Variability, Language Change, and the History of English

JAMES MILROY
University of Michigan

ABSTRACT
In historical language scholarship, it has been usual to assume that the transmission of language from generation to generation is itself a linguistic, rather than a social, process, and that the focus should be on uniform language states. Here it is argued that transmission is necessarily social and that the history of a language is necessarily a history of variation. First, it is shown that the history of British Received Pronunciation is not one of direct descent from a single uniform ancestral variety. It is then demonstrated that pre-vocalic [h] and [hw] in English have a long history as variables and that loss of [h] in these combinations is not a recent event. Finally, it is suggested that closely similar variants of certain variables, such as [w] for (wh), have most probably recurred independently at various points in history and that we therefore need to review the methods used for dating sound changes.

KEYWORDS: transmission, variability, genetic linguistics, sound changes, continuity.

1. INTRODUCTION
In this paper I am concerned with the manner in which historical linguists and textual scholars have interpreted evidence from the past. My approach to this, however, is sociolinguistic. I am taking the view that social factors are necessarily involved in historical change: since a language
is passed down in the social and situational contexts in which speech events take place, the transmission of language must be a social process—that is to say that a language is transmitted from person to person, from group to group and from generation to generation — within social and situational contexts. Furthermore, if language is passed down in this social way, linguistic changes must be passed down in this social way also. Linguistic change takes place in the activities of human users of language in social and situational contexts and would not otherwise take place at all.

This is emphatically not how traditional language historians have normally treated the topic of language change. For most of these linguists, the chronological history of a language has been, and remains, a history of sounds, grammatical forms, and words, all of which are treated as though they have had an existence independent of society and speakers. That is to say that accounts of language change have normally been confined to language-internal description and language-internal explanation. As I have pointed out elsewhere (Milroy 2003a:148-153), the discourse of the subject has encouraged this also; in this discourse, it is speech sounds, grammatical forms and lexical items that change, and not speakers who bring about changes in sounds, grammar and lexicon; thus, one can speak of, for example, "general principles of vowel-shifting" (Labov 1994:115-154) as though the sounds themselves were endowed with the potential to shift and with the capacity to follow out 'general principles'. Thus, 'internal/external' has been seen as a dichotomy, with one side of the dichotomy favoured at the expense of the other and treated as exclusive of the other. I do not think that this can be an adequate basis for explaining how linguistic structures move from one state to the next.

One important characteristic of language in use has been neglected or under-represented in these language-internal accounts. This characteristic is variability, and in this paper I am crucially concerned with variability. In social and situational contexts, language is normally highly variable, and variation is distributed in several different social and contextual dimensions. This is what we find in synchronic studies of language in speech communities at the present day, and we must presume that similar kinds of variation have existed at all points in history. Thus, the social history of a language is also a history of variability, and it is this emphasis on society that has enabled us to incorporate variationist studies into historical language studies. The methods of conventional historical linguistics, however, have favoured categorical statements in which any variation encountered is stripped away. Descriptive statements of language changes are normally of the form: A (categorically) > B (categorically), and not of the form: A (variably) > B (variably). In this paper, I will attempt to show how recognizing the importance of variability can lead to new ways of interpreting evidence from the past.

The traditional internalist view is stated very clearly by Roger Lass (1997:324), who explicitly rejects a social or cognitive basis for the methodology (which of course would take the speaker to be central): the historian's approach, he says, should be 'structuralist' in that its basis should be "neither 'cognitive' nor 'social': communication and meaning [...] are not at the centre of change, or at least of major structural change".
To judge by their research methods, however, some scholars have not agreed in principle with this opinion. Indeed, it should be acknowledged here that some traditional language historians (e.g., Wyld 1927, 1936) were keenly aware of the importance of social factors, even though they generally lacked a systematic framework of sociolinguistic description. More recently, William Labov (1972:3), noting that scholars have tended to "explain linguistic events only by other linguistic events" has insisted that "... one cannot understand the development of language change apart from the social life of the community in which it occurs". Most historical sociolinguists now place a high value on evidence of social variation in earlier centuries. Thomason & Kaufman (1988:4) also have put the point rather clearly: they mention their "conviction that the history of a language is a function of the history of its speakers, and not an independent phenomenon that can be thoroughly studied without reference to the social context in which it is embedded".

These enlightened views, however, are not always realized in practice quite as fully as one might wish. Much of Labov's work, for example, appears to accept a sharp dichotomy in which the internal structure of language is methodologically separate from external factors: he gives much attention to general principles of change, independently of social factors (on the latter see: Labov 2001). Labov also states (1994:115) that his empirical approach has the same aims as the 'universalistic' approach. This conforms with what has been called autonomous linguistics, which is non-social, and it implies that, just as there are synchronic universals, so there are also diachronic universals of language independent of society, including, for example, the universals of chain-shifting, which Labov discusses at length. Of course, we are free to agree or to disagree with this, but it is not my purpose here to argue about it —I merely want to call attention to the fact that these approaches still seem in practice to be depending on the internal/external dichotomy and emphasizing the internal structure of language in explaining language changes.

Thomason & Kaufman, also, comment that their own treatment of the subject is not a sociolinguistic one. It certainly is not, and there appear to be inconsistencies or logical difficulties in some of the claims that are made by them, particularly in their intralinguistic definition of 'normal' and 'abnormal' transmission (which we have characterized above as a social process). Their definition (1988:10) depends ultimately on the genetic metaphor (the idea that languages are genetically related to one another), and this is certainly not a social approach. According to them, 'imperfect' or 'abnormal' transmission occurs "when the label 'genetic relationship' does not properly apply". However, if we accept that transmission is a function of speakers and social groups, and not inherently linguistic, it follows that normal and abnormal transmission must be defined in social or socio-political terms. I have discussed this point more fully elsewhere (Milroy 1997:317): here I am concerned only with the fact that Thomason & Kaufman's approach depends on language-internal criteria, viz., the genetic metaphor of language descent and, within that metaphor, the idea that normally transmitted languages are of single parentage. It is the idea of single parentage that we especially need
explore further.

In most historical linguistics, the idea that a language is descended (quasi-genetically) from a single ancestor continues to be taken for granted without any justification being required. This is perhaps the ultimate in non-social and non-variationist language-centred thinking, and it seems to me that it is sometimes applied in inappropriate ways; therefore, I now turn to some examples of what I think are inappropriate appeals to the 'single ancestor principle' (henceforth SAP). I will conclude the paper by considering some of the ways in which evidence of sociolinguistic variation in past centuries has tended to be explained away and expunged from the account.

II. 'GENETIC' LINGUISTICS AND THE PRINCIPLE OF SINGLE PARENTAGE: THE HISTORY OF ENGLISH RECEIVED PRONUNCIATION

Traditionally, the history of English has been presented as mainly single-stranded, as though the language since about 1550 had been a single dialect or variety transmitted in a continuous line, with some superficial influence from other varieties. Typically, this direct line is presented as the history of what is usually called 'Standard English'. Aside from the fact that the term 'standard' begs many questions that need to be answered in social or ideological terms, the history of the chosen variety itself is not usually presented in a realistic socio-historical context. Specifically, the assumptions made about the social history of this variety are also single-stranded. It is commonly presented as though it was handed down from generation to generation within a single social group — generally thought to be the upper classes — more or less as though it had existed in a social and linguistic vacuum in which there was little or no direct influence from other social groups or from other varieties of language. It does not appear to matter that such a view of society is extremely naive. Sometimes, other varieties have been recognized as having had a valid existence, but when they are so recognized, their features have commonly been devalued or dismissed as 'vulgar' or 'dialectal'. This is one of the ways in which the historical account is simplified. As an example, let us consider E. J. Dobson's account of the merger between /hw/ and /w/.

Dobson (1968:974) notes that simplification of [hw] to [w] occurred in Middle English (ME) in the south and midlands and states that this also had currency in 'vulgar London speech'. However, he adds that "[i]n [Early Modern English], educated speech appears invariably to have [hw] ... ; but during the eighteenth century the previously vulgar [w] became increasingly current in good speech ...". Thus, [w] for [hw] is not really admitted to the legitimate language until the eighteenth century, even though it certainly existed in ME. It is striking that the ME Bestiary (13th century), for example, has categorical w throughout, except for one instance of qu (not wh). Dobson does not explain why he thinks that the ME evidence can be dismissed (was it 'vulgar'?), or how he knows that the early London variety in which [hw] became [w] was a 'vulgar' variety: there must have been variability throughout these centuries. It is considered sufficient to assume that 'good' English was passed down in a single
line, with the ‘good’ speech of one generation being descended directly from the good speech of the same social class in the previous generation, except that after several centuries (for some unexplained reason) the originally ‘vulgar’ variant was finally admitted to ‘good’ speech. It is not considered necessary to justify these assumptions, even though the evidence (in the form of orthoepic descriptions and occasional spellings) is not conclusive. In the background of all this, the SAP is taken for granted.

Now let us consider what difference it makes when we take the fact of variability into account. As language is always variable, it is clear that different variants of the same phoneme (or morpheme or syntactic construction) can co-exist in a single variety at a given time. Thus, [w] and [hw] could both have existed in either ‘vulgar’ speech or educated speech, or both, at any time. The genuine social history of a language involves variation in many different dimensions at all times, and it cannot possibly be based solely on the history of a single group. Although this point is valid for all levels of language — phonology, grammar and lexicon — it is particularly important in the case of pronunciation, as pronunciation is the level that is the most likely to be highly variable. However, phonological history is conventionally presented as the history of a single variety: from EModE onward the main interest has been in the development of the ‘Received Pronunciation’ of British English (RP), which is usually thought to be the same thing as ‘Standard English’ pronunciation.

RP is normally presented as having a single ancestor, having been handed down in a straight line from the usage of the Elizabethan Royal Court — the language of the upper classes. A.C. Gimson (1970:84-5) puts it this way:

The speech of the Court[...], phonetically largely that of the London area, increasingly acquired a prestige value and in time, lost some of the local characteristics of London speech. It may be said to have been finally fixed, as the speech of the ruling class, through the conformist influence of the public schools in the nineteenth century.

John Honey (1989:15) comments that “[i]t is crucial to realize that the direct ancestor of British English’s present-day standard accent (RP) was not simply a particular regional one; it was also the property of a limited social group within that region”. Although Honey is right about the limited currency of high-status speech, it is not at all clear that he is justified in postulating a single ancestry for modern RP, or in assuming that the single ancestor was necessarily a high-status variety. As for Gimson’s claims about the origin in courtly language — these seem to be based on speculation. We do not know whether the alleged pre-standard variety actually ‘lost’ some of the characteristics of London speech (as he alleges), or whether common London characteristics were developed independently of the high-status variety. Some recent research (e.g., Mugglestone 1995:194-199) has shown that some modern characteristics of RP (such as the back [a] in e.g. fast) could actually be stigmatized until little over a century ago. Thus, the true history of RP is certainly much more complicated than the accounts given by Gimson and Honey.
From a sociolinguistic point of view a unilinear history is intrinsically unlikely: varieties of language are not uniform states, and they are not insulated in airtight containers. In fact, the evidence that RP is a direct descendant of Elizabethan courtly language is quite flimsy. The favourite citation of language historians is a statement attributed to George Puttenham, presumed author of the Arte of Poesie (1589) (see, e.g., Gorlach (1999:483-486), Wyld (1936:10)), who defines the best speech as "... the usuall speech of the Court, and that of London and the shires lying about London within lx myles, and not much aboue". There are comments about 'true' pronunciation from quite early in the century, and this was associated with the Court and restricted to southern England. Puttenham goes on to comment that the usage of "Northern-men, whether they be noblemen or gentlemen [...] is not so Courtly nor so currant as our Southerne English is". This 'true' pronunciation seems to have had prestige in part of the country, it is probably not correct to label it at this stage as a 'standard'.

The greatest weakness in the argument, however, is that Puttenham's comments do not in themselves demonstrate the historical continuity that is claimed: they apply only to a given time in history. Continuity in a straight line for three centuries is an assumption, not a fact, and it is based entirely on inferences drawn selectively from occasional comments by writers, dominated by the genetic metaphor and driven by the SAP —the assumption that RP must necessarily have a single ancestor and that this ancestor can be identified — and, of course, also influenced by unacknowledged social-class bias. None of these inferences actually demonstrates the single-stranded continuity that is claimed. It can reasonably be assumed that RP, which arose in the nineteenth century, was influenced by other varieties, some of them used by low-status speakers and that the courtly language of Elizabethan times could have bequeathed some of its features to low prestige varieties rather than directly to RP. It is now a truism that high prestige features can lose prestige over time and that low-prestige features can be elevated to higher prestige. RP as a focussed variety is associated with nineteenth-century changes in the power-structures of Britain, rather than with the Elizabethan Court, and it is in that socio-political dimension that we need to investigate its origins.

III. THE PERMANENCE OF VARIABILITY AND THE DATING OF CHANGES

Here I propose to formulate an important question that has been implicit in our discussion so far: What difference does it make to our historical accounts if, instead of assuming single ancestry and single-stranded continua, we give full recognition to the fact of variability at all times? A sound change has traditionally been assumed to be a single event that takes place at a particular time in a language that is envisaged as a uniform state phenomenon. Sometimes, as we have seen, evidence that the change took place early in some dialects and later in others is discounted on the grounds that the dialects with the early change are 'vulgar' or non-standard. In such accounts, the evidence is recognized as indicating a change only at the point at which the new form enters the standard, and earlier evidence from other dialects is argued away. Sometimes,
However, 'colloquialisms' or 'vulgarisms' are considered to be of interest, but when this is so, they too tend to be seen as originating in single events at particular times, and evidence that the 'new' pronunciations might have been around for centuries before that time is discounted. Essentially, strong evidence that the change had already happened at some particular date is seen as a terminus ante quem non: it could not have happened significantly earlier, and any evidence that it might have happened earlier must be argued away as unreliable. Although this kind of reasoning is particularly characteristic of older scholars such as Wyld and Jespersen, it is still in use, as in the following (on the change from [hw] to [w]):

There is sporadic /x/-loss in ME, but spellings like with for which, etc. are rare before the sixteenth century, and then common only in prosodically weak words. The first good evidence for general loss appears to be Jones (1701: 118); what, when etc. sounded wai, wen, etc. by some.

Roger Lass (1999:123-4)

Here, 1701 (or, presumably, slightly earlier) is the terminus ante quem non. Occurrences in ME are said to be rare and largely in unstressed words; furthermore, no reason is given for first mentioning the sixteenth century and then apparently rejecting sixteenth-century evidence.

The kind of reasoning used here is familiar: yet, it is not logically watertight. Let me explain. If the change is attested in 1701, it does not logically follow that it might not have taken place much earlier — perhaps even centuries earlier: if it is not attested in earlier centuries, that is not in itself proof that it did not occur during those centuries. Lass's view seems to depend on his own predilection for accepting late dating of sound changes (a predilection he comments on (1997:289)), even though there is likely to be a time-lag between the implementation of a change and the representation of that change in writing. However, in the case of (wh) and other variables, the argument is not merely an argument ex silentio: there actually is spelling evidence from those earlier centuries, and this is not reasonably to be described as 'rare'. We have already noted that in the ME Bestiary, w is effectively categorical: i.e., in a work of more than 800 lines there is not a single instance of wh. The commonest English words with wh are the WH pronouns; thus, it is possible for Lass to suggest that these may be 'prosodically weak'. However, in The Bestiary, w for wh occurs also in stressed nouns (wete ‘wheat’—line 190; wile ‘while’, i.e., ‘time’ — line 200), in other parts of speech, and frequently in places where the metre indicates that the WH-words are stressed. Furthermore, w for wh is not 'rare' in ME: it is common in a number of other texts — four to five centuries before Lass's 'first good evidence'. The most reasonable interpretation of these facts is that [w] for [hw] was a variant which had come into use in some places, but possibly not in others, and that it may well have been categorical — a completed change — in some varieties of ME. I am inclined to the view that it was quite well established in parts of the South-East and much of the East Midlands of England, and recent work by Minkova (2004), who has cited many instances of hw alliterating with w in Old English, strongly supports an early dating. As Minkova shows, [w] for [hw] was probably a variant in OE.
My earliest interest in variation studies was triggered by ME spelling: it seemed to me that spelling variation in some ME texts was not simply a result of scribal indiscipline, but that it might often bear an orderly relation to the phonology, and, further, that it could be interpreted as indicating a long history for some present-day pronunciations that are not generally considered 'standard'. In some cases it might even indicate an earlier dating than is generally accepted for sound changes that have since been adopted in 'standard' English (an example is loss of the velar fricative in words of the type right, ought, which may have been quite advanced by around 1300 in some East Midland locations (see Milroy 1992: 134-136)). What was particularly interesting, however, is the kind of reasoning used by textual scholars in order to accept or reject particular readings—especially those that might indicate the progress of sound changes. In particular, the history of initial [h] seemed to stand out. Although [h] before vowels was quite unstable in many ME texts, sometimes being omitted and sometimes added 'unhistorically' or 'inorganically' (the choice of words is interesting), scholars did not accept that pre-vocalic initial [h] was lost in ME. They normally considered that what was known as 'the present-day vulgarism' (Wyld 1936:296) first occurred in the late eighteenth century. I have discussed the history of initial [h] very fully elsewhere and have commented also on final-stop deletion and, recently (Milroy 2003b), the history of (th) fronting (as indicated by early spellings of the type erffor 'earth'). Essentially, the story is that scholars have diligently searched for reasons to reject the evidence that would give such variants time-depth, the favourite argument being that ME scribes were Anglo-Normans with a poor command of English. However, this is by no means the only argument used, and I would like to conclude this paper by discussing certain other types of argumentation that have been used to exclude evidence for variability that occurs in the texts.

There is evidence that 'aitch-dropping' might have occurred in OE also. There is an excellent study by Scragg (1970), which deliberately sets out to find reasons why instability of [h] in OE spelling and alliterative practice might not be reliable evidence of (variable) [h]-deletion. The arguments that Scragg uses are aimed at disproof: it is as though /h/ is on trial, being required to prove that it really was omitted, the default assumption being that it was not omitted. Actually, we do not know whether it was or was not omitted until we have considered the evidence. Scragg cites MS forms in the Poetic Codices in which pre-vocalic [h] allirates with a vowel; however, he is prepared to reject these as evidence because "ernendations for all these lines have been suggested by editors", and because "corruptions [...] can occur in the course of transmission of a poem" (1970:173). That is to say that alliterations that are clearly indicated in the text cannot be relied on, even though they must have been acceptable to the copyists who were happy to alliterate [h] with vowels. The modern editors are assumed to know better than the medieval scribes did, and in a world of corruption, the editors' emendations themselves are not seen as possible corruptions of the text, even though, in a sense, they are. As for spelling variation, one of the reasons given for rejecting instances of addition of 'unhistoric' [h] is the phenomenon of 'dittography': the scribe mistakenly adds initial [h] because a prominent word closely preceding begins with [h]. Yet, this kind of explanation is no more likely to be correct than
the much more obvious explanation—i.e., that $h$ was added hypercorrectly: if initial prevocalic [h] was not pronounced, then the scribes may not have been sure whether the letter was required by the spelling conventions or not. A third argument is that instability of $h$ is due to the fact that it was also unstable in Latin texts; yet, even if Latin conventions had an effect, it actually does not follow that [h] was not also unstable in OE. It might have been. Thus, having used these and other arguments to disqualify a large number of relevant instances, Scragg (1970:192-195) concludes that there are very few instances that “cannot be explained in scribal, as opposed to phonological, terms”. In this way most of the prima facie evidence for [h]-loss in OE is explained away.

4. CONCLUSIONS
In this paper I have chosen to look at some fairly traditional matters. I am not suggesting that every modern historical descriptivist has been dominated by a non-variationist approach to history with exclusive emphasis on a single variety descended from a single ancestor. There are now many exceptions to this, including, for example, the work of Nevalainen and Raumolin-Brumberg and many of the contributions to the *Cambridge History of the English Language*. However, it seems that I am not (yet) flogging a dead horse. There are influential voices that insist on a sharp dichotomy between internal and external approaches and on the view that historical linguistics must concentrate on the former.

As for the variables discussed in this paper, I would like to conclude with a general observation, which is that closely similar variants of certain variables (such as [w] for (wh), [ʃ, v] for (th) and [ɔ] for (h)) have most probably appeared independently at many times throughout history, sometimes merely sporadically, sometimes being adopted by groups or communities of speakers and sometimes not, sometimes diffusing widely and sometimes not, sometimes advancing and sometimes retreating. As I have implied in my discussion of (th) (Milroy 2003b:218), the variants of the variables discussed in this paper represent what we might regard as ‘natural’ or ‘easy’ changes, which can occur at any time. Some of these may have been completed in some varieties in early English and may have historical continuity from that time onward. In other cases, this may not be so. What I have tried to show in this paper is that we still need to re-consider much of the traditional dating of sound changes in English, and, more generally, examine very critically the kind of reasoning that is used in arguments about sound change in history.

NOTES:
1. The SAP is a variant of what Jonathan Hope (2000:49) has called the ‘Single ancestor dialect’ principle (SAD). My examples in this paper refer mostly to dialects. I have preferred a more generalized label, as the principle originates in Stammbaumtheorie, which has been traditionally applied to whole languages.
2. This occurs in the word *qual* (‘whale’), which is a common northern ME spelling (probably taken over from the scribe’s exemplar). The East Anglian dialect of the *Bestiary* plainly had [w] for [wh].

3. I believe that the language of the court was probably recessive, much as conservative RP is recessive today, and that it bequeathed few, if any, phonological variants to any subsequent variety of the language (see Milroy 2001:27).

4. This is a simple insight, but traditional historians seem to have been totally unaware of it.

5. *w* for *hw* is admittedly uncommon in a number of texts, but it is common in, for example, the second continuation of the *Peterborough Chronicle*, in the Otho MS of Layamon’s *Brut* and in the *Caligula* MS of The *Owl and the Nightingale*. It is occasional in many other texts—and back-spellings (*wh* for *w*) also occur.

REFERENCES:


Variability, Language Change and the History of English

Milroy, J. (2001). Received Pronunciation: Who receives it and how long will it be 'received'? Studia Anglica Posnaniensia 36, 15-33.


