historical Sociolinguistics*, Edinburgh: Edinburgh University Press.
- Milroy, J. 1992. *Linguistic Variation and Change*, Oxford: Blackwell.
- Ogura, M. 1987. *Historical English Phonology: A Lexical Perspective*, Tokyo: Kenkyusha.
- Ogura, M. 1990. *Dynamic Dialectology: A Study of Language in Time and Space*. Tokyo: Kenkyusha.
- Ohala, J. 1993. The phonetics of sound change. In: *Historical Linguistics: Problems and Perspectives*, C. Jones (ed.), 237–278, London: Longman.
- Pinker, S. 1994. *The Language Instinct* Harmondsworth: Penguin.
- Romaine, S. 1982. *Sociohistorical Linguistics*, Cambridge: Cambridge University Press.

- Samuels, M.L. 1972. *Linguistic Evolution, with Special Reference to English*, Cambridge: Cambridge University Press.
- Smith, J.J. 2000. The letters *s* and *z* in South-Eastern Middle English. *Neuphilologische Mitteilungen* 101: 403–413.
- Smith, J.J. 2007. *Sound Change and the History of English*. Oxford: Oxford University Press.
- Stuart-Smith, J. 2004. *Phonetics and Philology*, Oxford: Oxford University Press.
- Wang, W.S.-Y. 1969. Competing changes as a cause of residue. *Language* 45: 9–5.
- Wells, J. 1982. *Accents of English*, Cambridge: Cambridge University Press.
- White, D. 2004. Why we should not believe in short diphthongs. In: *Studies in the History of the English Language II: Unfolding Conversations*, A. Curzan and K. Emmons (eds), 57–84, Berlin and New York: Mouton de Gruyter.